

ulm university universität ulm

Non-local Bakry-Émery theory and related functional inequalities

Dissertation zur Erlangung des Doktorgrades Dr. rer. nat. der Fakultät für Mathematik und Wirtschaftswissenschaften der Universität Ulm

Vorgelegt von Frederic Weber

geboren in Berlin

- April 2021 -

Amtierender Dekan: Prof. Dr. Martin Müller

Erster Gutachter: Prof. Dr. Rico Zacher

Zweiter Gutachter: Prof. Dr. Anna Dall'Acqua

Dritter Gutachter: Prof. Dr. Ansgar Jüngel

Tag der Promotion: 21.09.2021

Abstract

This cumulative thesis establishes a non-local variant of the Bakry-Émery Γ-calculus and investigates related functional inequalities and their consequences. We focus on non-local operators like the fractional Laplacian or, in the space-discrete setting, on generators of continuous-time Markov chains.

In the first article we define the condition $CD_\Upsilon(\kappa,\infty)$, $\kappa \in \mathbb{R}$, by identifying the respective substitutes for the carré du champ operator and the iterated carré du champ operator in the first and second derivative of the entropy along the heat flow. As a consequence, $CD_\Upsilon(\kappa,\infty)$ with $\kappa > 0$ implies the modified logarithmic Sobolev inequality by means of the entropy method, which is consistent to the role of $CD(\kappa,\infty)$ in the classical Bakry-Émery theory. We further derive fundamental properties of $CD_\Upsilon(\kappa,\infty)$ like tensorization principles for two non-local operators, but also for the sum of a non-local operator with an operator that fits to the classical framework of the Bakry-Émery theory. The latter can be applied to study the long-time behavior of linear reaction-drift-diffusion systems.

Moreover, we compare our approach with several alternative notions of synthetic curvature. Most notably, there exist strong links to other curvature-dimension conditions resulting in the fact that CD_Υ with non-negative curvature together with a finite dimension term implies Li-Yau type inequalities. Hence, the CD_Υ condition serves as the first curvature-dimension inequality in the non-local framework that is suitable to imply both, modified logarithmic Sobolev inequalities under positive curvature bounds and Li-Yau type estimates under finite dimension bounds and non-negative curvature.

In the second article we get more precise about the dimension parameter in the CD_Υ condition. We express the dimension part by so called CD-functions and introduce the condition $CD_\Upsilon(\kappa,F)$, where $\kappa \in \mathbb{R}$ and F is a CD-function. In the special case of $F(r) = \frac{r^2}{d}$, $d \geqslant 1$, we use the notation $CD_\Upsilon(\kappa,d)$ motivated by the classical case. By a refinement of the entropy method we can show that $CD_\Upsilon(\kappa,F)$ implies for $\kappa > 0$ and suitable F a functional inequality comparing the entropy to the Fisher information which we call, accordingly, entropy-information inequality. We study properties that are known consequences of entropy-energy inequalities in the classical framework. More precisely, we show finite diameter bounds, ultracontractive estimates and a modified version of the Nash inequality for a large class of entropy-information inequalities.

The third article focuses on long-range jump operators on the one-dimensional lattice $\mathbb{Z}$. We show that the Bakry-Émery condition $CD(0,d)$, which is necessary for the validity of $CD_\Upsilon(0,d)$, holds true whenever the associated kernel to the Markov generator L has finite second moment. Contrariwise, we are able to prove that $CD(0,d)$ is impossible to hold (with finite $d < \infty$) for the class of power type-kernels with correspondingly small powers. As a consequence, we show that fractional powers of the discrete Laplacian fail to satisfy $CD(0,d)$.

Following that, we extend in the fourth article the latter result to the space-continuous setting and show by a rather delicate approximation procedure that even $CD(\kappa,d)$ for $\kappa \in \mathbb{R}$ and $d < \infty$ is impossible to hold for the fractional Laplacian on $\mathbb{R}^d$.

Driven by the failure of the Bakry-Émery condition, we pursue in our fifth article an alternative approach to solve the long-standing open question whether positive solutions to the fractional heat equation satisfy a Li-Yau inequality. We develop an abstract method that is based on heat kernel representations for positive solutions to the corresponding non-local diffusion problem. With this method at hand, we are able to answer the above question in the affirmative by deriving a Li-Yau inequality for positive solutions to the fractional heat equation whose time-behavior is like in the Euclidean case. Further, our method also applies to the discrete setting. We demonstrate the latter by proving optimal Li-Yau estimates for diffusion on a complete graph.

Kurzzusammenfassung in deutscher Sprache

Diese kumulative Dissertation führt eine nicht-lokale Variante des Bakry-Émery Γ-Kalküls ein und untersucht zugehörige Funktionalungleichungen sowie deren Konsequenzen. Wir fokussieren uns hierbei auf nicht-lokale Operatoren wie den fraktionellen Laplace Operator und, im raumdiskreten Fall, auf Generatoren von zeit-stetigen Markov Ketten.

Im ersten Artikel definieren wir die Bedingung $CD_\Upsilon(\kappa,\infty)$, $\kappa \in \mathbb{R}$, indem wir jeweils einen geeigneten Ersatz für den carré du champ Operator und den iterierten carré du champ Operator in der ersten und zweiten Ableitung der Entropie entlang des Wärmeleitungsflusses finden. Als Konsequenz erhalten wir, dass die Bedingung $CD_\Upsilon(\kappa,\infty)$ für $\kappa > 0$ die modifizierte logarithmische Sobolev Ungleichung impliziert und dies, so wie es von der Bakry-Émery Bedingung $CD(\kappa,\infty)$ in der klassischen Situation bekannt ist, unter Verwendung der Entropie Methode. Außerdem leiten wir grundlegende Eigenschaften der Bedingung $CD_\Upsilon(\kappa,\infty)$ her, wie zum Beispiel Tensorisierungsprinzipien für zwei nicht-lokale Operatoren, oder aber für die Summe eines nicht-lokalen Operators mit einem zu der klassischen Bakry-Émery Theorie passenden Operator. Der zuletzt genannte Fall wird etwa angewandt um Aussagen über das Langzeitverhalten von Systemen linearer Reaktions-Drift-Diffusionsgleichungen zu treffen.

Zudem vergleichen wir unseren Ansatz auch mit einigen bereits existierenden Arbeiten, in welchen ebenfalls ein nicht-lokaler Ersatz für Krümmmung und Dimension aufgestellt worden ist. Insbesondere existieren starke Verbindungen zu anderen Krümmungs-Dimensions Ungleichungen in der nicht-lokalen Situation, welche dazu führen, dass CD_Υ im Falle nichtnegativer Krümmung und eines endlichen Dimensionsterm auch Li-Yau Ungleichungen impliziert. Somit stellt die CD_Υ Bedingung die erste Krümmungs-Dimensions Ungleichung in der nicht-lokalen Situation dar, welche sowohl dazu geeignet ist modifizierte logarithmische Sobolev Ungleichungen im Falle positiver Krümmung, als auch Li-Yau Ungleichungen im Falle eines endlichen Dimensionsterms und nicht-negativer Krümmung zur Folge zu haben.

Im zweiten Artikel präzisieren wir den Dimensionsparameter in der CD_Υ Bedingung. Genauer drücken wir diesen mithilfe sogenannter CD-Funktionen aus und führen die Bedingung $CD_\Upsilon(\kappa,F)$ ein, wobei $\kappa \in \mathbb{R}$ und F eine CD-Funktion ist. Motiviert durch die klassische Theorie nutzen wir im konkreten Fall von $F(r) = \frac{r^2}{d}$, $d \geqslant 1$, die Notation $CD_\Upsilon(\kappa,d)$. Mittels einer Verfeinerung der Entropie Methode können wir beweisen, dass $CD_\Upsilon(\kappa,F)$ mit $\kappa > 0$ und geeignet gewählter CD-Funktion F eine Funktionalungleichung zur Folge hat, welche die Entropie mit der Fisher Information vergleicht und welche wir dementsprechend Entropie-Informations-Ungleichung taufen. Wir untersuchen Eigenschaften, welche in der klassischen Situation als Konsequenzen sogenannter Entropie-Energie Ungleichung bekannt sind. Genauer gesagt zeigen wir, dass endliche Schranken an den Durchmesser, ultrakontraktive Abschätzungen und eine modifizierte Version der Nash Ungleichung für eine große Klasse von Entropie-Informations-Ungleichungen gültig sind.

Unser dritter Artikel fokussiert sich auf Operatoren, welche beliebig weite Sprünge auf dem Zustandsraum der ganzen Zahlen zulassen. Wir zeigen, dass die Bakry-Émery Bedingung $CD(0,d)$, welche notwendig für die Gültigkeit von $CD_\Upsilon(0,d)$ ist, gilt, falls der zum Operator gehörende Kern zweites endliches Moment besitzt. Andererseits können wir auch beweisen, dass $CD(0,d)$ für kein $d < \infty$ gültig sein kann im Falle von Potenzkernen mit entsprechend klein gewählter Potenz. Aus dieser Beobachtung ergibt sich zudem, dass fraktionelle Potenzen des diskreten Laplace Operators ebenfalls $CD(0,d)$ für kein $d < \infty$ erfüllen können.

Basierend auf den zuletzt beschriebenen Erkenntnissen, untersuchen wir in unserem vierten Artikel den fraktionellen Laplace Operator auf $\mathbb{R}^d$. Wir zeigen unter Verwendung eines technischen Approximationsargument, dass in diesem Fall $CD(\kappa,d)$ für kein $\kappa \in \mathbb{R}$ und $d < \infty$ erfüllt sein kann.

Motiviert durch das Fehlschlagen der Bakry-Émery Bedingung im Falle des fraktionellen Laplace Operators folgen wir in unserem fünften Artikel einem alternativen Ansatz um der be-

deutenden offenen Frage nachzugehen, ob positive Lösungen der fraktionellen Wärmeleitungsgleichung eine Li-Yau Ungleichung erfüllen. Wir entwickeln eine abstrakte Methode, welche auf Darstellungsformeln mittels der Fundamentallösung für positive Lösungen der zugehörigen Diffusionsgleichung beruht. Mithilfe der genannten Methode können wir die zuvor erwähnte offene Frage positiv beantworten, indem wir eine Li-Yau Ungleichung für positive Lösungen der fraktionellen Wärmeleitungsgleichung herleiten, deren Verhalten in der Zeit mit dem aus der euklidischen Situation bekannten Verhalten übereinstimmt. Außerdem demonstrieren wir, dass unsere abstrakte Methode auch Anwendung im raum-diskreten Fall findet, indem wir diese anwenden um eine optimale Li-Yau Ungleichung für Diffusion auf vollständigen Graphen zu zeigen.

Acknowledgements

It is my pleasure to express gratitude to the people that have accompanied and supported me during the time of my PhD studies.

First and foremost, I want to thank my supervisor Prof. Dr. Rico Zacher, who has guided me through all my mathematical works, starting with my bachelor thesis in 2016. The way I work mathematically is significantly influenced by his constant support and guidance over the years. In particular, I am thankful for the fact that I was directly involved in research projects from the beginning of my PhD, which served as the optimal start for me. At this point, I would also like to thank Dr. Adrian Spener, with whom the three of us tackled our research projects along many enduring discussions in the first year of my PhD studies. These experiences laid the basis for fostering my research in the sequel with great enthusiasm.

The formidable work environment at the Institute of Applied Analysis clearly also played a crucial role and therefore I want to thank all my colleagues at the institute.

In particular, I thank Prof. Dr. Anna Dall'Acqua for accepting to be the second referee of this thesis. For me this feels like coming full circle, since she has been the first to inspire me for Analysis as my Analysis I lecturer many years before.

Furthermore, I want to express my gratitude to Lukas Niebel and Anja Schmidt for proofreading my thesis very carefully and all their valuable suggestions.

I feel also highly grateful to the Studienstiftung des deutschen Volkes for their backing for so many years, first in form of a student scholarship and later by means of a PhD scholarship. Besides the financial support I am also thankful for the opportunity to meet interesting individuals and broaden my horizon apart from my field of studies.

Above all I want to thank my family, my friends and my girlfriend for their enduring support during my time in Ulm and the fact that I can truly rely on them also in the future, wherever that may lead me.

Contents

1. Introduction **3**
1.1. The classical Bakry-Émery theory for Markov diffusion operators 3
1.2. Non-local operators 11
1.3. Towards a non-local Bakry-Émery theory 13
1.4. Summary 17

2. The entropy method under curvature-dimension conditions in the spirit of Bakry-Émery in the discrete setting of Markov chains **29**

3. Entropy-information inequalities under curvature-dimension conditions for continuous-time Markov chains **91**

4. Curvature-dimension inequalities for non-local operators in the discrete setting **123**

5. The fractional Laplacian has infinite dimension **155**

6. Li-Yau inequalities for non-local diffusion equations via reduction to the heat kernel **171**

A. Formalities **187**

1. Introduction

Markov semigroups serve as an important object in analysis and probability theory. Understanding their long-time behavior is a fundamental aspect that can be investigated, for instance, by so called convex Sobolev inequalities. In the classical Bakry-Émery theory one focuses on Markov diffusion operators as generators of Markov semigroups. If a Markov diffusion operator satisfies certain curvature-dimension conditions, then various functional inequalities can be derived, like, e.g., convex Sobolev inequalities. Markov diffusion operators include a large class of differential operators. Since taking derivatives is a property that depends only on arbitrary small neighborhoods of the point of consideration, differential operators are sometimes called local operators. This has to be understood in contrast to so called non-local operators, where the evaluation at some point does also depend on the behavior of the function farther away. Those non-local operators, however, do not fit into the framework of the classical Bakry-Émery theory, which is, most notably, due to the lack of corresponding chain rules. Nevertheless, very important models among many disciplines like biology, finance or physics rely on non-local operators such as, for instance, the fractional Laplacian or, in the space-discrete setting, generators of continuous-time Markov chains. At the same time functional inequalities and their consequences have been studied a lot also for such operators since they induce similar strong properties as in the local case.

Therefore, it is natural to ask for a non-local version of the Bakry-Émery theory. This thesis is mainly dedicated to this aim. The remaining part of this introductory chapter is organized as follows. In Section 1.1 we introduce the classical Γ-calculus of Bakry and Émery. In particular, we motivate and define the original curvature-dimension condition $CD(\kappa, n)$ and present functional inequalities as its main application. Further, we demonstrate consequences of the respective functional inequalities and, in case of Li-Yau inequalities, provide an alternative approach for proving them.

Section 1.2 is dedicated to the operators of our interest, i.e. operators that are non-local and hence do not fit into the framework of Section 1.1. As important special cases, we present the fractional Laplacian and also generators of continuous-time Markov chains in more detail.

Subsequently, in Section 1.3, we explain how to combine the content of the previous two sections into one new framework. In the first part of Section 1.3 we present existing approaches for finding substitutes for CD inequalities, but also, more generally, for Ricci-curvature in the non-local setting. In the second part of Section 1.3, we present the main contributions of this thesis.

We finalize this chapter with Section 1.4 where we summarize each of our five research articles [31], [32], [34], [35] and [36].

1.1. The classical Bakry-Émery theory for Markov diffusion operators

The Γ-calculus of Bakry and Émery dates back to their seminal work [3]. The content of this section mainly follows the extensive monograph [4]. We note that for the sake of simplicity our description remains on a formal level, in particular, taking into account the respective assumptions on the function spaces underlying the theory.

Given a σ-finite measure space $\big(X, \Sigma, \mu\big)$, we consider a suitable algebra of real-valued bounded functions $\mathcal{A}_0$ which is dense in $L^p(\mu)$, for any $p \in [1, \infty)$. Then one introduces the so called carré

du champ operator as a symmetric, bilinear map $\Gamma : \mathcal{A}_0 \times \mathcal{A}_0 \to \mathcal{A}_0$ which is positive in the sense that $\Gamma(f) := \Gamma(f,f) \geqslant 0$ for any $f \in \mathcal{A}_0$ and satisfies the so called diffusion property

$$\Gamma\big(H(f_1,...,f_k),g\big) = \sum_{i=1}^{k} \partial_i H(f_1,...,f_k)\,\Gamma(f_i,g) \tag{1.1}$$

for any smooth function $H : \mathbb{R}^k \to \mathbb{R}$ with $H(0) = 0$ and $f_1,...,f_k,g \in \mathcal{A}_0$. The operator $L : \mathcal{A}_0 \to \mathcal{A}_0$ is then defined as the (unique) operator satisfying the integration by parts formula

$$\int_X fLg\,\mathrm{d}\mu = -\mathcal{E}(f,g),$$

where we write $\mathcal{E}(f,g) = \int_X \Gamma(f,g)\mathrm{d}\mu$. In fact, further assumptions on $\mathcal{A}_0$ guarantee that L is well-defined. The symmetric form $\mathcal{E}$ can be extended to a Dirichlet form whose generator is an extension of L (we denote the extensions also by $\mathcal{E}$ resp. L in the sequel). Consequently, L generates a symmetric submarkov semigroup $\big(P_t\big)_{t\geqslant 0}$ on $L^2(\mu)$ (see e.g. [10]). Further, μ serves as an invariant and reversible measure for $\big(P_t\big)_{t\geqslant 0}$.

One frequently calls L a Markov diffusion operator. This is due to the fact that L is the generator of a Markov process $\big(Z_t\big)_{t\geqslant 0}$ whose state space is given by X and the semigroup that is represented on suitable functions by

$$P_t f(x) = \mathbb{E}\big(f(Z_t)|Z_0 = x\big), \tag{1.2}$$

is typically assumed to be a Markov semigroup (i.e. we also have $P_t\mathbb{1} = \mathbb{1}$, for any $t > 0$) and, moreover, that the diffusion property (1.1) implies also a chain rule formula for the generator L, which reads as

$$L\big(H(f_1,...,f_k)\big) = \sum_{i=1}^{k} \partial_i H(f_1,...,f_k)\,Lf_i + \sum_{i,j=1}^{k} \partial^2_{ij} H(f_1,...,f_k)\,\Gamma(f_i,f_j), \tag{1.3}$$

for any smooth $H : \mathbb{R}^k \to \mathbb{R}$ with $H(0) = 0$ and $f_1,...,f_k \in \mathcal{A}_0$. Note that the assumption $H(0) = 0$ is to ensure that H leaves $\mathcal{A}_0$ invariant. Later, we will extend the operators L and Γ to a larger algebra where the chain rules (1.1) and (1.3) also hold without this technical assumption.

Choosing $H(r,s) = rs$ in (1.3), where $r,s \in \mathbb{R}$, we get the important relation

$$\Gamma(f,g) = \frac{1}{2}\big(L(fg) - f\,Lg - g\,Lf\big). \tag{1.4}$$

The one-dimensional version of (1.3) is of such importance that it is worth to mention it separately

$$LH(f) = H'(f)Lf + H''(f)\Gamma(f). \tag{1.5}$$

Applying (1.1) in the case of $H = \log$, the latter translates to

$$L\log(f) = \frac{Lf}{f} + \Gamma(\log(f)). \tag{1.6}$$

In order to illustrate these abstract objects, we consider the example of the Laplace-Beltrami operator on a Riemannian manifold M. Taking $\mathcal{A}_0 = C_c^\infty(M)$ (the space of compactly supported

and smooth functions on M) and $\Gamma(f,g) = \langle \nabla f, \nabla g \rangle$, it is well known that

$$\int_M f \Delta g \, \mathrm{d}\mu = -\int_M \Gamma(f,g) \, \mathrm{d}\mu, \tag{1.7}$$

where Δ denotes the corresponding Laplace-Beltrami operator on M and μ the canonical Riemannian measure. Clearly, the Laplace-Beltrami operator and also Γ can be defined on a larger class of functions. We extend these operators to the space of smooth functions $C^\infty(M)$.

The celebrated Bochner-Lichnerowicz formula states that

$$\frac{1}{2}\Delta(|\nabla f|^2) = \langle \nabla f, \nabla \Delta f \rangle + \|\nabla^2 f\|_{HS}^2 + \mathrm{Ric}(\nabla f, \nabla f), \quad f \in C^\infty(M), \tag{1.8}$$

where $\|\nabla^2 f\|_{HS}$ denotes the Hilbert-Schmidt norm of the Hessian of f and $\mathrm{Ric}(\cdot,\cdot)$ the Ricci-curvature tensor. This identity motivates to consider an iteration of formula (1.4).

Indeed, returning to the abstract framework one extends Γ and L to a larger algebra of functions $\mathcal{A}$ and defines the iterated carré du champ operator as

$$\Gamma_2(f,g) = \frac{1}{2}\big(L\Gamma(f,g) - \Gamma(f, Lg) - \Gamma(g, Lf)\big), \tag{1.9}$$

where $f, g \in \mathcal{A}$. We write $\Gamma_2(f) := \Gamma_2(f,f)$ and say that L satisfies the curvature-dimension condition $CD(\kappa, n)$ for $\kappa \in \mathbb{R}$ and $n \in [1, \infty]$, if

$$\Gamma_2(f) \geqslant \kappa \Gamma(f) + \frac{1}{n}\big(Lf\big)^2, \ \mu\text{-a.e.}, \tag{1.10}$$

holds for any $f \in \mathcal{A}$.

In case of the Laplace-Beltrami operator, the Bochner-Lichnerowicz formula (1.8) yields that $\Gamma_2(f) = \|\nabla^2 f\|_{HS}^2 + \mathrm{Ric}(\nabla f, \nabla f)$, where $f \in C^\infty(M)$. From this, one can now see that the Laplace-Beltrami operator satisfies $CD(\kappa, n)$ if and only if the Ricci-curvature of M is bounded from below by κ and $\dim M \leqslant n$, where $\dim M$ denotes the topological dimension of M. This justifies to call $CD(\kappa, n)$ a curvature-dimension condition.

The above abstract framework applies, more generally, to the case when L is a second order differential operator without zero-th order term and with sufficiently smooth coefficients. For instance, let us consider the generalized Ornstein-Uhlenbeck operator on $\mathbb{R}^d$, which is given by

$$Lf = \Delta f - \langle \nabla V, \nabla f \rangle$$

for $f \in C^\infty(\mathbb{R}^d)$ and where $V \in C^\infty(\mathbb{R}^d)$ denotes the so called confining potential. In this case, the carré du champ operator reads as $\Gamma(f,g) = \langle \nabla f, \nabla g \rangle$, $f, g \in C^\infty(\mathbb{R}^d)$, and the invariant and reversible measure is given by $\mu(\mathrm{d}x) = e^{-V(x)}\mathrm{d}x$ up to a normalizing constant. Further, one calculates

$$\Gamma_2(f) = \|\nabla^2 f\|_{HS} + \langle \nabla f, \nabla^2 V \nabla f \rangle,$$

which shows that $CD(\rho, \infty)$ holds true for some $\rho > 0$ if $\nabla^2 V \geqslant \rho$ in the sense of positive-definite matrices. The latter condition is commonly called the Bakry-Émery condition in the theory of Fokker-Planck equations, where it serves as a key assumption in order to show convergence with an exponential rate of non-negative solutions to the steady state with an exponential rate via the entropy method, see [2, 9, 21].

The most standard example among the class of generalized Ornstein-Uhlenbeck operators is given by the choice $V(x) = \frac{|x|^2}{2}$, in which case μ coincides with the standard gaussian measure. While $CD(1, \infty)$ holds, one can show that this example cannot satisfy $CD(\rho, n)$ for any $\rho \in \mathbb{R}$ and $n < \infty$ by applying the respective operators to coordinate functions. In this sense, the

standard Ornstein-Uhlenbeck operator has infinite dimension.

1.1.1. $CD(\kappa,\infty)$ and convex Sobolev inequalities

As indicated above, positive curvature bounds with regard to the Bakry-Émery curvature-dimension condition lead to certain estimates regarding the long-time behavior of the semigroup. In the following lines we make this more precise. For this purpose we introduce the so called Φ-entropies

$$\mathrm{Ent}_\mu^\Phi(f) = \int_X \Phi(f)\mathrm{d}\mu - \Phi\Big(\int_X f\mathrm{d}\mu\Big),$$

where $\Phi : J \to \mathbb{R}$ is a smooth and convex mapping and f belongs to a sufficiently rich class of J-valued functions. The corresponding Φ-entropy inequality with constant $C > 0$ is then formulated as

$$\mathrm{Ent}_\mu^\Phi(f) \leqslant \frac{1}{2C}\mathcal{E}(\Phi'(f), f). \tag{1.11}$$

In order to understand in which way this functional inequality is related to the long-time behavior of the semigroup $\big(P_t\big)_{t\geqslant 0}$ we use the important identity

$$\frac{d}{dt}\mathrm{Ent}_\mu^\Phi(P_t f) = -\mathcal{E}(\Phi'(P_t f), P_t f), \tag{1.12}$$

which can be derived from μ being invariant and the integration by parts formula (1.7) as follows

$$\frac{d}{dt}\mathrm{Ent}_\mu^\Phi(P_t f) = \frac{d}{dt}\Big(\int_X \Phi(P_t f)\mathrm{d}\mu - \Phi\Big(\int_X f\mathrm{d}\mu\Big)\Big) = \int_X \Phi'(P_t f) L P_t f \,\mathrm{d}\mu = -\mathcal{E}(\Phi'(P_t f), P_t f). \tag{1.13}$$

Assuming that (1.11) holds, we observe from (1.12) that

$$\frac{d}{dt}\mathrm{Ent}_\mu^\Phi(P_t f) \leqslant -2C\mathrm{Ent}_\mu^\Phi(P_t f).$$

In particular, due to

$$\frac{d}{dt}\big(e^{2Ct}\mathrm{Ent}_\mu^\Phi(P_t f)\big) = e^{2Ct}\Big(\frac{d}{dt}\mathrm{Ent}_\mu^\Phi(P_t f) + 2C\,\mathrm{Ent}_\mu^\Phi(P_t f)\Big),$$

the mapping $t \mapsto e^{2Ct}\mathrm{Ent}_\mu^\Phi(P_t f)$ is decreasing and thus the estimate

$$\mathrm{Ent}_\mu^\Phi(P_t f) \leqslant e^{-2Ct}\mathrm{Ent}_\mu^\Phi(f) \tag{1.14}$$

follows. This describes an exponential decay behavior of the entropy.

Conversely, subtracting $\mathrm{Ent}_\mu^\Phi(f)$ in (1.14), dividing by t and sending $t \to 0$ yields, by using (1.12) again, the reverse implication, i.e. the functional inequality (1.11) is equivalent to the estimate (1.14) describing the long-time behavior of the semigroup.

We continue with discussing the most important examples among Φ-entropy inequalities. When choosing $\Phi(r) = r\log r$, the Φ-entropy results in the Boltzmann entropy, which is of tremendous importance not only among mathematics but also in thermodynamics or information theory (resp. with opposite sign), where it is better known as the Shannon entropy. We simply write Ent_μ to denote the Boltzmann entropy from now on. The corresponding functional inequality (1.11) is the logarithmic Sobolev inequality

$$\mathrm{Ent}_\mu(f) \leqslant \frac{1}{2C}\mathcal{E}(\log f, f). \tag{1.15}$$

In fact, as $\Gamma(\mathbb{1}, f) = 0$ (which follows from the underlying assumptions), we have $\mathcal{E}(\Phi'(f), f) = \mathcal{E}(\log f, f)$. The quantity $\mathcal{E}(\log f, f)$ is also called Fisher information and is denoted by $\mathcal{I}(f)$. By the diffusion property (1.1) we deduce

$$\mathcal{I}(f^2) = \mathcal{E}(\log f^2, f^2) = 2\int_X \Gamma(\log f, f^2)\mathrm{d}\mu = 4\int_X \Gamma(f, f)\mathrm{d}\mu = 4\mathcal{E}(f), \tag{1.16}$$

where we write $\mathcal{E}(f) := \mathcal{E}(f, f)$. Consequently, (1.15) is equivalent to

$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{2}{C}\mathcal{E}(f), \tag{1.17}$$

which is the more well-known formulation of the logarithmic Sobolev inequality. In fact, the inequality (1.15) is often called modified logarithmic Sobolev inequality. While being equivalent to the logarithmic Sobolev inequality in the setting of the classical Bakry-Émery theory, the inequalities (1.15) and (1.17) differ in situations where the chain rule (1.1) is not available.

Choosing $\Phi(r) = r^2$, $r \in \mathbb{R}$, the corresponding Φ-entropy translates to another well-known object, the variance, which we denote henceforth by Var_μ. The resulting functional inequality (1.11) is called Poincaré inequality and reads as

$$\mathrm{Var}_\mu(f) \leqslant \frac{1}{C}\mathcal{E}(f).$$

There also exists an important familiy of functional inequalities that interpolate between logarithmic Sobolev inequalities and Poincaré inequalities and which have been first mentioned in [7]. For the choice of $\Phi_p = \frac{r^p - r}{p(p-1)}$, where $r > 0$ and $p \in (1, 2)$, the related functional inequalities are called Beckner inequalities and are given by

$$\mathrm{Ent}_\mu^{\Phi_p}(f) \leqslant \frac{1}{2C}\mathcal{E}(\Phi_p'(f), f). \tag{1.18}$$

Beckner inequalities, logarithmic Sobolev inequalities and Poincaré inequalities are commonly called convex Sobolev inequalities. Further, there also exist a hierarchy between the three mentioned functional inequalities. Beckner inequalities are implied by the logarithmic Sobolev inequality and imply the Poincaré inequality (resp. with the same constant), see e.g. [22].

The curvature-dimension condition $CD(\kappa, \infty)$, $\kappa > 0$, implies the logarithmic Sobolev inequality with constant κ (and hence also implies Beckner inequalities and the Poincaré inequality with constant κ, respectively). A way of proving this makes use of the entropy method, which relies on comparing the entropy production with the second-time derivative of the entropy along the heat flow and which we now outline in brief.

If an estimate of the form

$$\frac{d^2}{dt^2}\mathrm{Ent}_\mu(P_t f) \geqslant -2\alpha\frac{d}{dt}\mathrm{Ent}_\mu(P_t f), \tag{1.19}$$

holds for some $\alpha > 0$, then Gronwall's lemma yields $-\frac{d}{dt}\mathrm{Ent}_\mu(P_t f) \leqslant -e^{-2\alpha t}\,\frac{d}{dt}\mathrm{Ent}_\mu(P_t f)\big|_{t=0}$. Integrating this inequality, we observe

$$\mathrm{Ent}_\mu(f) - \mathrm{Ent}_\mu(P_t f) \leqslant -\left.\frac{d}{dt}\mathrm{Ent}_\mu(P_t f)\right|_{t=0}\int_0^t e^{-2\alpha s}\mathrm{d}s.$$

Letting now $t \to \infty$, the logarithmic Sobolev inequality in the form of (1.15) follows from (1.12) provided that $\mathrm{Ent}_\mu(P_t f) \to 0$ as $t \to \infty$. The latter property is a consequence of an ergodicity assumption being used in the Bakry-Émery theory.

The condition $CD(\kappa, \infty)$ now ensures the validity of an estimate of the form (1.19). Indeed, a

lengthy calculation shows that the second derivative of the entropy along the heat flow is given by

$$\frac{d^2}{dt^2}\mathrm{Ent}_\mu(P_t f) = 2\int_X P_t f\,\Gamma_2(\log P_t f)\mathrm{d}\mu. \tag{1.20}$$

Moreover, applying the diffusion property to identity (1.12), we observe that

$$\frac{d}{dt}\mathrm{Ent}_\mu(P_t f) = -\int_X P_t f\,\Gamma(\log P_t f)\mathrm{d}\mu \tag{1.21}$$

holds and hence (1.19) is valid with $\alpha = \kappa$ whenever $CD(\kappa,\infty)$ is satisfied.

We remark that the above procedure of the entropy method does not only apply to the Boltzmann entropy, but also, more generally, to suitable Φ-entropies. In particular, one can also derive the Poincaré inequality and the Beckner inequalities from $CD(\kappa,\infty)$, $\kappa > 0$, by means of the entropy method applied to the corresponding Φ-entropies in a similar way as presented above.

1.1.2. $CD(\kappa, n)$ and logarithmic entropy-energy inequalities

If the curvature-dimension condition not only allows for a positive curvature bound but also for a finite dimension term, a refinement of the entropy method leads to a stronger functional inequality, the so called logarithmic entropy-energy inequality. More precisely, the condition $CD(\kappa,n)$, with $\kappa > 0$ and $n < \infty$, yields

$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{n}{2}\log\Big(1 + \frac{4}{\kappa n}\mathcal{E}(f)\Big). \tag{1.22}$$

Note in fact that the logarithmic entropy-energy inequality (1.22) implies the logarithmic Sobolev inequality (1.17) with constant $\kappa > 0$ due to the estimate $\log(1+r) \leqslant r$, $r \geqslant 0$. For later purposes we remark that (1.22) also implies the celebrated Nash inequality

$$\|f\|^{n+2}_{L^2(\mu)} \leqslant \Big(\|f\|^2_{L^2(\mu)} + \frac{\mathcal{E}(f)}{\kappa n}\Big)^2 \|f\|^2_{L^1(\mu)}.$$

The inequality (1.22) is an example from the more general class of so called entropy-energy inequalities

$$\mathrm{Ent}_\mu(f^2) \leqslant \Phi\big(\mathcal{E}(f)\big), \tag{1.23}$$

where $\Phi : (0,\infty) \to \mathbb{R}$ is a strictly increasing concave C^1-function that is commonly called growth function. Note that the logarithmic Sobolev inequality (1.17) is also included in this class of functional inequalities. However, the logarithmic entropy-energy inequality has significantly stronger consequences than the logarithmic Sobolev inequality. For instance, it is known by the Gross' hypercontractivity theorem that the logarithmic Sobolev inequality with constant $\kappa > 0$ can be characterized by the inequality

$$\|P_t f\|_{L^q(\mu)} \leqslant \|f\|_{L^p(\mu)}, \tag{1.24}$$

where $1 < p < q < \infty$ are such that $e^{2t\kappa} \geqslant \frac{q-1}{p-1}$. The semigroup is said to be hypercontractive if estimates of this form hold. When $q = \infty$ can be reached in (1.24) one says that the semigroup is ultracontractive. Note that in the above formulation of the logarithmic Sobolev inequality ultracontractivity can not be reached. In the setting of entropy-energy inequalities with growth function Φ, however, ultracontractive bounds are satisfied provided that the mapping $r \mapsto \frac{\Phi'(r)}{r}$ is integrable at ∞. While this is not true for the growth function $\Phi(r) = r$ that is related to the logarithmic Sobolev inequality, the growth function corresponding to the logarithmic

entropy-energy inequality (1.22) satisfies this integrability condition.

Another fundamental difference between logarithmic Sobolev inequalities and logarithmic entropy-energy inequalities concerns the notion of the diameter. In the general setting of the Bakry-Émery theory, a function $f \in \mathcal{A}$ is said to be a Lipschitz function if $\Gamma(f)$ is bounded. This definition is motivated by the Laplace-Beltrami case, where $\Gamma(f) = |\nabla f|^2$, and by Rademacher's Theorem, which tells us in the Euclidean case that Lipschitz functions coincide with functions that have bounded gradients almost everywhere. One defines the Lipschitz semi-norm as $\|f\|_{\mathrm{Lip}} := \sqrt{\|\Gamma(f)\|_\infty}$. The diameter of X (w.r.t. Γ and μ) is then defined by

$$\mathrm{diam}_\varrho = \sup_{x,y \in X} \varrho(x,y),$$

where $\varrho(x,y) = \sup\{f(y) - f(x) : \|f\|_{\mathrm{Lip}} \leqslant 1\}$. For the generalized Ornstein-Uhlenbeck operator on $\mathbb{R}^d$, we have that $\Gamma(f) = |\nabla f|^2$. Clearly, there exist unbounded Lipschitz functions (e.g. linear functions) in this context which yields, in particular, that the logarithmic Sobolev inequality is not sufficient to obtain a finite diameter in the above sense. Conversely, the logarithmic entropy-energy inequality implies that diam_ϱ is finite. As a consequence, we observe on the level of CD-inequalities that $CD(\kappa, n)$ is strong enough to imply finite diameter bounds while this observation is not true when assuming solely $CD(\kappa, \infty)$.

1.1.3. $CD(0,n)$ and Li-Yau inequalities

Let us now investigate the case of the $CD(0,n)$ condition. Even without a positive curvature bound, very important functional inequalities can still be deduced. More precisely, it has been shown in [5] that the Li-Yau inequality

$$-L \log P_t f \leqslant \frac{n}{2t}$$

holds true, provided that L satisfies $CD(0,n)$. Parabolic Harnack inequalities serve as a main application of such estimates.

We discuss the content a bit more detailed in the case of the Euclidean Laplacian on $\mathbb{R}^n$, which satisfies, as we have seen further above, the curvature-dimension condition $CD(0,n)$. In their seminal work [24], Li and Yau showed that on a complete Riemannian manifold M with topological dimension n and non-negative Ricci curvature the estimate

$$-\Delta(\log u) \leqslant \frac{n}{2t}, \qquad t > 0, \tag{1.25}$$

holds true for any positive solution u to the heat equation $\partial_t u = \Delta u$ on $(0,\infty) \times M$, where Δ denotes the corresponding Laplace-Beltrami operator. Although it can not be expected that Li and Yau, back then in the mid-80ies, were aware of the young probabilistic school of Bakry and Émery, the original approach of [24] has a Γ-calculus interpretation. For motivational purposes, we briefly give a formal argument of how to use the $CD(0,n)$ condition to derive the Li-Yau inequality for the heat equation on $\mathbb{R}^n$ with the original approach.

Recall that $\Gamma(f) = |\nabla f|^2$ for $f \in C^\infty(\mathbb{R}^n)$. By the chain rule (1.6) we have

$$\partial_t \log u = \frac{\Delta u}{u} = \Delta \log u + |\nabla \log u|^2,$$

where u is a positive solution to the heat equation on $(0,\infty) \times \mathbb{R}^n$. Applying the Laplace operator, we get

$$\partial_t \Delta \log u = \Delta\Delta \log u + \Delta\big(|\nabla \log u|^2\big).$$

For $T > 0$ being arbitrarily given, we consider the mapping $G(s,y) = -\frac{2s}{n}\Delta \log u(s,y)$, where

$(s,y) \in (0,T] \times \mathbb{R}^n$ and assume that we can find $(t,x) \in (0,T] \times \mathbb{R}^n$ such that G achieves its maximum in (t,x). Hence, we have $\partial_t G(t,x) \geqslant 0$, $\nabla G(t,x) = 0$ and $\Delta G(t,x) \leqslant 0$. Since it suffices to show that $G(t,x) \leqslant 1$, we may assume that $G(t,x) > 0$. With (1.9) we observe that

$$\begin{aligned}
0 \leqslant \partial_t G(t,x) &= -\frac{2\Delta \log u(t,x)}{n} - \frac{2t\,\partial_t \Delta \log u(t,x)}{n} \\
&= -\frac{2\Delta \log u(t,x)}{n} - \frac{2t}{n}\Big(\Delta\big(|\nabla \log u|^2\big)(t,x) + \Delta\Delta \log u(t,x)\Big) \\
&= -\frac{2\Delta \log u(t,x)}{n} - \frac{2t}{n}\Big(2\Gamma_2(\log u)(t,x) + 2\langle \nabla \log u(t,x), \nabla \Delta \log u(t,x)\rangle + \Delta\Delta \log u(t,x)\Big) \\
&= -\frac{2\Delta \log u(t,x)}{n} - \frac{4t}{n}\Gamma_2(\log u)(t,x) + 2\langle \nabla \log u(t,x), \nabla G(t,x)\rangle + \Delta G(t,x) \\
&\leqslant -\frac{2\Delta \log u(t,x)}{n} - \frac{4t}{n}\Gamma_2(\log u)(t,x).
\end{aligned}$$

Using the $CD(0,n)$ condition we conclude

$$\frac{2t}{n}\big(\Delta \log u(t,x)\big)^2 \leqslant -\Delta \log u(t,x),$$

which shows that $G(t,x) \leqslant 1$.

The assumption that G achieves its maximum on $(0,T] \times \mathbb{R}^n$ is clearly not satisfied in general. In order to make the above argument rigorous one has to multiply G with a suitably chosen cut-off function. We refer to [23] for a rigorous proof in this spirit.

For later purposes, we also present another approach for deriving the Euclidean Li-Yau inequality that works via reduction to the heat kernel and which can be found in [18]. This strategy relies on the fact that positive solutions to the heat equation are given by the convolution of the initial datum with the Euclidean heat kernel $H(t,x) = \big(4\pi t\big)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}$, which is known as the Widder theorem [37]. Note that the Li-Yau inequality can be reformulated by using the chain rule as

$$\frac{|\nabla u(t,x)|^2}{u(t,x)} - \partial_t u(t,x) \leqslant \frac{n}{2t} u(t,x), \tag{1.26}$$

which is also called differential Harnack inequality. One readily checks that the Euclidean heat kernel satisfies this estimate with equality at any $(t,x) \in (0,\infty) \times \mathbb{R}^n$. The idea is to make use of the latter information by transferring the derivatives from u to the heat kernel in the usual way, that is

$$\begin{aligned}
\partial_{x_i} u(t,x) &= \int_{\mathbb{R}^n} \partial_{x_i} H(t,x-y) u_0(y) \mathrm{d}y \\
&\leqslant \Big(\int_{\mathbb{R}^n} \frac{\partial_{x_i} H(t,x-y)^2}{H(t,x-y)} u_0(y) \mathrm{d}y\Big)^{\frac{1}{2}} \Big(\int_{\mathbb{R}^n} H(t,x-y) u_0(y) \mathrm{d}y\Big)^{\frac{1}{2}},
\end{aligned}$$

where we have applied the Cauchy-Schwarz inequality. Consequently, we observe that

$$\begin{aligned}
\frac{|\nabla u(t,x)|^2}{u(t,x)} &\leqslant \int_{\mathbb{R}^n} \frac{|\nabla_x H(t,x-y)|^2}{H(t,x-y)} u_0(y) \mathrm{d}y \\
&= \int_{\mathbb{R}^n} \big(\partial_t H(t,x-y) + \frac{n}{2t} H(t,x-y)\big) u_0(y) \mathrm{d}y \\
&= \partial_t u(t,x) + \frac{n}{2t} u(t,x).
\end{aligned}$$

This shows (1.26) and, equivalently, the desired Li-Yau inequality.

1.2. Non-local operators

The scope of this thesis covers operators of the form

$$Lf(x) = \int_{X\setminus\{x\}} \big(f(y) - f(x)\big)k(x, \mathrm{d}y), \tag{1.27}$$

where the kernel is such that $k(x,\cdot)$ defines a σ-finite measure on the Borel sets $\mathcal{B}(X\setminus\{x\})$ for any $x \in X$ and suitable $f : X \to \mathbb{R}$. Here, X denotes some metric space. The integral in the right-hand side of (1.27) may have to be interpreted in the principal value sense. An operator of this form is called non-local since the value of Lf at x depends not only on an arbitrary small neighborhood (like in the case of differential operators) but instead also on the behavior of the function f farther away. This entails that the theory of non-local PDE differs a lot from the classical local counterpart. A quite significant difference is, for instance, that operators of the above form do not satisfy in general chain rules like (1.5), where the carré du champ operator Γ is defined by formula (1.4).

From the viewpoint of stochastic processes, non-local operators of the form (1.27) appear as generators of jump processes. We focus in the sequel on our two main cases of consideration. The first one deals with the case of the state space X being a continuum (we will consider $X = \mathbb{R}^n$). In the second case X is a discrete set.

1.2.1. Generators of pure-jump Lévy processes

Lévy processes constitute an important subclass of Markov processes. They are uniquely determined by their so called characteristic triplet $\big(b, Q, \nu\big)$, where $b \in \mathbb{R}^d$, $Q \in \mathbb{R}^{d\times d}$ is a positive semi-definite matrix and ν is a measure on the Borel σ-algebra of $\mathbb{R}^d\setminus\{0\}$ with $\int_{y\neq 0} \min\{|y|^2, 1\}\mathrm{d}\nu < \infty$. The corresponding infinitesimal generator L is given by

$$Lf(x) = \langle b, \nabla f(x)\rangle + \frac{1}{2}\mathrm{div}(Q\nabla f)(x) + \int_{\mathbb{R}^d\setminus\{0\}} \big(f(x+y) - f(x) - \langle \nabla f(x), y\rangle \mathbb{1}_{(0,1)}(|y|)\big)\mathrm{d}\nu(y), \tag{1.28}$$

where $f \in C_c^\infty(\mathbb{R}^d)$. Note that the assumption $\int_{y\neq 0} \min\{|y|^2, 1\}\mathrm{d}\nu < \infty$ ensures that the integral in the right-hand side of (1.28) is well-defined (which can be seen by a classical Taylor argument).

The matrix Q represents the diffusion part, the vector b the drift part and the measure ν the jump part of the process. In the spirit of the discussion above, b and Q determine the local part while ν stands for the non-local part of the operator L. In the case of $\nu \equiv 0$, L fits into the framework of the classical Bakry-Émery theory which includes (as it has been mentioned in Section 1.1) second order differential operators without zero-th order term and with sufficiently smooth coefficients. In the concrete case of $\nu \equiv 0$, $b = 0$ and Q being given by the identity matrix, the resulting Lévy processes coincides with the classical Brownian motion. However, in this section we are less interested in the objects b and Q than in the pure jump part of the process. If ν is an isotropic measure, that is there exists a density function $\nu_0 : (0,\infty) \to [0,\infty)$ (not necessarily integrable) such that $\nu(\mathrm{d}z) = \nu_0(|z|)\mathrm{d}z$, then we have for any $r > 0$ that

$$\begin{aligned}\int_{\mathbb{R}^d\setminus B_r(0)} \big(f(x+y) - f(x) - \langle \nabla f(x), y\rangle \mathbb{1}_{(0,1)}(|y|)\big)\mathrm{d}\nu(y) &= \int_{\mathbb{R}^d\setminus B_r(0)} \big(f(x+y) - f(x)\big)\mathrm{d}\nu(y)\\ &= \int_{\mathbb{R}^d\setminus B_r(x)} \big(f(z) - f(x)\big)k(x, \mathrm{d}z),\end{aligned}$$

where $k(x, \mathrm{d}z) = \nu_0(|x - z|)\mathrm{d}z$. Sending $r \to 0$ yields that the jump part is an operator of the form (1.27) provided that the measure ν is isotropic.

The most prominent such example is given by the fractional Laplacian $-\big(-\Delta\big)^{\frac{\beta}{2}}$, $\beta \in (0,2)$,

which is the generator of the β-stable isotropic Lévy process. This operator consists only of a jump part that is given by

$$-\big(-\Delta\big)^{\frac{\beta}{2}} f(x) = c_{\beta,d} \lim_{\varepsilon \to 0} \int_{\mathbb{R}^d \setminus B_\varepsilon(x)} \frac{f(y) - f(x)}{|x-y|^{d+\beta}} \mathrm{d}y, \tag{1.29}$$

where $\beta \in (0,2)$, with normalizing constant $c_{\beta,d} = \frac{2^\beta \Gamma(\frac{d+\beta}{2})}{\pi^{\frac{d}{2}} |\Gamma(-\frac{\beta}{2})|}$ (here Γ denotes the Gamma-function) and f can be chosen, e.g., as an element of $C^2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. One naturally asks if and in which way properties of the Euclidean Laplace operator carry over to the fractional Laplacian. We will shed light on this question with respect to curvature-dimension inequalities and related functional inequalities in this thesis.

For further information on Lévy processes we refer the reader, for instance, to the extensive monograph [1].

1.2.2. Generators of continuous-time Markov chains

If X is a discrete set, then we refer to the Markov process $(Z_t)_{t \geqslant 0}$ with state space X as a continuous-time Markov chain. The generator of this process can be written in the form of (1.27) as

$$Lf(x) = \sum_{y \in X} k(x,y)\big(f(y) - f(x)\big),$$

where f belongs to a class of suitable functions and where we set $k(x,y) := k(x,\{y\})$ if $x \neq y$ and $k(x,x) := -\sum_{y \neq x} k(x,y)$ else, where the latter sum is assumed to be finite. The value of $k(x,y)$ describes the transition rate to jump from state x to state y within the next jump.

In a natural way the Markov generator L induces a graph structure, where the vertex set is given by the elements of the state space X and edge weights given by the transition rates $k(x,y)$ for $x, y \in X$ with $x \neq y$. We refer to this graph as the underlying graph to L.

A Markov chain is called irreducible if the probability to reach an arbitrary state is positive no matter at which state the chain starts and no matter how many jumps are needed. More precisely, given two fixed distinct states $x, y \in X$, we find intermediary states $z_1, ..., z_m \in X$, $m \geqslant 2$, with $z_1 = x$ and $z_m = y$ such that $k(z_i, z_{i+1}) > 0$ for any $i \in \{1, ..., n-1\}$.

An irreducible Markov chain is called positive recurrent if and only if the semigroup, which is given by (1.2), is stochastically complete (that is $P_t \mathbb{1} = \mathbb{1}$) and a (unique) invariant measure μ exists. This notion is crucial for the long-time behavior of the chain. In fact, positive recurrence implies ergodicity, which means that the transition probabilities converge to the invariant measure as $t \to \infty$. Using this property, one concludes also that the entropy along the heat flow $\mathrm{Ent}_\mu(P_t f)$ does converge to 0 under respective assumptions on the initial condition f. Recall that this property is crucial when applying the entropy method as explained in Section 1.1.1.

Another quite important assumption will be that the chain is reversible. In the discrete setting this property is characterized by the detailed balance condition

$$\mu(\{x\})k(x,y) = \mu(\{y\})k(y,x).$$

Taking the sum over this identity, one immediatly observes that $\sum_{y \in X} k(x,y)\mu(\{y\}) = 0$, which is known to be equivalent to the measure μ being invariant (see e.g. [21]).

We refer to [30] for a detailed account on the theory of continuous-time Markov chains.

For motivational purposes, we briefly discuss the example of long-range jump operators on the state space $\mathbb{Z}$. We assume that the transition rates are given by $k(x,y) = k_*(y-x)$, where

$x, y \in \mathbb{Z}$ with $x \neq y$ and $k_* \in \ell^1(\mathbb{Z}\backslash\{0\})$ is non-negative and symmetric. Then we can rewrite the generator as

$$Lf(x) = \sum_{h\in\mathbb{Z}\backslash\{0\}} k_*(h)\big(f(x+h) - f(x)\big). \tag{1.30}$$

In this context we often speak of long-range jump operators, which is motivated from the case where k_* has full support on $\mathbb{Z}$ (or at least unbounded support), i.e. jumps of arbitrary length can occur at any step. The resulting underlying graph in this case is a locally infinite graph, which means that each vertex is adjacent to infinitely many vertices.

Of particular importance is the specific choice $k_*(j) = \frac{1}{|j|^{1+\beta}}$, $j \in \mathbb{Z}\backslash\{0\}$, $\beta > 0$, to which we refer as power-type kernels. We observe that these kernels appear in their corresponding continuous version in the above formula for the fractional Laplacian (cf. (1.29)). Moreover, in [12] it is shown that the operator $-\big(-\Delta\big)^{\frac{\beta}{2}}$, $\beta \in (0,2)$, defined by means of the semigroup method and where Δ denotes the discrete Laplacian on $\mathbb{Z}$, is of the form (1.30) with kernel k_β and constants $C(\beta) \geqslant c(\beta) > 0$ such that

$$\frac{c(\beta)}{|j|^{1+\beta}} \leqslant k_\beta(j) \leqslant \frac{C(\beta)}{|j|^{1+\beta}}$$

holds for any $j \in \mathbb{Z}$ and $\beta \in (0,2)$, which is an interesting link to Section 1.2.1.

1.3. Towards a non-local Bakry-Émery theory

The core aim of this thesis is to establish a theory in the spirit of Bakry-Émery for operators of the form (1.27). Clearly, one can define Γ and Γ_2 by the formulas (1.4) and (1.9), and also the condition $CD(\kappa, n)$ can be carried over to the non-local framework on a suitably chosen function space. However, there is one major difficulty that prevents the Bakry-Émery curvature-dimension condition from having such strong implications in the non-local situation as in the classical framework: The lack of chain rules of the form (1.1) and (1.5).

Also on the level of functional inequalities, this leads to significant differences to the classical setting. For instance, since the calculation in (1.16) is not valid in the non-local setting, the functional inequalities (1.15) and (1.17) are not equivalent anymore (here and in the sequel, the entropy, the energy and the Fisher information in the non-local setting are defined as for the classical objects in Section 1.1). Instead, the logarithmic Sobolev inequality is known to be stronger than the modified logarithmic Sobolev inequality. We will give an example below in Section 1.3.1, for which (1.15) holds with some positive constant and (1.17) does not hold for any positive constant. For a detailed discussion of the hierarchy between convex Sobolev inequalities in the discrete setting, we refer to [8]. In particular, it is also shown there that modified logarithmic Sobolev inequalities can be characterized by a hypercontractivity formulation for $e^{P_t f}$ instead of the formulation that we have discussed in Section 1.1.1, while the latter ist still valid for the logarithmic Sobolev inequality in the context of finite Markov chains (cf. [13]).

In the following subsection we will describe existing approaches for finding substitutes for curvature and dimension in the non-local setting. Since the research related to synthetic notions of curvature on discrete structures is a quite active research topic, it is too ambitiuous to give an exhaustive account here. Instead, we focus on those approaches which seemingly have the strongest links to the content of this thesis. We refer to the recent monograph [29] for a more detailed account. In the second part of this section we give an overview about the main contributions of this thesis.

1.3.1. Existing approaches

We have emphasized the high importance of the chain rules throughout Section 1.1 for the classical theory. As a substitute of (1.5) for non-local operators, the authors showed in [14] that for $H \in C^1(D;\mathbb{R})$, $f : X \to D$, where $D \subset \mathbb{R}$ is some open set, we have

$$L\big(H(f)\big)(x) = H'(f(x))Lf(x) + \int_{X\setminus\{x\}} \Lambda_H\big(f(y), f(x)\big)k(x,\mathrm{d}y), \tag{1.31}$$

where $\Lambda_H(w,z) := H(w) - H(z) - H'(z)(w-z)$ for any $w, z \in D$ and the integral on the right-hand side may has to be understood in the principal value sense.

The choice of $H(r) = \frac{r^2}{2}$ implies that

$$\Gamma(f)(x) = \frac{1}{2}\int_{X\setminus\{x\}} \big(f(y) - f(x)\big)^2 k(x,\mathrm{d}y), \tag{1.32}$$

where Γ is defined by formula (1.4).

When taking $H(r) = \sqrt{r}$, $r > 0$, (1.31) translates to an identity which serves as the key in [6] for showing Li-Yau inequalities under their so called exponential curvature-dimension condition. This was the first article in which Li-Yau inequalities under curvature-dimension conditions in the discrete setting of locally finite graphs have been studied.

In the context of this thesis, the most important situation in which we apply (1.31) concerns the specific choice of $H(r) = \log(r)$, $r \in (0,\infty)$. In this case we have

$$\Lambda_{\log}(w,z) = \log w - \log z - \frac{w-z}{z} = -\Upsilon(\log w - \log z),\ w, z \in (0,\infty),$$

where $\Upsilon(r) := e^r - r - 1$, $r \in \mathbb{R}$. Introducing the operator

$$\Psi_\Upsilon(f)(x) = \int_{X\setminus\{x\}} \Upsilon\big(f(y) - f(x)\big)k(x,\mathrm{d}y),$$

we deduce from (1.31) that

$$L(\log f) = \frac{Lf}{f} - \Psi_\Upsilon(\log f). \tag{1.33}$$

In this identity, the operator $\Psi_\Upsilon(\log f)$ thus serves as the replacement of $\Gamma(\log f)$ in (1.6). The expression $\Psi_\Upsilon(\log f)$ already appeared (although not being used but only mentioned in a remark) in a rather implicit form in the work of [28], where a quite general curvature-dimension condition $CD\psi$ has been introduced in the setting of unweighted finite graphs for some concave function ψ. The choice of $\psi = \log$ has interesting links with this thesis and will be discussed below.

Another important concept of [14], which is, in particular, different to the approach of [28], concerns the notion of CD-functions. We call a continuous function $F : [0,\infty) \to [0,\infty)$ a CD-function if $F(0) = 0$, $r \mapsto \frac{F(r)}{r}$ is strictly increasing on $(0,\infty)$ and $1/F$ is integrable at ∞. This notion generalizes the square function in the dimension term for their CD-inequality. One important advantage of this approach is that situations in which the corresponding CD-inequality with the classical quadratic dimension term does not hold can still be treated by taking a different CD-function. Another important aspect is that the behavior of the corresponding Li-Yau inequality can be described more precisely. In fact, to any CD-function one associates the so called relaxation function φ, which is the (unique) strictly positive solution to the ordinary differential equation $\varphi'(t) + F(\varphi(t)) = 0$. Then, one can show that under the CD-inequality $CD(F;0)$ of [14], a Li-Yau type inequality of the form $-L\log u \leqslant \varphi(t)$ holds true for finite graphs. For instance, in case of the unweighted two-point space the method of [14] leads to a

sharp Li-Yau inequality $-L\log u \leqslant \varphi(t)$, where the relaxation function φ behaves like $-\log t$ as $t \to 0$ and like $2e^{-2t}$ as $t \to \infty$. Note that such a precise analysis is not possible, when being restricted to the quadratic CD-function as in [6] and [28], in which case the corresponding relaxation function is always of the form $t \mapsto \frac{c}{t}$, $c > 0$.

In the discrete setting, the entropy method as described in Section 1.1.1 has been investigated in [11]. While the first derivative of the entropy along the heat flow is also relatively simple to calculate in the discrete setting, the second derivative leads to major difficulties due to the lack of chain rule. The authors of [11] circumvent this problem by using a discrete version of the Bochner-Lichnerowicz formula (1.8) whose validity relies on certain strong symmetry assumptions of the process and on reversibility. The approach of [11] consists of estimating the whole integral of the second derivative instead of finding a representation formula that allows for pointwise estimates, like in the classical approach, via a curvature-dimension condition (cf. Section 1.1.1). Using their identity, they provide a criterion for estimating the second derivative of the entropy along the heat flow which translates for some important examples such as zero range processes or birth-death processes into conditions that are easy to check. We illustrate this in case of the latter by considering the state space $X = \mathbb{N}_0$ (the case of $X = \{0, ..., N\}$, $N \in \mathbb{N}$, works similar) and transition rates being given by $k(x, y) = 0$ whenever $|x - y| > 1$ and $a(x) := k(x, x + 1) > 0$, $b(x) := k(x, x - 1) > 0$ with $b(0) = 0$. Provided that a and b are monotone in the sense that $a(x) \leqslant a(x - 1)$ and $b(x - 1) \leqslant b(x)$ for any $x \in \mathbb{N}$, the condition

$$a(x - 1) - a(x) + b(x) - b(x - 1) \geqslant \kappa, \quad \forall x \in \mathbb{N}, \tag{1.34}$$

where $\kappa > 0$, yields by the method of [11] the estimate (1.19) with $\alpha = \frac{\kappa}{2}$ and consequently, by means of the entropy method, the modified logarithmic Sobolev inequality with constant $\frac{\kappa}{2}$. An important special case is the so called Poisson case of a birth-death process. Here, the rates are given by $a(x) = \lambda$, $\lambda > 0$, and $b(x) = x$. While the modified logarithmic Sobolev inequality is valid with optimal constant $\frac{1}{2}$ by the above method, it is known that a logarithmic Sobolev inequality can not be satisfied, see [11]. This demonstrates the difference between the discrete setting and the setting of Section 1.1 for the functional inequalities (1.15) and (1.17).

Besides the characterization via the Bakry-Émery curvature-dimension condition there is also a notion of Ricci-curvature lower bounds for a Riemannian manifold M relying on the powerful theory of optimal transport. More precisely, in [33] it has been observed that the Ricci-curvature of M is bounded from below by some constant $\kappa \in \mathbb{R}$ if and only if the Boltzmann entropy is κ-convex along geodesics in the Wasserstein space $P_2(M)$. This deep result opened up the synthetic theory of Ricci-curvature notions in the setting of so called geodesic metric measure spaces. However, this theory does not apply to the non-local setting since the corresponding L^2-Wasserstein space does not contain geodesics in this case.

Erbar and Maas and, independently, Mielke did overcome this problem by defining a different metric that is suitable in the setting of finite Markov chains, cf. [16, 27] (see [15] for the space-continuous setting). With their new metric taking over the role of the L^2-Wasserstein metric, [16] introduced the definition of Ricci-curvature in the setting of finite Markov chains via geodesic convexity of the entropy. In order to compare this curvature notion to our approach later, let us consider for ρ being a probability density with respect to the invariant and reversible measure μ and $\psi : X \to \mathbb{R}$ some function, the objects

$$\mathcal{A}(\rho, \psi) = \frac{1}{2} \sum_{x,y \in X} \big(\psi(x) - \psi(y)\big)^2 \theta(\rho(x), \rho(y)) k(x, y) \mu(\{x\})$$

and

$$\begin{aligned}\mathcal{B}(\rho,\psi) =&\frac{1}{4}\sum_{x,y\in X}\big(\psi(x)-\psi(y)\big)^2\hat{L}\rho(x,y)k(x,y)\mu(\{x\})\\&-\frac{1}{2}\sum_{x,y\in X}\big(L\psi(x)-L\psi(y)\big)(\psi(x)-\psi(y))\,\theta(\rho(x),\rho(y))k(x,y)\mu(\{x\}).\end{aligned}$$

Here, $\hat{L}\rho(x,y) = \partial_1\theta(\rho(x),\rho(y))L\rho(x) + \partial_2\theta(\rho(x),\rho(y))L\rho(y)$ and θ denotes the logarithmic mean, which is given by $\theta(s,t) := \int_0^1 s^{1-p}t^p\mathrm{d}p = \frac{s-t}{\log s-\log t}$, $s,t>0$, with the latter identity being valid whenever $s \neq t$. The entropic Ricci curvature of [16] is bounded from below by $\kappa \in \mathbb{R}$ if and only if $\mathcal{B}(\rho,\psi) \geqslant \kappa\mathcal{A}(\rho,\psi)$ holds for any strictly positive probability density ρ (w.r.t μ) and function ψ (cf. [16]). Again we observe that, in contrast to the Bakry-Émery approach, the entropic Ricci curvature also compares certain integral expressions rather than respective operators in the pointwise sense.

1.3.2. Contribution of this thesis

The goal of this thesis is to develop a non-local version of the Bakry-Émery theory. Despite the lack of the classical chain rule, we want to find a curvature-dimension condition for operators of the form (1.27) that has similar consequences as the Bakry-Émery CD condition in the classical setting (see Section 1.1). Thereby, we also want to identify the non-local variants of functional inequalities that are related to the Bakry-Émery theory and study their consequences. In this sense, we can divide the topic of this thesis in two main parts.

Part A: We identify a curvature-dimension condition $CD_\Upsilon(\kappa,F)$, where $\kappa \in \mathbb{R}$ and F is a CD-function, that serves as a natural non-local replacement of the Bakry-Émery curvature-dimension condition from the diffusive setting. More precisely, $CD_\Upsilon(\kappa,F)$ satisfies the following properties:

A1 The condition relies merely on the difference structure of the respective operators. In particular, it is applicable in a general framework that includes both, the discrete and continuous setting that we have outlined in Section 1.2.

A2 The condition $CD_\Upsilon(\kappa,\infty)$ can be used, analogously as the classical Bakry-Émery condition (cf. Section 1.1.1), to deduce modified logarithmic Sobolev inequalities by means of the entropy method. In particular, the respective operators appear also in the first and second derivative of the corresponding entropy along the heat flow.

A3 Regarding Beckner inequalities we also identify a corresponding curvature-dimension inequality that implies those types of convex Sobolev inequalities by means of the entropy method with power-type entropies taking over the role of the Boltzmann entropy. In contrast to the classical setting, it is natural that different Φ-entropies lead to different conditions depending on the respective variant of the chain rule which comes into consideration.

A4 Our curvature-dimension condition is flexible enough to employ the concept of CD-functions to express the dimension term.

A5 $CD_\Upsilon(\kappa,F)$ implies natural non-local variants of entropy-energy inequalities. In particular, using a quadratic CD-function a natural substitute of the logarithmic entropy-energy inequality can be obtained.

A6 Li-Yau inequalities can be deduced as a consequence of $CD_\Upsilon(0,F)$. In particular, to the best knowledge of the author, CD_Υ serves as the first curvature-dimension condition in the discrete setting that is suitable to deduce both, modified logarithmic Sobolev inequalities under positive curvature and Li-Yau inequalities under finite dimension and non-negative curvature.

A7 Tensorization principles are satisfied. First, our curvature-dimension condition is stable under tensorization of two Markov chain generators. Second, studying hybrid processes, i.e. products of diffusive Markov processes with space-discrete ones, we observe that curvature in the classical sense of Bakry-Émery for the diffusive part is compatible with $CD_\Upsilon(\kappa,\infty)$.

A8 As we have mentioned above, there exist a variety of approaches to define synthetic notions of Ricci-curvature in non-local settings and, in particular, quite a few proposals for replacements of the Bakry-Émery CD-condition. Therefore, we provide and discuss links from several of these approaches to $CD_\Upsilon(\kappa,F)$.

Part B: We identify functional inequalities that correspond to the CD_Υ condition, study their consequences and develop alternative methods for proving them. To be precise, we examine the following items:

B1 While modified logarithmic Sobolev inequalities in the discrete setting are relatively well understood due to the work of [8], discrete variants of entropy-energy inequalities are not studied elsewhere in the literature to the best knowledge of the author. We introduce a class of functional inequalities that we call entropy-information inequalities and establish ultracontractive bounds, finite diameter bounds and a modified version of the celebrated Nash inequality as some of their consequences.

B2 We establish a general reduction principle for deriving Li-Yau type inequalities that works similar to the corresponding proof of the classical Li-Yau inequality that we have presented in Section 1.1.3 in the case of the Euclidean heat equation and reduces the problem to the corresponding heat kernel.

B3 In his recent survey article [18] about the fractional Laplacian, Garofalo dedicates two whole chapters to the open problem whether positive solutions to the fractional heat equation $\partial_t u + \big(-\Delta\big)^{\frac{\beta}{2}} u = 0$ satisfy a Li-Yau inequality. In particular, Garofalo conjectured that the Bakry-Émery condition could play a key role for studying the validity of a Li-Yau inequality for the fractional heat equation. On the one hand, we show that the fractional Laplacian does not satisfy $CD(\kappa,N)$ with $\kappa \in \mathbb{R}$ and $N < \infty$. On the other hand, using our findings described in B2, we can solve the above long-standing open problem by showing that positive solutions to the fractional heat equation satisfy a Li-Yau inequality which has the same time-behavior as the classical one for the Euclidean heat equation (cf. (1.25)).

1.4. Summary

In this section we give a summary of the five research articles of this thesis. To that purpose, we structure the following lines like Section 1.1 with respect to the curvature-dimension condition. We start with positive curvature and related functional inequalities, then discus the improvement of additional dimension terms and, finally, consider the case of non-negative curvature with a finite dimension part and related Li-Yau inequalities.

1.4.1. The entropy method under curvature-dimension conditions in the spirit of Bakry-Émery in the discrete setting of Markov chains

In this article we investigate the research objectives A1, A2, A3 and A7 in the setting of reversible continuous-time Markov chains with state space X and reversible invariant measure μ (cf. Section 1.2.2). Using the non-local chain rule (1.33), we observe that the operator Ψ_Υ is a suitable substitute for Γ since we have

$$\frac{d}{dt}\mathrm{Ent}_\mu(P_t f) = -\int_X P_t f \Psi_\Upsilon(\log(P_t f))\mathrm{d}\mu, \tag{1.35}$$

where the right hand side of the latter identity is nothing else than the negative Fisher information $\mathcal{I}(P_t f)$. As in the classical case, calculating the second derivative of the entropy along the heat flow is much more involved. Using again the chain rule formula (1.33) and that μ is a reversible invariant measure, we show that

$$\frac{d^2}{dt^2}\mathrm{Ent}_\mu(P_t f) = 2\int_X P_t f \Psi_{2,\Upsilon}(\log P_t f)\mathrm{d}\mu, \tag{1.36}$$

where the operator $\Psi_{2,\Upsilon}$ is defined as

$$\Psi_{2,\Upsilon}(f)(x) = \frac{1}{2}\big(L\Psi_\Upsilon(f) - B_{\Upsilon'}\big(f, Lf\big)\big)$$

with

$$B_{\Upsilon'}(u,v)(x) = \sum_{y\in X} k(x,y)\Upsilon'(u(y)-u(x))\big(v(y)-v(x)\big).$$

Defining the curvature-dimension condition $CD_\Upsilon(\kappa,\infty)$ by

$$\Psi_{2,\Upsilon}(f) \geqslant \kappa\Psi_\Upsilon(f),$$

where $\kappa > 0$ and f is a bounded function, we have found a natural substitute of the Bakry-Émery curvature-dimension condition in the discrete setting of Markov chains in terms of the entropy method. Indeed, recalling the general procedure of the entropy method from Section 1.1.1, the modified logarithmic Sobolev inequality (1.15) with constant $\kappa > 0$, or equivalently an exponential decay behavior of the entropy with constant $\kappa > 0$ (cf. (1.14)), follows from $CD_\Upsilon(\kappa,\infty)$ together with the assumption that the chain is positive recurrent (cf. Section 1.2.2) as a corollary from the identities (1.35) and (1.36). We also point out that our method rather depends on the difference structure of (1.27) than on the specific state-space. On a formal level we show that our method also works in the space-continuous setting.

Moreover, we observe that curvature bounds are preserved under tensorization. Indeed, if $CD_\Upsilon(\kappa_i,\infty)$ holds for the Markov generators L_i, $i \in \{1,2\}$, then the generator of the product chain satisfies $CD_\Upsilon(\min\{\kappa_1,\kappa_2\},\infty)$. This important property leads to numerous examples for $CD_\Upsilon(\kappa,\infty)$. For instance, we can show that the Markov generator whose underlying graph is given by the hypercube H_n satisfies $CD_\Upsilon(2,\infty)$ for any $n \in \mathbb{N}$. This is in fact optimal, since $CD(2,\infty)$ is optimal for the hypercube H_n and $CD(\kappa,\infty)$ is necessary for the validity of $CD_\Upsilon(\kappa,\infty)$ (see the lines below). Further, the resulting modified logarithmic Sobolev inequality is best possible in case of the hypercube, which shows that our method is sharp.

We also establish a tensorization principle for hybrid processes. Here, the generator L is the sum of the generator of a positive recurrent reversible continuous-time Markov chain with invariant measure μ_d and state spcae X and the generator of a diffusive Markov process that fits into the classical setting of Section 1.1 with invariant measure μ_c and state space E. We show that $CD(\kappa_1,\infty)$ for the diffusive part and $CD_\Upsilon(\kappa_2,\infty)$ for the discrete part together imply that

the product process satisfies a modified logarithmic Sobolev inequality

$$\mathrm{Ent}_\mu(f) \leqslant \frac{1}{2\min\{\kappa_1,\kappa_2\}}\mathcal{I}(f)$$

for suitably chosen f and with entropy

$$\mathrm{Ent}_\mu(\rho) = \int_E \int_X \rho \log \rho \,\mathrm{d}\mu$$

and Fisher information

$$\mathcal{I}(\rho) = \int_E \int_X \rho(\Gamma \oplus \Psi_\Upsilon)(\log \rho)\mathrm{d}\mu,$$

where μ is the product measure $\mu = \mu_c \otimes \mu_d$ and ρ denotes, respectively, a probability density with respect to μ. This method can be applied to study the long-time behavior of systems of linear PDEs, such as linear reaction-drift-diffusion systems.

As we have discussed in Section 1.3.1, there exist various approaches to formulate discrete notions of Ricci-curvature. With our curvature-dimension condition and its role concerning the entropy method, we have formulated another such candidate. Therefore, it is highly desirable to compare $CD_\Upsilon(\kappa,\infty)$ with existing approaches.

First and foremost, $CD_\Upsilon(\kappa,\infty)$ is stronger than $CD(\kappa,\infty)$ which can be seen as a consequence of a scaling argument and the quadratic behavior of the mapping $r \mapsto \Upsilon(r)$ as $r \to 0$. This simple fact has the important outcome that it is often useful to study first the Bakry-Émery condition in order to get a feeling whether $CD_\Upsilon(\kappa,\infty)$ does hold. We further compare $CD_\Upsilon(\kappa,\infty)$ with $CD(\kappa,\infty)$ along many examples. While for certain underlying graphs, such as for Ricci-flat graphs (e.g. the lattice) and (R)-Ricci flat graphs (e.g. the hypercube), the respective best possible curvature constants coincide, we also provide examples where both concepts differ. For complete graphs, for instance, the optimal curvature constants behave significantly different (see below). Furthermore, the Poisson case of a birth-death process serves as an important example where positive curvature is achieved with regard to the Bakry-Émery condition, while $CD_\Upsilon(0,\infty)$ is best possible on the other hand. Of quite remarkable contrast is the fact that even in a large class of situations with finite state space X, no lower curvature bound exists regarding the CD_Υ condition. Indeed, we establish that if $x, y \in X$ with $k(x,y) > 0$ exist such that

$$\sum_{z \notin \{x,y\}} \big(k(x,z) + k(y,z)\big) > k(x,y) + k(y,x) \tag{1.37}$$

hold, then $CD_\Upsilon(\kappa,\infty)$ fails in x (and due to symmetry also in y) for any $\kappa \in \mathbb{R}$. This criterion has a quite intuitive motivation: If a particle is located in x or y, then (1.37) says that within the next move it is more likely that the particle will use a different edge instead of the one connecting x and y.

In the unweighted case, (1.37) implies that if the underlying graph to L satisfies $CD_\Upsilon(0,\infty)$ and has girth of at least 5 (which means that the smallest cycle in the graph has length at least 5) then it is either a finite or infinite path or a cycle of length at least 5. The corresponding result for the Bakry-Émery condition reads slightly different since also the 3-star falls into the class of unweighted graphs with girth at least 5 and non-negative Bakry-Émery curvature, see [19]. Without having this single outlier, the situation for CD_Υ seems to be more natural. However, the 3-star is a very interesting example on its own as well. We show that there exists a constant $\kappa < 0$ such that the 3-star satisfies $CD_\Upsilon(\kappa,\infty)$ while at the same time $CD_\Upsilon(0,\infty)$ is impossible to hold.

Moreover, using the negative criterion (1.37) we can also demonstrate that CD_Υ is not robust against small perturbations. In fact, adding a single edge to a birth-death process (see the lines

below) with arbitrary small weights can turn positive curvature into a situation where (1.37) applies.

For birth-death processes on the state space $\mathbb{N}_0$ or $\{0, ..., N\}$, $N \in \mathbb{N}$, we show that if monotonicity of the rates $a(x) \leqslant a(x-1)$ and $b(x-1) \leqslant b(x)$ and the condition

$$\sqrt{2\min\{a(x-1)-a(x), b(x)-b(x-1)\}\big(a(x-1)-a(x)+b(x)-b(x-1)\big)} \geqslant \kappa \qquad (1.38)$$

holds true for some $\kappa > 0$ and any $x \in X$, then $CD_\Upsilon(\kappa, \infty)$ is valid. Here we have employed the notation which we have used in our description of the approach of [11] in Section 1.3.1. It is easy to see that the condition (1.38) is stronger than (1.34). This is no surprise since we require the pointwise condition $CD_\Upsilon(\kappa, \infty)$ to be true, which is more than aiming just to estimate the respective integrals as in [11]. Interestingly, our method seems to be finer compared to the approach of [11] since the corresponding constant improves by a factor of 2 if the conditions (1.34) and (1.38) coincide.

Regarding the entropic Ricci-curvature of [16], we recover the objects Ψ_Υ and $\Psi_{2,\Upsilon}$ in the operators $\mathcal{A}$ and $\mathcal{B}$ that characterize their curvature notion (see Section 1.3.1). More precisely, the identities

$$\mathcal{A}(\rho, \log\rho) = \int_X \rho\Psi_\Upsilon(\log\rho)\,\mathrm{d}\mu$$

and

$$\mathcal{B}(\rho, \log\rho) = \int_X \rho\Psi_{2,\Upsilon}(\log\rho)\,\mathrm{d}\mu,$$

show that on one hand, the entropic notion of curvature is stronger since it is characterized by a more general function ψ instead of $\log\rho$, while on the other hand it is weaker since it is of integral form and not a pointwise condition like $CD_\Upsilon(\kappa, \infty)$. It is an interesting open question, that has also been raised in [17], whether $\psi = \log\rho$ is an extreme choice for the validity of $\mathcal{B}(\rho, \psi) \geqslant \kappa\mathcal{A}(\rho, \psi)$. If so, then $CD_\Upsilon(\kappa, \infty)$ would imply the entropic Ricci-curvature. While this question is yet unsolved, we know from the concrete example where the underlying graph is given by the (unweighted) complete graph that the entropic curvature of [16] is not stronger than the CD_Υ condition. In fact, while $CD_\Upsilon(\sqrt{2n}, \infty)$ holds for the Markov generator with the complete graph K_n, $n \geqslant 2$, as the corresponding underlying graph, we can also show that the curvature constant with regard to CD_Υ can not grow linearly as n gets large. This is in contrast to Bakry-Émery curvature, but also to entropic Ricci-curvature after normalization to the setting of [16].

Although the dimension term was not part of this research article, we have investigated important connections to curvature-dimension inequalities that have been established in order to give sufficient conditions for Li-Yau inequalities in the discrete setting. We can show that our condition is equivalent to $CD\psi$ of [28] for the special choice of $\psi = \log$. This is by no means obvious since the operators appearing in Münch's condition are defined in a rather implicit form. Further, neither in [28] nor elsewhere in the literature, the condition of Münch has been studied with respect to consequences of positive curvature bounds. Apart from that, we can also compare our curvature-dimension condition to the approach of [14]. The counterpart of $\Psi_{2,\Upsilon}$ in [14] is in fact nothing else as the sum of $\Psi_{2,\Upsilon}$ and some term that is non-negative if $x \in X$ and f are such that $Lf(x) \leqslant Lf(y)$ for any y in the support of $k(x, \cdot)$ (and this can be assumed in order to deduce Li-Yau estimates by the method of [14]). Defining $CD_\Upsilon(0, F)$ in the natural way (see also the summary of the research article below), these connections to [28] and [14] answer our research goal A6.

Regarding other types of convex Sobolev inequalities, we investigate our research objective A3 by identifying a curvature-dimension condition that implies Beckner inequalities by means of the entropy method with respect to a certain power-type entropy functional (cf. Section 1.1.1).

We show that

$$\frac{d}{dt}\mathrm{Ent}_\mu^{(p)}(P_t f) = -\frac{1}{2-p}\int_X P_t f \Psi_\Upsilon^{(p)}(P_t f)\mathrm{d}\mu \tag{1.39}$$

and

$$\frac{d^2}{dt^2}\mathrm{Ent}_\mu(P_t f) = 2\int_X P_t f \Psi_{2,\Upsilon}^{(p)}(P_t f)\mathrm{d}\mu, \tag{1.40}$$

where the operators $\Psi_\Upsilon^{(p)}$ and $\Psi_{2,\Upsilon}^{(p)}$ are defined for any f that is bounded and bounded away from zero as

$$\Psi_\Upsilon^{(p)}(f)(x) = -\sum_{y\in X} k(x,y)\Lambda_{\Phi_p'}(f(y), f(x)),$$
$$\Psi_{2,\Upsilon}^{(p)}(f) = \frac{1}{2}\Big(L\Psi_\Upsilon^{(p)}(f) - B_{\Upsilon'}(\log f, L\Phi_p'(f))\Big).$$

Here, the expression $\Lambda_{\Phi_p'}$ is defined as above in Section 1.3.1 and $\Phi_p = \frac{r^p - r}{p(p-1)}$, where $r > 0$ and $p \in (1,2)$ (cf. Section 1.1.1). If $\Psi_{2,\Upsilon}^{(p)}(f) \geqslant \frac{\kappa}{2-p}\Psi_\Upsilon^{(p)}(f)$ holds for some $\kappa \in \mathbb{R}$ and any f that is bounded and bounded away from zero, then we say that L satisfies $CD_\Upsilon^{(p)}(\kappa,\infty)$. As a consequence of the identities (1.39) and (1.40), the Beckner inequality (1.18) with constant $\kappa > 0$ follows from $CD_\Upsilon^{(p)}(\kappa,\infty)$ by means of the entropy method, where the power-type entropies take over the role of the Boltzmann entropy.

1.4.2. Entropy-information inequalities under curvature-dimension conditions for continuous-time Markov chains

This work addresses the research objectives A4, A5 and B1 in the setting of reversible continuous-time Markov chains with state space X and reversible invariant measure μ. We introduce the condition $CD_\Upsilon(\kappa, F)$, where F denotes a CD-function (see Section 1.3.1 for the definition) and $\kappa \in \mathbb{R}$ by

$$\Psi_{2,\Upsilon}(f) \geqslant \kappa\Psi_\Upsilon(f) + F_0(-Lf),$$

where F_0 is the trivial extension of F, i.e. $F_0(r) = F(r)$ if $r \geqslant 0$ and 0 otherwise. Since the quadratic CD-function is of particular importance, we use the notation $CD_\Upsilon(\kappa, d)$ if $F(r) = \frac{r^2}{d}$, $r > 0$, for $\kappa \in \mathbb{R}$ and $d \geqslant 1$, to be in line with the classical theory.

We give examples of how to trade-off from the curvature term in order to get finite dimension bounds, e.g., if the complete graph K_n is the underlying graph to L, then L satisfies $CD_\Upsilon\big(\sqrt{2n(1-2\alpha)}, \frac{n}{\alpha}\big)$ for any $\alpha \in \big(0, \frac{1}{2}\big)$. Interestingly, we can observe that for the complete graphs K_n it is not possible to find a dimension bound that is uniform in n while having non-negative curvature. This is in contrast to the Bakry-Émery curvature-dimension condition (see [20]). Apart from the quadratic case we also study examples with other types of CD-functions. (R) Ricci-flat graphs, for instance, satisfy $CD_\Upsilon(2, F)$ where the CD-function F behaves quadratically near zero and grows exponentially at ∞. This example is very interesting, since we do not need a trade off from the (optimal) curvature term in order to get a dimension bound.

We study power-type CD-functions of the form $F_\delta(r) = \frac{r^{1+\delta}}{n}$, where $n > 0$ and $\delta \geqslant 1$ in more detail. It is due to the scaling behavior of $\Psi_{2,\Upsilon}$ that only the range of $\delta \geqslant 1$ comes into consideration. The quadratic case of $\delta = 1$ is seemingly an extreme case, since we can show that under the assumption of positive recurrence the state space is necessarily finite when $CD_\Upsilon(\kappa, F_\delta)$ holds with $\kappa > 0$ and $\delta > 1$.

Regarding research objective B1, we investigate functional inequalities of the form

$$\mathrm{Ent}_\mu(f) \leqslant \Phi(\mathcal{I}(f)), \tag{1.41}$$

where $\Phi : (0,\infty) \to (0,\infty)$ is a strictly increasing and concave C^1-function, to which we refer as the growth function. Since it relates the entropy with the Fisher information we call inequalities of this kind entropy-information inequalities, which is in line with calling the functional inequality (1.23) entropy-energy inequality in the classical context. The general formulation of (1.41) also includes the modified logarithmic Sobolev inequality (cf. (1.15)).

Now, let us bring together the $CD_\Upsilon(\kappa, F)$ condition with the notion of entropy-information inequalities. Under the assumption that L generates a positive recurrent and reversible Markov chain and satisfies $CD_\Upsilon(\kappa, F)$ with a convex CD-function which is C^1 on $(0,\infty)$ and such that there exist some $\delta > 0$ such that $r \mapsto \frac{F(r)}{r^{1+\delta}}$ is increasing, then

$$\mathrm{Ent}_\mu(f) \leqslant \int_0^\infty G\Big(\frac{\kappa}{e^{2\delta\kappa t}\big(1 + \frac{\kappa\mathcal{I}(f)}{F(\mathcal{I}(f))}\big) - 1}\Big)\mathrm{d}t \tag{1.42}$$

holds, where $G : (0,\infty) \to (0,\infty)$ denotes the inverse function of $r \mapsto \frac{F(r)}{r}$ and where f is a positive probability density with respect to the invariant and reversible measure μ with finite entropy and Fisher information. For general CD-functions the integral on the right-hand side of (1.42) can not be calculated explicitly. However, in the class of power-type CD-functions this can be done and we are able to show that (1.42) falls into the class of entropy-information inequalities in this case. For example, if $CD_\Upsilon(\kappa, n)$ holds true, then (1.42) translates to

$$\mathrm{Ent}_\mu(f) \leqslant \frac{n}{2}\log\Big(1 + \frac{\mathcal{I}(f)}{\kappa n}\Big), \tag{1.43}$$

which is apparently a suitable replacement for the logarithmic entropy-energy inequality of the classical Bakry-Émery theory (cf. (1.22)). Indeed, in the classical setting (1.43) and (1.22) are equivalent due to the calculation presented in (1.16).

As we have mentioned above, general entropy-information inequalities have not been studied elsewhere in the literature to the best knowledge of the author. Hence, it is natural to ask whether they have similar consequences as their diffusive counterparts. Having Section 1.1 in mind, we address questions for ultracontractive bounds, finite diameter bounds and Nash inequalities. As for the hypercontractive formulation of modified logarithmic Sobolev inequalites from [8], our ultracontractive bounds come also in the formulation for $e^{P_t f}$ and not for the semigroup itself. In particular, if $CD_\Upsilon(\kappa, n)$ holds true with $\kappa > 0$ and $n \in [1,\infty)$ we deduce that

$$\|e^{P_t f}\|_\infty \leqslant \Big(1 + \frac{1}{2\kappa t}\Big)^{\frac{n}{2}} \|e^f\|_1$$

is valid for any bounded f and $t > 0$. Here and in the remainder of this subsection $\|\cdot\|_p$ denotes the $\ell^p(\mu)$-norm, $p \in [1,\infty]$.

Further, we show that if an entropy-information inequality with growth function Φ holds, then any 1-Lipschitz function f (i.e. f fulfills $\|\Gamma(f)\|_\infty \leqslant 1$) that is integrable with respect to μ satisfies

$$\int_X e^{tf}\mathrm{d}\mu \leqslant \exp\Big(t\Big(\int_0^t \frac{\Phi(s^2)}{s^2}\mathrm{d}s + \int_X f\mathrm{d}\mu\Big)\Big),\ t \neq 0.$$

From that, we deduce the diameter bound $\mathrm{diam}_\varrho \leqslant 2\int_0^\infty \frac{\Phi(s^2)}{s^2}\mathrm{d}s$ (where diam_ϱ is defined as in Section 1.1.2) provided that the growth function is such that the latter integral exists. Most notably, in case of the logarithmic growth function that is related to the $CD_\Upsilon(\kappa, n)$ condition, the resulting diameter bound coincides with the one of [25]. On the one hand, Liu, Münch and Peyerimhoff have only assumed in [25] the weaker condition $CD(\kappa, n)$. On the other hand, their assumptions reduce to the case that the underlying graph is finite, which is seemingly more restrictive.

As in the classical setting, our logarithmic entropy-information inequality also implies a version of a modified Nash inequality, which naturally reads as

$$\|f\|_2^{n+2} \leqslant \Big(\|f\|_2^2 + \frac{\mathcal{I}(f^2)}{\kappa n}\Big)^{\frac{n}{2}} \|f\|_1^2$$

and holds for any non-vanishing function f that is square integrable with respect to μ.

1.4.3. Curvature-dimension inequalities for non-local operators in the discrete setting

Now, we focus on the dimension part only. Analogously as for curvature, the Bakry-Émery condition $CD(0,d)$ is necessary for $CD_\Upsilon(0,d)$. In this article we examine long-range jump operators with state space $\mathbb{Z}$ (see Section 1.2.2) and derive positive and negative results for the validity of $CD(0,d)$. In a wider sense, this article is concerned with the research objectives A6 and A8.

A big advantage of operators of the form (1.30) is that their structure allows for a useful representation formula for the iterated carré du champ operator that reads as

$$\Gamma_2(f)(x) = \frac{1}{4}\sum_{j\neq 0}\sum_{l\neq 0} k_*(j)k_*(l)\big(f(x+j+l) - f(x+j) - f(x+l) + f(x)\big)^2. \tag{1.44}$$

We observe from this formula alone that operators of the form (1.30) satisfy $CD(0,\infty)$. Our aim is now to analyze if and under which conditions these Markov generators satisfy the Bakry-Émery condition $CD(0,d)$ for some finite $d \geqslant 1$. For operators whose kernel has finite support, the (sharp) bound $CD(0,2N)$, $N \geqslant 1$, when $2N$ is the cardinality of the support of k, follows relatively simple. Therefore, our main focus lies on long-range jump operators, i.e. the corresponding Markov chain can reach states that are arbitrary far away from the current state within the next jump. From the graph-theoretical point of view, the resulting underlying graph is locally infinite. In most of the literature that we have discussed in Section 1.3.1 the standing assumptions require the underlying graph to be locally finite, if not finite. As we have already described in Section 1.2.2, the power-type kernel

$$k_*(j) = \frac{1}{|j|^{1+\beta}},$$

where $\beta > 0$, is strongly related to the fractional powers of the discrete Laplacian if $\beta \in (0,2)$. We show that power-type kernels fail to satisfy $CD(0,d)$ if $\beta \in (0,2)$. It follows that fractional powers of the discrete Laplacian do not satisfy $CD(0,d)$ for finite $d < \infty$.

Contrariwise, we show that the condition $CD(0,d)$ is met for some $d < \infty$ if the associated kernel is eventually non-increasing on $\mathbb{N}$ and has a finite second moment (that is $\sum_{j\in\mathbb{N}} k_*(j)j^2 < \infty$). In particular, power-type kernels satisfy $CD(0,d)$ for some $d < \infty$ whenever $\beta > 2$. In this sense, power-type kernels decay sufficiently fast when $\beta > 2$ and are too non-local if $\beta < 2$ for the Bakry-Émery condition $CD(0,d)$ to be valid. The critical case of $\beta = 2$ is yet unsolved.

If we drop monotonicity of the kernel, some interesting phenomena occur. We establish a certain number-theoretic criterion, saying that a sufficiently thin support of the kernel prevents $CD(0,d)$ from being valid if $d < \infty$. This criterion says in particular that the gaps between positive elements of the support of the kernel grow faster than the sequence of the powers of three. For the case when the support of the kernel is exactly given by the set $\{\pm 3^l : l \in \mathbb{N}\}$, we can give, on the other hand, a condition on the kernel such that $CD(0,d)$ holds true for some finite $d < \infty$. Thus, this constitutes an extreme case.

We also address the question of positive Bakry-Émery curvature bounds for Markov generators

of the form (1.30). Formally, positive curvature can not be expected since applying the operators Γ_2 and Γ (cf. (1.32)) to the identity function contradicts positive curvature bounds. However, this argument is only rigorous when the kernel allows for a finite second moment. We are able to extend this result to monotonic kernels by using a delicate approximation argument.

1.4.4. The fractional Laplacian has infinite dimension

The space-continuous variants of the power-type kernels that we have considered in our research article before appear in the pointwise definition of the fractional Laplace operator on $\mathbb{R}^d$ (cf. Section 1.2.1). Being identified as a key open problem towards research objective B3 in Garofalo's recent survey article [18], we aim in this research article to find out whether the negative result for the fractional powers of the discrete Laplacian, that we have described above, can be transferred also to the space-continuous setting. A representation formula like (1.44) can be derived for the iterated carré du champ operator with respect to the fractional Laplacian and reads as

$$\Gamma_2(u)(x) = c_{\beta,d}^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\big(u(x+h+\sigma) - u(x+h) - u(x+\sigma) + u(x)\big)^2}{|h|^{d+\beta}|\sigma|^{d+\beta}} \mathrm{d}h\mathrm{d}\sigma, \tag{1.45}$$

where u belongs to the space of Schwartz functions. We remark that sufficient regularity of the function u is in fact important in order to guarantee that integrals in the intermediary steps towards the derivation of (1.45) do exist. This is one among a few more technical difficulties that appear in the continuous but not in the discrete setting. One of the most significant such differences is due to the singularity of the kernel at 0.

We show that the multidimensional case can be reduced to the one-dimensional situation by considering functions that only depend on one variable. Following the idea of our previously discussed research article, we consider in the one-dimensional case for sufficiently small $\varepsilon > 0$ the function $u_\varepsilon(x) = |x|^{\beta-\varepsilon}$ if $|x| \geqslant 1$ and, in order to treat the singularity, with u_ε vanishing in a neighborhood of 0. We observe that $\Gamma_2(u_\varepsilon)(0) \sim \frac{1}{\varepsilon}$ while $-\big(-\Delta\big)^{\frac{\beta}{2}} u_\varepsilon(0) \sim \frac{1}{\varepsilon}$ as $\varepsilon \to 0$. Consequently, by sending $\varepsilon \to 0$ we see that there can not exist $N < \infty$ such that $CD(0,N)$ holds.

However, in order to calculate $\Gamma_2(u_\varepsilon)$, we have used the representation formula (1.45) while it is not clear whether this formula is in fact valid since u_ε is not a Schwartz function. Therefore, it is important to approximate $\Gamma_2(u_\varepsilon)(0)$ by $\Gamma_2(w_\varepsilon)(0)$ and $-\big(-\Delta\big)^{\frac{\beta}{2}} u_\varepsilon(0)$ by $-\big(-\Delta\big)^{\frac{\beta}{2}} w_\varepsilon(0)$, where w_ε is sufficiently regular.

By a delicate cut-off procedure, we can give a counterexample to the validity of $CD(0,N)$ in the class of compactly supported and smooth functions. A continuity argument and a scaling property of the fractional Laplacian then extends the failure of $CD(0,N)$ to arbitrary large balls around zero. This certainly matters in the continuous setting, since showing that $CD(0,N)$ only fails at one single point is like showing that it fails on a null-set with respect to the invariant measure, which is in case of the fractional Laplacian given by the Euclidean Lebesgue measure on $\mathbb{R}^d$.

Further, we prove that even $CD(-\kappa,N)$ with $\kappa > 0$ and $N < \infty$ is impossible to hold. Summing up, the fractional Laplacian does not satisfy $CD(\kappa,N)$ for any $\kappa \in \mathbb{R}$ and finite $N < \infty$, while $CD(0,\infty)$ holds by (1.45). This justifies to say that the fractional Laplacian has infinite dimension.

1.4.5. Li-Yau inequalities for non-local diffusion equations via reduction to the heat kernel

The result that $CD(0,N)$ fails for the fractional Laplacian is a quite surprising property, since it is in sharp contrast to the Euclidean Laplacian (see Section 1.1.3). Regarding the important

question whether positive solutions to the fractional heat equation satisfy a Li-Yau inequality, this is a setback for the approach via CD-inequalities. Contrariwise, it encourages to pursue the different approach via heat kernel representation of the solution, as presented in Section 1.1.3 for the Euclidean heat equation. This is the content of this research article. More precisely, we investigate research objectives B2 and B3.

We work in the very general context of an operator of the form (1.27), where the state-space X is a metric space. When pursuing the heat kernel approach as presented above in Section 1.1.3, a certain integral estimate involving the operator Ψ_Υ takes the role of the Cauchy-Schwarz inequality. More precisely, if $Pf(x) = \int_X H(x,y)f(y)\mathrm{d}\nu(y)$ for some given function $f : X \to (0,\infty)$, a kernel $H : M \times M \to (0,\infty)$ and a σ-finite Borel measure ν, then we establish the key estimate

$$\int_X \Psi_\Upsilon(\log H(\cdot,y))(x)H(x,y)f(y)\mathrm{d}\nu(y) \geqslant \Psi_\Upsilon(\log Pf)(x)Pf(x). \tag{1.46}$$

For some arbitrary $T > 0$, we consider $u : [0,T) \times X \to (0,\infty)$ with $u(\cdot,x) \in C^1((0,T))$ such that

$$\partial_t u(t,x) = L(u(t,\cdot))(x)$$

holds true for any $(t,x) \in (0,T) \times X$. We set $u_0(\cdot) = u(0,\cdot)$ and assume that u is given by a kernel representation of the form

$$u(t,x) = \int_X p(t,x,y)u_0(y)\mathrm{d}y,$$

where $p : (0,T) \times X \times X \to (0,\infty)$ satisfies certain measurablity conditions and is differentiable with respect to time. Representation formulas of this kind are quite common in the theory of Markov processes, where the kernel $p(t,x,y)$ then plays the role of the transition density.

We apply the key estimate (1.46) with $u(t,x) = Pf(x)$ and $p(t,x,y) = H(x,y)$ for any $t \in (0,T)$. Using the chain rule (1.33), we show by an argument that is similar to the idea presented in Section 1.1.3, that the kernel $p(t,x,y)$ determines the Li-Yau inequality. More precisely, it suffices to show a Li-Yau estimate for the heat kernel, like in the case of the Euclidean heat equation, in order to deduce it for positive solutions to the corresponding non-local diffusion problem.

As an application, we are able to solve the long-standing open problem whether positive solutions to the fractional heat equation satisfy a Li-Yau inequality. By showing the respective estimate for the fractional heat kernel and employing our abstract method, we conclude that for a strong solution u to the fractional heat equation $\partial_t u + (-\Delta)^{\frac{\beta}{2}}u = 0$ on $(0,T) \times \mathbb{R}^d$, where $\beta \in (0,2)$ and $T > 0$, the Li-Yau type estimate

$$\big(-\Delta\big)^{\frac{\beta}{2}}(\log u(t,\cdot))(x) \leqslant \frac{C(\beta,d)}{t},$$

holds true, where $C(\beta,d) > 0$ is some constant. In particular, the time-behavior is like in the Euclidean case.

In order to demonstrate that our method is widely applicable, we make use of it to improve known results for the diffusion problem on a complete graph. We derive the optimal Li-Yau inequality

$$-L\big(\log u(t,\cdot)\big)(x) \leqslant (n-1)\log\Big(\frac{1+(n-1)e^{-nt}}{1-e^{-nt}}\Big), \tag{1.47}$$

where $u : [0,\infty) \times X \to (0,\infty)$ solves $\partial_t u = Lu$ on $(0,\infty) \times X$, X is some finite set consisting of at least two elements and L denotes the Markov generator whose underlying graph is given by the

unweighted complete graph with vertex set X. This estimate in fact improves the corresponding result from [14]. Interestingly, the right-hand side of (1.47) is the relaxation function (see Section 1.3.1) to a function F that satisfies the inequality which defines the curvature-dimension condition of [14], where the latter already has been observed in [14]. However, the function F is not a CD-function and, in order to apply the method of [14], the authors had to find a suitable lower bound for F that is a CD-function. At this step optimality gets lost in [14]. Our result shows that the function F and its relaxation function is tailor-made for deriving Li-Yau estimates in the context of diffusion on a complete graph.

Bibliography

[1] Applebaum, D.: Lévy processes and stochastic calculus. Second edition. Cambridge Studies in Advanced Mathematics, **116**. Cambridge University Press, Cambridge, 2009.

[2] Arnold, A.; Markowich, P.; Toscani, G.; Unterreiter, A.: On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. Comm. Partial Differential Equations **26** (2001), 43–100.

[3] Bakry, D.; Émery, M.: *Diffusions hypercontractives.* Séminaire de probabilités, XIX, 1983/84, 177–206, Lecture Notes in Math., Springer, Berlin, 1985.

[4] Bakry, D.; Gentil, I.; Ledoux, M.: *Analysis and geometry of Markov diffusion operators.* Springer, 2014.

[5] Bakry, Dominique; Ledoux, Michel: A logarithmic Sobolev form of the Li-Yau parabolic inequality. Rev. Mat. Iberoam. **22** (2006)

[6] Bauer, F.; Horn, P.; Lin, Y.; Lippner, G.; Mangoubi, D.; Yau, S.-T.: Li-Yau inequality on graphs. J. Differential Geom. **99** (2015), 359–405.

[7] Beckner, W.: A generalized Poincaré inequality for Gaussian measures. Proc. Amer. Math. Soc. **105** (1989), 397–400.

[8] Bobkov, S. G.; Tetali, P.: Modified logarithmic Sobolev inequalities in discrete settings. J. Theoret. Probab. **19** (2006), 289–336.

[9] Bolley, F.; Gentil, I.: Phi-entropy inequalities for diffusion semigroups. J. Math. Pures Appl. (9) **93** (2010), 449–473.

[10] Bouleau, N.; Hirsch, F.: Dirichlet forms and analysis on Wiener space. De Gruyter Studies in Mathematics, **14**. Walter de Gruyter & Co., Berlin, 1991.

[11] Caputo, P.; Dai Pra, P.; Posta, G.: Convex entropy decay via the Bochner-Bakry-Émery approach. Ann. Inst. Henri Poincaré Probab. Stat. **45** (2009), 734–753.

[12] Ciaurri, Ó.; Roncal, L.; Stinga, P.; Torrea, J. ; Varona, J.: Nonlocal discrete diffusion equations and the fractional discrete Laplacian, regularity and applications Adv. Math. **330** (2018), 688–738.

[13] Diaconis, P.; Saloff-Coste, L.: Logarithmic Sobolev inequalities for finite Markov chains. Ann. Appl. Probab. **6** (1996), 695–750.

[14] Dier, D.; Kassmann, M.; Zacher, R.: Discrete versions of the Li-Yau gradient estimate. To appear in Ann. Sc. Norm. Super. Pisa Cl. Sci. (5), DOI: 10.2422/2036-2145.201807-005.

[15] Erbar, M.: Gradient flows of the entropy for jump processes. Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), 920–945.

[16] Erbar, M.; Maas, J.: Ricci curvature of finite Markov chains via convexity of the entropy. Arch. Ration. Mech. Anal. **206** (2012), 997–1038.

[17] Fathi, M.; Maas, J.: Entropic Ricci curvature bounds for discrete interacting systems. Ann. Appl. Probab. **26** (2016), 1774–1806.

[18] Garofalo, N.: Fractional thoughts. New developments in the analysis of nonlocal operators, 1–135, Contemp. Math., **723**, Amer. Math. Soc., RI, 2019.

[19] Hua, B.; Lin, Y.: Graphs with large girth and nonnegative curvature dimension conditions. Comm. Anal. Geom. **27** (2019), 619–638.

[20] Jost, J.; Liu, S.: Ollivier's Ricci curvature, local clustering and curvature-dimension inequalities on graphs. Discrete Comput. Geom. **51** (2014), 300–322.

[21] Jüngel, A.: *Entropy methods for diffusive partial differential equations*. SpringerBriefs in Mathematics. Springer, 2016.

[22] Latała, R.; Oleszkiewicz, K.: Between Sobolev and Poincaré. Geometric aspects of functional analysis, 147–168, Lecture Notes in Math., **1745**, Springer, 2000.

[23] Li, P.: *Geometric analysis*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2012.

[24] Li, P.; Yau, S.-T.: On the parabolic kernel of the Schrödinger operator. Acta Math. **156** (1986), 153–201.

[25] Liu, S.; Münch, F.; Peyerimhoff, N.: Bakry-Émery curvature and diameter bounds on graphs. Calc. Var. Partial Differential Equations, **57** (2018), no. 2, Art. 67, 9 pp.

[26] Maas, J.: Gradient flows of the entropy for finite Markov chains. J. Funct. Anal. **261** (2011), 2250–2292.

[27] Mielke, A.: Geodesic convexity of the relative entropy in reservible Markov chains. Calc. Var. Partial Diff. Equ. **48** (2013), 1–31.

[28] Münch, F.: Li-Yau inequality on finite graphs via non-linear curvature dimension conditions. J. Math. Pures Appl. (9) **120** (2018), 130–164.

[29] Najman, L., Romon, P. (eds.): *Modern approaches to discrete curvature*. Lecture Notes in Mathematics. Springer, Cham, 2017.

[30] Norris, J.: *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

[31] Spener, A.; Weber, F.; Zacher, R.: Curvature-dimension inequalities for non-local operators in the discrete setting. Calc. Var. Partial Diff. Equ. **58** (2019), no. 5, Art. 171, 30 pp.

[32] Spener, A.; Weber, F.; Zacher, R.: The fractional Laplacian has infinite dimension. Comm. Partial Differential Equations **45** (2020), 57–75.

[33] von Renesse, M.-K.; Sturm, K.-T.: Transport inequalities, gradient estimates, entropy, and Ricci curvature. Comm. Pure Appl. Math. **58** (2005), 923–940.

[34] Weber, F.: Entropy-information inequalities under curvature-dimension conditions for continuous-time Markov chains. Preprint(2020)

[35] Weber, F.; Zacher, R.: The entropy method under curvature-dimension conditions in the spirit of Bakry-Émery in the discrete setting of Markov chains. Preprint (2020).

[36] Weber, F.; Zacher, R.: Li-Yau inequalities for non-local diffusion equations via reduction to the heat kernel. Preprint (2020).

[37] Widder, D. V. Positive temperatures on an infinite rod. Trans. Amer. Math. Soc. **55** (1944), 85–95.

2. The entropy method under curvature-dimension conditions in the spirit of Bakry-Émery in the discrete setting of Markov chains

Authors: Frederic Weber and Rico Zacher

Journal: accepted pending minor revision in Journal of Functional Analysis

arXiv: 2007.01264, arXiv preprint, 2020

Copyright: *This is an Accepted Manuscript of an article published by Elsevier in Journal of Functional Analysis on 1. September 2021, available at: https://doi.org/10.1016/j.jfa.2021.109061.*

THE ENTROPY METHOD UNDER CURVATURE-DIMENSION CONDITIONS IN THE SPIRIT OF BAKRY-ÉMERY IN THE DISCRETE SETTING OF MARKOV CHAINS

FREDERIC WEBER AND RICO ZACHER*

Abstract. We consider continuous-time (not necessarily finite) Markov chains on discrete spaces and identify a curvature-dimension inequality, the condition $CD_\Upsilon(\kappa,\infty)$, which serves as a natural analogue of the classical Bakry-Émery condition $CD(\kappa,\infty)$ in several respects. In particular, it is tailor-made to the classical approach of proofing the modified logarithmic Sobolev inequality via computing and estimating the second time derivative of the entropy along the heat flow generated by the generator of the Markov chain. We prove that curvature bounds in the sense of CD_Υ are preserved under tensorization, discuss links to other notions of discrete curvature and consider a variety of examples including complete graphs, the hypercube and birth-death processes. We further consider power type entropies and determine, in the same spirit, a natural CD condition which leads to Beckner inequalities. The CD_Υ condition is also shown to be compatible with the diffusive setting, in the sense that corresponding hybrid processes enjoy a tensorization property.

Keywords: Markov chain, discrete space, discrete curvature, curvature-dimension inequalities, entropy, modified logarithmic Sobolev inequality, tensorization, Beckner inequality, non-local operator
MSC(2010): 60J27 (primary), 47D07, 39A12 (secondary).

1. Introduction and main results

1.1. **The Γ-calculus by Bakry and Émery.** Curvature-dimension (CD) conditions play a central role in the study of Markov semigroups and operators and related functional inequalities like, e.g., the spectral gap or logarithmic Sobolev inequality (cf. [7]). They encode important information about the geometric properties (curvature and dimension) of the underlying state space X, e.g. a Riemannian manifold, and thus also constitute an important tool in geometric analysis ([37]).

The classical CD condition has been introduced by Bakry and Émery in the pioneering work [6] and is formulated by means of the so-called Γ-calculus. Denoting by L the infinitesimal generator of a Markov semigroup, the carré du champ operator Γ and the iterated carré du champ operator Γ_2 are defined on a suitable algebra of functions as

$$\Gamma(f,g) = \frac{1}{2}\big(L(fg) - fLg - gLf\big),$$
$$\Gamma_2(f,g) = \frac{1}{2}\big(L\Gamma(f,g) - \Gamma(f,Lg) - \Gamma(Lf,g)\big),$$

(Frederic Weber, Rico Zacher) Institut für Angewandte Analysis, Universität Ulm, Helmholtzstrasse 18, 89081 Ulm, Germany.
E-mail addresses: frederic.weber@uni-ulm.de, rico.zacher@uni-ulm.de.
Date: April 10, 2021.
*Corresponding author. F. W. is supported by a PhD-scholarship of the "Studienstiftung des deutschen Volkes", Germany. R. Z. is supported by the DFG (project number 355354916, GZ ZA 547/4-2).

and one sets $\Gamma(f) = \Gamma(f,f)$ and $\Gamma_2(f) = \Gamma_2(f,f)$. The operator L is said to satisfy the curvature-dimension inequality $CD(\kappa, d)$, for $\kappa \in \mathbb{R}$ and $d \in [1, \infty]$, if for every f in a sufficiently rich class of functions $\mathcal{A}$,

$$\Gamma_2(f) \geqslant \frac{1}{d}\left(Lf\right)^2 + \kappa\Gamma(f), \quad \mu - \text{a.e.}, \tag{1}$$

where μ is a fixed invariant and reversible measure on X for the semigroup e^{Lt}.

It is well known that various fundamental results for Markov semigroups can be obtained under a certain CD condition in the Bakry-Émery sense. For example, positive curvature, i.e. $CD(\kappa, \infty)$ with $\kappa > 0$, implies a logarithmic Sobolev inequality, which is equivalent to both exponential decay in entropy and hypercontractivity of the Markov semigroup and can be also used to show the Gaussian measure concentration property for the measure μ, see [7, Chapter 5]. Furthermore, finite dimension and nonnegative curvature, i.e. $CD(0, d)$ with $d < \infty$, imply the Li-Yau inequality, which in turn leads to the parabolic Harnack inequality, cf. [7, Chapter 6].

A key assumption in this rather general and powerful theory, for which the monograph [7] is an excellent source, is the so-called *diffusion property*, which means that the generator L satisfies the chain rule

$$LH(f) = H'(f)Lf + H''(f)\Gamma(f) \tag{2}$$

for every $H \in C^2(\mathbb{R})$ and every sufficiently smooth function f. The diffusion property can be also expressed equivalently via a certain chain rule of the corresponding carré du champ operator. The typical example is a second order differential operator on $\mathbb{R}^d$ (or on a manifold) without zeroth order term. For example, if $L = \Delta_g$ is the Laplace-Beltrami operator on a Riemannian manifold $X = (M, g)$ and μ_g the canonical Riemannian measure, then using the Bochner-Lichnerowicz formula one can show that $CD(\kappa, d)$ is equivalent to $\mathrm{Ric}_g(x) \geqslant \kappa g(x)$ for almost every $x \in M$ and $\dim M \leqslant d$.

If one wants to extend the theory from [7] to the discrete setting of Markov chains, which is highly desirable, then the *failure* of the diffusion property creates a major difficulty. The same problem occurs in case of non-local operators like, e.g., the fractional Laplacian on $\mathbb{R}^d$, where the setting is continuous but the chain rule is violated. Moreover, even though it is still possible to define the operators Γ and Γ_2 and to deduce positive results (like e.g. eigenvalue or diameter estimates) by means of the Γ-calculus (see e.g. [40, 41, 42]), the Bakry-Émery condition $CD(\kappa, d)$, in general, does not have such strong implications or is not applicable as in the diffusion setting, see e.g. [9] and [32, Chapter 5] concerning the Li-Yau inequality on graphs and discrete entropy methods, respectively. These facts suggest that the classical Bakry-Émery CD inequality is not the natural condition in the discrete setting, in particular in situations where the *logarithm* plays a crucial role like in Li-Yau inequalities, entropy or the logarithmic Sobolev inequality.

The problem of finding an appropriate substitute for the classical Γ-calculus in the discrete setting has stimulated a lot of research in the last decade and a half. Several notions of generalized Ricci curvature (lower bounds) for metric measure spaces including discrete spaces have been proposed and studied, see e.g. the recent book [53].

1.2. **Notions of Ricci curvature via optimal transport.** One important approach is based on geodesic convexity of the entropy. It relies on the crucial observation made in [65] that on a Riemannian manifold M, the Ricci curvature is bounded from below by $\kappa \in \mathbb{R}$ if and only if the Boltzmann entropy $\mathcal{H}(\rho) = \int_M \rho \log \rho \, d\mu_g$ is κ-convex along geodesics in the Wasserstein space $P_2(M)$; we also refer to the earlier important contributions [16, 46, 57]. This characterization opened up the synthetic theory of Ricci curvature bounds in geodesic metric measure spaces established independently by Lott-Villani [43] and Sturm [61, 62], see also the monograph [64] and [1, 2, 3, 21].

However, this theory does not apply directly to discrete spaces, since the L^2-Wasserstein space over a discrete space does not contain geodesics. This problem was solved for finite Markov chains by Erbar and Maas [22] and Mielke [47] by replacing the L^2-Wasserstein metric by a different metric $\mathcal{W}$, which had been introduced before by Maas [45] (see also Mielke [47]). The metric $\mathcal{W}$ is constructed in such a way that the heat flow associated with the Markov kernel of the chain is the gradient flow of the entropy with respect to the metric $\mathcal{W}$, thereby obtaining a discrete analogue of the celebrated Wasserstein gradient flow interpretation of the heat flow in $\mathbb{R}^n$ by Jordan, Kinderlehrer and Otto [30]. It is shown that, together with the metric $\mathcal{W}$, the space of probability measures on the finite state space X becomes a geodesic space. Using this property, Erbar and Maas [22] (see also Maas [45]) define lower Ricci curvature bounds by means of geodesic convexity of the entropy and prove a variety of discrete analogues of important results from the continuous setting including that positive curvature implies the (modified) logarithmic Sobolev and Talagrand inequality as well as the Poincaré inequality. They also show that Ricci curvature bounds enjoy a tensorization principle and obtain, as a special case, sharp Ricci curvature lower bounds for the discrete hypercube. Recently, various functional inequalities have also been proved under nonnegative discrete Ricci curvature and some suitable additional conditions ([20]). By means of different methods, discrete entropic Ricci curvature bounds have been derived for several examples, e.g. one-dimensional birth-death processes, Bernoulli-Laplace models, zero-range processes and random transposition models, see [23, 24, 47]. We also refer to the survey [44] and the references given therein.

A different approach to adapt the Lott-Sturm-Villani theory to the discrete setting has been proposed by Bonciocat and Sturm in [13] and is based on the notion of approximate midpoints w.r.t. the L^2-Wasserstein metric. This notion is also studied in [56] in connection with a Brunn-Minkowski inequality on the discrete hypercube.

Another important notion of Ricci curvature including the discrete case has been introduced by Ollivier [55]. This approach also relies on optimal transport and makes use of the L^1-Wasserstein metric. Positive Ollivier Ricci curvature also implies certain functional inequalities. In contrast to the entropic bounds in [22], Ollivier's condition does not require reversibility. As shown in [55], his criterion can be checked without much effort in many examples. It does not seem to be known whether Ollivier's notion is connected to the entropic one. For further studies of Ollivier's Ricci curvature in the graphical setting we refer to [27, 31, 40, 52].

1.3. **CD inequalities in the context of discrete Li-Yau gradient estimates.** Recently, motivated mainly by the problem of finding a discrete version of the celebrated Li-Yau inequality ([38]), various modifications of the classical Bakry-Émery CD condition have been introduced ([9, 19, 50]). We recall that positive solutions u of the heat equation $\partial_t u - \Delta u = 0$ on $(0,\infty) \times \mathbb{R}^d$ satisfy the (sharp) Li-Yau inequality $-\Delta \log u \leqslant \frac{d}{2t}$, which is equivalent to the *differential Harnack inequality*

$$\partial_t(\log u) \geqslant |\nabla(\log u)|^2 - \frac{d}{2t} \quad \text{in } (0,\infty) \times \mathbb{R}^d. \tag{3}$$

By integration along a suitable path in space-time, the gradient estimate (3) yields a sharp parabolic Harnack estimate. Corresponding results hold true for the heat equation on a manifold with nonnegative Ricci curvature, see [37, 38].

First results concerning Li-Yau inequalities on graphs were obtained by Bauer, Horn, Lin, Lippner, Mangoubi, and Yau [9]. Their key idea to circumvent the problem of the failure of the chain rule for the logarithm is to consider the *square root* instead of the logarithm and to use

the identity

$$L\sqrt{f} = \frac{Lf}{2\sqrt{f}} - \frac{\Gamma(\sqrt{f})}{\sqrt{f}}, \tag{4}$$

which still holds true for the graph Laplacian L. Furthermore, the authors of [9] introduce a new notion of CD-inequality, the so-called *exponential curvature-dimension inequality* $CDE(\kappa, d)$, which means that for any vertex x and any positive function f defined on the set of vertices X satisfying $(Lf)(x) < 0$ one has

$$\tilde{\Gamma}_2(f)(x) := \Gamma_2(f)(x) - \Gamma\Big(f, \frac{\Gamma(f)}{f}\Big)(x) \geqslant \frac{1}{d}(Lf)(x)^2 + \kappa\,\Gamma(f)(x). \tag{5}$$

By means of (4) and their new notion of CD-inequality the authors of [9], using maximum principle arguments, obtain various Li-Yau type results for graphs. They also introduce a different type of exponential CD-inequality, the condition $CDE'(\kappa, d)$, which holds if for all positive functions f on X,

$$\tilde{\Gamma}_2(f) \geqslant \frac{1}{d} f^2 (L \log f)^2 + \kappa\,\Gamma(f).$$

In [26], a family of Li-Yau type estimates similar to the ones in [9] (involving again the square root) are derived under the condition $CDE'(0, d)$, now using semigroup techniques as in [7]. Moreover, volume doubling, the Poincaré inequality and Gaussian heat kernel estimates are proved for non-negatively curved graphs.

A few years later, Münch [50, 51] introduced a new calculus, called Γ^ψ-calculus, where ψ is a concave function. Assuming that a *finite* graph satisfies the so-called $CD\psi(d, 0)$-condition, Münch proves a Li-Yau estimate which for $\psi = \sqrt{\ }$ becomes one of those obtained in [9] and which in case $\psi = \log$ leads to the more natural logarithmic estimate $-L \log u \leqslant \frac{d}{2t}$, for positive solutions u of the heat equation on the graph. Münch shows that in some examples (e.g. Ricci-flat graphs) the latter choice gives better estimates than the square-root approach.

Very recently, Dier, Kassmann and Zacher [19] proposed another CD inequality in the graphical setting, the condition $CD(F; 0)$, which allows to significantly improve the Li-Yau estimates in [9, 50]. For example, they obtain a sharp logarithmic Li-Yau estimate for the unweighted two-vertex graph and prove a Li-Yau inequality for the lattice $\tau\mathbb{Z}$ with grid size τ which for $\tau \to 0$ leads to the classical sharp Li-Yau inequality on $\mathbb{R}$.

One of the new key ideas in [19] is to replace the square $(Lf)^2$ in the CD-inequality by a more general term of the form $F(-Lf)$, where F is a so-called *CD-function*. Their Li-Yau inequality for positive solutions u of the heat equation on $[0, \infty) \times X$ (in case of a finite graph satisfying $CD(F; 0)$) takes the form $-L \log u \leqslant \varphi(t)$, where the *relaxation function* φ is the unique strictly positive solution of the ODE $\varphi'(t) + F(\varphi(t)) = 0$ which has $(0, \infty)$ as its maximal interval of existence. In all examples studied in [19], φ exhibits a logarithmic behaviour near $t = 0$ and thus is integrable at $t = 0$ (in contrast to the $\frac{c}{t}$ estimates in [9, 50]), which makes it possible to deduce Harnack inequalities without time gap at $t = 0$.

1.4. **A discrete analogue of the $CD(\kappa, \infty)$ condition that fits to the entropy method.** The main purpose of the present paper is to identify a CD inequality in the discrete setting of Markov chains which serves as a natural analogue of the classical condition $CD(\kappa, \infty)$ in several respects, most notably with regard to the Bakry-Émery strategy to prove decay of the entropy (and the (modified) logarithmic Sobolev inequality) via computing and estimating the second time derivative of the entropy along the heat flow generated by the generator L of the Markov chain.

Before summarizing the main results, we describe the setting and some notation used in this paper. We consider a time-homogeneous Markov process $(Z_t)_{t\geqslant 0}$ on a finite or countable state space X, equipped with the maximal σ-algebra. The generator L of the Markov process, defined for a suitable class of functions $f : X \to \mathbb{R}$, is given by

$$Lf(x) = \sum_{y\in X} k(x,y)f(y) = \sum_{y\in X} k(x,y)\big(f(y)-f(x)\big), \quad x \in X, \tag{6}$$

where $k(x,y) \geqslant 0$ is the transition rate for jumping from x to y ($\neq x$) and $\sum_{y\in X} k(x,y) = 0$ for all $x \in X$, that is $k(x,x) = -\sum_{y\in X\setminus\{x\}} k(x,y) < 0$, $x \in X$. The Markov chain is called irreducible if for any $x, y \in X$ we find $z_1, ..., z_m \in X$, $m \geqslant 2$, with $x = z_1$ and $y = z_m$ such that $k(z_i, z_{i+1}) > 0$ for any $i \in \{1, ..., m-1\}$. We denote the associated (Sub-)Markov semigroup on the space of bounded functions on X by $\mathbf{P} = (P_t)_{t\geqslant 0}$, which is given by

$$P_t f(x) = \mathbb{E}\big(f(Z_t)|Z_0 = x\big). \tag{7}$$

We recall that a σ-finite measure μ on X is said to be invariant for the semigroup $\mathbf{P}$ if for any bounded function $f : X \to [0,\infty)$, $\int_X P_t f \, d\mu = \int_X f \, d\mu$, $t \geqslant 0$. In the recurrent case this is known to be equivalent to $\sum_{x\in X} \mu(\{x\})k(x,y) = 0$. Furthermore, a σ-finite measure μ on X is said to be reversible if the so-called *detailed balance condition*

$$\mu(\{x\})k(x,y) = \mu(\{y\})k(y,x), \quad x, y \in X, \tag{8}$$

is satisfied. It is known that reversibility implies invariance. In particular, under reversibility $\int_X Lf d\mu = 0$ holds true on a suitable class of functions (e.g. on bounded functions if μ is finite). For more details on the general theory of continuous-time Markov chains we refer to [4] and [54].

The starting point in our search for suitable substitutes for Γ and Γ_2 is the identity

$$L(\log f) = \frac{1}{f} Lf - \Psi_\Upsilon(\log f) \tag{9}$$

for positive functions f, which has been observed in [19]. Here the operator Ψ_Υ is defined by

$$\Psi_\Upsilon(f)(x) = \sum_{y\in X} k(x,y)\Upsilon\big(f(y)-f(x)\big), \quad x \in X,$$

with the function $\Upsilon(r) = e^r - 1 - r$, $r \in \mathbb{R}$, see also Lemma 2.2 below. Comparing (9) with the corresponding identity in the diffusion setting, one sees that $\Psi_\Upsilon(\log f)$ replaces $\Gamma(\log f)$ in the discrete setting. In other words, the quadratic function $\frac{1}{2}r^2$, which represents the (discrete) Γ-operator, is replaced with the function $\Upsilon(r)$, which interestingly for $r \to 0$ behaves like $\frac{1}{2}r^2$. The same phenomenon can be observed with the Fisher information $\mathcal{I}(\rho)$ of a probability density ρ w.r.t. to a fixed invariant and reversible probability measure μ on X. In fact, in the discrete case we have $\mathcal{I}(\rho) = \int_X \rho\Psi_\Upsilon(\log\rho)\, d\mu$, cf. Section 3 below.

Following the Bakry-Émery strategy, i.e. computing the second time derivative of the entropy, it turns out, as we show, that one can proceed further and, by means of detailed balance and certain *algebraic identities for differences*, find a natural analogue of Γ_2, too, namely

$$\Psi_{2,\Upsilon}(f) = \frac{1}{2}\big(L\Psi_\Upsilon(f) - B_{\Upsilon'}(f, Lf)\big),$$

where

$$B_{\Upsilon'}(f,g)(x) = \sum_{y\in X} k(x,y)\Upsilon'\big(f(y)-f(x)\big)\big(g(y)-g(x)\big).$$

Motivated by these findings, we introduce the CD condition $CD_\Upsilon(\kappa,\infty)$ with $\kappa \in \mathbb{R}$ via the inequality $\Psi_{2,\Upsilon}(f) \geqslant \kappa\Psi_\Upsilon(f)$ on X, $f \in \ell^\infty(X)$. That this notion is indeed a suitable discrete version of the classical $CD(\kappa,\infty)$ condition will be demonstrated in this paper by proving a

variety of important results that parallel those in the diffusion setting. In particular, we will show the following.

- $CD_\Upsilon(\kappa,\infty)$ with $\kappa>0$ implies the *modified logarithmic Sobolev inequality* with constant κ, see Corollary 3.5.
- $CD_\Upsilon(\kappa,\infty)$ is characterized by the *gradient bounds* $\Psi_\Upsilon(P_t f)\leqslant e^{-2\kappa t}P_t(\Psi_\Upsilon(f))$, cf. Remark 2.10.
- Curvature bounds are preserved under *tensorization*, more precisely $CD_\Upsilon(\kappa_i,\infty)$ for two chains ($i=1,2$) implies $CD_\Upsilon(\kappa,\infty)$ with $\kappa=\min\kappa_i$ for the product chain (Theorem 4.2).

See also the beginning of Section 2 for the underlying basic assumptions.

Moreover, by a scaling argument, one can also see that $CD_\Upsilon(\kappa,\infty)$ implies $CD(\kappa,\infty)$ (defined as in Section 1.1 via (1)), see Proposition 2.11. If the state space X is *finite* and $CD(\kappa_0,\infty)$ and $CD_\Upsilon(\kappa,\infty)$ hold true with $\kappa_0>\kappa>0$, then for any $\varepsilon>0$ we will prove that the entropy $\mathcal{H}(P_t f)$ is bounded above by $C\exp(-2(\kappa_0-\varepsilon)t)$ for all $t\geqslant 0$, where the (explicitly known) constant C only depends on ε, κ_0, κ, the kernel k, the invariant measure and the entropy of f, cf. Theorem 3.6. So the rate of exponential convergence is 'almost' that given by the Bakry-Émery constant κ_0.

In Section 5 we discuss several examples. We show positive curvature bounds with respect to the CD_Υ condition, e.g. for the two-point space, the complete graph, the hypercube and some birth-death processes on $\mathbb{N}_0$. In case of the hypercube, we will further see that the CD_Υ condition yields the optimal constant in the resulting modified logarithmic Sobolev inequality. Moreover, we compare the CD_Υ condition with Bakry-Émery curvature along many examples. While, for instance, both curvature notions coincide for the hypercube or complete bipartite graphs, the Poisson case of the birth-death process on $\mathbb{N}_0$ serves as an example with positive Bakry-Émery curvature for which $CD_\Upsilon(0,\infty)$ is best possible. We give a quite applicable condition to show $CD_\Upsilon(\kappa,\infty)$, with $\kappa>0$, for birth-death processes which is on one hand stronger than the assumption of [14], but improves on the other hand the constant in the corresponding modified logarithmic Sobolev inequality by a factor 2. Further, we give a quite general criterion that ensures that no (negative) lower curvature bound exists regarding the CD_Υ condition for a large class of examples among which are, for instance, unweighted trees that are no (finite or infinite) paths and have at least 5 vertices. This is in huge contrast to the Bakry-Émery case. The latter criterion allows to show that the CD_Υ condition is not robust with regard to small perturbations of the transition rates and, moreover, to characterize unweighted graphs with girth at least 5 satisfying $CD_\Upsilon(0,\infty)$.

Another striking feature of the CD_Υ condition is its compatibility with the diffusion setting. To be more precise, we will establish a *tensorization principle* for *hybrid processes*, that is, tensor products of a process in the diffusion setting and a (discrete) Markov chain. If $CD(\kappa_c,\infty)$ holds for the continuous part and $CD_\Upsilon(\kappa_d,\infty)$ for the Markov chain, then for the product process we will derive the lower curvature bound $\kappa=\min\{\kappa_d,\kappa_c\}$ meaning that the inequality

$$\big(\Gamma\oplus\Psi_\Upsilon\big)_2(f)\geqslant\kappa\big(\Gamma\oplus\Psi_\Upsilon\big)(f)$$

is satisfied for all f in a sufficiently rich class of functions defined on the product space. Here $\Gamma\oplus\Psi_\Upsilon$ and $\big(\Gamma\oplus\Psi_\Upsilon\big)_2$ are natural analogues of Γ and Γ_2, respectively, in the hybrid continuous-discrete setting, see Section 7. We will further show that positive curvature in the hybrid sense implies a corresponding modified logarithmic Sobolev inequality and exponential decay of the entropy. As a possible application we discuss systems of linear reaction-diffusion systems.

The logarithmic Sobolev and the Poincaré inequality are in some sense extreme cases of a more general family of *convex Sobolev inequalities*, see [5, 32]. Whereas in the diffusive setting, both

inequalities follow from $CD(\kappa,\infty)$ with $\kappa > 0$, this does not seem to be known in the discrete setting for the (modified) logarithmic Sobolev inequality; but it is true for the Poincaré inequality (see e.g. [36, Section 3]). Among those convex Sobolev inequalities, Beckner inequalities are also of high significance. Here the corresponding entropy is generated by a power type function with exponent between 1 and 2. Having identified a natural CD condition which is sufficient for the modified logarithmic Sobolev inequality we will also determine such a CD condition for power type entropies and derive the corresponding Beckner inequalities, cf. Section 6. In order to succeed, we follow again the Bakry-Émery strategy and derive a suitable representation for the second time derivative of the power type entropy along the heat flow.

As our approach to discrete CD inequalities is merely based on certain algebraic rules for differences, it can be naturally extended to a large class of *non-local operators* in a continuous setting. We will illustrate this in Section 8 for operators of the form

$$\mathcal{L}f(x) = \int_\Omega \big(f(y) - f(x)\big)k(x,dy), \quad x \in \Omega,$$

where $\Omega \subset \mathbb{R}^d$ is a domain and the integral may possibly have to be understood in the principal value sense. A prominent example for such operators is given by the fractional Laplacian on $\mathbb{R}^d$, which satisfies the condition $CD_\Upsilon(0,\infty)$.

1.5. **Links to other notions of (discrete) curvature bounds.** It turns out that there are quite a few close relations of the $CD_\Upsilon(\kappa,\infty)$ condition to other notions of (discrete) curvature and CD inequalities. Firstly, as we will prove, $CD_\Upsilon(\kappa,\infty)$ is equivalent to Münch's condition $CD\psi(\infty,\kappa)$ with $\psi = \log$, which is introduced in a side remark in [50, Remark 3.12] in the special case of the discrete Laplacian on a finite unweighted graph and which seemingly has not been used at all neither in [50] nor elsewhere in the literature. Actually, at the same place, Münch also defines the condition $CD\psi(d,\kappa)$ with finite dimension parameter $d > 0$, which for $\psi = \log$, is equivalent to the inequality

$$\Psi_{2,\Upsilon}(f) \geqslant \frac{1}{d}(Lf)^2 + \kappa\Psi_\Upsilon(f), \tag{10}$$

which we will refer to as condition $CD_\Upsilon(\kappa,d)$, cf. Remark 2.9. We emphasize that the equivalence of (10) and $CD\log(\infty,\kappa)$ is by no means easy to see as the involved quantities in Münch's condition are defined in a highly implicit fashion, see Remark 2.9(i),(ii) and Section 9 for more details.

Further, condition $CD_\Upsilon(0,d)$ is also closely related to the condition $CD(F;0)$ by Dier, Kassmann and Zacher [19] with a quadratic CD-function F. Instead of $\Psi_{2,\Upsilon}(f)(x)$, the authors of [19] use a sum of the latter term and an additional term which is nonnegative if one restricts the class of functions f to those where Lf has a 'local' minimum at the vertex x, see Remark 2.9(iii) for the precise formulation.

Note that these references show that the condition $CD_\Upsilon(0,d)$ (and more general versions with a CD-function F in the dimension term) is suitable to derive *Li-Yau and Harnack inequalities*. Thus, like the classical Bakry-Émery Γ-calculus in the diffusive setting, the CD_Υ-calculus is powerful enough to obtain *both* modified logarithmic Sobolev inequalities and entropy decay as well as Li-Yau and Harnack inequalities. To our knowledge, this is not known for other approaches to (discrete) CD conditions in particular for those that rely on optimal transport.

There is also an interesting connection to the generalized Γ-calculus which has been recently introduced by Monmarché in the context of *degenerate Markov processes* like e.g. *kinetic* Fokker-Planck diffusion in $\mathbb{R}^{2d}$ and chains of interacting particles ([48, Section 2], see also [49]). It is known that for such processes the classical methods based on the Γ-calculus do not work in the usual way. To overcome this problem, Villani [63] proposed to study certain *distorted quantities*

when the classical quantities like entropy or variance do not behave well under the action of the semigroup, see also [8]. Monmarché's basic strategy also follows the Bakry-Émery approach, but now allowing for a wider class of functionals and operators, respectively. His motivation is to provide a general framework where the computations can be done more easily. We will show that (in the discrete setting) for a special choice of Monmarché's Φ-operator, his general CD inequality amounts to the condition $CD_\Upsilon(\kappa,\infty)$, see Remark 2.9(iv). However, operators like Ψ_Υ, $\Psi_{2,\Upsilon}$ or the discrete Boltzmann entropy are apparently not the subject of the investigations in [48, 49] as the author's focus lies on degenerate processes in a continuous setting.

In the case of finite Markov chains, it is not clear whether there is a direct implication between the condition $CD_\Upsilon(\kappa,\infty)$ and the notion of lower Ricci curvature bounds by Erbar and Maas [22] and Mielke [47] defined via geodesic convexity. But what we can say is that there is no $\alpha > 0$ and $\beta > -\frac{\alpha}{2}$ such that the entropic Ricci curvature of [22] with constant $\kappa > 0$ does imply $CD_\Upsilon(\alpha\kappa+\beta,\infty)$, see Remark 5.5. Further, we know that both conditions, those of [22] with constant κ and $CD_\Upsilon(\kappa,\infty)$, imply the classical Bakry-Émery condition $CD(\kappa,\infty)$, see [53, Chapter 5] respectively Proposition 2.11. Moreover, there is an interesting link between both notions. Erbar and Maas introduce two functionals $\mathcal{A}(\rho,\psi)$ and $\mathcal{B}(\rho,\psi)$ for probability densities ρ on the state space X and functions $\psi \in \mathbb{R}^X$, cf. also Section 9. They show that the chain has curvature bounded from below by κ (in their sense) if and only if the inequality $\mathcal{B}(\rho,\psi) \geqslant \kappa\mathcal{A}(\rho,\psi)$ holds for all admissible ρ and ψ, see [22, Theorem 4.5]. It turns out that

$$\mathcal{A}(\rho,\log\rho) = \int_X \rho\Psi_\Upsilon(\log\rho)\,d\mu$$

and

$$\mathcal{B}(\rho,\log\rho) = \int_X \rho\Psi_{2,\Upsilon}(\log\rho)\,d\mu,$$

see Section 9. This shows that on one hand, the entropic condition is stronger with regard to the general function ψ instead of $\log\rho$. On the other hand, it is weaker than $CD_\Upsilon(\kappa,\infty)$ in the sense that it is of integral form, whereas $CD_\Upsilon(\kappa,\infty)$ is a *pointwise condition*, which has to hold at any $x \in X$.

The paper is organized as follows. In Section 2 we motivate and introduce the curvature-dimension condition $CD_\Upsilon(\kappa,\infty)$. We demonstrate that this notion fits well to the discrete entropy method and deduce modified logarithmic Sobolev inequalities or, equivalently, an entropy decay of exponential rate in Section 3. Afterwards, we show in Section 4 that a tensorization property holds for $CD_\Upsilon(\kappa,\infty)$. Section 5 then illustrates the content with several examples. The remaining part of the paper is mainly concerned with related situations to what we have studied before. In Section 6, we replace the Boltzmann entropy by power type entropies, introduce corresponding curvature-dimension conditions and deduce Beckner inequalities. After that, in Section 7, we study hybrid processes, that is, products of processes satisfying the diffusion property and those with discrete state space. Section 8 is devoted to an outlook how the CD_Υ-calculus can be extended to non-local operators in the space-continuous setting. Finally, we conclude with Section 9, where we provide more details on some links between the $CD_\Upsilon(\kappa,\infty)$ condition and other existing approaches in the literature.

2. Setting, basic identities and CD inequalities

Let us begin this section by introducing our setting. We consider a time-homogeneous irreducible continuous-time Markov chain on a countable state space X with unique (up to positive multiples) invariant measure μ whose generator is given by (6). Let $\pi : X \to (0,\infty)$ denote the density for μ with respect to the counting measure, i.e. $d\mu = \pi d\#$. We assume that the detailed balance condition (8) holds true.

Let $\mathbb{R}^X$ denote the space of real-valued functions on X. We say that $f \in \ell^p(\mu)$, $1 \leqslant p < \infty$, if $f \in \mathbb{R}^X$ is p-summable with respect to μ and $f \in \ell^\infty(X)$ if $f \in \mathbb{R}^X$ is bounded. Further, $(P_t)_{t\geqslant 0}$ denotes the (Sub-)Markov semigroup given by (7).

Throughout this paper we assume that for any $x \in X$

$$M_1(x) := \sum_{y\in X\setminus\{x\}} k(x,y) < \infty \tag{11}$$

and

$$M_2(x) := \sum_{y\in X\setminus\{x\}} k(x,y) \sum_{z\in X\setminus\{y\}} k(y,z) < \infty \tag{12}$$

holds true.

In Section 3, 6 and 7 we will assume that the Markov chain is positive recurrent, i.e. we assume in addition that μ is a probability measure and that stochastic completeness holds (that is $P_t\mathbb{1} = \mathbb{1}$) or in other words that the Markov chain is non-explosive. It is known that the latter holds automatically whenever $\sup_{x\in X} M_1(x) < \infty$, see e.g. [54].

Due to the detailed balance condition, the generator of the Dirichlet form given by

$$\mathcal{E}(f,g) = \frac{1}{2}\sum_{x\in X}\sum_{y\in X} k(x,y)\big(f(y)-f(x)\big)\big(g(y)-g(x)\big)\pi(x) \tag{13}$$

for f, g lying in the form domain, coincides with L given by (6) on bounded functions which are contained in the domain of the form generator, see [34]. Therefore, we will also denote the $\ell^2(\mu)$ operator with L in the sequel. The corresponding $\ell^2(\mu)$-semigroup generated by L is an extension of the semigroup given by (7) restricted to $\ell^\infty(X)\cap\ell^2(\mu)$. We will also use the notation $(P_t)_{t\geqslant 0}$ for the corresponding $\ell^2(\mu)$-semigroup.

L determines in a natural way a graph structure with vertex set X and edge weights given by $k(x,y)$ for $x, y \in X$, $x \neq y$. We say that x, y are adjacent if $k(x,y) > 0$. Note that irreducibility and the detailed balance condition together imply that $k(x,y) > 0$ if and only if $k(y,x) > 0$. We will refer to the described graph structure as the underlying or associated graph to L. If $k(x,y) \in \{0,1\}$ for any $x, y \in X$ with $x \neq y$, then the underlying graph to L is an unweighted graph.

We bring back to mind the definitions of the carré du champ operator Γ and the iterated carré du champ operator Γ_2. Due to the conditions (11) and (12), we can define Γ and Γ_2 on the class of bounded functions by the formulas given in Section 1.1. Moreover, a straightforward computation leads to the well known representation formula

$$\Gamma(f,g)(x) = \frac{1}{2}\sum_{y\in X} k(x,y)\big(f(y)-f(x)\big)\big(g(y)-g(x)\big), \tag{14}$$

where $x \in X$ and $f, g \in \ell^\infty(X)$.

Now we recall a basic identity, which can be viewed as a kind of chain rule for the discrete operator L and we refer to as the *(first) fundamental identity*, cf. Section 2 in [19].

Lemma 2.1. *Let $\Omega \subset \mathbb{R}$ be an open set and $f : X \to \mathbb{R}$ such that the range of f is contained in Ω. Let further $H \in C^1(\Omega;\mathbb{R})$ and $f, H(f) \in \ell^1\big(k(x,\cdot)\big)$ for $x \in X$. Then there holds*

$$L\big(H(f)\big)(x) = H'(f(x))Lf(x) + \sum_{y\in X} k(x,y)\Lambda_H\big(f(y),f(x)\big) \tag{15}$$

where

$$\Lambda_H(w,z) := H(w) - H(z) - H'(z)(w-z), \quad w, z \in \mathbb{R}. \tag{16}$$

This follows from a straightforward computation, cf. [19, Lemma 2.1]. In the case of a strictly convex function H, the quantity $\Lambda_H(w,z)$ is called *Bregman distance* associated with H for the points w and z. Note that identity (15) can be regarded as an analogue in the discrete setting of the chain rule for Markov diffusion operators which reads

$$LH(f) = H'(f)Lf + H''(f)\Gamma(f). \tag{17}$$

It turns out that for several important examples of the function H, the second term on the right-hand side of (15) can be rewritten in a form that provides more insights into the structure of the term. To this purpose it is useful to define, for a given function $H : \mathbb{R} \to \mathbb{R}$, the operator

$$\Psi_H(f)(x) = \sum_{y\in X} k(x,y)H\big(f(y) - f(x)\big), \quad x \in X, \tag{18}$$

where $f \in \mathbb{R}^X$ is such that the latter is well defined. Clearly, $\Psi_H(f) = Lf$ if H is the identity. Observe as well that in case of $\Omega = \mathbb{R}$ and $H(y) = \frac{1}{2}y^2$ we have $\Psi_H(f) = \Gamma(f)$, and thus (15) leads to the well known identity

$$\frac{1}{2}\,L(f^2) = fLf + \Gamma(f). \tag{19}$$

The following identity plays a central role in our approach. It is a discrete version of the identity

$$L(\log f) = \frac{1}{f}\,Lf - \Gamma(\log f) \tag{20}$$

from the diffusion setting and can be found in [19, Example 2.4].

Lemma 2.2. *For positive functions $f \in \mathbb{R}^X$ such that $f, \log f \in \ell^1\big(k(x,\cdot)\big)$ for any $x \in X$, there holds*

$$L(\log f) = \frac{1}{f}\,Lf - \Psi_\Upsilon(\log f), \tag{21}$$

where

$$\Upsilon(y) := e^y - 1 - y, \quad y \in \mathbb{R}.$$

Proof. For the reader's convenience we repeat the short calculation, in which we apply Lemma 2.1 with $\Omega = (0,\infty)$ and $H(y) = \log y$. In this case we have

$$\Lambda_H(w,z) = \log w - \log z - \frac{1}{z}\,(w - z) = -\Upsilon(\log w - \log z),$$

and thus (21) follows from (15). □

Given a function $H : \mathbb{R} \to \mathbb{R}$, we next define the operator B_H by

$$B_H(f,g)(x) = \sum_{y\in X} k(x,y)H\big(f(y) - f(x)\big)\big(g(y) - g(x)\big), \quad x \in X, \tag{22}$$

where $f, g \in \mathbb{R}^X$ are such that the latter is well defined. Provided that H is continuous this is the case if $f, g \in \ell^\infty(X)$ by (11), as well as for the quite important choice of $g = Lf$ with $f \in \ell^\infty(X)$ due to (12). Observe further that $B_H(f,g) = 2\Gamma(f,g)$ in case of H being the identity.

The next result provides a fundamental identity for the state space $X = \mathbb{Z}^d$ in case of translation invariant kernels.

Lemma 2.3. *Let $d \in \mathbb{N}$, $H \in C^1(\mathbb{R})$ and suppose that the kernel $k : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}$ is of the form*

$$k(x,y) = k_*(y - x), \quad x, y \in \mathbb{Z}^d,\ x \neq y, \quad k(x,x) = -\sum_{y\in\mathbb{Z}^d\setminus\{0\}} k_*(y), \quad x \in \mathbb{Z}^d,$$

where k_* is nonnegative on $\mathbb{Z}^d\backslash\{0\}$ and $k_* \in \ell^1(\mathbb{Z}^d\backslash\{0\})$. Let L be the Markov generator associated with the kernel k, the state space being $X = \mathbb{Z}^d$, and let $f \in \ell^\infty(\mathbb{Z}^d)$. Then for all $x \in \mathbb{Z}^d$,

$$L\big(\Psi_H(f)\big)(x) = B_{H'}(f, Lf)(x) + \sum_{h,\sigma\in\mathbb{Z}^d\backslash\{0\}} k_*(h)k_*(\sigma)\Lambda_H\big(f(x+h+\sigma) - f(x+h), f(x+\sigma) - f(x)\big). \tag{23}$$

In particular, if H is convex,

$$L\big(\Psi_H(f)\big)(x) - B_{H'}(f, Lf)(x) \geqslant 0, \quad x \in \mathbb{Z}^d. \tag{24}$$

Proof. Using the special form of the kernel k we have

$$\begin{aligned} L\big(\Psi_H(f)\big)(x) &= \sum_{h\in\mathbb{Z}^d\backslash\{0\}} k_*(h)\big(\Psi_H(f)(x+h) - \Psi_H(f)(x)\big) \\ &= \sum_{h,\sigma\in\mathbb{Z}^d\backslash\{0\}} k_*(h)k_*(\sigma)\big[H\big(f(x+h+\sigma) - f(x+h)\big) - H\big(f(x+\sigma) - f(x)\big)\big] \end{aligned}$$

and

$$\begin{aligned} B_{H'}(f, Lf)(x) &= \sum_{h\in\mathbb{Z}^d\backslash\{0\}} k_*(h)H'\big(f(x+h) - f(x)\big)\big(Lf(x+h) - Lf(x)\big) \\ &= \sum_{h,\sigma\in\mathbb{Z}^d\backslash\{0\}} k_*(h)k_*(\sigma)H'\big(f(x+h) - f(x)\big)\big(f(x+h+\sigma) - f(x+h) - f(x+\sigma) + f(x)\big). \end{aligned}$$

Recalling the definition of Λ_H (see (16)), identity (23) follows immediately from the last two relations. The second assertion is evident by positivity of k_* and the definition of Λ_H. □

It is interesting to note that identity (23) is structurally very similar to the fundamental identity (15). In contrast to (15), the Λ_H-term in (23) involves a difference of second order. We also remark that (23) can be viewed as a generalized discrete version of Bochner's identity. In fact, taking $H(y) = \frac{1}{2}y^2$ we have

$$\Lambda_H(w, z) = \frac{1}{2}(w - z)^2, \quad \Psi_H(f) = \Gamma(f), \quad B_{H'}(f, g) = 2\Gamma(f, g),$$

and thus obtain

$$\begin{aligned} L\Gamma(f)(x) &= L\Psi_H(f)(x) \\ &= B_{H'}(f, Lf)(x) + \sum_{h,\sigma\in\mathbb{Z}^d\backslash\{0\}} k_*(h)k_*(\sigma)\Lambda_H\big(f(x+h+\sigma) - f(x+h), f(x+\sigma) - f(x)\big) \\ &= 2\Gamma(f, Lf)(x) + \frac{1}{2}\sum_{h,\sigma\in\mathbb{Z}^d\backslash\{0\}} k_*(h)k_*(\sigma)\big(f(x+h+\sigma) - f(x+h) - f(x+\sigma) + f(x)\big)^2. \end{aligned}$$

This is a discrete variant of Bochner's identity for the Laplacian Δ on $\mathbb{R}^d$, which states that

$$\Delta|\nabla f|^2 = 2\nabla f \cdot \nabla\Delta f + 2\sum_{i,j=1}^d (\partial_i\partial_j f)^2.$$

Since identity (23) plays an important role with regard to curvature-dimension inequalities for non-local operators in the discrete setting we will also refer to it as *second fundamental identity*.

The following identity is of crucial importance in the proof of Theorem 3.4 on the second time derivative of the entropy.

Lemma 2.4. *There holds*

$$\frac{1}{2}\int_X e^g B_{\Upsilon'}(g,h)\,d\mu = -\int_X e^g Lh\,d\mu. \tag{25}$$

for any functions $g, h \in \mathbb{R}^X$ such that the expressions in (25) are well defined.

Proof. Inserting $\Upsilon'(z) = e^z - 1$ into the definition of the operator $B_{\Upsilon'}$ we obtain

$$\begin{aligned}\int_X e^g B_{\Upsilon'}(g,h)\,d\mu &= \sum_{x,y\in X} k(x,y)e^{g(x)}\big(e^{g(y)-g(x)}-1\big)\big(h(y)-h(x)\big)\pi(x)\\ &= \sum_{x,y\in X} k(x,y)e^{g(y)}\big(h(y)-h(x)\big)\pi(x) - \sum_{x\in X} e^{g(x)}Lh(x)\pi(x).\end{aligned}$$

By detailed balance,

$$\begin{aligned}\sum_{x,y\in X} k(x,y)e^{g(y)}\big(h(y)-h(x)\big)\pi(x) &= \sum_{x,y\in X} k(y,x)e^{g(y)}\big(h(y)-h(x)\big)\pi(y)\\ &= \sum_{y\in X} e^{g(y)}\big(-Lh(y)\big)\pi(y),\end{aligned}$$

which proves the assertion. □

Remark 2.5. Identity (25) can be viewed as discrete analogue of the formula

$$\int_X e^g \Gamma\big(g,h\big)\,d\mu = -\int_X e^g Lh\,d\mu, \tag{26}$$

which is valid in the diffusion setting (with μ being an invariant and reversible measure on X) and follows from the chain rule $\Gamma(H(g),h) = H'(g)\Gamma(g,h)$ for the carré du champ operator and the identity

$$\int_X \Gamma(f,g)\,d\mu = -\int_X fLg\,d\mu,$$

see [7, Section 3.4.1].

Lemma 2.2 suggests to regard $\Psi_\Upsilon(f)$ as a suitable substitute for the carré du champ operator $\Gamma(f)$. Further, Lemma 2.4 shows that identity (26) has a discrete version where Γ is replaced by the operator $2B_{\Upsilon'}$. These two key observations are also reflected in the next definition, where we introduce the operator which will serve as appropriate substitute for the gamma deux $\Gamma_2(f)$ in our curvature-dimension condition.

Definition 2.6. *For any $f \in \ell^\infty(X)$, we define the operator*

$$\Psi_{2,\Upsilon}(f) = \frac{1}{2}\big(L\Psi_\Upsilon(f) - B_{\Upsilon'}(f,Lf)\big). \tag{27}$$

We want to emphasize here the great structural similarity to the definition of $\Gamma_2(f)$, which is given by

$$\Gamma_2(f) = \frac{1}{2}\big(L\Gamma(f) - 2\Gamma(f,Lf)\big).$$

To better understand the connection between both objects, let us introduce the more general operator

$$\Psi_{2,H}(f) = \frac{1}{2}\big(L\Psi_H(f) - B_{H'}(f,Lf)\big), \tag{28}$$

where $H \in C^1(\mathbb{R})$ is an arbitrary function and $f \in \ell^\infty(X)$. Now observe that the choice $H(y) = \frac{1}{2}y^2$ leads to the classical object, the gamma deux operator $\Gamma_2(f)$, whereas $H = \Upsilon$ gives the new object from Definition 2.6. So the quadratic function is replaced with $H(y) = \Upsilon(y) = e^y - 1 - y$!

Definition 2.7. *The Markov generator L is said to satisfy the curvature-dimension inequality $CD_\Upsilon(\kappa,\infty)$ at $x \in X$ for $\kappa \in \mathbb{R}$, if for every function $f \in l^\infty(X)$ there holds*

$$\Psi_{2,\Upsilon}(f)(x) \geqslant \kappa \Psi_\Upsilon(f)(x). \tag{29}$$

If L satisfies $CD_\Upsilon(\kappa,\infty)$ at any $x \in X$, then we say that L satisfies $CD_\Upsilon(\kappa,\infty)$.

Remark 2.8. (i) The choice of the suitable class of functions is a quite subtle question in the classical Bakry-Émery theory. In our case, the assumptions (11) and (12) ensure that the operators Ψ_Υ and $\Psi_{2,\Upsilon}$ are both well defined on $\ell^\infty(X)$. This is enough for the applications we have in mind, as we will see below. However, one may choose alternative function spaces that possibly depend on the kernel.

(ii) If the associated graph to L is locally finite, (29) holds true for any $f \in \mathbb{R}^X$ if and only it holds on $\ell^\infty(X)$. This is a consequence of the fact that for any $x \in X$ the expressions $\Psi_{2,\Upsilon}(f)(x)$ and $\Psi_\Upsilon(f)(x)$ only depend on finitely many values of f.

Remark 2.9. (i) In [50, Remark 3.12], Münch introduces the CD condition $CD\psi(d,K)$ for the discrete Laplacian on a finite graph without weights, where ψ is a concave function on $(0,\infty)$, $d > 0$ is the dimension parameter and $K \in \mathbb{R}$ the lower curvature bound. It turns out that for $\psi = \log$ the condition $CD\psi(\infty,\kappa)$ is equivalent to $CD_\Upsilon(\kappa,\infty)$, see the proof in Section 9. This observation is by no means obvious, as the quantity Γ_2^ψ appearing in $CD\psi(d,K)$ is defined in a rather implicit way. We point out that in [50], the condition $CD\psi(d,K)$ with $K > 0$ only crops up in Remark 3.12., in form of a definition. It is not used at all in [50] as this paper focuses on Li-Yau inequalities, for which, assuming nonnegative curvature, the term involving the dimension parameter plays the decisive role.

(ii) We can also define a more general $CD_\Upsilon(\kappa,d)$ condition with $\kappa \in \mathbb{R}$ and a dimension parameter $d \in [1,\infty]$ via the inequality

$$\Psi_{2,\Upsilon}(f) \geqslant \frac{1}{d}(Lf)^2 + \kappa \Psi_\Upsilon(f). \tag{30}$$

This condition is equivalent to $CD\psi(d,\kappa)$ from [50] with $\psi = \log$, cf. Section 9 and the comments in [50] concerning the case $\psi = \log$. The condition $CD_\Upsilon(\kappa,d)$ with $d < \infty$ is not so much in the focus of this paper. It will appear in Remark 2.13 on the relation to the classical Bakry-Émery condition and in Remark 4.3 in connection with a generalization of our main tensorization result.

(iii) The condition $CD_\Upsilon(0,d)$ is also closely linked to the condition $CD(F;0)$ with a quadratic CD-function F. These notions were introduced in the very recent work [19, Definitions 3.1 and 3.8], which is also concerned with discrete versions of the Li-Yau gradient estimate. The operator Ψ_Υ is also already used in [19]. Translating the situation from [19], where the authors consider a generalized graph Laplacian with vertex weights and symmetric edge weights on a locally finite graph, the condition $CD(F;0)$ at $x \in X$ with CD-function F says that

$$L\Psi_{\Upsilon'}(f)(x) \geqslant F\big(-Lf(x)\big) \tag{31}$$

holds for any function $f : X \to \mathbb{R}$ satisfying the conditions

$$Lf(x) < 0 \quad \text{and} \quad Lf(x) \leqslant Lf(y) \quad \text{if } k(x,y) > 0. \tag{32}$$

Using $\Upsilon'(z) = \Upsilon(z) + z$, $e^z = \Upsilon'(z) + 1$, $\Psi_{Id}(f) = Lf$ and denoting the function with constant value 1 by $\mathbf{1}$ we have

$$\begin{aligned} L\Psi_{\Upsilon'}(f)(x) &= L\Psi_\Upsilon(f)(x) + L^2 f(x) \\ &= 2\Psi_{2,\Upsilon}(f)(x) + B_{\Upsilon'}(f,Lf)(x) + B_{\mathbf{1}}(f,Lf)(x) \\ &= 2\Psi_{2,\Upsilon}(f)(x) + B_{\exp}(f,Lf)(x). \end{aligned}$$

This shows that under (32) (which suffices for the Li-Yau inequality when using the maximum principle), it is advantageous to work with $L\Psi_{\Upsilon'}(f)(x)$ instead of $2\Psi_{2,\Upsilon}(f)(x)$, since one has the additional nonnegative term $B_{\exp}(f, Lf)(x)$. Employing this term and allowing also for nonquadratic CD-functions may lead to significantly better (even sharp) Li-Yau estimates ([19]).

(iv) The condition $CD_\Upsilon(\kappa,\infty)$ is further closely connected to the generalized Γ-calculus described in Monmarché [48, Section 2]. Here the author extends the classical Γ-calculus to study functional inequalities, hypoellipticity and hypocoercivity for degenerate Markov processes like, e.g., generalized Ornstein-Uhlenbeck processes, kinetic Fokker-Planck diffusion in $\mathbb{R}^{2d}$ or chains of interacting particles. We also refer to [49, Section 4], where Monmarché already formulates the basic ingredients of his calculus for Markov semigroups on the Euclidean space $\mathbb{R}^d$.

Assuming that the domain $D(L)$ of the generator L contains a core $\mathcal{A}$ (which is also an algebra) and setting $\mathcal{A}_+ := \{f \in \mathcal{A} : f \geqslant 0\}$ he considers maps $\Phi : \mathcal{A}_+ \to \mathcal{A}$ differentiable in the sense that $\frac{1}{\tau}(\Phi(f+\tau g) - \Phi(f))(x)$ converges as $\tau \to 0$ to $(d\Phi(f).g)(x)$ (this is the notation used in [48]) for all $x \in X$, $f, g \in \mathcal{A}_+$, where $d\Phi$ is the corresponding derivation operator. For $f \in \mathcal{A}_+$ Monmarché then defines the operator

$$\Gamma_{L,\Phi}(f) = \frac{1}{2}\big(L\Phi(f) - d\Phi(f).Lf\big).$$

For $\Phi(f) = f^2$ one obtains $\Gamma_{L,\Phi}(f) = \Gamma(f)$. With the choice $\Phi(f) = \Gamma(f)$ one recovers the Γ_2-operator. In this general framework, Monmarché introduces the curvature condition $\Gamma_{L,\Phi}(f) \geqslant \kappa\Phi(f)$ with some $\kappa \in \mathbb{R}$.

If we now take $\Phi(f) = \Psi_\Upsilon(f)$ in the discrete setting, we obtain that

$$\frac{1}{\tau}(\Phi(f+\tau g) - \Phi(f))(x) \to B_{\Upsilon'}(f,g)(x) \quad \text{as } \tau \to 0,$$

under suitable conditions on f, g and the kernel k, and consequently, $\Gamma_{L,\Phi}(f) = \Psi_{2,\Upsilon}(f)$ so that Monmarché's curvature condition is formally equivalent to $CD_\Upsilon(\kappa,\infty)$. Note that here, there is no reason to restrict Φ to nonnegative functions as in [48].

We point out that objects like Ψ_Υ, $\Psi_{2,\Upsilon}$ or the discrete Boltzmann entropy (see Section 3 below) do not seem to appear at all in [48], [49], as the focus of the author does not lie on the pure discrete setting but mainly on degenerate problems in a continuous setting.

(v) As already mentioned in the introduction, the operators Ψ_Υ and $\Psi_{2,\Upsilon}$ also appear implicitly in the context of entropic Ricci curvature of Erbar and Maas in [22]. More precisely, in Section 9 we will demonstrate that

$$\mathcal{A}(\rho, \log\rho) = \int_X \rho\Psi_\Upsilon(\log\rho)\, d\mu \tag{33}$$

and

$$\mathcal{B}(\rho, \log\rho) = \int_X \rho\Psi_{2,\Upsilon}(\log\rho)\, d\mu, \tag{34}$$

hold respectively, where the functionals $\mathcal{A}$ and $\mathcal{B}$ play a key role in [22], see also Section 9 for a definition.

On suitable functions the Bakry-Émery condition is known to be equivalent to respective gradient bounds in the diffusive setting of [7] as well as in the discrete setting, for instance, of [39] and [29]. The following remark indicates that we can characterize the CD_Υ condition in a similar spirit.

Remark 2.10. Due to the work of [48], which we have discussed in Remark 2.9(iv), (29) is equivalent to the gradient bound

$$\Psi_\Upsilon(P_t f) \leqslant e^{-2\kappa t} P_t(\Psi_\Upsilon(f)), \quad t > 0, \tag{35}$$

in a class of sufficiently good functions, cf. Lemma 1 in [48]. We revisit the instructive argument on a formal level for the reader's convenience.

Let $t > 0$ and $\varphi(s) := e^{-2\kappa s} P_s\big(\Psi_\Upsilon(P_{t-s}f)\big)$, $s \in [0,t]$. Then we have

$$\varphi'(s) = e^{-2\kappa s}\Big(- 2\kappa P_s(\Psi_\Upsilon(P_{t-s}f)) + LP_s(\Psi_\Upsilon(P_{t-s}f)) + P_s\big(\frac{d}{d_s}\Psi_\Upsilon(P_{t-s}f)\big)\Big).$$

Now, for any $\tau > 0$

$$\frac{d}{d\tau}\Psi_\Upsilon(P_\tau f)) = B_{\Upsilon'}(P_\tau f, LP_\tau f) \tag{36}$$

holds true and thus

$$\varphi'(s) = 2e^{-2\kappa s} P_s\big(\Psi_{2,\Upsilon}(P_{t-s}f) - \kappa\Psi_\Upsilon(P_{t-s}f)\big).$$

Thus, we infer from $CD_\Upsilon(\kappa,\infty)$ and the positivity of the semigroup that φ is increasing and consequently we observe (35) by

$$e^{-2\kappa t}P_t(\Psi_\Upsilon(f)) - \Psi_\Upsilon(P_t f) = \varphi(t) - \varphi(0) \geqslant 0.$$

Conversely, (35) implies for $t > 0$

$$\frac{G(t) - G(0)}{t} \geqslant 0,$$

where $G(t) = e^{-2\kappa t}P_t(\Psi_\Upsilon(f)) - \Psi_\Upsilon(P_t f)$. Using (36), we have further

$$\frac{d}{dt}G(t) = e^{-2\kappa t}\big(- 2\kappa P_t(\Psi_\Upsilon(f)) + LP_t(\Psi_\Upsilon(f))\big) - B_{\Upsilon'}(P_t f, LP_t f).$$

In particular, we deduce from $G'(0) \geqslant 0$ the estimate

$$0 \leqslant -2\kappa\Psi_\Upsilon(f) + L\Psi_\Upsilon(f) - B_{\Upsilon'}(f, Lf),$$

which yields $CD_\Upsilon(\kappa,\infty)$.

The above calculation is rigorous for any $f \in \mathbb{R}^X$ if X is finite. In the general case, regularity assumptions on f are needed, particularly, to guarantee that the generator and the semigroup do commute in the above lines.

Proposition 2.11. *If the Markov generator L satisfies $CD_\Upsilon(\kappa,\infty)$, then L also satisfies $CD(\kappa,\infty)$.*

Remark 2.12. We remark that the statement of Proposition 2.11 is implicitly contained in [50] in the special case of finite graphs without weights, cf. Remark 2.9(i) above. Since our setting is more general, we present a self-contained proof.

Proof. Let $f \in \ell^\infty(X)$ and $x \in X$. In order to calculate $\lim\limits_{\lambda\to 0}\frac{\Psi_\Upsilon(\lambda f)(x)}{\lambda^2}$ and $\lim\limits_{\lambda\to 0}\frac{\Psi_{2,\Upsilon}(\lambda f)(x)}{\lambda^2}$ we apply l' Hospital's rule, respectively. We obtain

$$\frac{d}{d\lambda}\Psi_\Upsilon(\lambda f)(x) = \sum_{y\in X\setminus\{x\}} k(x,y)\Upsilon'\big(\lambda(f(y)-f(x))\big)(f(y)-f(x))$$

and

$$\frac{d^2}{d\lambda^2}\Psi_\Upsilon(\lambda f)(x) = \sum_{y\in X\setminus\{x\}} k(x,y)e^{\lambda(f(y)-f(x))}\big(f(y)-f(x)\big)^2. \tag{37}$$

Hence, we have

$$\lim_{\lambda\to 0}\frac{\Psi_\Upsilon(\lambda f)(x)}{\lambda^2} = \frac{1}{2}\sum_{y\in X\setminus\{x\}} k(x,y)\big(f(y)-f(x)\big)^2 = \Gamma(f)(x).$$

Above, we were able to interchange the order of differentation and summation by (11). In the subsequent lines we use (12) for that purpose. We obtain

$$\frac{d}{d\lambda}L\Psi_\Upsilon(\lambda f)(x) = \sum_{y\in X\setminus\{x\}} k(x,y)\frac{d}{d\lambda}\big(\Psi_\Upsilon(\lambda f)(y) - \Psi_\Upsilon(\lambda f)(x)\big)$$

and by (37),

$$\begin{aligned}\frac{d^2}{d\lambda^2}L\Psi_\Upsilon(\lambda f)(x) = \sum_{y\in X\setminus\{x\}} k(x,y)\Big(\sum_{z\in X\setminus\{y\}} k(y,z)e^{\lambda(f(z)-f(y))}\big(f(z)-f(y)\big)^2 \\ - \sum_{z\in X\setminus\{x\}} k(x,z)e^{\lambda(f(z)-f(x))}\big(f(z)-f(x)\big)^2\Big).\end{aligned}$$

Consequently,

$$\lim_{\lambda\to 0}\frac{L\Psi_\Upsilon(\lambda f)(x)}{\lambda^2} = \sum_{y\in X\setminus\{x\}} k(x,y)(\Gamma(f)(y)-\Gamma(f)(x)) = L\Gamma(f)(x). \tag{38}$$

Further, by the product rule

$$\begin{aligned}\frac{d}{d\lambda}B_{\Upsilon'}(\lambda f, L(\lambda f))(x) = \sum_{y\in X\setminus\{x\}} k(x,y)\Big(\Upsilon'\big(\lambda(f(y)-f(x))\big)\big(Lf(y)-Lf(x)\big) \\ + \lambda e^{\lambda(f(y)-f(x))}\big(Lf(y)-Lf(x)\big)\big(f(y)-f(x)\big)\Big)\end{aligned}$$

and thus

$$\begin{aligned}\frac{d^2}{d\lambda^2}B_{\Upsilon'}(\lambda f, L(\lambda f))(x) = \sum_{y\in X\setminus\{x\}} k(x,y)\Big(2e^{\lambda(f(y)-f(x))}\big(Lf(y)-Lf(x)\big)\big(f(y)-f(x)\big) \\ + \lambda e^{\lambda(f(y)-f(x))}\big(Lf(y)-Lf(x)\big)\big(f(y)-f(x)\big)^2\Big).\end{aligned}$$

We deduce

$$\lim_{\lambda\to 0}\frac{B_{\Upsilon'}(\lambda f, L(\lambda f))(x)}{\lambda^2} = 2\Gamma(f,Lf)(x). \tag{39}$$

Combining (38) and (39), yields

$$\lim_{\lambda\to 0}\frac{\Psi_{2,\Upsilon}(\lambda f)(x)}{\lambda^2} = \Gamma_2(f)(x).$$

Now, $CD_\Upsilon(\kappa,\infty)$ implies

$$\begin{aligned}\Gamma_2(f)(x) &= \lim_{\lambda\to 0}\frac{\Psi_{2,\Upsilon}(\lambda f)(x)}{\lambda^2} \\ &\geqslant \kappa\lim_{\lambda\to 0}\frac{\Psi_\Upsilon(\lambda f)(x)}{\lambda^2} = \kappa\Gamma(f)(x).\end{aligned}$$

□

Remark 2.13. (i) The proof of Proposition 2.11 shows that the classical Bakry-Émery condition $CD(\kappa,\infty)$ is obtained from $CD_\Upsilon(\kappa,\infty)$ by taking the scaling limit at zero, the crucial property being that $\Upsilon(y)$ behaves like $\frac{1}{2}y^2$ as $y\to 0$.

(ii) Since clearly $(L(\lambda f))^2 = \lambda^2(Lf)^2$, the statement of Proposition 2.11 (and its proof) extends to the finite dimension case, that is, $CD_\Upsilon(\kappa,d)$ (as defined in Remark 2.9(ii)) implies $CD(\kappa,d)$ for any $\kappa\in\mathbb{R}$ and $d\in[1,\infty]$.

We next want to derive estimates which allow us to compare $\Psi_\Upsilon(f)$ and $\Gamma(f)$ as well as $\Psi_{2,\Upsilon}(f)$ and $\Gamma_2(f)$, respectively, in a quantitative way provided that f has small values. We begin with a useful representation of the operator $\Psi_{2,H}(f)$, which was introduced in (28). By definition, we have

$$\begin{aligned}
2\Psi_{2,H}(f)(x) &= L\Psi_H(f)(x) - B_{H'}(f, Lf)(x) \\
&= \sum_{y\in X\setminus\{x\}} k(x,y)\Big(\Psi_H(f)(y) - \Psi_H(f)(x) - H'\big(f(y)-f(x)\big)\big(Lf(y)-Lf(x)\big)\Big) \\
&= \sum_{y\in X\setminus\{x\}} k(x,y)\sum_{z\in X} k(y,z)\Big(H\big(f(z)-f(y)\big) - H'\big(f(y)-f(x)\big)\big(f(z)-f(y)\big)\Big) \\
&\quad + \sum_{y\in X\setminus\{x\}} k(x,y)H'\big(f(y)-f(x)\big)\sum_{z\in X} k(x,z)\big(f(z)-f(x)\big) \\
&\quad - \sum_{y\in X\setminus\{x\}} k(x,y)\sum_{z\in X} k(x,z)H\big(f(z)-f(x)\big). \qquad (40)
\end{aligned}$$

For the choice of $H(r) = \frac{1}{2}r^2$, it follows easily from (40) that $CD(-M,\infty)$ holds provided that $\sup_{x\in X} M_1(x) \leqslant M < \infty$. Interestingly, this simple property is not satisfied in general by the CD_Υ condition, as will be demonstrated in Example 5.14.

Given $\varepsilon \in (0, \frac{1}{2})$ there exists a $\delta_\varepsilon > 0$ such that

$$\frac{1-\varepsilon}{2}\, r^2 \leqslant \Upsilon(r) \leqslant \frac{1+\varepsilon}{2}\, r^2, \quad r \in [-\delta_\varepsilon, \delta_\varepsilon]. \qquad (41)$$

Suppose that $f \in \mathbb{R}^X$ satisfies $2|f|_\infty \leqslant \delta \leqslant \delta_\varepsilon$. Applying (40) with $H = \Upsilon$ and using the shorter notation f_x in place of $f(x)$ as well as the relation $\Upsilon'(r) = \Upsilon(r) + r$ we obtain

$$\begin{aligned}
2\Psi_{2,\Upsilon}(f)(x) &= \sum_{y\in X\setminus\{x\}} k(x,y)\sum_{z\in X\setminus\{y\}} k(y,z)\Big(\Upsilon(f_z-f_y) - \big[(f_y-f_x)+\Upsilon(f_y-f_x)\big](f_z-f_y)\Big) \\
&\quad + \sum_{y\in X\setminus\{x\}} k(x,y)\big[(f_y-f_x)+\Upsilon(f_y-f_x)\big]\sum_{z\in X} k(x,z)(f_z-f_x) \\
&\quad - \sum_{y\in X\setminus\{x\}} k(x,y)\sum_{z\in X} k(x,z)\Upsilon(f_z-f_x).
\end{aligned}$$

Employing (41), Young's inequality, and formula (40) with $H(r) = \frac{1}{2}r^2$ we infer that

$$\begin{aligned}
2\Psi_{2,\Upsilon}(f)(x) &\geqslant \sum_{y\in X\setminus\{x\}} k(x,y)\sum_{z\in X\setminus\{y\}} k(y,z)\Big(\frac{1-\varepsilon}{2}(f_z-f_y)^2 - (1-2\varepsilon)(f_y-f_x)(f_z-f_y) \\
&\qquad - 2\varepsilon(f_y-f_x)(f_z-f_y) - \Upsilon(f_y-f_x)(f_z-f_y)\Big) \\
&\quad + \sum_{y\in X\setminus\{x\}} k(x,y)\sum_{z\in X} k(x,z)\Big((1-2\varepsilon)(f_y-f_x)(f_z-f_x) \\
&\qquad + 2\varepsilon(f_y-f_x)(f_z-f_x) + \Upsilon(f_y-f_x)(f_z-f_x)\Big) \\
&\quad - \sum_{y\in X\setminus\{x\}} k(x,y)\sum_{z\in X} k(x,z)\Big(\frac{1-2\varepsilon}{2}(f_z-f_x)^2 + \frac{3\varepsilon}{2}(f_z-f_x)^2\Big) \\
&\geqslant 2(1-2\varepsilon)\Gamma_2(f) \\
&\quad - \sum_{y\in X\setminus\{x\}} k(x,y)\sum_{z\in X\setminus\{y\}} k(y,z)\Big(2\varepsilon(f_y-f_x)^2 + \frac{\delta(1+\varepsilon)}{2}(f_y-f_x)^2\Big)
\end{aligned}$$

$$- \sum_{y\in X\setminus\{x\}} k(x,y) \sum_{z\in X\setminus\{x\}} k(x,z)\Big(\varepsilon(f_y-f_x)^2 + \frac{5\varepsilon}{2}(f_z-f_x)^2 + \frac{\delta(1+\varepsilon)}{2}(f_y-f_x)^2\Big).$$

Assuming that $\sum_{y\in X\setminus\{x\}} k(x,y) \leqslant M$ for all $x \in X$, the last estimate gives

$$\begin{aligned} 2\Psi_{2,\Upsilon}(f)(x) &\geqslant 2(1-2\varepsilon)\Gamma_2(f)(x) - M\sum_{y\in X\setminus\{x\}} k(x,y)\big(3\varepsilon+\delta(1+\varepsilon)\big)(f_y-f_x)^2 \\ &\quad - M\sum_{z\in X\setminus\{x\}} k(x,z)\frac{5\varepsilon}{2}(f_z-f_x)^2 \\ &= 2(1-2\varepsilon)\Gamma_2(f)(x) - M\big(11\varepsilon+2\delta(1+\varepsilon)\big)\Gamma(f)(x). \end{aligned}$$

In analogous fashion one can show that

$$2\Psi_{2,\Upsilon}(f)(x) \leqslant 2(1+2\varepsilon)\Gamma_2(f)(x) + M\big(11\varepsilon+2\delta(1+\varepsilon)\big)\Gamma(f)(x).$$

Furthermore,

$$\begin{aligned} (1-\varepsilon)\Gamma(f)(x) &= \sum_{y\in X\setminus\{x\}} k(x,y)\frac{1-\varepsilon}{2}(f_y-f_x)^2 \leqslant \sum_{y\in X\setminus\{x\}} k(x,y)\Upsilon(f_y-f_x) \\ &\leqslant \sum_{y\in X\setminus\{x\}} k(x,y)\frac{1+\varepsilon}{2}(f_y-f_x)^2 = (1+\varepsilon)\Gamma(f)(x). \end{aligned}$$

This proves the following proposition.

Proposition 2.14. *Let $\varepsilon \in (0,\frac{1}{2})$ and $\delta_\varepsilon > 0$ be such that the two-sided quadratic estimate* (41) *for the Υ-function is satisfied. Suppose further that there is a constant $M > 0$ such that*

$$\sup_{x\in X} M_1(x) \leqslant M. \tag{42}$$

Then the following estimates hold for any $x \in X$ and $f \in \mathbb{R}^X$ with $2|f|_\infty \leqslant \delta \leqslant \delta_\varepsilon$.

$$(1-\varepsilon)\Gamma(f)(x) \leqslant \Psi_\Upsilon(f)(x) \leqslant (1+\varepsilon)\Gamma(f)(x), \tag{43}$$

$$\begin{aligned} (1-2\varepsilon)\Gamma_2(f)(x) - 6M(\varepsilon+\delta)\Gamma(f)(x) &\leqslant \Psi_{2,\Upsilon}(f)(x) \\ &\leqslant (1+2\varepsilon)\Gamma_2(f)(x) + 6M(\varepsilon+\delta)\Gamma(f)(x). \end{aligned} \tag{44}$$

From Proposition 2.14 we immediately deduce the following result.

Corollary 2.15. *Let $\varepsilon \in (0,\frac{1}{2})$ and $\delta_\varepsilon > 0$ be such that the two-sided quadratic estimate* (41) *for the Υ-function is satisfied. Suppose further that* (42) *holds and that the Markov generator L satisfies the condition $CD(\kappa,\infty)$ with some $\kappa > 0$. Then for any $x \in X$ and any $f \in \mathbb{R}^X$ with $2|f|_\infty \leqslant \delta \leqslant \delta_\varepsilon$ there holds*

$$\Big(1-2\varepsilon-\frac{6M}{\kappa}(\varepsilon+\delta)\Big)\Gamma_2(f)(x) \leqslant \Psi_{2,\Upsilon}(f)(x) \leqslant \Big(1+2\varepsilon+\frac{6M}{\kappa}(\varepsilon+\delta)\Big)\Gamma_2(f)(x). \tag{45}$$

We come back to the important case where $X = \mathbb{Z}^d$ and k is a translation invariant kernel.

Proposition 2.16. *Let $d \in \mathbb{N}$ and suppose that the kernel $k : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}$ is of the form described in Lemma 2.3, that is $k(x,y) = k_*(y-x)$, $x,y \in \mathbb{Z}^d$, $x \neq y$, where $k_* \in \ell^1(\mathbb{Z}^d\setminus\{0\})$ is nonnegative. Let $f \in \ell^\infty(\mathbb{Z}^d)$ and $x \in \mathbb{Z}^d$. Then*

$$\Gamma_2(f)(x) = \frac{1}{4}\sum_{h,\sigma\in\mathbb{Z}^d\setminus\{0\}} k_*(h)k_*(\sigma)\big(f(x+h+\sigma)-f(x+h)-f(x+\sigma)+f(x)\big)^2 \tag{46}$$

and

$$\Psi_{2,\Upsilon}(f)(x) = \frac{1}{2}\sum_{h,\sigma\in\mathbb{Z}^d\setminus\{0\}} k_*(h)k_*(\sigma)e^{f(x+\sigma)-f(x)}\Upsilon\big(f(x+h+\sigma)-f(x+h)-f(x+\sigma)+f(x)\big). \tag{47}$$

In particular, the CD conditions $CD(0,\infty)$ and $CD_\Upsilon(0,\infty)$ are satisfied.

Proof. The assertion on Γ_2 and the classical Bakry-Émery condition are well known in the literature. See also the discussion above after Lemma 2.3.

Concerning $\Psi_{2,\Upsilon}$ we apply Lemma 2.3 with $H = \Upsilon$ and use that

$$\begin{aligned}\Lambda_\Upsilon(w,z) &= \Upsilon(w) - \Upsilon(z) - \Upsilon'(z)(w-z)\\ &= \big(e^w - 1 - w\big) - \big(e^z - 1 - z\big) - \big(e^z - 1\big)(w-z)\\ &= e^w - e^z - e^z(w-z)\\ &= \Lambda_{\exp}(w,z) = e^z\Upsilon(w-z).\end{aligned}$$

□

Remark 2.17. The identity (46) has been used in the one-dimensional setting of $X = \mathbb{Z}$ in the quite recent work [59], where positive and negative results concerning the $CD(0,n)$ condition with $n < \infty$ were established. According to Remark 2.9(i) and Remark 2.13(ii), this is important in the context of Li-Yau inequalities since the $CD(0,n)$ condition is necessary for $CD_\Upsilon(0,n)$, which has been introduced in Remark 2.9(ii). However, as to curvature bounds, we observe that, formally, $CD(\kappa,\infty)$ (and thus $CD_\Upsilon(\kappa,\infty)$) can not be valid by plugging the identity function in (46). See [59, Theorem 6.1] for a rigorous proof that applies to a large class of kernels using approximative arguments.

3. Entropy decay and modified log-Sobolev inequality

Recall that we assume throughout this section that the invariant and reversible measure μ is a probability measure on X and $(P_t)_{t\geqslant 0}$ is a Markov semigroup. Also recall that π denotes the (probability) density of μ with respect to the counting measure on X. Let

$$\mathcal{P}(X) := \{\rho : X \to [0,\infty) \text{ such that } \sum_{x\in X}\rho(x)\pi(x) = 1\}$$

denote the set of probability densities with respect to μ on X. Further, let $\mathcal{P}_*(X)$ denote the set of strictly positive probability densities with respect to μ and $\mathcal{P}_*^+(X) := \mathcal{P}_*(X) \cap \ell^{\infty,+}(X)$, where

$$\ell^{\infty,+}(X) = \{f \in \ell^\infty(X) : \exists c > 0 \text{ s.t. } f(x) \geqslant c > 0, \forall x \in X\}.$$

We consider the entropy

$$\mathcal{H}(\rho) = \int_X \rho\log\rho\, d\mu = \sum_{x\in X}\rho(x)\log(\rho(x))\pi(x), \quad \rho\in\mathcal{P}(X),$$

where we set $0\log 0 = 0$ in this definition (and $\mathcal{H}(f) = \infty$ is allowed), and the discrete Fisher information

$$\mathcal{I}(\rho) = \int_X \rho\Psi_\Upsilon(\log\rho)\, d\mu = \sum_{x\in X}\rho(x)\Psi_\Upsilon(\log\rho)(x)\pi(x), \quad \rho\in\mathcal{P}_*^+(X). \tag{48}$$

Using the identity

$$a\Upsilon(\log b - \log a) = b - a - a(\log b - \log a), \quad a,b > 0,$$

and detailed balance we can rewrite the Fisher information as

$$\begin{aligned}
\mathcal{I}(\rho) &= \sum_{x\in X}\rho(x)\sum_{y\in X}k(x,y)\Upsilon\big(\log\rho(y)-\log\rho(x)\big)\pi(x)\\
&= \frac{1}{2}\sum_{x\in X}\rho(x)\sum_{y\in X}k(x,y)\Upsilon\big(\log\rho(y)-\log\rho(x)\big)\pi(x)\\
&\quad+\frac{1}{2}\sum_{y\in X}\rho(y)\sum_{x\in X}k(y,x)\Upsilon\big(\log\rho(x)-\log\rho(y)\big)\pi(y)\\
&= \frac{1}{2}\sum_{x,y\in X}k(x,y)\Big(\rho(y)-\rho(x)-\rho(x)\big(\log\rho(y)-\log\rho(x)\big)\Big)\pi(x)\\
&\quad+\frac{1}{2}\sum_{x,y\in X}k(y,x)\Big(\rho(x)-\rho(y)-\rho(y)\big(\log\rho(x)-\log\rho(y)\big)\Big)\pi(y)\\
&= \frac{1}{2}\sum_{x,y\in X}k(x,y)\big(\rho(y)-\rho(x)\big)\big(\log\rho(y)-\log\rho(x)\big)\pi(x). \qquad (49)
\end{aligned}$$

The representation (49) for the Fisher information has been frequently used in the existing literature, see e.g. [22], [32, Section 5.1]. Note also that $\mathcal{I}(\rho)=\mathcal{E}(\rho,\log\rho)$, where $\mathcal{E}$ is the Dirichlet form given by (13). In the sequel we use this relation to extend $\mathcal{I}$ to functions $\rho\in\mathcal{P}_*(X)$, where $\mathcal{I}(\rho)=\infty$ is allowed.

Definition 3.1. *We say that L satisfies the modified logarithmic Sobolev inequality* MLSI(α) *with $\alpha>0$, if*

$$\mathcal{H}(f)\leqslant\frac{1}{2\alpha}\mathcal{I}(f) \qquad (50)$$

holds for any $f\in P_*(X)$ with $\mathcal{H}(f)<\infty$.

It is a classical fact that exponential decay of the entropy

$$\mathcal{H}(P_tf)\leqslant e^{-2\alpha t}\mathcal{H}(f),\quad t>0, \qquad (51)$$

is equivalent to (50) on a suitable class of functions, see e.g. [14]. With regard to the latter we will work with the function space $\mathcal{P}_*^+(X)$.

Lemma 3.2. *Assume that* (50) *holds for $\alpha>0$ and any $f\in\mathcal{P}_*^+(X)$. Then L satisfies* MLSI(α).

Proof. Clearly, there is nothing to show in the case of $\mathcal{I}(f)=\infty$. If $\mathcal{I}(f)<\infty$ the claim can be established by a standard truncation argument. More precisely, one considers the sequence $(f_n)_{n\in\mathbb{N}}\subset P_*^+(X)$ given by $f_n(x)=\frac{f(x)}{C_n}$ if $\frac{1}{n}\leqslant f(x)\leqslant n$, $f_n(x)=\frac{1}{C_nn}$ if $f(x)>\frac{1}{n}$ and $f_n(x)=\frac{n}{C_n}$ if $f(x)>n$, where C_n denotes the corresponding normalizing constant. Now, it is straightforward to check that $\mathcal{H}(f_n)\to\mathcal{H}(f)$ and $\mathcal{I}(f_n)\to\mathcal{I}(f)$ as $n\to\infty$. □

In the classical diffusive setting the modified logarithmic Sobolev inequality coincides with the logarithmic Sobolev inequality by means of classical chain rules. This is no longer true in the discrete setting, where the logarithmic Sobolev inequality is stronger than (50) (cf. [18, Lemma 2.7]). We refer to [11] and [18] for an extensive overview on these functional inequalities in the discrete setting of Markov chains.

As a motivation, we briefly outline the Bakry-Émery approach to establish the functional inequality (50). The key step is to achieve the differential inequality

$$\frac{d^2}{dt^2}\mathcal{H}(P_tf)\geqslant-2\alpha\frac{d}{dt}\mathcal{H}(P_tf), \qquad (52)$$

which is established in the classical diffusive setting by means of the Bakry-Émery condition $CD(\alpha,\infty)$. Starting from (52) the strategy is quite general. By Gronwall's lemma one deduces $-\frac{d}{dt}\mathcal{H}(P_tf) \leqslant -e^{-2\alpha t}\,\frac{d}{dt}\mathcal{H}(P_tf)\big|_{t=0}$. Integrating this inequality, yields

$$\mathcal{H}(f) - \mathcal{H}(P_tf) \leqslant -\,\frac{d}{dt}\mathcal{H}(P_tf)\Big|_{t=0}\,\int_0^t e^{-2\alpha s}ds.$$

As $\frac{d}{dt}\mathcal{H}(P_tf) = -\mathcal{I}(P_tf)$ holds (see Proposition 3.3 below), (50) follows by sending $t\to\infty$ if $\mathcal{H}(P_tf)\to 0$ as $t\to\infty$. If $f\in\mathcal{P}_*^+(X)$, the latter property follows from ergodicity of the Markov chain, the dominated convergence theorem and $(P_t)_{t\geqslant 0}$ being a Markov semigroup.

Note that this strategy not only applies to the entropy functional $\mathcal{H}$. Therefore we will use it not only in the present context to deduce Corollary 3.5, but also in the setting of power-type entropies (Corollary 6.9) and those of hybrid processes (Corollary 7.4).

In order to apply the pointwise condition (29), we need to be able to interchange the order of differentation and integration in the strategy explained above. Concerning the finite state-space case this is of course clear. For the general case, we will impose respective integrability conditions on the functions M_1 and M_2 (see (11) and (12)) such that the dominated convergence theorem applies, see Proposition 3.3, Theorem 3.4, Corollary 3.5, Theorem 6.8 and Corollary 6.9.

It is well known that the time-derivative of the entropy equals the negative Fisher information along the heat flow associated with the operator L, see e.g. [22], [32, Section 5.1].

Proposition 3.3. *Let $M_1\in\ell^1(\mu)$. Then for any $f\in\mathcal{P}_*^+(X)$,*

$$\frac{d}{dt}\,\mathcal{H}(P_tf) = -\mathcal{I}(P_tf),\quad t\geqslant 0 \tag{53}$$

holds true.

Even though this result is well known it is instructive to revisit the argument in terms of our notation. Since $\frac{d}{dt}P_tf = LP_tf$ we have

$$\begin{aligned}\frac{d}{dt}\,\mathcal{H}(P_tf) &= \frac{d}{dt}\int_X P_tf\log\big(P_tf\big)\,d\mu\\ &= \int_X (LP_tf)\log\big(P_tf\big)\,d\mu + \int_X LP_tf\,d\mu.\end{aligned}$$

The last integral vanishes by invariance of μ. Since $L=L^*$ w.r.t. μ (by reversibility) and applying Lemma 2.2, the first integral can be rewritten as

$$\begin{aligned}\int_X (LP_tf)\log\big(P_tf\big)\,d\mu &= \int_X P_tfL\log\big(P_tf\big)\,d\mu\\ &= \int_X P_tf\Big(\frac{1}{P_tf}\,LP_tf - \Psi_\Upsilon\big(\log(P_tf)\big)\Big)\,d\mu\\ &= -\int_X P_tf\,\Psi_\Upsilon\big(\log(P_tf)\big)\,d\mu = -\mathcal{I}(P_tf),\end{aligned} \tag{54}$$

where we use again invariance of μ. This shows (53).

Observe that the only difference of the previous proof to the one in the diffusion setting is the usage of the fundamental identity from Lemma 2.2 in line (54). In the diffusion setting, at this place, one employs the chain rule

$$L\log\big(P_tf\big) = \frac{1}{P_tf}\,LP_tf - \Gamma\big(\log(P_tf)\big),$$

which then leads to the desired result as the Fisher information is defined as

$$\mathcal{I}(\rho) = \int_X \rho\,\Gamma\big(\log\rho\big)\,d\mu = \int_X \frac{1}{\rho}\,\Gamma(\rho)\,d\mu. \tag{55}$$

We turn now to derive an estimate for the second time derivative of the entropy, which is the key idea in the Bakry-Émery approach as described above. We will prove a representation formula for the second time derivative of the entropy which is a discrete version of the crucial identity

$$\frac{d^2}{dt^2}\,\mathcal{H}(P_t f) = 2\int_X P_t f\,\Gamma_2\big(\log(P_t f)\big)\,d\mu \tag{56}$$

from the diffusion setting and is tailor-made for applying a CD_Υ condition. We will also see that the whole line of arguments is a natural discrete analogue of a suitably chosen proof of (56) in the diffusion setting. To make this explicit and to motivate our discrete computations we will first give such a proof in the diffusion setting and treat the discrete case afterwards. The following argument is also interesting in its own, regardless of the question of adaptability to the discrete setting. It does not seem to be included in standard references such as [7] and [32], at least in the form we present here.

The Fisher information being given by (55), we have in the diffusion setting

$$\begin{aligned}
\frac{d^2}{dt^2}\,\mathcal{H}(P_t f) &= -\frac{d}{dt}\,\mathcal{I}(P_t f) = -\frac{d}{dt}\int_X P_t f\,\Gamma\big(\log(P_t f)\big)\,d\mu \\
&= -\int_X (LP_t f)\,\Gamma\big(\log(P_t f)\big)\,d\mu - 2\int_X P_t f\,\Gamma\Big(\log(P_t f), \frac{d}{dt}\,\log(P_t f)\Big)\,d\mu \\
&= -\int_X P_t f\,L\Gamma\big(\log(P_t f)\big)\,d\mu - 2\int_X P_t f\,\Gamma\Big(\log(P_t f), \frac{1}{P_t f} LP_t f\Big)\,d\mu.
\end{aligned}$$

Inserting the identity

$$\frac{1}{P_t f}\,LP_t f = L\log(P_t f) + \Gamma\big(\log(P_t f)\big),$$

cf. (20), we obtain

$$\begin{aligned}
2\int_X P_t f\,\Gamma\Big(\log(P_t f), \frac{1}{P_t f} LP_t f\Big)\,d\mu &= 2\int_X P_t f\,\Gamma\big(\log(P_t f), L\log(P_t f)\big)\,d\mu \\
&\quad + 2\int_X P_t f\,\Gamma\Big(\log(P_t f), \Gamma\big(\log(P_t f)\big)\Big)\,d\mu.
\end{aligned}$$

Applying to the last integral the identity (26) with $g = \log(P_t f)$ and $h = \Gamma(\log(P_t f))$, we find that

$$2\int_X P_t f\,\Gamma\Big(\log(P_t f), \Gamma\big(\log(P_t f)\big)\Big)\,d\mu = -2\int_X P_t f\,L\Gamma\big(\log(P_t f)\big)\,d\mu.$$

Combining the previous relations yields

$$\begin{aligned}
\frac{d^2}{dt^2}\,\mathcal{H}(P_t f) &= \int_X P_t f\,L\Gamma\big(\log(P_t f)\big)\,d\mu - 2\int_X P_t f\,\Gamma\Big(\log(P_t f), L\log(P_t f)\Big)\,d\mu \\
&= 2\int_X P_t f\,\Gamma_2\big(\log(P_t f)\big)\,d\mu.
\end{aligned}$$

In the discrete case, we have now the following fundamental result.

Theorem 3.4. *Let $M_1 \in \ell^2(\mu)$ and $M_2 \in \ell^1(\mu)$. Then for any $f \in \mathcal{P}_*^+(X)$*

$$\frac{d^2}{dt^2}\,\mathcal{H}(P_t f) = 2\int_X P_t f\,\Psi_{2,\Upsilon}\big(\log(P_t f)\big)\,d\mu, \quad t \geqslant 0 \tag{57}$$

holds true.

Proof.

$$\begin{aligned}\frac{d^2}{dt^2}\,\mathcal{H}(P_tf) &= -\frac{d}{dt}\,\mathcal{I}(P_tf) = -\frac{d}{dt}\int_X P_tf\,\Psi_\Upsilon\big(\log(P_tf)\big)\,d\mu\\ &= -\int_X (LP_tf)\,\Psi_\Upsilon\big(\log(P_tf)\big)\,d\mu - \int_X P_tf\,B_{\Upsilon'}\Big(\log(P_tf),\frac{d}{dt}\,\log(P_tf)\Big)\,d\mu\\ &= -\int_X P_tf\,L\Psi_\Upsilon\big(\log(P_tf)\big)\,d\mu - \int_X P_tf\,B_{\Upsilon'}\Big(\log(P_tf),\frac{1}{P_tf}LP_tf\Big)\,d\mu\end{aligned}$$

Using the fundamental identity from Lemma 2.2 and the linearity of the operator $B_{\Upsilon'}$ in the second argument we obtain

$$\begin{aligned}\int_X P_tf\,B_{\Upsilon'}\Big(\log(P_tf),\frac{1}{P_tf}LP_tf\Big)\,d\mu &= \int_X P_tf\,B_{\Upsilon'}\big(\log(P_tf),L\log(P_tf)\big)\,d\mu\\ &\quad+\int_X P_tf\,B_{\Upsilon'}\Big(\log(P_tf),\Psi_\Upsilon\big(\log(P_tf)\big)\Big)\,d\mu.\end{aligned}$$

Invoking Lemma 2.4 with $g=\log(P_tf)$ and $h=\Psi_\Upsilon(\log(P_tf))$, the last integral can be rewritten as

$$\int_X P_tf\,B_{\Upsilon'}\Big(\log(P_tf),\Psi_\Upsilon\big(\log(P_tf)\big)\Big)\,d\mu = -2\int_X P_tf\,L\Psi_\Upsilon\big(\log(P_tf)\big)\,d\mu.$$

Combining the previous relations yields

$$\begin{aligned}\frac{d^2}{dt^2}\,\mathcal{H}(P_tf) &= \int_X P_tf\,L\Psi_\Upsilon\big(\log(P_tf)\big)\,d\mu - \int_X P_tf\,B_{\Upsilon'}\Big(\log(P_tf),L\log(P_tf)\Big)\,d\mu\\ &= 2\int_X P_tf\,\Psi_{2,\Upsilon}\big(\log(P_tf)\big)\,d\mu.\end{aligned}$$

This proves the theorem. □

We conclude the following result.

Corollary 3.5. *If the Markov generator L satisfies $CD_\Upsilon(\kappa,\infty)$, $\kappa>0$, $M_1\in\ell^2(\mu)$ and $M_2\in\ell^1(\mu)$, then L satisfies* MLSI(κ). *In particular, the entropy decay estimate* (51) *(with $\alpha=\kappa$) holds true for any $f\in\mathcal{P}_*^+(X)$.*

Proof. Combining Proposition 3.3 with Theorem 3.4, we conclude that the estimate (52) holds with $\alpha=\kappa$ for any $f\in\mathcal{P}_*^+(X)$. Hence, the claim follows by means of the entropy method (as described in the lines following to equation (52)) and Lemma 3.2. □

Note that the $CD_\Upsilon(\kappa,\infty)$ condition has been applied to $\log(P_tf)$ in order to deduce Corollary 3.5. However, due to ergodicity $\log(P_tf)$ is relatively small for large t, which motivates the following improvement of (51).

Theorem 3.6. *Let X be finite and assume that the CD conditions $CD(\kappa_0,\infty)$ and $CD_\Upsilon(\kappa,\infty)$ are satisfied with $\kappa_0>\kappa>0$. Let $\pi_0=\min_{x\in X}\pi(x)$. Then for any $\varepsilon_0\in(0,1-\frac{\kappa}{\kappa_0})$ there exist constants $h=h(\varepsilon_0,\kappa_0,|M_1|_\infty,\pi_0)>0$ and $C=C(\varepsilon_0,\kappa_0,\kappa,|M_1|_\infty,\pi_0)>0$ such that for any $f\in\mathcal{P}_*^+(X)$ the following holds true.*

(i) If $\mathcal{H}(f)\leqslant h$ then

$$\mathcal{H}(P_tf)\leqslant e^{-2(1-\varepsilon_0)\kappa_0 t}\mathcal{H}(f),\quad t\geqslant 0. \tag{58}$$

(ii) If $\mathcal{H}(f)>h$ then

$$\mathcal{H}(P_tf)\leqslant Ce^{-2(1-\varepsilon_0)\kappa_0 t}\mathcal{H}(f)^{\frac{\kappa_0}{\kappa}},\quad t\geqslant 0. \tag{59}$$

Proof. Recall the Csiszár-Kullback-Pinsker inequality

$$|\nu-\mu|_{TV}^2 \leqslant \frac{1}{2}H(\nu|\mu),$$

where $H(\nu|\mu)$ denotes the relative entropy of the probability measure ν with respect to the reference probability measure μ. If $\nu \ll \mu$ and $g = \frac{d\nu}{d\mu}$, we have $H(\nu|\mu) = \int_X g \log g \, d\mu$ and

$$|\nu-\mu|_{TV} = \frac{1}{2}\int_X |g-1|\, d\mu.$$

For $d\nu = P_t f \, d\mu$ this gives

$$|P_t f(x) - 1|\pi(x) \leqslant \int_X |P_t f - 1|\, d\mu \leqslant \sqrt{2\mathcal{H}(P_t f)}, \quad t \geqslant 0,\ x \in X,$$

and thus

$$|P_t f(x) - 1| \leqslant \frac{1}{\pi_0}\, e^{-\kappa t}\sqrt{2\mathcal{H}(f)}, \quad t \geqslant 0,\ x \in X. \tag{60}$$

Let $c \geqslant 1$ such that $|\log y| \leqslant c|y-1|$ for all $y \in [\frac{1}{2}, \frac{3}{2}]$. Given $\varepsilon_0 \in (0, 1 - \frac{\kappa}{\kappa_0})$ we choose $\varepsilon \in (0, \frac{1}{2})$ so small that

$$\frac{(1-2\varepsilon)\kappa_0 - 12M\varepsilon}{1+\varepsilon} \geqslant (1-\varepsilon_0)\kappa_0. \tag{61}$$

We then fix $\delta_\varepsilon \in (0, \varepsilon]$ such that the two-sided quadratic estimate (41) holds. Next, define $T_\varepsilon = 0$ if $2c\sqrt{2\mathcal{H}(f)} \leqslant \delta_\varepsilon \pi_0$ and set

$$T_\varepsilon = -\frac{1}{\kappa}\log\Big(\frac{\delta_\varepsilon \pi_0}{2c\sqrt{2\mathcal{H}(f)}}\Big) \tag{62}$$

otherwise. Employing (60), one can easily check that

$$|P_t f(x) - 1| \leqslant \frac{\delta_\varepsilon}{2c} \leqslant \frac{1}{2}, \quad t \geqslant T_\varepsilon,\ x \in X,$$

which further entails that

$$2|\log\big(P_t f(x)\big)| \leqslant \delta_\varepsilon, \quad t \geqslant T_\varepsilon,\ x \in X. \tag{63}$$

Thanks to (63) we may apply Proposition 2.14 with $M = |M_1|_\infty$ to the function $\log(P_t f)$, which together with the condition $CD(\kappa_0, \infty)$ and (61) implies that for all $t \geqslant T_\varepsilon$,

$$\begin{aligned}
\Psi_{2,\Upsilon}\big(\log(P_t f)\big) &\geqslant (1-2\varepsilon)\Gamma_2\big(\log(P_t f)\big) - 6M(\varepsilon+\delta_\varepsilon)\Gamma\big(\log(P_t f)\big) \\
&\geqslant \big((1-2\varepsilon)\kappa_0 - 12M\varepsilon\big)\Gamma\big(\log(P_t f)\big) \\
&\geqslant \frac{1}{1+\varepsilon}\big((1-2\varepsilon)\kappa_0 - 12M\varepsilon\big)\Psi_\Upsilon\big(\log(P_t f)\big) \\
&\geqslant (1-\varepsilon_0)\kappa_0 \Psi_\Upsilon\big(\log(P_t f)\big).
\end{aligned}$$

Inserting the latter estimate in the representation formula (57) from Theorem 3.4 then yields that for all $t \geqslant T_\varepsilon$,

$$\begin{aligned}
\frac{d^2}{dt^2}\,\mathcal{H}(P_t f) &= 2\int_X P_t f\, \Psi_{2,\Upsilon}\big(\log(P_t f)\big)\, d\mu \\
&\geqslant 2(1-\varepsilon_0)\kappa_0 \int_X P_t f\, \Psi_\Upsilon\big(\log(P_t f)\big)\, d\mu \\
&= 2(1-\varepsilon_0)\kappa_0\, \mathcal{I}(P_t f) = -2(1-\varepsilon_0)\kappa_0 \frac{d}{dt}\,\mathcal{H}(P_t f).
\end{aligned}$$

cf. (48). Setting $\lambda = (1-\varepsilon_0)\kappa_0\,(>\kappa)$ and using $\mathcal{H}(P_t f)\to 0$ as $t\to\infty$, this implies

$$\mathcal{H}(P_t f) \leqslant e^{-2\lambda(t-T_\varepsilon)}\mathcal{H}(P_{T_\varepsilon} f), \quad t \geqslant T_\varepsilon. \tag{64}$$

This shows assertion (i) with

$$h = \frac{(\delta_\varepsilon \pi_0)^2}{8c^2},$$

since $\mathcal{H}(f)\leqslant h$ is equivalent to $T_\varepsilon = 0$.

By the entropy decay estimate resulting from $CD_\Upsilon(\kappa,\infty)$ (see Corollary 3.5),

$$\mathcal{H}(P_{T_\varepsilon} f) \leqslant e^{-2\kappa T_\varepsilon}\mathcal{H}(f).$$

Combining this with (64) gives

$$\mathcal{H}(P_t f) \leqslant e^{-2\lambda t} e^{2T_\varepsilon(\lambda-\kappa)}\mathcal{H}(f), \quad t \geqslant T_\varepsilon.$$

Suppose that $2c\sqrt{2\mathcal{H}(f)} > \delta_\varepsilon\pi_0$, which is equivalent to $\mathcal{H}(f) > h$ and means that $T_\varepsilon > 0$ is given by (62). Then

$$e^{2T_\varepsilon(\lambda-\kappa)} = \Big(\frac{2c\sqrt{2\mathcal{H}(f)}}{\delta_\varepsilon\pi_0}\Big)^{\frac{2}{\kappa}(\lambda-\kappa)} = \Big(\frac{2c\sqrt{2}}{\delta_\varepsilon\pi_0}\Big)^{\frac{2}{\kappa}(\lambda-\kappa)} \mathcal{H}(f)^{\frac{\lambda}{\kappa}-1}, \tag{65}$$

and thus

$$\mathcal{H}(P_t f) \leqslant e^{-2\lambda t}\Big(\frac{2c\sqrt{2}}{\delta_\varepsilon\pi_0}\Big)^{\frac{2}{\kappa}(\lambda-\kappa)} \mathcal{H}(f)^{\frac{\lambda}{\kappa}}, \quad t \geqslant T_\varepsilon.$$

Since $\mathcal{H}(f) > h$, we have $\mathcal{H}(f)^{\frac{\lambda}{\kappa}} \leqslant \mathcal{H}(f)^{\frac{\kappa_0}{\kappa}} h^{-\varepsilon_0\frac{\kappa_0}{\kappa}}$ and therefore

$$\mathcal{H}(P_t f) \leqslant Ce^{-2\lambda t}\mathcal{H}(f)^{\frac{\kappa_0}{\kappa}}, \quad t\geqslant T_\varepsilon, \quad \text{with } C = \Big(\frac{2c\sqrt{2}}{\delta_\varepsilon\pi_0}\Big)^{\frac{2}{\kappa}(\lambda-\kappa)} h^{-\varepsilon_0\frac{\kappa_0}{\kappa}}.$$

Finally, for $0\leqslant t\leqslant T_\varepsilon$, we may estimate

$$\mathcal{H}(P_t f) \leqslant e^{-2\kappa t}\mathcal{H}(f) \leqslant e^{-2\lambda t} e^{2T_\varepsilon(\lambda-\kappa)}\mathcal{H}(f)$$

and use again (65) as before to see that $\mathcal{H}(P_t f)\leqslant Ce^{-2\lambda t}\mathcal{H}(f)^{\frac{\kappa_0}{\kappa}}$. This proves assertion (ii). □

4. Tensorization property

In this section we investigate the tensorization property for the CD_Υ condition, which is fundamental for the classical Bakry-Émery curvature-dimension condition and has been extended (see [7] for the diffusive setting) to the discrete setting in [42].

Let L_1 and L_2 generate a Markov chain on state spaces X_1 and X_2 respectively. The operator $L := L_1\oplus L_2$ acts on functions $f: X_1\times X_2\to\mathbb{R}$ as $(L_1\oplus L_2)(f)(x,y) = L_1(f_y)(x) + L_2(f^x)(y)$, where we set $f_y(\cdot) := f(\cdot,y)$, $f^x(\cdot) = f(x,\cdot)$. The transition rates of the Markov process generated by $L_1\oplus L_2$ (on the state space $X_1\times X_2$) are given by

$$k_\oplus((x_1,y_1),(x_2,y_2)) = \begin{cases} k_1(x_1,x_2), & \text{if } y_1 = y_2,\\ k_2(y_1,y_2), & \text{if } x_1 = x_2,\\ 0, & \text{else,}\end{cases}$$

where k_ν denote the transition rates for the process generated by L_ν, $\nu\in\{1,2\}$. In particular, the operator in (18) and respectively that in (22), both corresponding to the transition rates $k_\oplus$, are given by

$$\Psi_H(f)(x,y) = \Psi_H^{(1)}(f_y)(x) + \Psi_H^{(2)}(f^x)(y),$$
$$B_H(f,g)(x,y) = B_H^{(1)}(f_y,g_y)(x) + B_H^{(2)}(f^x,g^x)(y),$$

where $H : \mathbb{R} \to \mathbb{R}$, $f, g : X_1 \times X_2 \to \mathbb{R}$ and upper indices were used to indicate the corresponding state space.

We are now ready to analyze the operator $\Psi_{2,H}$ corresponding to $k_{\oplus}$. To that aim, we first provide the following crucial estimate.

Lemma 4.1. *Let $H \in C^1(\mathbb{R})$ be convex. Then we have*

$$\begin{aligned} L\Psi_H(f)(x,y) - B_{H'}(f, Lf)(x,y) \geqslant L_1\Psi_H^{(1)}(f_y)(x) - B_{H'}^{(1)}(f_y, L_1 f_y)(x) \\ + L_2\Psi_H^{(2)}(f^x)(y) - B_{H'}^{(2)}(f^x, L_1 f^x)(y) \end{aligned} \tag{66}$$

for any $f \in \ell^\infty(X)$.

Proof. We observe that

$$\begin{aligned} L\Psi_H(f)(x,y) &= \sum_{x_i \in X_1} k_1(x,x_i)(\Psi_H(f)(x_i,y) - \Psi_H(f)(x,y)) \\ &\quad + \sum_{y_j \in X_2} k_2(y,y_j)(\Psi_H(f)(x,y_j) - \Psi_H(f)(x,y)) \\ &= L_1\Psi_H^{(1)}(f_y)(x) + L_2\Psi_H^{(2)}(f^x)(y) \\ &\quad + \sum_{x_i \in X_1}\sum_{y_j \in X_2} k_1(x,x_i)k_2(y,y_j)\big(H(f(x_i,y_j) - f(x_i,y)) - H(f(x,y_j) - f(x,y))\big) \\ &\quad + \sum_{x_i \in X_1}\sum_{y_j \in X_2} k_1(x,x_i)k_2(y,y_j)\big(H(f(x_i,y_j) - f(x,y_j)) - H(f(x_i,y) - f(x,y))\big). \end{aligned}$$

Furthermore,

$$\begin{aligned} B_{H'}(f, Lf)(x,y) &= \sum_{x_i \in X_1} k_1(x,x_i)H'(f(x_i,y) - f(x,y))(Lf(x_i,y) - Lf(x,y)) \\ &\quad + \sum_{y_j \in X_2} k_2(y,y_j)H'(f(x,y_j) - f(x,y))(Lf(x,y_j) - Lf(x,y)) \\ &= B_{H'}^{(1)}(f_y, L_1 f_y)(x) + B_{H'}^{(2)}(f^x, L_2 f^x)(y) \\ &\quad + \sum_{x_i \in X_1}\sum_{y_j \in X_2} k_1(x,x_i)k_2(y,y_j)H'(f(x_i,y) - f(x,y))\big(f(x_i,y_j) - f(x_i,y) - (f(x,y_j) - f(x,y))\big) \\ &\quad + \sum_{x_i \in X_1}\sum_{y_j \in X_2} k_1(x,x_i)k_2(y,y_j)H'(f(x,y_j) - f(x,y))\big(f(x_i,y_j) - f(x,y_j) - (f(x_i,y) - f(x,y))\big). \end{aligned}$$

Hence, we deduce

$$\begin{aligned} &L\Psi_H(f)(x,y) - B_{H'}(f, Lf)(x,y) \\ &\quad = L_1\Psi_H^{(1)}(f_y)(x) - B_{H'}^{(1)}(f_y, L_1 f_y)(x) + L_2\Psi_H^{(2)}(f^x)(y) - B_{H'}^{(2)}(f^x, L_2 f^x)(x) \\ &\qquad + \sum_{x_i \in X_1}\sum_{y_j \in X_2} k_1(x,x_i)k_2(y,y_j)\Big[H(f(x_i,y_j) - f(x_i,y)) - H(f(x_i,y) - f(x,y)) \\ &\qquad\qquad - H'(f(x,y_j) - f(x,y))\big(f(x_i,y_j) - f(x,y_j) - (f(x_i,y) - f(x,y))\big)\Big] \\ &\qquad + \sum_{x_i \in X_1}\sum_{y_j \in X_2} k_1(x,x_i)k_2(y,y_j)\Big[H(f(x_i,y_j) - f(x,y_j)) - H(f(x_i,y) - f(x,y)) \end{aligned}$$

$$- H'(f(x_i,y)-f(x,y))\big(f(x_i,y_j)-f(x_i,y)-(f(x,y_j)-f(x,y))\big)\Big].$$

The convexity of H implies that $H(a)-H(b)\geqslant H'(b)(a-b)$ for all $a,b\in\mathbb{R}$. Inspecting above expressions, we thus conclude the claim. □

For the choice of $H(r)=\frac{r^2}{2}$ the statement of Lemma 4.1 has been obtained in a similar way in [42] and served as the key estimate in order to show the tensorization property for the Bakry-Émery curvature-dimension condition. For our purposes, we choose $H(r)=\Upsilon(r)$.

Theorem 4.2. *If L_1 satisfies $CD_\Upsilon(\kappa_1,\infty)$ and L_2 satisfies $CD_\Upsilon(\kappa_2,\infty)$ for $\kappa_1,\kappa_2\in\mathbb{R}$, then $L_1\oplus L_2$ satisfies $CD_\Upsilon(\kappa,\infty)$ with $\kappa=\min\{\kappa_1,\kappa_2\}$.*

Proof. From Lemma 4.1, we obtain

$$\begin{aligned}\Psi_{2,\Upsilon}(f)(x,y)&\geqslant\Psi^{(1)}_{2,\Upsilon}(f_y)(x)+\Psi^{(2)}_{2,\Upsilon}(f^x)(y)\\&\geqslant\kappa_1\Psi^{(1)}_{\Upsilon}(f_y)(x)+\kappa_2\Psi^{(2)}_{\Upsilon}(f^x)(y)\\&\geqslant\kappa\,\big(\Psi^{(1)}_{\Upsilon}(f_y)(x)+\Psi^{(2)}_{\Upsilon}(f^x)(y)\big)\\&=\kappa\,\Psi_\Upsilon(f)(x,y).\end{aligned}$$

□

Remark 4.3. The above tensorization property extends to the $CD_\Upsilon(\kappa,d)$ condition as defined in Remark 2.9(ii). Indeed, if L_i satisfies $CD_\Upsilon(\kappa_i,d_i)$ with $\kappa_i\in\mathbb{R}$ and $d_i\in[1,\infty]$, $i\in\{1,2\}$, then $L_1\oplus L_2$ satisfies $CD_\Upsilon(\kappa,d)$ with $\kappa=\min\{\kappa_1,\kappa_2\}$ and $d=d_1+d_2$. This follows analogously as in [42], using Young's inequality to treat the dimension terms.

5. Examples

Example 5.1 (*The two-point space*)**.** We consider $X=\{0,1\}$ with $k(0,1)=a$ and $k(1,0)=b$, where $a,b>0$. Here, the invariant and reversible probability measure μ is given by $d\mu=\pi d\#$ with $\pi(0)=\frac{b}{a+b}$ and $\pi(1)=\frac{a}{a+b}$. Writing $\tilde{x}=1-x$ for $x\in X$ we have

$$\Psi_\Upsilon(f)(x)=k(x,\tilde{x})\Upsilon\big(f(\tilde{x})-f(x)\big)$$

and

$$\begin{aligned}2\Psi_{2,\Upsilon}(f)(x)&=L\Psi_\Upsilon(f)(x)-B_{\Upsilon'}(f,Lf)(x)\\&=k(x,\tilde{x})\Big(\big(\Psi_\Upsilon(f)(\tilde{x})-\Psi_\Upsilon(f)(x)\big)-\Upsilon'\big(f(\tilde{x})-f(x)\big)\big(Lf(\tilde{x})-Lf(x)\big)\Big)\\&=k(x,\tilde{x})\,k(\tilde{x},x)\big[\Upsilon\big(f(x)-f(\tilde{x})\big)-\Upsilon'\big(f(\tilde{x})-f(x)\big)\big(f(x)-f(\tilde{x})\big)\big]\\&\quad+k(x,\tilde{x})^2\big[-\Upsilon\big(f(\tilde{x})-f(x)\big)+\Upsilon'\big(f(\tilde{x})-f(x)\big)\big(f(\tilde{x})-f(x)\big)\big].\end{aligned}$$

Setting $t=f(\tilde{x})-f(x)$, we thus have $\Psi_{2,\Upsilon}(f)(x)\geqslant\kappa\Psi_\Upsilon(f)(x)$ for all $f\in\mathbb{R}^X$ if and only if

$$k(\tilde{x},x)\big(e^{-t}-1+te^t\big)+k(x,\tilde{x})\big(te^t-e^t+1\big)\geqslant 2\kappa\Upsilon(t),\quad t\in\mathbb{R}. \tag{67}$$

Note that

$$\omega(t):=te^t-e^t+1=\Upsilon'(t)t-\Upsilon(t)>0,\quad t\in\mathbb{R}\backslash\{0\},$$

since for any $t\in\mathbb{R}\backslash\{0\}$ there is some ξ between 0 and t such that

$$0=\Upsilon(0)=\Upsilon(t)+\Upsilon'(t)(-t)+\frac{1}{2}e^\xi t^2>\Upsilon(t)+\Upsilon'(t)(-t).$$

Further, $\omega(t)\sim\frac{1}{2}t^2$ as $t\to0$ and thus $\omega(t)/\Upsilon(t)\to1$ as $t\to0$. Note also that

$$e^{-t}-1+te^t=\omega(t)+\big(e^t+e^{-t}-2\big)=\omega(t)+\big(e^{t/2}-e^{-t/2}\big)^2>0,\quad t\in\mathbb{R}\backslash\{0\},$$

and $e^{-t} - 1 + te^t \sim \frac{3}{2}t^2$ as $t \to 0$. Consequently, for $x = 0$, the condition (67) is equivalent to

$$b\,\frac{\omega(t) + \big(e^{t/2} - e^{-t/2}\big)^2}{\Upsilon(t)} + a\,\frac{\omega(t)}{\Upsilon(t)} \geqslant 2\kappa, \quad t \in \mathbb{R}\backslash\{0\}. \tag{68}$$

Observe that $\omega(t)/\Upsilon(t) \to 0$ as $t \to -\infty$ and $(e^{t/2} - e^{-t/2})^2/\Upsilon(t) \to 1$ as $t \to \infty$. On the other hand, $\omega(t)/\Upsilon(t) \to \infty$ as $t \to \infty$ and $(e^{t/2} - e^{-t/2})^2/\Upsilon(t) \to \infty$ as $t \to -\infty$. This implies that for any $a, b > 0$ there exists a maximal $\kappa = \kappa_0(a, b) > 0$ such that (68) holds true, and for arbitrarily fixed $a > 0$ we have $\kappa_0(a, b) \to 0$ as $b \to 0+$. Since $(e^{t/2} - e^{-t/2})^2/\Upsilon(t) > 1$ for all $t \in \mathbb{R}\backslash\{0\}$ (with limit 2 at $t = 0$), it follows that $\kappa_0(a, b) \geqslant \frac{b}{2}$. By symmetry, one obtains analogous statements concerning the validity of (67) at $x = 1$, i.e. for any $a, b > 0$ there is a maximal $\kappa = \kappa_1(a, b) > 0$ such that this condition is satisfied, and $\kappa_1(a, b) \to 0$ as $a \to 0+$ as well as $\kappa_1(a, b) \geqslant \frac{a}{2}$. Consequently, $CD_\Upsilon(\kappa, \infty)$ is satisfied with optimal $\kappa = \min(\kappa_0(a, b), \kappa_1(a, b))$.

As to the classical Bakry-Émery condition, one finds that

$$\Gamma(f)(x) = \frac{1}{2}\,k(x, \tilde{x})\big(f(\tilde{x}) - f(x)\big)^2$$

and

$$\begin{aligned}
2\Gamma_2(f)(x) &= L\Gamma(f)(x) - 2\Gamma(f, Lf)(x) \\
&= k(x, \tilde{x})\Big(\big(\Gamma(f)(\tilde{x}) - \Gamma(f)(x)\big) - \big(f(\tilde{x}) - f(x)\big)\big(Lf(\tilde{x}) - Lf(x)\big)\Big) \\
&= \frac{3}{2}k(x, \tilde{x})\,k(\tilde{x}, x)\big(f(x) - f(\tilde{x})\big)^2 + \frac{1}{2}k(x, \tilde{x})^2\big(f(\tilde{x}) - f(x)\big)^2
\end{aligned}$$

which shows that $\Gamma_2(f)(x) \geqslant \kappa\Gamma(f)(x)$ for all $f \in \mathbb{R}^X$ if and only if

$$\kappa \leqslant \frac{1}{2}\big(3k(\tilde{x}, x) + k(x, \tilde{x})\big).$$

Consequently, $CD(\kappa, \infty)$ holds true with optimal

$$\kappa = \frac{a + b}{2} + \min(a, b).$$

Comparing the two estimates, we see that keeping one of the two parameters a and b fixed and sending the other to zero, the optimal curvature constant κ in $CD_\Upsilon(\kappa, \infty)$ tends to zero as well whereas the optimal curvature bound in $CD(\kappa, \infty)$ stays bounded away from zero.

Example 5.2 (*Weighted complete graphs*)**.** Here we consider an arbitrary countable set X with at least two elements. Let $l : X \to (0, \infty)$ be integrable on X and $k(x, y) = l(y)$ for all $x, y \in X$ with $x \neq y$. Then μ given by $d\mu = \pi d\#$ with $\pi(x) = l(x)$, $x \in X$, is an invariant and reversible measure. We have

$$\Psi_\Upsilon(f)(x) = \sum_{y \in X} l(y)\Upsilon\big(f(y) - f(x)\big),$$

and by (40)

$$\begin{aligned}
2\Psi_{2,\Upsilon}(f)(x) = &\sum_{y \in X\backslash\{x\}}\sum_{z \in X} l(y)l(z)\Big(\Upsilon\big(f(z) - f(y)\big) - \Upsilon'\big(f(y) - f(x)\big)\big(f(z) - f(y)\big)\Big) \\
&+ \sum_{y \in X\backslash\{x\}}\sum_{z \in X} l(y)l(z)\Upsilon'\big(f(y) - f(x)\big)\big(f(z) - f(x)\big) \\
&- \sum_{y \in X\backslash\{x\}}\sum_{z \in X} l(y)l(z)\Upsilon\big(f(z) - f(x)\big)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{y\in X\setminus\{x\}} \sum_{z\in X} l(y)l(z)\Big(\Upsilon'\big(f(y)-f(x)\big)\big(f(y)-f(x)\big) - \Upsilon\big(f(y)-f(x)\big)\Big) \\
&\quad + \sum_{y\in X\setminus\{x\}} \sum_{z\in X} l(y)l(z)\Upsilon\big(f(z)-f(y)\big) + l(x)\Psi_\Upsilon(f)(x) \\
&= \sum_{y\in X\setminus\{x\}} l(y)\Big(|l|_1\Upsilon'(f(y)-f(x))(f(y)-f(x)) + l(x)\Upsilon(f(x)-f(y)) \\
&\qquad\qquad - \big(|l|_1 - l(x)\big)\Upsilon(f(y)-f(x))\Big) + \sum_{\substack{y,z\in X\\ y,z\neq x}} l(y)l(z)\Upsilon(f(z)-f(y)).
\end{aligned}
$$

We introduce the functions $\nu_{c,d}:\mathbb{R}\to\mathbb{R}$, given by

$$
\nu_{c,d}(r) = c\Upsilon'(r)r + \Upsilon(-r) - d\Upsilon(r), \tag{69}
$$

with constants $c,d\in\mathbb{R}$. In Appendix A, we collect several basic properties of these functions, which will be of great importance throughout this section.

Now, $\Psi_{2,\Upsilon}(f)(x)\geqslant\kappa\Psi_\Upsilon(f)(x)$ is equivalent to

$$
l(x)\sum_{y\in X\setminus\{x\}} l(y)\nu_{\frac{|l|_1}{l(x)},\frac{|l|_1}{l(x)}+\frac{2\kappa}{l(x)}-1}(f(y)-f(x)) + \sum_{\substack{y,z\in X\\ y,z\neq x}} l(y)l(z)\Upsilon(f(z)-f(y))\geqslant 0. \tag{70}
$$

Lemma 5.3. *In the setting of the present example, $CD_\Upsilon(\kappa,\infty)$ holds true for $\kappa\in\mathbb{R}$ in $x\in X$ if and only if*

$$
\nu_{\frac{|l|_1}{l(x)},\frac{|l|_1}{l(x)}+\frac{2\kappa}{l(x)}-1}(r)\geqslant 0 \tag{71}
$$

holds for any $r\in\mathbb{R}$.

Proof. If $\nu_{\frac{|l|_1}{l(x)},\frac{|l|_1}{l(x)}+\frac{2\kappa}{l(x)}-1}$ is non-negative, then $CD_\Upsilon(\kappa,\infty)$ in $x\in X$ follows by (70). Conversely, assume that there exists some $r_x\in\mathbb{R}$ such that $\nu_{\frac{|l|_1}{l(x)},\frac{|l|_1}{l(x)}+\frac{2\kappa}{l(x)}-1}(r_x)<0$. We define $f\in\mathbb{R}^X$ by $f(y)=r_x$ for any $y\neq x$ and $f(x)=0$. For this choice, (70) fails to be true. □

Continuing with Example 5.2, we apply Lemma A.4(ii) to (71) and observe that (70) holds true if $|l|_1\geqslant 2l(x)$ and

$$
\frac{2\kappa}{l(x)} - 1 \leqslant 2^{\frac{3}{2}}\sqrt{\frac{|l|_1}{l(x)}} - 1,
$$

where the latter is equivalent to $\kappa\leqslant\sqrt{2|l|_1 l(x)}$. In the case of $|l|_1<2l(x)$, Lemma A.4(ii) applied to (71) implies that (70) is valid if

$$
\frac{2\kappa}{l(x)} - 1 \leqslant 2\frac{|l|_1}{l(x)} - 1,
$$

which is equivalent to $\kappa\leqslant|l|_1$. Note, that $|l|_1\geqslant\sqrt{2|l|_1 l(x)}$ holds true if and only if $|l|_1\geqslant 2l(x)$, which is clearly satisfied for at least one $x\in X$.

As a consequence we obtain that $CD_\Upsilon(\sqrt{2|l|_1 l_*},\infty)$ holds true, where $l_*:=\inf_{x\in X} l(x)$. If X is of finite size, we thus deduce a positive global curvature bound. In the case of X being of infinite size, we deduce $CD_\Upsilon(0,\infty)$, which is in fact best possible. Indeed, for $\kappa>0$ we write $\nu_{\frac{|l|_1}{l(x)},\frac{|l|_1}{l(x)}+\frac{2\kappa}{l(x)}-1} = \nu_{\frac{|l|_1}{l(x)},\frac{|l|_1}{l(x)}(1+\frac{2\kappa}{|l|_1})-1}$ and apply Lemma A.4(i) to deduce from Lemma 5.3 that $CD_\Upsilon(\kappa,\infty)$ can not hold if we choose $x\in X$ such that $l(x)$ is small enough, which is possible since l is integrable on X.

Example 5.4 (*The complete graph*)**.** Now we examine a quite important special case of Example 5.2. Let X be an arbitrary set of n elements, $n \geqslant 2$, and $k(x,y) = 1$ for all $x, y \in X$ with $x \neq y$. The graph associated to the Markov chain generated by L is then given by the complete graph K_n. The uniform measure serves as the invariant and reversible measure. In the setting of Example 5.2 this translates to $l(x) = 1$, $x \in X$, i.e. L satisfies $CD_\Upsilon(\sqrt{2n}, \infty)$ by Example 5.2.

In [36] it is shown that the optimal Bakry-Émery curvature constant for the Markov generator associated to the complete graph is given by $1 + \frac{n}{2}$. Due to Proposition 2.11 this is the best possible constant one can hope for regarding the CD_Υ condition. By Lemma 5.3, the constant $1 + \frac{n}{2}$ is achieved if and only if $\nu_{n,2n+1}$ is non-negative, which is only true for the case of $n = 2$ by Lemma A.3.

We collect a few implications of the properties presented in Appendix A:

- In the case of $n = 2$, the optimal constant is achieved, i.e. L satisfies $CD_\Upsilon(2, \infty)$ (which is best possible since it is best possible for the Bakry-Émery curvature-dimension condition). This follows from Lemma A.3.
- If $n > 2$, the optimal constant for the CD_Υ condition is strictly less than the optimal Bakry-Émery curvature constant by Lemma A.3. Moreover, the constant does not grow linearly in n, which is in contrast to Bakry-Émery curvature. Indeed, setting $\kappa = \alpha n + \beta$ with $\alpha > 0$ and $\beta \in \mathbb{R}$ we find by Lemma A.4(i) that there exists some $r \in \mathbb{R}$ and $N \in \mathbb{N}$ such that $\nu_{n,(1+2\alpha)n+2\beta-1}(r) < 0$ for any $n \geqslant N$. Thus, $CD_\Upsilon(\kappa, \infty)$ can not hold true by Lemma 5.3.
- Let $f \in \mathbb{R}^X$ and $x_* \in X$ such that f achieves its global minimum at $x_* \in X$. By (70) and Lemma A.2, the estimate

$$\Psi_{2,\Upsilon}(f)(x_*) \geqslant \big(1 + \frac{n}{2}\big)\Psi_\Upsilon(f)(x_*)$$

  holds true. This estimate is best possible, as for the function $f_*(y) = \mathrm{dist}(y, x_*)$, where $\mathrm{dist} : X \times X \to \mathbb{N}_0$ denotes the combinatorical graph distance on the underlying graph, $\Gamma_2(f_*)(x_*) = \big(1 + \frac{n}{2}\big)\Gamma(f_*)(x_*)$ is valid (cf. the proof of Theorem 1.2 in [36], where, more generally, locally finite graphs are considered).

Remark 5.5. Concerning the link between the $CD_\Upsilon(\kappa, \infty)$ condition and the entropic Ricci curvature from [22], Example 5.4 has the interesting consequence that the curvature notion of Erbar and Maas does not imply in general the $CD_\Upsilon(\kappa, \infty)$ condition. More precisely, it is not true in general that there exist $\alpha > 0$ and $\beta > -\frac{\alpha}{2}$ such that the entropic Ricci curvature of [22] with constant $\kappa > 0$ does imply $CD_\Upsilon(\alpha\kappa + \beta, \infty)$. Indeed, let us denote for the moment by $\Psi_{2,\Upsilon}^{(n)}$ and $\Psi_\Upsilon^{(n)}$ the respective operators corresponding to the normalized complete graph, i.e. in the setting of Example 5.2 we have $l(x) = \frac{1}{n}$ for any $x \in X$, where X is a set of n elements. Note that this corresponds to the setting which has been considered in [22]. Due to [22, Example 5.1], the normalized complete graph has entropic curvature constant $\frac{1}{2} + \frac{1}{2n}$. We assume for contradiction that we find $\alpha > 0$ and $\beta > -\frac{\alpha}{2}$ such that the normalized complete graph satisfies $CD_\Upsilon(\kappa, \infty)$ with $\kappa = \alpha\big(\frac{1}{2} + \frac{1}{2n}\big) + \beta$ for any $n \geqslant 2$. Then we have

$$\Psi_{2,\Upsilon}(f) = n^2 \Psi_{2,\Upsilon}^{(n)}(f) \geqslant n^2 \kappa \Psi_\Upsilon^{(n)}(f) = n\kappa\, \Psi_\Upsilon(f),$$

where $\Psi_{2,\Upsilon}$ and Ψ_Υ correspond to the non-normalized setting of Example 5.4. Now, we get a contradiction to the fact that the curvature constant does not grow linearly in n in Example 5.4.

Example 5.6 (*The hypercube*)**.** We have seen in Example 5.4 that the Markov generator associated to K_2 satisfies $CD_\Upsilon(2, \infty)$. Now, since K_2 corresponds to the one-dimensional hypercube,

we denote the respective Markov generator by L_1. Iteratively, we define for $n \geqslant 2$

$$L_n := L_{n-1} \oplus L_{n-1}$$

in the sense of Section 4. L_n is then defined on the state space X^n, where X is a set of two elements. The transition rates for L_n are given by $k_n(x,y) = 1$, $x, y \in X^n$, if the Hamming distance between x and y equals 1 and 0 otherwise. The respective invariant and reversible measure is the n-fold product measure of X and hence the uniform measure on X^n and the underlying graph to L_n is the hypercube H_n.

Combining Theorem 4.2 with the fact that L_1 satisfies $CD_\Upsilon(2,\infty)$ we conclude that L_n satisfies $CD_\Upsilon(2,\infty)$ for any $n \in \mathbb{N}$. Note that this constant is optimal as it is optimal with respect to Bakry-Émery curvature (cf. [36]).

Remark 5.7. Following Example 3.7 in [11], the constant $\kappa = 2$ yields the optimal constant in the respective modified logarithmic Sobolev inequality in case of the hypercube (note that the Dirichlet form considered in [11, Example 3.7] equals $2\mathcal{E}(f,g)$, with $\mathcal{E}$ given by (13)).

Example 5.8 (*Weighted 4-cycles*)**.** Let $X = \{x_1, x_2, x_3, x_4\}$ with $k(x_i, x_j) = 0$, if i and j are both even or both odd, and $k(x_1,x_2) = k(x_4,x_3) = a_+$, $k(x_2,x_1) = k(x_3,x_4) = a_-$, $k(x_1,x_4) = k(x_2,x_3) = b_+$ and $k(x_4,x_1) = k(x_3,x_2) = b_-$, with $a_+, a_-, b_+, b_- > 0$. Hence, the underlying graph to L is a weighted 4-cycle.

Let $f \in \mathbb{R}^X$, then we observe from (40)

$$\begin{aligned}
2\Psi_{2,\Upsilon}(f)(x_1) = a_+b_+\Big(&\Upsilon(f(x_3)-f(x_2)) - \Upsilon'(f(x_2)-f(x_1))(f(x_3)-f(x_2)) \\
&+ \Upsilon(f(x_3)-f(x_4)) - \Upsilon'(f(x_4)-f(x_1))(f(x_3)-f(x_4))\Big) \\
&+ a_+a_-\big(\Upsilon(f(x_1)-f(x_2)) + \Upsilon'(f(x_2)-f(x_1))(f(x_2)-f(x_1))\big) \\
&+ b_+b_-\big(\Upsilon(f(x_1)-f(x_4)) + \Upsilon'(f(x_4)-f(x_1))(f(x_4)-f(x_1))\big) \\
&+ a_+^2\Upsilon'(f(x_2)-f(x_1))(f(x_2)-f(x_1)) + b_+^2\Upsilon'(f(x_4)-f(x_1))(f(x_4)-f(x_1)) \\
&+ a_+b_+\Big(\Upsilon'(f(x_2)-f(x_1))(f(x_4)-f(x_1)) + \Upsilon'(f(x_4)-f(x_1))(f(x_2)-f(x_1))\Big) \\
&- (a_+ + b_+)\big(a_+\Upsilon(f(x_2)-f(x_1)) + b_+\Upsilon(f(x_4)-f(x_1))\big).
\end{aligned}$$

Now, we minimize this expression with respect to the value of $f(x_3)$. For that purpose, we consider the mapping $r \mapsto \Upsilon(r-c_1) - \Upsilon'(c_1-c_0)(r-c_1) + \Upsilon(r-c_2) - \Upsilon'(c_2-c_0)(r-c_2)$, with $c_0, c_1, c_2 \in \mathbb{R}$, which attains its minimum at $r = c_1 + c_2 - c_0$. Henceforth we will use the notation $t = f(x_2) - f(x_1)$ and $s = f(x_4) - f(x_1)$. We deduce

$$\begin{aligned}
2\Psi_{2,\Upsilon}(f)(x_1) &\geqslant a_+\Big(\Upsilon'(t)t(a_- + a_+) + \Upsilon(-t)a_- - \Upsilon(t)a_+\Big) \\
&\quad + b_+\Big(\Upsilon'(s)s(b_- + b_+) + \Upsilon(-s)b_- - \Upsilon(s)b_+\Big) \\
&= a_+a_-\,\nu_{1+\frac{a_+}{a_-},\frac{a_+}{a_-}}(t) + b_+b_-\,\nu_{1+\frac{b_+}{b_-},\frac{b_+}{b_-}}(s).
\end{aligned}$$

Due to the symmetric structure, we can perform analogous computations at x_2, x_3 and x_4. Then, we deduce that $CD_\Upsilon(\kappa,\infty)$ is satisfied for some $\kappa > 0$ provided that the function $\nu_{\lambda_1,\lambda_2}$ is non negative for any pair of parameters

$$(\lambda_1,\lambda_2) \in \Big\{\big(1+\frac{a_+}{a_-}, \frac{a_+ + 2\kappa}{a_-}\big), \big(1+\frac{a_-}{a_+}, \frac{a_- + 2\kappa}{a_+}\big), \big(1+\frac{b_+}{b_-}, \frac{b_+ + 2\kappa}{b_-}\big), \big(1+\frac{b_-}{b_+}, \frac{b_- + 2\kappa}{b_+}\big)\Big\}.$$

We employ Lemma A.4(ii) to conclude after a short computation that this is satisfied if

$$\kappa \leqslant \min\{\sqrt{2a_{\min}(a_{\min}+a_{\max})}, a_{\min}+a_{\max}\} = \sqrt{2a_{\min}(a_{\min}+a_{\max})}$$

and

$$\kappa \leqslant \min\{\sqrt{2b_{\min}(b_{\min}+b_{\max})}, b_{\min}+b_{\max}\} = \sqrt{2b_{\min}(b_{\min}+b_{\max})},$$

where $a_{\min} = \min\{a_-, a_+\}, a_{\max} = \max\{a_-, a_+\}, b_{\min} = \min\{b_-, b_+\}$ and $b_{\max} = \max\{b_-, b_+\}$.

Consequently, $CD_\Upsilon(\min\{\sqrt{2a_{\min}(a_{\min}+a_{\max})}, \sqrt{2b_{\min}(b_{\min}+b_{\max})}\}, \infty)$ holds true. Note, that this generalizes the result on the (unweighted) 2-dimensional hypercube, which satisfies the optimal bound $CD_\Upsilon(2,\infty)$ by Example 5.6.

A class of examples having non-negative Bakry-Émery curvature is given by Ricci-flat graphs, which were originally introduced in [15]. We recall the definition for the reader's convenience.

Definition 5.9. *Let $G=(V,E)$ be an unweighted d-regular graph. We call G Ricci-flat at $x \in V$ if there exist maps $\eta_i : B_1(x) \to V$ (where $B_1(x)$ is the closed ball with radius 1 and center x with respect to the combinatorical graph distance) for $1 \leqslant i \leqslant d$ satisfying the following properties:*

(i) *$\eta_i(u) \in B_1(u)\backslash\{u\}$ for any $u \in B_1(x)$,*
(ii) *$\eta_i(u) \neq \eta_j(u)$, whenever $i \neq j$,*
(iii) *$\bigcup_j \eta_j(\eta_i(x)) = \bigcup_j \eta_i(\eta_j(x))$ for any $i \in \{1,...,d\}$.*

We say G is Ricci-flat if G is Ricci-flat at each $x \in V$.

In [50] it was shown already that Ricci-flat graphs satisfy $CD\log(d,0)$ with some finite $d>0$. As a consequence, $CD_\Upsilon(0,\infty)$ holds, cf. Remark 2.9(i). See also [19]. Nevertheless we revisit this example by giving an argument that uses the convexity of the mapping $r \mapsto \Upsilon(r)$. This argument will also be crucial below to treat the case of (R)-Ricci-flat graphs.

Example 5.10 (*Ricci-flat graphs*)**.** Let the transition rates are chosen in such a way that the underlying graph to L is Ricci-flat with vertex set X. The invariant and reversible measure is given by the uniform measure on X. Then we can write

$$\begin{aligned}2\Psi_{2,\Upsilon}(f)(x) &= L\Psi_\Upsilon(f)(x) - B_{\Upsilon'}(f,Lf)(x)\\ &= \sum_{i=1}^d \Big(\Psi_\Upsilon(f(\eta_i(x)) - \Psi_\Upsilon(f(x))) - \Upsilon'\big(f(\eta_i(x))-f(x)\big)\big(Lf(\eta_i(x))-Lf(x)\big)\Big).\end{aligned}$$

Applying property (iii) from Definition 5.9, yields

$$\begin{aligned}&\sum_{i=1}^d \Upsilon'\big(f(\eta_i(x))-f(x)\big)\big(Lf(\eta_i(x))-Lf(x)\big)\\ &\quad= \sum_{j=1}^d\sum_{i=1}^d \Upsilon'\big(f(\eta_i(x))-f(x)\big)\big(f(\eta_j(\eta_i(x)))-f(\eta_i(x))-(f(\eta_j(x))-f(x))\big)\\ &\quad= \sum_{j=1}^d\sum_{i=1}^d \Upsilon'\big(f(\eta_i(x))-f(x)\big)\big(f(\eta_i(\eta_j(x)))-f(\eta_i(x))-(f(\eta_j(x))-f(x))\big).\end{aligned}$$

Hence, we can write

(72)

$$\begin{aligned}2\Psi_{2,\Upsilon}(f)(x) = \sum_{j=1}^d\sum_{i=1}^d \Big(&\Upsilon\big(f(\eta_j(\eta_i(x)))-f(\eta_i(x))\big) - \Upsilon\big(f(\eta_j(x))-f(x)\big)\\ &- \Upsilon'\big(f(\eta_j(x))-f(x)\big)\big(f(\eta_j(\eta_i(x)))-f(\eta_i(x))-(f(\eta_j(x))-f(x))\big)\Big).\end{aligned}$$

By convexity of $r \mapsto \Upsilon(r)$ we have $\Upsilon(a)-\Upsilon(b) \geqslant \Upsilon'(b)(a-b)$, $a,b \in \mathbb{R}$. Setting $a = f(\eta_j(\eta_i(x)))-f(\eta_i(x))$ and $b = f(\eta_j(x))-f(x)$, yields that L satisfies $CD_\Upsilon(0,\infty)$. This is optimal in general,

due to the existence of Ricci-flat graphs which do not allow for positive Bakry-Émery curvature, e.g. the discrete Laplacian on the lattice $\mathbb{Z}^d$, $d \in \mathbb{N}$.

The situation of the latter example improves if we restrict on a smaller class of graphs. Very recently the notion of (R)-Ricci flatness was introduced in [17]. Again, we repeat the definition.

Definition 5.11. *We call a graph $G = (V, E)$ (R)-Ricci-flat in $x \in V$ if it is Ricci-flat in $x \in V$ and there exist mappings η_i, $1 \leqslant i \leqslant d$, satisfying the conditions from Definition 5.9 and additionally the reflexivity condition $\eta_i(\eta_i(x)) = x$ for any $i \in \{1, ..., d\}$. We say G is (R)-Ricci-flat if G is (R)-Ricci-flat at each $x \in V$.*

The complete bipartite graphs $K_{d,d}$ or the Shrikhande graph are important examples of (R)-Ricci-flat graphs (cf. [17]). Furthermore, in [17] it is shown that the Bakry-Émery condition $CD(2, \infty)$ holds given that the underlying graph is (R)-Ricci-flat. The following example shows an analogue behaviour regarding the CD_Υ condition.

Example 5.12 (*(R)-Ricci-flat graphs*)**.** Let the transition rates be chosen in such a way that the underlying graph to L is (R)-Ricci-flat with vertex set X. The invariant and reversible measure is given by the uniform measure on X.

The convexity argument from Example 5.10 yields in particular that each summand in (72) is non-negative. Applying the reflexivity condition leads to

$$\begin{aligned} 2\Psi_{2,\Upsilon}(f)(x) &\geqslant \sum_{i=1}^{d} \Big(\Upsilon\big(f(x) - f(\eta_i(x))\big) - \Upsilon\big(f(\eta_i(x)) - f(x)\big) \\ &\qquad + 2\Upsilon'\big(f(\eta_i(x)) - f(x)\big)\big(f(\eta_i(x)) - f(x)\big)\Big) \\ &= \sum_{i=1}^{d} \nu_{2,1}(f(\eta_i(x)) - f(x)) \\ &= 4\Psi_\Upsilon(f)(x) + \sum_{i=1}^{d} \nu_{2,5}(f(\eta_i(x)) - f(x)). \end{aligned}$$

We conclude from Lemma A.3 that L satisfies $CD_\Upsilon(2, \infty)$. This is optimal in general, since $CD(2, \infty)$ is optimal for any (unweighted) triangle free graph (see Theorem 1.2 in [36]) and each $K_{d,d}$ is both, triangle free and (R)-Ricci-flat.

Example 5.13 (*Birth-death processes*)**.** Here, we consider a birth-death process with state space given by $X = \{0, ..., N\}$ for some $N \in \mathbb{N}$ or even the infinite case of $X = \mathbb{N}_0$. Following the notation of [14], we introduce the functions $a, b : X \to [0, \infty)$ with $a(x) = k(x, x + 1)$, $b(x) = k(x, x - 1)$, $b(0) = 0$, $b(x) > 0$ otherwise, and $a(x) > 0$ for any $x \in X$ if X is infinite and for any $x \in X\backslash\{N\}$ if X is finite, in which case $a(N) = 0$. Moreover, we set $k(x, y) = 0$ whenever $|x - y| > 1$. The detailed balance condition now reads as

$$a(x)\pi(x) = b(x + 1)\pi(x + 1) \tag{73}$$

for any $x \in X$ in the infinite case respectively for any $x \in X\backslash\{N\}$ if X is of finite size. Note that in the infinite case, the measure μ given by $d\mu = \pi d\#$ is a finite measure if and only if

$$\sum_{x=1}^{\infty} \frac{a(x - 1) \cdots a(0)}{b(x) \cdots b(1)} < \infty.$$

The method of [14] applies quite successfully to birth-death processes if one assumes monotonicity of the rates in the sense that

$$a(x) \leqslant a(x - 1),\ b(x - 1) \leqslant b(x), \tag{74}$$

for any $x \in X\backslash\{0\}$ and, additionally, the bound

$$a(x-1) - a(x) + b(x) - b(x-1) \geqslant \kappa, \quad \forall x \in X\backslash\{0\}, \tag{75}$$

for some $\kappa > 0$. Due to [14] this suffices to deduce the modified logarithmic Sobolev inequalitiy with constant $\frac{\kappa}{2}$. For our purposes it will be crucial to assume that the condition

$$\sqrt{2\min\{a(x-1)-a(x), b(x)-b(x-1)\}\big(a(x-1)-a(x)+b(x)-b(x-1)\big)} \geqslant \kappa \tag{76}$$

holds true for some $\kappa > 0$ and any $x \in X\backslash\{0\}$. Note that the condition (76) is stronger than (75). This is not surprising as the $CD_\Upsilon(\kappa,\infty)$ condition (for which (76) is suitable) yields pointwise estimates with regard to the entropy method (cf. Section 3) and hence is a stronger assumption compared to the non-local approach of [14].

In the subsequent lines we will chronologically show the following assertions:

(i) Assuming (74) and (75), $CD(\frac{\kappa}{2},\infty)$ holds.
(ii) The condition (74) implies $CD_\Upsilon(0,\infty)$.
(iii) $CD_\Upsilon(0,\infty)$ is best possible for the Poisson case on $\mathbb{N}_0$, even though (74) and (75) are satisfied.
(iv) If strict monotonicity holds in (74), for both a and b, and (76) is valid, then $CD_\Upsilon(\kappa,\infty)$ holds true.

We emphasize that (iv) combined with Corollary 3.5 yields MLSI(κ) while the method of [14] (assuming (74) and (75)) implies the modified logarithmic Sobolev inequality with constant $\frac{\kappa}{2}$. For instance, if the left-hand sides of (75) and (76) coincide, which can only happen in the finite state space case, our method thus improves the corresponding constant by a factor 2.

Now, we prove (i). We have by (40)

$$\begin{aligned}
2\Gamma_2(f)(x) &= L\Gamma(f)(x) - 2\Gamma(f, Lf)(x)\\
&= \frac{1}{2}\Big[a(x)a(x+1)\Big((f(x+2)-f(x+1))^2 - 2(f(x+1)-f(x))(f(x+2)-f(x+1))\Big)\\
&\quad + a(x)b(x+1)\Big((f(x)-f(x+1))^2 - 2(f(x+1)-f(x))(f(x)-f(x+1))\Big)\\
&\quad + b(x)a(x-1)\Big((f(x)-f(x-1))^2 - 2(f(x-1)-f(x))(f(x)-f(x-1))\Big)\\
&\quad + b(x)b(x-1)\Big((f(x-2)-f(x-1))^2 - 2(f(x-1)-f(x))(f(x-2)-f(x-1))\Big)\\
&\quad + 2\Big(a(x)(f(x+1)-f(x)) + b(x)(f(x-1)-f(x))\Big)^2\Big]\\
&\quad - (a(x)+b(x))\Gamma(f)(x).
\end{aligned}$$

Here and in the sequel, we set $a(x-1) = b(x-1) = 0$ if $x = 0$ and in the case of finite X, $a(x+1) = b(x+1) = 0$ for $x = N$. Note that in those cases the respective choice of extension for $f \in \mathbb{R}^X$ does not play a role.

Viewing the above expression as a function of $f(x+2)$ and $f(x-2)$, it can be minimized by setting $f(x+2) - f(x+1) = f(x+1) - f(x)$ and $f(x-2) - f(x-1) = f(x-1) - f(x)$. Henceforth, we will use the notation $t = f(x+1) - f(x)$ and $s = f(x-1) - f(x)$. We have

$$\begin{aligned}
4\Gamma_2(f)(x) &\geqslant a(x)t^2\big(3b(x+1) - a(x+1)\big) + b(x)s^2\big(3a(x-1) - b(x-1)\big) + 2\big(a(x)t + b(x)s\big)^2\\
&\quad - (a(x)+b(x))\big(a(x)t^2 + b(x)s^2\big)\\
&= a(x)t^2\big(b(x+1) - b(x) + a(x) - a(x+1)\big)\\
&\quad + b(x)s^2\big(a(x-1) - a(x) + b(x) - b(x-1)\big)
\end{aligned}$$

$$
\begin{aligned}
&\quad + 2\big(b(x+1)a(x)t^2 + a(x-1)b(x)s^2 + 2a(x)b(x)ts\big)\\
&\geqslant \kappa\big(a(x)t^2 + b(x)s^2\big) + 2a(x)b(x)(s+t)^2\\
&\geqslant 2\kappa\Gamma(f)(x),
\end{aligned}
$$

where we have applied (74) and (75). Hence, $CD(\frac{\kappa}{2},\infty)$ holds true.

Regarding (ii), we will follow a similar approach. In this case, we observe by (40)

$$
\begin{aligned}
2\Psi_{2,\Upsilon}(f)(x) &= L\Psi_\Upsilon(f)(x) - B_{\Upsilon'}(f,Lf)(x)\\
&= a(x)a(x+1)\Big(\Upsilon(f(x+2)-f(x+1)) - \Upsilon'(f(x+1)-f(x))(f(x+2)-f(x+1))\Big)\\
&\quad + a(x)b(x+1)\Big(\Upsilon(f(x)-f(x+1)) - \Upsilon'(f(x+1)-f(x))(f(x)-f(x+1))\Big)\\
&\quad + b(x)a(x-1)\Big(\Upsilon(f(x)-f(x-1)) - \Upsilon'(f(x-1)-f(x))(f(x)-f(x-1))\Big)\\
&\quad + b(x)b(x-1)\Big(\Upsilon(f(x-2)-f(x-1)) - \Upsilon'(f(x-1)-f(x))(f(x-2)-f(x-1))\Big)\\
&\quad + \big(a(x)\Upsilon'(f(x+1)-f(x)) + b(x)\Upsilon'(f(x-1)-f(x)\big)Lf(x)\\
&\quad - (a(x)+b(x))\Psi_\Upsilon(f)(x).
\end{aligned}
$$

Again, we minimize the above expression by choosing the values of $f(x+2)$ and $f(x-2)$ accordingly. To that aim, it can be easily checked that the mapping $r \mapsto \Upsilon(r-c_1) - \Upsilon'(c_1-c_2)(r-c_1)$ with parameters $c_1, c_2 \in \mathbb{R}$ achieves the global minimum for $r-c_1 = c_1-c_2$. Recalling the notation $t = f(x+1)-f(x)$ and $s = f(x-1)-f(x)$, we obtain

$$
\begin{aligned}
2\Psi_{2,\Upsilon}(f)(x) &\geqslant a(x)a(x+1)\big(\Upsilon(t) - \Upsilon'(t)t\big) + a(x)b(x+1)\big(\Upsilon(-t) + \Upsilon'(t)t\big)\\
&\quad + b(x)a(x-1)\big(\Upsilon(-s) + \Upsilon'(s)s\big) + b(x)b(x-1)\big(\Upsilon(s) - \Upsilon'(s)s\big)\\
&\quad + \big((a(x)\Upsilon'(t) + b(x)\Upsilon'(s)\big)\big(a(x)t + b(x)s\big) - (a(x)+b(x))\big(a(x)\Upsilon(t) + b(x)\Upsilon(s)\big)\\
&= a(x)\big[\Upsilon(t)(a(x+1)-a(x)-b(x)) + \Upsilon(-t)b(x+1)\\
&\qquad\quad + \Upsilon'(t)t(b(x+1)+a(x)-a(x+1))\big]\\
&\quad + b(x)\big[\Upsilon(s)(b(x-1)-b(x)-a(x)) + \Upsilon(-s)a(x-1)\\
&\qquad\quad + \Upsilon'(s)s(a(x-1)+b(x)-b(x-1))\big]\\
&\quad + a(x)b(x)\big(\Upsilon'(t)s + \Upsilon'(s)t\big).
\end{aligned}
$$

The last summand, which is of mixed type with respect to s and t, can be treated by the inequality $e^\beta\beta - e^\beta\alpha - e^\beta + e^\alpha \geqslant 0$, $\alpha,\beta\in\mathbb{R}$, which follows from convexity of the exponential function. In fact, we conclude

$$
\begin{aligned}
\Upsilon'(t)s + \Upsilon'(s)t = e^t s - s + e^s t - t &\geqslant e^t - e^t t - e^{-t} - t + e^s - e^s s - e^{-s} - s\\
&= \Upsilon(t) - \Upsilon(-t) - \Upsilon'(t)t + \Upsilon(s) - \Upsilon(-s) - \Upsilon'(s)s.
\end{aligned}
\tag{77}
$$

The major advantage of the right-hand side of (77) is that no mixed terms occur anymore, which allows to consider the terms involving s resp. t separately. We have

$$
\begin{aligned}
&\Upsilon(t)\big(a(x+1)-a(x)-b(x)\big) + \Upsilon(-t)b(x+1) + \Upsilon'(t)t\big(b(x+1)+a(x)-a(x+1)\big)\\
&\qquad + b(x)\big(\Upsilon(t) - \Upsilon(-t) - \Upsilon'(t)t\big)\\
&\quad = b_{x,+}\big(\Upsilon'(t)t + \Upsilon(-t)\big) + a_{x,+}\big(\Upsilon'(t)t - \Upsilon(t)\big),
\end{aligned}
$$

where $a_{x,+} = a(x) - a(x+1)$ and $b_{x,+} = b(x+1) - b(x)$ and analogously

$$
\Upsilon(s)\big(b(x-1)-b(x)-a(x)\big) + \Upsilon(-s)a(x-1) + \Upsilon'(s)s\big(a(x-1)+b(x)-b(x-1)\big)
$$

$$
\begin{aligned}
&+ a(x)\big(\Upsilon(s) - \Upsilon(-s) - \Upsilon'(s)s\big) \\
&= b_{x,-}\big(\Upsilon'(s)s - \Upsilon(s)\big) + a_{x,-}\big(\Upsilon'(s)s + \Upsilon(-s)\big),
\end{aligned}
$$

where $a_{x,-} = a(x-1) - a(x)$ and $b_{x,-} = b(x) - b(x-1)$.

In total, we thus have

$$
\begin{aligned}
2\Psi_{2,\Upsilon}(f)(x) \geqslant\ & a(x)\big(b_{x,+}(\Upsilon'(t)t + \Upsilon(-t)) + a_{x,+}(\Upsilon'(t)t - \Upsilon(t))\big) \\
&+ b(x)\big(b_{x,-}(\Upsilon'(s)s - \Upsilon(s)) + a_{x,-}(\Upsilon'(s)s + \Upsilon(-s))\big).
\end{aligned} \tag{78}
$$

Whenever $b_{x,-}, a_{x,-}, b_{x,+}, a_{x,+} \geqslant 0$ we deduce that $CD_\Upsilon(0,\infty)$ is valid in x. Hence, the global monotonicity condition (74) implies $CD_\Upsilon(0,\infty)$, which shows (ii).

The Poisson case on $\mathbb{N}_0$ shows that this bound is in general sharp, even if the condition (75) is additionally assumed. Indeed, let $a(x) = \lambda$, $\lambda > 0$, and $b(x) = x$ for any $x \in \mathbb{N}_0$. Then the corresponding Markov chain coincides with the Poisson case of the birth-death process on $\mathbb{N}_0$ with invariant probability measure given by $d\mu = \pi_\lambda d\#$ and $\pi_\lambda(x) = \frac{\lambda^x}{x!}e^{-\lambda}$, $x \in \mathbb{N}_0$. The convexity inequality (77) becomes an equality if $s = -t$. For $n \geqslant 2$ and $\tau > 0$, we define the functions $f_{n,\tau} : \mathbb{N}_0 \to \mathbb{R}$ by $f_{n,\tau}(x) = \tau(x-n)$ for $x \in \{n-2, n-1, n+1, n+2\}$ and 0 otherwise. For this particular choice we have equality in (78) at $x = n$. More precisely,

$$
2\Psi_{2,\Upsilon}(f_{n,\tau})(n) = \lambda\big(\Upsilon'(\tau)\tau + \Upsilon(-\tau)\big) + n\big(\Upsilon'(-\tau)(-\tau) - \Upsilon(-\tau)\big).
$$

Consequently, for $\kappa > 0$ the estimate $2\Psi_{2,\Upsilon}(f_{n,\tau})(n) \geqslant \kappa\Psi_\Upsilon(f_{n,\tau})(n)$ is equivalent to

$$
\lambda\big(\Upsilon'(\tau)\tau + \Upsilon(-\tau) - \kappa\Upsilon(\tau)\big) + n\big(\Upsilon'(-\tau)(-\tau) - (1+\kappa)\Upsilon(-\tau)\big) \geqslant 0. \tag{79}
$$

Choosing $\tau > 0$ large enough such that $\Upsilon'(-\tau)(-\tau) - (1+\kappa)\Upsilon(-\tau) < 0$ and sending $n \to \infty$ yields a contradiction to (79). This shows assertion (iii). Interestingly, since here the assumptions (74) and (75) are satisfied, this gives an example of a Markov chain satisfying $CD(\frac{1}{2},\infty)$ by assertion (i), where a positive CD_Υ curvature condition is impossible to hold.

Next we aim for positive CD_Υ curvature bounds. We assume that $b_{x,+}, a_{x,-} > 0$ for any $x \in X$. Note that this does not violate the conditions on the boundary, as we have $b_{0,-} = 0$ and, in the finite case, $a_{N,+} = 0$. Recall the inequality (78). The estimate

$$
a(x)\big(b_{x,+}(\Upsilon'(t)t + \Upsilon(-t)) + a_{x,+}(\Upsilon'(t)t - \Upsilon(t))\big) \geqslant 2\kappa a(x)\Upsilon(t), \quad t \in \mathbb{R},
$$

(which is clear for $x = N$ in the finite case) is equivalent for $x \in X$ resp. $x \in X\backslash\{N\}$ in the finite case to

$$
b_{x,+}\nu_{1+\frac{a_{x,+}}{b_{x,+}}, \frac{a_{x,+}}{b_{x,+}} + \frac{2\kappa}{b_{x,+}}}(t) \geqslant 0, \quad t \in \mathbb{R}, \tag{80}
$$

and analogously

$$
b(x)\big(b_{x,-}(\Upsilon'(s)s - \Upsilon(s)) + a_{x,-}(\Upsilon'(s)s + \Upsilon(-s))\big) \geqslant 2\kappa b(x)\Upsilon(s), \quad s \in \mathbb{R},
$$

(which is clear for $x = 0$) is equivalent for $x \in X\backslash\{0\}$ to

$$
a_{x,-}\nu_{1+\frac{b_{x,-}}{a_{x,-}}, \frac{b_{x,-}}{a_{x,-}} + \frac{2\kappa}{a_{x,-}}}(s) \geqslant 0, \quad s \in \mathbb{R}. \tag{81}
$$

We want to employ Lemma A.4(ii) (with $\lambda = 1 + \frac{a_{x,+}}{b_{x,+}}$ for (80) and $\lambda = 1 + \frac{b_{x,-}}{a_{x,-}}$ for (81)). A short computation (as in Example 5.8) shows that if

$$
\kappa \leqslant \sqrt{2\min\{a_{x,-}, b_{x,-}\}(a_{x,-} + b_{x,-})} \tag{82}
$$

holds for $x \in X\backslash\{0\}$ then (81) is satisfied. But since $a_{x-1,+} = a_{x,-}$ and $b_{x-1,+} = b_{x,-}$, (80) also follows from (82). Clearly, (76) and (82) are the same and hence (iv) is established.

We conclude this example with a remark on the infinite state space case. Note that (74) implies $\inf_{x\in\mathbb{N}} a_{x,-} = 0$. Hence, in order to achieve a global positive curvature bound by (76),

we need to choose the sequence $(b_{x,-})_{x\in\mathbb{N}}$ such that $\inf_{x\in\mathbb{N}} a_{x,-}b_{x,-} \geqslant c$ for some constant $c > 0$. This implies in particular that $b(x) \to \infty$, and hence $M_1(x) \to \infty$, as $x \to \infty$. Concerning the modified logarithmic Sobolev inequality the assumptions $M_1 \in \ell^2(\mu)$ and $M_2 \in \ell^1(\mu)$, which were imposed in Theorem 3.4 and Corollary 3.5, boil down to assume that $b \in \ell^2(\mu)$. This is immediate for $M_1 \in \ell^2(\mu)$ and can be easily deduced from the detailed balance condition (73) for $M_2 \in \ell^1(\mu)$. The ratio test implies that $b \in \ell^2(\mu)$ is satisfied if there exists some constant $q \in (0,1)$ such that

$$\frac{a(x)b(x+1)}{b(x)^2} \leqslant q < 1, \tag{83}$$

for all sufficiently large x, where again (73) has been used. Finally, we illustrate the infinite state space case by the following concrete example. We choose the transition rates as

$$a(x) = \sum_{n=x+1}^{\infty} n^{-\alpha},\ b(x) = \sum_{n=0}^{x} n^{\beta}$$

for any $x \in \mathbb{N}_0$ and parameters $\beta \geqslant \alpha > 1$. Then (83) is apparently satisfied. Moreover, we have $a(x-1) - a(x) = x^{-\alpha}$ and $b(x) - b(x-1) = x^{\beta}$ for any $x \in \mathbb{N}$ and thus (76) turns to

$$\sqrt{2(x^{-2\alpha} + x^{\beta-\alpha})} \geqslant \kappa$$

for any $x \in \mathbb{N}$. Clearly, κ can at least be chosen as $\sqrt{2}$ and hence MLSI($\sqrt{2}$) holds. On the other hand, the method of [14] yields at most a modified logarithmic Sobolev inequality with constant 1.

Example 5.14 (*Weighted graphs with large girth*)**.** Now we consider the situation, where the underlying graph has girth at least 5, i.e. if there exist a subset $\{x_1, ..., x_g\} \subset X$, of pairwise distinct elements, with $k(x_i, x_{i+1}) > 0$ for any $i \in \{1, ..., g-1\}$ and $k(x_g, x_1) > 0$, then $g \geqslant 5$ must hold. Further, we assume the existence of $x, y \in X$ with $k(x,y) > 0$ (and hence $k(y,x) > 0$) and such that

$$M_1(x) + M_1(y) - 2\big(k(x,y) + k(y,x)\big) > 0 \tag{84}$$

holds true. We claim that this implies that $CD_\Upsilon(\kappa, \infty)$ fails in x (and by symmetry also in y) for all $\kappa \in \mathbb{R}$.

We remark that any (connected) unweighted graph (here $M_1(x)$ and $M_1(y)$ equal the degree of x and y respectively) with $|X| \geqslant 5$ and girth at least 5 (in particular trees) that is neither a path nor a cycle satisfies (84) for some $x, y \in X$. Indeed, choose $x \in X$ such that x has at least 3 neighbors (which is possible since the underlying graph is neither a path nor a cycle) and y a neighbor of x that has at least one neighbor distinct from x (which exists as $|X| \geqslant 5$). Then (84) holds for x and y.

For the discrete Laplacian on $\mathbb{Z}$ we have equality in (84) at any $x \in \mathbb{Z}$ and $y = x \pm 1$, as well as $CD_\Upsilon(0, \infty)$ holds true by Example 5.10. In this sense, the condition (84) is sharp.

Note that (84) also applies to birth-death processes, which have been discussed in Example 5.13. In fact, using the notation of Example 5.13, (84) translates to

$$a(x) + b(x) + a(x \pm 1) + b(x \pm 1) - 2\big(k(x, x \pm 1) + k(x \pm 1, x)\big) > 0$$

and hence to the condition that $0 > a_{x,+} + b_{x,+}$ resp. $0 > a_{x,-} + b_{x,-}$ holds true for some $x \in X$.

Let us now prove the claim. For $z \in X$, we denote by $S_1(z) = \{v \in X \backslash \{z\} : k(z,v) > 0\}$ the neighborhood of z. Let $x, y \in X$ be given as in (84), $S_1(x) = \{y_1, ..., y_m\}$, $m \geqslant 1$, and let w.l.o.g. $y_1 = y$. We remark that the arguments below also apply to the case that $S_1(x)$ consists of infinitely many elements, as well as to the special case of $m = 1$ (where the situation is easier

as the corresponding sums over the index set $\{2,...,m\}$ are empty). Since the girth of the graph is at least 5, we have $y_j \notin S_1(y_i)$ and $S_1(y_i) \cap S_1(y_j) = \{x\}$ if $i \neq j$. Thus, we can choose for some $\tau \in \mathbb{R}$ a function $f_\tau \in \mathbb{R}^X$ such that $\tau = f_\tau(y_i) - f_\tau(x)$ for any $i \in \{2,...,m\}$, $f_\tau(y_1) - f_\tau(x) = -\tau$ and $f_\tau(z) - f_\tau(y_j) = f_\tau(y_j) - f_\tau(x)$ for any $z \in S_1(y_j)\backslash\{x\}$, $j \in \{1,...,m\}$. For this particular choice, (40) turns to

$$\begin{aligned}
&2\Psi_{2,\Upsilon}(f_\tau)(x)\\
&\quad= \sum_{i=1}^m k(x,y_i)\big(M_1(y_i) - k(y_i,x)\big)\big(\Upsilon(f_\tau(y_i) - f_\tau(x)) - \Upsilon'(f_\tau(y_i) - f_\tau(x))(f_\tau(y_i) - f_\tau(x))\big)\\
&\qquad+ \sum_{i=1}^m k(x,y_i)k(y_i,x)\big(\Upsilon(f_\tau(x) - f_\tau(y_i)) + \Upsilon'(f_\tau(y_i) - f_\tau(x))(f_\tau(y_i) - f_\tau(x))\big)\\
&\qquad+ \sum_{i=1}^m k(x,y_i)^2\Upsilon'(f_\tau(y_i) - f_\tau(x))(f_\tau(y_i) - f_\tau(x)) - M_1(x)\sum_{i=1}^m k(x,y_i)\Upsilon(f_\tau(y_i) - f_\tau(x))\\
&\qquad+ \sum_{\substack{i,j\in\{1,...,m\}\\ i\neq j}} k(x,y_i)k(x,y_j)\big(\Upsilon'(f_\tau(y_i) - f_\tau(x))(f_\tau(y_j) - f_\tau(x))\big).
\end{aligned}$$

Further, we have

$$\begin{aligned}
&\sum_{\substack{i,j\in\{1,...,m\}\\ i\neq j}} k(x,y_i)k(x,y_j)\big(\Upsilon'(f_\tau(y_i) - f_\tau(x))(f_\tau(y_j) - f_\tau(x))\big)\\
&\quad= \sum_{j=2}^m k(x,y_1)k(x,y_j)\big(\Upsilon'(-\tau)\tau - \Upsilon'(\tau)\tau\big) + \sum_{\substack{i,j\in\{2,...,m\}\\ i\neq j}} k(x,y_i)k(x,y_j)\Upsilon'(\tau)\tau\\
&\quad= k(x,y_1)\Upsilon'(-\tau)\tau\big(M_1(x) - k(x,y_1)\big) + \sum_{i=2}^m k(x,y_i)\Upsilon'(\tau)\tau\big(M_1(x) - k(x,y_i) - 2k(x,y_1)\big).
\end{aligned}$$

Combininig the previous identities yields

$$\begin{aligned}
&2\Psi_{2,\Upsilon}(f_\tau)(x)\\
&= \sum_{i=2}^m k(x,y_i)\Big[\Upsilon'(\tau)\tau\big(M_1(x) - M_1(y_i) + 2(k(y_i,x) - k(x,y_1))\big)\\
&\qquad\qquad+ \Upsilon(-\tau)k(y_i,x) + \Upsilon(\tau)\big(M_1(y_i) - M_1(x) - k(y_i,x)\big)\Big]\\
&\quad+ \Upsilon'(-\tau)\tau k(x,y_1)\big(M_1(x) + M_1(y_1) - 2(k(x,y_1) + k(y_1,x))\big) + \Upsilon(\tau)k(x,y_1)k(y_1,x)\\
&\quad+ k(x,y_1)\Upsilon(-\tau)\big(M_1(y_1) - M_1(x) - k(y_1,x)\big)\\
&= C_{1,x}\Upsilon'(\tau)\tau + C_{2,x}\Upsilon(-\tau) + C_{3,x}\Upsilon(\tau)\\
&\qquad+ k(x,y_1)\Upsilon'(-\tau)\tau\big(M_1(x) + M_1(y_1) - 2(k(x,y_1) + k(y_1,x))\big),
\end{aligned}$$

where $C_{1,x}, C_{2,x}, C_{3,x} \in \mathbb{R}$ are constants that depend on x and the transition rates, but not on τ. Now, we observe

$$\begin{aligned}
&2\big(\Psi_{2,\Upsilon}(f_\tau)(x) - \kappa\Psi_\Upsilon(f_\tau)(x)\big)\\
&\quad= C_{1,x}\Upsilon'(\tau)\tau + \big(C_{2,x} - 2\kappa k(x,y_1)\big)\Upsilon(-\tau) + \big(C_{3,x} - 2\kappa(M_1(x) - k(x,y_1))\big)\Upsilon(\tau)\\
&\qquad+ k(x,y_1)\Upsilon'(-\tau)\tau\big(M_1(x) + M_1(y_1) - 2(k(x,y_1) + k(y_1,x))\big),
\end{aligned} \tag{85}$$

where $\kappa \in \mathbb{R}$. Sending $\tau \to -\infty$, the dominating term in the right hand side of (85) is given by $\Upsilon'(-\tau)\tau$, which tends to $-\infty$. By (84), we thus can find for any $\kappa \in \mathbb{R}$ some $\tau < 0$ such that (85) is negative. Hence, $CD_\Upsilon(\kappa, \infty)$ is impossible to hold in x for any $\kappa \in \mathbb{R}$.

Remark 5.15. The condition (84) also appears in the context of Ollivier curvature. Indeed, in [52, Example 2.3] it is shown that if an edge is not contained in a 3-, 4- or 5-cycle, then the condition (84) means exactly that the Ollivier curvature of this edge is negative. Note that the transition rate $k(x, y)$ equals $\frac{\omega(x,y)}{m(x)}$ in the notation of a graph laplacian as in [52], where $\omega(x, y)$ denotes the edge weight between vertices x and y and m is the so called vertex measure.

The latter example has the very interesting consequence that the CD_Υ condition is not robust against small perturbations, which will be demonstrated in the following example.

Example 5.16 (*Perturbed birth-death process*)**.** Let L be the generator of a birth-death process on the state space $X = \{0, ..., N\}$ with $N \geqslant 4$, resp. $X = \mathbb{N}_0$, with invariant and reversible measure μ given by $d\mu = \pi d\#$. The transitions rates $\big(k(x, y)\big)_{x,y\in X}$ are given as explained in Example 5.13 and the notation is taken from there. Now, we add an edge to the underlying graph. Precisely, let $x_0, y_0 \in X$ with $y_0 \geqslant x_0 + 4$ be fixed and let $\lambda_{x_0,y_0} = \frac{\pi(x_0)}{\pi(y_0)}$. Let L_ε denote the Markov generator whose transition rates are given by $k_\varepsilon(x, y) = k(x, y)$, $x \neq y$, if $(x, y) \notin \{(x_0, y_0), (y_0, x_0)\}$, $k_\varepsilon(x_0, y_0) = \varepsilon$ and $k_\varepsilon(y_0, x_0) = \lambda_{x_0,y_0}\varepsilon$. Note that μ is also reversible (and hence invariant) for L_ε since

$$k_\varepsilon(y_0, x_0)\pi(y_0) = \varepsilon\lambda_{x_0,y_0}\pi(y_0) = k_\varepsilon(x_0, y_0)\pi(x_0)$$

holds. Besides that we obtain

$$\begin{aligned} M_1(x_0) + M_1(y_0) - 2\big(k_\varepsilon(x_0, y_0) + k_\varepsilon(y_0, x_0)\big) & \\ &= a(x_0) + b(x_0) + \varepsilon + a(y_0) + b(y_0) + \lambda_{x_0,y_0}\varepsilon - 2\big(\varepsilon + \varepsilon\lambda_{x_0,y_0}\big) \\ &= a(x_0) + b(x_0) + a(y_0) + b(y_0) - \varepsilon\big(1 + \lambda_{x_0,y_0}\big) \end{aligned}$$

and consequently there exists some $\varepsilon_0 = \varepsilon_0(x_0, y_0) > 0$ such that condition (84) is satisfied for x_0 and y_0 and any $\varepsilon \in (0, \varepsilon_0]$. In particular, we can choose for given $\kappa_0 > 0$ the transition rates a and b such that L satisfies $CD_\Upsilon(\kappa_0, \infty)$ (see Example 5.13) while L_ε fails to satisfy $CD_\Upsilon(\kappa, \infty)$ for any $\kappa \in \mathbb{R}$ and $\varepsilon > 0$ small enough by Example 5.14.

In view of Example 5.14 the natural question arises, whether there exists a Markov generator L that does not satisfy $CD_\Upsilon(0, \infty)$ but $CD_\Upsilon(-\kappa, \infty)$ for some $\kappa > 0$. We answer this question positively by the following example.

Example 5.17 (*The unweighted 3-star*)**.** We consider $X = \{x_*, a_1, a_2, a_3\}$ with $k(x_*, a_i) = k(a_i, x_*) = 1$ for any $i \in \{1, 2, 3\}$ and $k(a_i, a_j) = 0$ if $i \neq j$. The underlying graph is given by the unweighted 3-star. Note that this is the only unweighted graph with girth at least 5 that is neither a path nor a cycle to which (84) does not apply. Let $f \in \mathbb{R}^X$ be arbitrary chosen. We set $r = f(a_1) - f(x_*)$, $s = f(a_2) - f(x_*)$ and $t = f(a_3) - f(x_*)$. By (40), we have for $\kappa \geqslant 0$

$$\begin{aligned} &2\big(\Psi_{2,\Upsilon}(f)(x_*) + \kappa\Psi_\Upsilon(f)(x_*)\big) \\ &\quad= \sum_{i=1}^{3} \Upsilon'(f(a_i) - f(x_*))(f(a_i) - f(x_*)) + \Upsilon(f(x_*) - f(a_i)) - (3 - 2\kappa)\Upsilon(f(a_i) - f(x_*)) \\ &\qquad+ \sum_{i=1}^{3} \Upsilon'(f(a_i) - f(x_*)) \sum_{j=1}^{3} (f(a_j) - f(x_*)) \\ &\quad= \nu_{1,3-2\kappa}(r) + \nu_{1,3-2\kappa}(s) + \nu_{1,3-2\kappa}(t) + (r + s + t)(e^r + e^s + e^t - 3). \end{aligned} \tag{86}$$

First we aim to show that $CD_\Upsilon(0,\infty)$ is not satisfied in x_*. Having the approach of Example 5.14 in mind, we choose $s=t$ and $r=-t$. Then (86) (with $\kappa=0$) reads as

$$\begin{aligned}2\Psi_{2,\Upsilon}(f)(x_*) &= \nu_{1,3}(-t)+2\nu_{1,3}(t)+t(e^{-t}+2e^t-3)\\ &= 4\Upsilon'(t)t-5\Upsilon(t)-\Upsilon(-t).\end{aligned}$$

Thus, choosing $t<0$ with $|t|$ sufficiently large shows that the latter becomes negative and consequently $CD_\Upsilon(0,\infty)$ fails in x_*.

Next, we aim to show the existence of $\kappa>0$ such that $CD_\Upsilon(-\kappa,\infty)$ holds. We assume in the sequel that $\kappa\geqslant\frac{3}{2}$. In particular, $\nu_{1,3-2\kappa}$ is non-negative. By (86) there is nothing to show if r,s and t all have the same sign. In case that $r+s+t\geqslant 0$ it follows from Jensen's inequality that

$$e^r+e^s+e^t\geqslant 3e^{\frac{1}{3}(r+s+t)}\geqslant 3.$$

Hence, we are left with the case of $r+s+t<0$, which will be assumed from now on. We write

(87)
$$\begin{aligned}2\big(\Psi_{2,\Upsilon}(f)(x_*)+\kappa\Psi_\Upsilon(f)(x_*)\big) &= (2\kappa-3)\big(\Upsilon(r)+\Upsilon(s)+\Upsilon(t)\big)+\Upsilon(-r)+\Upsilon(-s)+\Upsilon(-t)\\ &\quad+\Upsilon'(r)(s+t+2r)+\Upsilon'(s)(r+t+2s)+\Upsilon'(t)(r+s+2t).\end{aligned}$$

We infer from (87) that for any $R>0$ there exists some $\kappa_R\geqslant\frac{3}{2}$ such that the expression $\Psi_{2,\Upsilon}(f)(x_*)+\kappa\Psi_\Upsilon(f)(x_*)$ is non-negative for any $\kappa\geqslant\kappa_R$ provided that $(r,s,t)\in B_R(0)$ (where $B_R(0)$ denotes the ball with center $(0,0,0)$ and radius R with respect to the euclidean distance in $\mathbb{R}^3$).

1. case: One element of r,s,t is positive while the other two are non-positive (w.l.o.g. $r>0$, $s,t\leqslant 0$).

Clearly, $r+s+2t, r+t+2s<0$ and $\Upsilon'(s),\Upsilon'(t)\leqslant 0$. Thus, the only summand in the right-hand side of (87) that may be negative is given by $\Upsilon'(r)(s+t+2r)$. By convexity we deduce

$$\begin{aligned}\Upsilon(-t)+\Upsilon(-s)+\Upsilon'(r)(s+t+2r) &\geqslant 2\Big(\Upsilon\big(-\frac{t+s}{2}\big)+\Upsilon'(r)\big(\frac{s+t}{2}+r\big)\Big)\\ &\geqslant 2\Big(\Upsilon(r)-\Upsilon'(r)\big(\frac{s+t}{2}+r\big)+\Upsilon'(r)\big(\frac{s+t}{2}+r\big)\Big)\\ &= 2\Upsilon(r)\geqslant 0\end{aligned}$$

and hence the term $\Psi_{2,\Upsilon}(f)(x_*)+\kappa\Psi_\Upsilon(f)(x_*)$ is non-negative by (87).

2. case: One element of r,s,t is negative while the other two are non-negative (w.l.o.g. $r<0$, $s,t\geqslant 0$).

Clearly, $s+t+2r<0$ and $\Upsilon'(r)<0$. Thus, the only summands that may be negative in the right-hand side of (87) are given by $\Upsilon'(t)(r+s+2t)$ and $\Upsilon'(s)(r+t+2s)$. As $\Upsilon'(t),\Upsilon'(s)\geqslant 0$, we are left with $-r>s+2t$ resp. $-r>t+2s$. Due to the asymptotic behavior of the following summands there exists a radius $R_0>0$ such that

$$\Upsilon(-r)+\Upsilon'(t)(r+s+2t)+\Upsilon'(s)(r+t+2s)\geqslant 0$$

for any $(r,s,t)\in B_{R_0}(0)^C$ with $r<0$, $s,t\geqslant 0$ and $-r>t+2s$ or $-r>s+2t$. Consequently, $\Psi_{2,\Upsilon}(f)(x_*)+\kappa\Psi_\Upsilon(f)(x_*)$ is non-negative by (87) for any $\kappa\geqslant\kappa_{R_0}$ in the present case.

Overall, we deduce that $CD_\Upsilon(-\kappa_{R_0},\infty)$ holds at x_*. In a_j, $j\in\{1,2,3\}$, the situation is better. W.l.o.g. let $j=1$. After minimizing $\Psi_{2,\Upsilon}(f)(a_1)$ with respect to $f(a_2)$ and $f(a_3)$ (analogously as explained in Example 5.13), we read from (40)

$$2\Psi_{2,\Upsilon}(f)(a_1)=\Upsilon(f(x_*)-f(a_1))+\Upsilon(f(a_1)-f(x_*))\geqslant 0$$

and hence $CD_\Upsilon(0,\infty)$ holds in a_1.

In [28] it was shown that an unweighted graph with girth at least 5 satisfies the Bakry-Émery condition $CD(0,\infty)$ if and only if it is either a (finite or infinite) path or a cycle of length at least 5 or the 3-star, which has been considered in Example 5.17. Combining Example 5.14 with Example 5.17 and Example 5.10 (paths and cycles are Ricci-flat), the analogue of the mentioned result regarding the CD_Υ condition is the following.

Corollary 5.18. *Let the underlying graph to L be unweighted with girth at least 5. Then L satisfies $CD_\Upsilon(0,\infty)$ if and only if the underlying graph is either a (finite or infinite) path or a cycle of length at least 5.*

Example 5.19 (*Weighted stars*)**.** This example is concerned with a weighted star. Precisely, let $X = \{x_*, a_1, ..., a_m\}$, $m \geqslant 3$, such that $k(x_*, a_i), k(a_i, x_*) > 0$ for any $i \in \{1, ..., m\}$ and $k(a_i, a_j) = 0$ whenever $i \neq j$. We assume that for some $\kappa > 0$

$$k(a_i, x_*) - (M_1(x_*) - k(x_*, a_i)) \geqslant \kappa \tag{88}$$

and

$$k(x_*, a_i) \geqslant \frac{\kappa}{1+\sqrt{3}} \tag{89}$$

hold respectively for any $i \in \{1, ..., m\}$.

First, we show $\Psi_{2,\Upsilon}(f)(x_*) \geqslant \kappa\Psi_\Upsilon(f)(x_*)$ for any $f \in \mathbb{R}^X$. We use the notation $z_i = f(a_i) - f(x)$ for any $i \in \{1, ..., m\}$. From (40) we observe

$$\begin{aligned} 2\Psi_{2,\Upsilon}(f)(x_*) = &\sum_{i=1}^m k(x_*, a_i)k(a_i, x_*)\big(\Upsilon(-z_i) + \Upsilon'(z_i)z_i\big) \\ &+ \sum_{i=1}^m k(x_*, a_i)\Upsilon'(z_i) \sum_{j=1}^m k(x_*, a_j)z_j - M_1(x_*)\sum_{i=1}^m k(x_*, a_i)\Upsilon(z_i). \end{aligned}$$

Applying the convexity inequality (77), yields

$$\sum_{i=1}^m \sum_{\substack{j=1 \\ j\neq i}}^m k(x_*, a_i)k(x_*, a_j)\Upsilon'(z_i)z_j \geqslant \sum_{i=1}^m k(x_*, a_i)(M_1(x_*) - k(x_*, a_i))\Big(\Upsilon(z_i) - \Upsilon(-z_i) - \Upsilon'(z_i)z_i\Big).$$

Consequently, we observe

$$\begin{aligned} 2\Psi_{2,\Upsilon}(f)(x_*) &\geqslant \sum_{i=1}^m k(x_*, a_i)\Big[\Upsilon'(z_i)z_i\big(k(a_i, x_*) - (M_1(x_*) - k(x_*, a_i)) + k(x_*, a_i)\big) \\ &\qquad + \Upsilon(-z_i)\big(k(a_i, x_*) - (M_1(x_*) - k(x_*, a_i))\big) - \Upsilon(z_i)k(x_*, a_i)\Big] \\ &\geqslant \kappa \sum_{i=1}^m k(x_*, a_i)\big(\Upsilon'(z_i)z_i + \Upsilon(-z_i)\big) \end{aligned}$$

by (88) and the convexity of $r \mapsto \Upsilon(r)$. Now, the last term is greater than or equal to $2\kappa\Psi_\Upsilon(f)(x_*)$ by Lemma A.4(ii).

Next, we aim to show $CD_\Upsilon(\kappa,\infty)$ at a_j for $j \in \{1, ..., m\}$. Let $z_* = f(x_*) - f(a_j)$. After minimization over the second neighborhood (cf. Example 5.13), we observe

$$\begin{aligned} 2\Psi_{2,\Upsilon}(f)(a_j) \geqslant &\sum_{\substack{i=1 \\ i\neq j}}^m k(a_j, x_*)k(x_*, a_i)\big(\Upsilon(z_*) - \Upsilon'(z_*)z_*\big) \\ &+ k(a_j, x_*)k(x_*, a_j)\big(\Upsilon(-z_*) + \Upsilon'(z_*)z_*\big) \end{aligned}$$

$$
\begin{aligned}
&+ k(a_j,x_*)^2\big(\Upsilon'(z_*)z_* - \Upsilon(z_*)\big)\\
&= k(a_j,x_*)\Big[\Upsilon(z_*)\big(M_1(x_*) - k(x_*,a_j) - k(a_j,x_*)\big) + \Upsilon(-z_*)k(x_*,a_j)\\
&\quad + \Upsilon'(z_*)z_*\big(k(a_j,x_*) - (M_1(x_*) - k(x_*,a_j)) + k(x_*,a_j)\big)\Big]\\
&\geqslant k(a_j,x_*)\big(\Upsilon'(z_*)z_*(k(x_*,a_j)+\kappa) + k(x_*,a_j)\Upsilon(-z_*) - \kappa\Upsilon(z_*)\big)
\end{aligned}
$$

by (88) and the convexity of $r \mapsto \Upsilon(r)$. Now, the last term is greater than or equal to $2\kappa\Psi_\Upsilon(f)(a_j)$ if and only if

$$
\nu_{1+\frac{\kappa}{k(x_*,a_j)},3\frac{\kappa}{k(x_*,a_j)}}(z_*) \geqslant 0. \tag{90}
$$

It is straightforward to deduce from Lemma A.4(ii) (by setting $\lambda = 1 + \frac{\kappa}{k(x_*,a_j)}$), that (90) is satisfied by the condition (89).

6. Power type entropies and Beckner inequalities

Recall that we assume throughout this section that the invariant and reversible measure μ is a probability measure on X and $(P_t)_{t\geqslant 0}$ is a Markov semigroup. Recall also the notation which have been introduced in the beginning of Section 3.

Beckner inequalities, originally introduced in [10], interpolate between logarithmic Sobolev inequalities and Poincaré inequalities. In the diffusive situation of [7] they can be derived from the classical Bakry-Émery condition with a positive curvature constant, see for instance [5]. As in the case of the Boltzmann entropy, which we have investigated in Section 3, the Bakry-Émery curvature-dimension condition is not suitable in the discrete setting due to the lack of chain rule. Following the approach of [14] and [24], in [33] discrete Beckner inequalities were derived in the setting of Markov chains with finite state space.

In this section we identify the corresponding operators and curvature-dimension conditions which enable us to argue similarly as in Section 3 to obtain Beckner-type inequalities. To that aim, we introduce the entropy functionals

$$
\mathcal{H}_p(\rho) = \int_X \Phi_p(\rho)d\mu = \sum_{x\in X}\Phi_p(\rho)(x)\pi(x), \quad \rho\in\mathcal{P}(X), \tag{91}
$$

where $\Phi_p : [0,\infty)\to\mathbb{R}$ is given by

$$
\Phi_p(r) = \frac{r^p - r}{p(p-1)}, \quad p\in(1,2).
$$

Those functions are included in the class of so called admissible functions in the context of Φ-entropy inequalities, cf. [12].

Definition 6.1. *For any $f\in\ell^{\infty,+}(X)$, we define the operators*

$$
\begin{aligned}
\Psi_\Upsilon^{(p)}(f) &= -\overline{\Lambda}_{\Phi_p'}(f),\\
\Psi_{2,\Upsilon}^{(p)}(f) &= \frac{1}{2}\Big(L\Psi_\Upsilon^{(p)}(f) - B_{\Upsilon'}(\log f, L\Phi_p'(f))\Big),
\end{aligned}
$$

where $\overline{\Lambda}_{\Phi_p'}(f) := \sum_{y\in X}k(x,y)\Lambda_{\Phi_p'}(f(y),f(x))$.

Remark 6.2. Note that the structural similarity of Ψ_Υ and $\Psi_\Upsilon^{(p)}$ is hidden in Lemma 2.2 since $\Psi_\Upsilon(\log f) = -\overline{\Lambda}_{\log}(f)$ holds for $f\in\ell^{\infty,+}(X)$. Further, as $\Phi_p\to\Phi$ resp. $\Phi_p'\to\Phi'$ as $p\to 1$ in the pointwise sense, where $\Phi(r) = r\log r$, $r>0$, is the function that generates the Boltzmann entropy, the operators Ψ_Υ and $\Psi_\Upsilon^{(p)}$ resp. $\Psi_{2,\Upsilon}$ and $\Psi_{2,\Upsilon}^{(p)}$ are related in the following sense. For

any $f \in \ell^{\infty,+}(X)$ we deduce by means of the dominated convergence theorem (recall (11) and (12)) that $L\Phi_p'(f) \to L\log(f)$, $\Psi_\Upsilon^{(p)}(f) \to \Psi_\Upsilon(\log f)$ and $\Psi_{2,\Upsilon}^{(p)}(f) \to \Psi_{2,\Upsilon}(\log f)$ as $p \to 1$ in the pointwise sense, respectively.

The functional

$$\mathcal{I}_p(\rho) = \frac{1}{2-p}\int_X \rho\, \Psi_\Upsilon^{(p)}(\rho) d\mu, \quad \rho \in \mathcal{P}_*^+(X)$$

will play the role of the Fisher information as in Section 3. We have the relation

$$\mathcal{E}(\rho, \Phi_p'(\rho)) = \mathcal{I}_p(\rho). \tag{92}$$

This will be shown in the proof of Theorem 6.8. As in Section 3, we will use (92) to extend $\mathcal{I}_p$ to functions $\rho \in \mathcal{P}_*(X)$, where $\mathcal{I}_p(\rho) = \infty$ is allowed.

Definition 6.3. *We say that L satisfies the Beckner inequality* $\mathrm{Bec}(\alpha)$ *with $\alpha > 0$, if*

$$\mathcal{H}_p(f) \leqslant \frac{1}{2\alpha}\mathcal{I}_p(f) \tag{93}$$

holds for any $f \in \mathcal{P}_*(X)$ with $\mathcal{H}_p(f) < \infty$.

Remark 6.4. Note that (93) is in fact equivalent to the discrete Beckner inequality considered in [33]. This follows from the relation (92).

Lemma 6.5. *Assume that* (93) *holds for $\alpha > 0$ and any $f \in P_*^+(X)$. Then L satisfies* $\mathrm{Bec}(\alpha)$.

Proof. This follows by the analogous truncation argument as it has been described in the proof of Lemma 3.2. □

Definition 6.6. *The Markov generator L is said to satisfy $CD_\Upsilon^{(p)}(\kappa, \infty)$ at $x \in X$ for $\kappa \in \mathbb{R}$, if*

$$\Psi_{2,\Upsilon}^{(p)}(f)(x) \geqslant \frac{\kappa}{2-p}\Psi_\Upsilon^{(p)}(f)(x) \tag{94}$$

holds for any $f \in \ell^{\infty,+}(X)$. If L satisfies $CD_\Upsilon^{(p)}(\kappa, \infty)$ at any $x \in X$, then we say that L satisfies $CD_\Upsilon^{(p)}(\kappa, \infty)$.

Remark 6.7. (i) The assumptions (11) and (12) guarantee that $\Psi_\Upsilon^{(p)}$ and $\Psi_{2,\Upsilon}^{(p)}$ are both well defined. We remark, as in Section 2, that one may choose different function spaces that possibly depend on the kernel. In particular, see Remark 6.10.

(ii) In case that the underlying graph to L is locally finite, (94) holds true on all positive functions $f \in \mathbb{R}^X$ if and only if it is satisfied on $\ell^{\infty,+}(X)$. This follows by an analogous reason as explained in Remark 2.8(ii).

Analogously to Theorem 3.4, the curvature-dimension condition of Definition 6.6 yields a differential inequality regarding the heat flow along the entropy functional defined in (91).

Theorem 6.8. *Let $M_1 \in \ell^2(\mu)$ and $M_2 \in l^1(\mu)$. Then for any $f \in \mathcal{P}_*^+(X)$, we have*

$$\frac{d}{dt}\mathcal{H}_p(P_t f) = -\mathcal{I}_p(P_t f)$$

and

$$\frac{d^2}{dt^2}\mathcal{H}_p(P_t f) = 2\int_X P_t f \Psi_{2,\Upsilon}^{(p)}(P_t f) d\mu.$$

In particular, if the Markov generator L satisfies $CD_\Upsilon^{(p)}(\kappa, \infty)$, then the differential inequality

$$\frac{d^2}{dt^2}\mathcal{H}_p(P_t f) \geqslant -2\kappa\frac{d}{dt}\mathcal{H}_p(P_t f)$$

holds for any $t > 0$ and $f \in \mathcal{P}_*^+(X)$.

Proof. As L is symmetric with respect to the $L^2(\mu)$-inner product,

$$\frac{d}{dt}\mathcal{H}_p(P_tf) = \int_X \Phi_p'(P_tf)LP_tf d\mu = \int_X P_tf L\Phi_p'(P_tf)d\mu$$

holds true. With regard to the identity (92), note that the last term equals $-\mathcal{E}(P_tf, \Phi_p'(P_tf))$. By Lemma 2.1, we have

$$L\Phi_p'(P_tf) = \Phi_p''(P_tf)LP_tf + \overline{\Lambda}_{\Phi_p'}(P_tf). \tag{95}$$

Besides that, we observe due to μ being invariant

$$\int_X P_tf\Phi_p''(P_tf)LP_tf d\mu = (p-1)\int_X \Phi_p'(P_tf)LP_tf d\mu$$

and consequently

$$(1-(p-1))\int_X \Phi_p'(P_tf)LP_tf d\mu = \int_X P_tf\overline{\Lambda}_{\Phi_p'}(P_tf)d\mu,$$

which yields

$$\frac{d}{dt}\mathcal{H}_p(P_tf) = \frac{1}{2-p}\int_X P_tf\overline{\Lambda}_{\Phi_p'}(P_tf)d\mu = -\frac{1}{2-p}\int_X P_tf\ \Psi_\Upsilon^{(p)}(P_tf)d\mu.$$

Regarding the second derivative, the following representation formula will be crucial

$$\begin{aligned}
\frac{d}{dt}\overline{\Lambda}_{\Phi_p'}(P_tf) &= \sum_{y\in X} k(x,y)\frac{d}{dt}\Big(\Phi_p'(P_tf)(y) - \Phi_p'(P_tf)(x) - \Phi_p''(P_tf)(x)(P_tf(y)-P_tf(x))\Big)\\
&= \sum_{y\in X} k(x,y)\big(\Phi_p''(P_tf)(y)LP_tf(y) - \Phi_p''(P_tf)(x)LP_tf(x)\big)\\
&\quad - \Phi_p'''(P_tf)(x)(LP_tf)(x)\sum_{y\in X} k(x,y)(P_tf(y)-P_tf(x))\\
&\quad - \Phi_p''(P_tf)(x)\sum_{y\in X} k(x,y)(LP_tf(y)-LP_tf(x))\\
&= L\big(\Phi_p''(P_tf)LP_tf\big)(x) - \Phi_p'''(P_tf)(x)(LP_tf(x))^2 - \Phi_p''(P_tf)(x)L(LP_tf)(x).
\end{aligned}$$

Then we have

$$\begin{aligned}
(2-p)\frac{d^2}{dt^2}\mathcal{H}_p(P_tf) &= \int_X LP_tf\ \overline{\Lambda}_{\Phi_p'}(P_tf)d\mu + \int_X P_tf\frac{d}{dt}\overline{\Lambda}_{\Phi_p'}(P_tf)d\mu\\
&= \int_X P_tf L(\overline{\Lambda}_{\Phi_p'}(P_tf))d\mu + \int_X \Phi_p''(P_tf)(LP_tf)^2 d\mu - \int_X P_tf\ \Phi_p'''(P_tf)(LP_tf)^2 d\mu\\
&\quad - \int_X P_tf\ \Phi_p''(P_tf)L(LP_tf)d\mu\\
&= \int_X P_tf L(\overline{\Lambda}_{\Phi_p'}(P_tf))d\mu + (3-p)\int_X \Phi_p''(P_tf)(LP_tf)^2 d\mu - (p-1)\int_X \Phi_p'(P_tf)L(LP_tf)d\mu.
\end{aligned}$$

Invoking (95), we obtain

$$\begin{aligned}
\int_X \Phi_p''(P_tf)(LP_tf)^2 d\mu &= \int_X L\Phi_p'(P_tf)LP_tf d\mu - \int_X LP_tf\ \overline{\Lambda}_{\Phi_p'}(P_tf)d\mu\\
&= \int_X P_tf\ L\big(L\Phi_p'(P_tf)\big)d\mu - \int_X P_tf\ L\overline{\Lambda}_{\Phi_p'}(P_tf)d\mu
\end{aligned}$$

and hence

$$(96)\quad \frac{d^2}{dt^2}\mathcal{H}_p(P_tf) = \frac{1}{2-p}\Big((p-2)\int_X P_tf\ L(\overline{\Lambda}_{\Phi_p'}(P_tf))d\mu + 2(2-p)\int_X P_tf\ L(L\Phi_p'(P_tf))d\mu\Big).$$

Further, the detailed balance property yields

$$\begin{aligned}
2\int_X P_tf\ L(L\Phi_p'(P_tf))d\mu &= 2\sum_{x\in X} P_tf(x)\sum_{y\in X} k(x,y)\big(L\Phi_p'(P_tf)(y) - L\Phi_p'(P_tf)(x)\big)\pi(x)\\
&= -\sum_{y\in X}\sum_{x\in X} k(y,x)P_tf(x)\big(L\Phi_p'(P_tf)(x) - L\Phi_p'(P_tf)(y)\big)\pi(y)\\
&\quad + \sum_{x\in X} P_tf(x)\sum_{y\in X} k(x,y)\big(L\Phi_p'(P_tf)(y) - L\Phi_p'(P_tf)(x)\big)\pi(x)\\
&= -\sum_{y\in X} P_tf(y)\sum_{x\in X} k(y,x)e^{\log P_tf(x)-\log P_tf(y))}\big(L\Phi_p'(P_tf)(x) - L\Phi_p'(P_tf)(y)\big)\pi(y)\\
&\quad + \sum_{x\in X} P_tf(x)\sum_{y\in X} k(x,y)\big(L\Phi_p'(P_tf)(y) - L\Phi_p'(P_tf)(x)\big)\pi(x)\\
&= -\sum_{y\in X} P_tf(y)\sum_{x\in X} k(y,x)(e^{\log P_tf(x)-\log P_tf(y))}-1)\big(L\Phi_p'(P_tf)(x) - L\Phi_p'(y)\big)\pi(y)\\
&= -\int_X P_tf B_{\Upsilon'}\big(\log P_tf, L\Phi_p'(P_tf)\big)d\mu.
\end{aligned}$$

Inserting this into (96), we end up with

$$\frac{d^2}{dt^2}\mathcal{H}_p(P_tf) = \int_X P_tf\big(L\Psi_\Upsilon^{(p)}(P_tf) - B_{\Upsilon'}\big(\log P_tf, L\Phi_p'(P_tf)\big)d\mu = 2\int_X P_tf\ \Psi_{2,\Upsilon}^{(p)}(P_tf))d\mu.$$

The additional claim now follows by applying $CD_\Upsilon^{(p)}(\kappa,\infty)$. □

Consequently, using the entropy method, as explained in Section 3, and Lemma 6.5, we deduce the following result.

Corollary 6.9. *If the Markov generator L satisfies $CD_\Upsilon^{(p)}(\kappa,\infty)$, $\kappa>0$, $M_1\in\ell^2(\mu)$ and $M_2\in l^1(\mu)$, then L satisfies* $\mathrm{Bec}(\kappa)$.

Remark 6.10. Note that we only applied the $CD_\Upsilon^{(p)}$ condition on functions belonging to $\mathcal{P}_*^+(X)$. Thus, in order to get the statements of Theorem 6.8 and Corollary 6.9 respectively, we could formulate Definition 6.6 on the smaller class of $\mathcal{P}_*^+(X)$.

7. Tensorization of the discrete and diffusion setting

Here, we revisit the tensorization procedure from Section 4 in a different spirit. More generally speaking, given two independent Markov processes $(Z_t)_{t\geqslant 0}$ on a state space E and $(Y_t)_{t\geqslant 0}$ on a state space F, respectively, each of which with a given unique invariant measure, the product $(Z_t,Y_t)_{t\geqslant 0}$ defines a Markov process on the state space $E\times F$. Further, the generator of this process is given by $L_Z\oplus L_Y$, where L_Z denotes the generator of $(Z_t)_{t\geqslant 0}$ and L_Y denotes the generator of $(Y_t)_{t\geqslant 0}$. The invariant measure is given by the product measure of the respective invariant measures and reversibility is valid provided that it is valid for both Markov processes respectively.

The interest in this section lies in the product of two Markov processes where one fits into the classical diffusive setting of [7] and the other one to the discrete Markov chain setting, which has been studied so far in this paper. In this case we will also speak of *hybrid processes*.

As a motivating example let us consider a linear reaction-drift-diffusion system of the form

$$\partial_t \rho_i - \Delta \rho_i - \operatorname{div}(\rho_i \nabla V) = \sum_{j=1}^{n} \alpha_{i,j} \rho_j \quad \text{in } (0,\infty) \times \mathbb{R}^d,\ i = 1, 2, \dots, n, \tag{97}$$

for the chemical concentrations $\rho_i(t,x)$. Here $V \in C^2(\mathbb{R}^d)$ denotes the confining potential. We assume that $e^{-V} \in L^1(\mathbb{R}^d)$. The reaction rates are such that the total mass $\int_{\mathbb{R}^d} \sum_{i=1}^n \rho_i(t,x)\,dx$ is conserved, that is we ask for $\sum_{i=1}^n \alpha_{i,j} = 0$ for all $j \in \{1,\dots,n\}$. Furthermore, we assume that $\alpha_{i,i} < 0$ for all i, and $\alpha_{i,j} \geqslant 0$ whenever $i \neq j$.

A key idea, which has also been used in [25] and [58], is now to introduce a new variable for the coordinates in (97), i.e. to interpret solutions to (97) as functions on the product space $[0,\infty) \times \mathbb{R}^d \times X$, where $X = \{1, 2, \dots, n\}$. Then we can reformulate (97) into

$$\partial_t \varrho(t,x,i) = (L_c^* + L_d^*)\varrho(t,x,i), \tag{98}$$

where $\varrho : [0,\infty) \times \mathbb{R}^d \times X \to \mathbb{R}$ is given by $\varrho(t,x,i) = \rho_i(t,x)$. Here the operator L_c^* acts w.r.t. the space variable $x \in \mathbb{R}^d$ and takes the form $L_c^* v = \Delta v + \operatorname{div}(v \nabla V)$, whereas L_d^* acts w.r.t. the component variable $i \in X$ and reads as $(L_d^* v)(i) = \sum_{j=1}^n \alpha_{i,j} v(j)$. Equation (98) can be viewed as the Fokker-Planck equation associated with a stochastic process which is the tensor product of a generalized Ornstein-Uhlenbeck process with generator $L_c v = \Delta v - \nabla V \cdot \nabla v$ and a finite Markov chain with generator

$$L_d v(i) = \sum_{j \in X} k(i,j) v(j) = \sum_{j \in X} k(i,j)\big(v(j) - v(i)\big), \quad i \in X,$$

with transition rates $k(i,j) = \alpha_{j,i}$. Note that L_c^* is the adjoint of L_c w.r.t. the Lebesgue measure in $\mathbb{R}^d$ (see e.g. [32]) and L_d^* is the adjoint of L_d w.r.t. the counting measure on X.

If V is uniformly convex with Hessian $\nabla^2 V \geqslant \lambda > 0$ (in the sense of positive definite matrices) then the generalized Ornstein-Uhlenbeck process generated by L_c satisfies $CD(\lambda,\infty)$, where the unique invariant measure has the density $Ce^{-V(x)}$ w.r.t. the Lebesgue measure with a normalizing constant $C > 0$. Let us suppose that the Markov chain satisfies the assumptions of Section 2 and has positive curvature in the CD_Υ sense. Is there a notion of curvature (bounds) for such hybrid processes which imply modified logarithmic Sobolev inequalities and exponential decay of the entropy and enjoys a natural tensorization principle similar to Section 4? In particular, can we expect decay of the (relative) entropy for the reaction-drift-diffusion system (97) under the described assumptions? We will answer both questions in the affirmative.

In what follows, let L_c be the generator of a Markov process on the state space E with invariant and reversible probability measure μ_c. We assume that L_c satisfies the diffusion property, cf. [7]. Let further Γ and Γ_2 denote the corresponding carré du champ and iterated carré du champ operator, respectively. Besides that, let L_d denote the generator of a positive recurrent Markov chain on the state space X as in (6), with invariant and reversible probability measure μ_d.

In the following lines we will remain on a formal level with respect to the class of (admissible) functions. Note however, that in case of finite X we could choose $u : E \times X \to \mathbb{R}$ such that $u(\cdot, i)$ is in the function space $\mathcal{A}$ resp. $\mathcal{A}_0^{const+}$, which are defined in [7], for any $i \in X$.

We now introduce operators which are natural substitutes for the carré du champ and iterated carré du champ operator in the hybrid case. Here we use the notation $A \oplus B$ analogously to Section 4, where now A acts on the continuous variable and B on the discrete one. It turns out that the natural replacement of the carré du champ operator in the hybrid case is just given by $\Gamma \oplus \Psi_\Upsilon$. The analogue of the iterated carré du champ operator is introduced in the following definition.

Definition 7.1. *For sufficiently regular functions $u : E \times X \to \mathbb{R}$ we define*

$$\big(\Gamma \oplus \Psi_\Upsilon\big)_2(u) = \frac{1}{2}\Big((L_c \oplus L_d)(\Gamma \oplus \Psi_\Upsilon)(u) - 2\Gamma\big(u, (L_c \oplus L_d)(u)\big) - B_{\Upsilon'}\big(u, (L_c \oplus L_d)(u)\big)\Big).$$

The next Lemma plays the role of Lemma 4.1 in Section 4.

Lemma 7.2. *We have*

$$\big(\Gamma \oplus \Psi_\Upsilon\big)_2(u) \geqslant \big(\Gamma_2 \oplus \Psi_{2,\Upsilon}\big)(u)$$

for sufficiently regular functions $u : E \times X \to \mathbb{R}$.

Proof. Throughout this proof, the symbols L_c, Γ and Γ_2 will always refer to the continuous variable, whereas L_d, Ψ_Υ, $\Psi_{2,\Upsilon}$ and $B_{\Upsilon'}$ will refer to the discrete variable.

We have

$$\begin{aligned}
&2\big(\Gamma \oplus \Psi_\Upsilon\big)_2(u) \\
&\quad = (L_c(\Gamma(u) + \Psi_\Upsilon(u)) + L_d(\Gamma(u) + \Psi_\Upsilon(u)) - 2\Gamma(u, L_c u + L_d u) - B_{\Upsilon'}(u, L_c u + L_d u) \\
&\quad = 2\Gamma_2(u) + 2\Psi_{2,\Upsilon}(u) + L_c(\Psi_\Upsilon(u)) + L_d(\Gamma(u)) - 2\Gamma(u, L_d u) - B_{\Upsilon'}(u, L_c u).
\end{aligned}$$

The diffusion property (17) yields

$$\begin{aligned}
L_c(\Psi_\Upsilon(u))(\cdot, i) &= \sum_{j \in X \setminus \{i\}} k(i,j) L_c\big(\Upsilon(u(\cdot, j) - u(\cdot, i))\big) \\
&= \sum_{j \in X \setminus \{i\}} k(i,j)\big[\Upsilon'(u(\cdot, j) - u(\cdot, i)) L_c(u(\cdot, j) - u(\cdot, i)) \\
&\qquad\qquad + \Upsilon''(u(\cdot, j) - u(\cdot, i))\Gamma(u(\cdot, j) - u(\cdot, i))\big] \\
&= B_{\Upsilon'}(u, L_c u)(\cdot, i) + \sum_{j \in X \setminus \{i\}} k(i,j)\Upsilon''(u(\cdot, j) - u(\cdot, i))\Gamma(u(\cdot, j) - u(\cdot, i)).
\end{aligned}$$

Further, bilinearity yields

$$\begin{aligned}
\Gamma(u, L_d u)(\cdot, i) &= \sum_{j \in X \setminus \{i\}} k(i,j)\Gamma(u(\cdot, i), u(\cdot, j) - u(\cdot, i)) \\
&= \sum_{j \in X \setminus \{i\}} k(i,j)\big(\Gamma(u(\cdot, i), u(\cdot, j)) - \Gamma(u(\cdot, i))\big).
\end{aligned}$$

Consequently, we deduce

$$\begin{aligned}
2\big(\Gamma \oplus \Psi_\Upsilon\big)_2(u)(\cdot, i) &= 2(\Gamma_2 \oplus \Psi_{2,\Upsilon})(u)(\cdot, i) \\
&\quad + \sum_{j \in X \setminus \{i\}} k(i,j)\big[e^{u(\cdot,j) - u(\cdot,i)}\Gamma(u(\cdot, j) - u(\cdot, i)) + \Gamma(u(\cdot, j)) + \Gamma(u(\cdot, i)) \\
&\qquad\qquad - 2\Gamma(u(\cdot, j), u(\cdot, i))\big] \\
&= 2(\Gamma_2 \oplus \Psi_{2,\Upsilon})(u)(\cdot, i) + \sum_{j \in X \setminus \{i\}} k(i,j)\Gamma(u(\cdot, j) - u(\cdot, i))\big(e^{u(\cdot,j) - u(\cdot,i)} + 1\big) \\
&\geqslant 2(\Gamma_2 \oplus \Psi_{2,\Upsilon})(u)(\cdot, i).
\end{aligned}$$

□

With Lemma 7.2 at hand, we are now able to prove the corresponding analogue to Theorem 3.4. The entropy associated to the invariant probability measure $\mu = \mu_c \otimes \mu_d$ is given by

$$\mathcal{H}(\rho) = \int_E \int_X \rho \log(\rho) d\mu, \quad \rho \in \mathcal{P}(E \times X), \tag{99}$$

where $\mathcal{P}(E \times X)$ now denotes all probability densities with respect to μ (and the convention $0 \log 0 = 0$ being assumed). Further, the Fisher information is defined on sufficiently regular positive functions $\rho \in \mathcal{P}(E \times X)$ as

$$\mathcal{I}(\rho) = \int_E \int_X \rho \, (\Gamma \oplus \Psi_\Upsilon)(\log \rho) d\mu.$$

Theorem 7.3. *We have for sufficiently regular positive functions $f \in \mathcal{P}(E \times X)$*

$$\frac{d}{dt} \mathcal{H}(P_t f) = -\mathcal{I}(P_t f)$$

and

$$\frac{d^2}{dt^2} \mathcal{H}(P_t f) = 2 \int_E \int_X P_t f (\Gamma \oplus \Psi_\Upsilon)_2 (\log P_t f) d\mu. \tag{100}$$

Moreover, if L_c satisfies $CD(\kappa_c, \infty)$ and L_d satisfies $CD_\Upsilon(\kappa_d, \infty)$, then the following differential inequality holds true

$$\frac{d^2}{dt^2} \mathcal{H}(P_t f) \geqslant -2\kappa \frac{d}{dt} \mathcal{H}(P_t f),$$

with $\kappa = \min\{\kappa_c, \kappa_d\}$.

Proof. As in the proof of Lemma 7.2, the symbols L_c, Γ and Γ_2 will refer to the continuous variable, while L_d, Ψ_Υ, $\Psi_{2,\Upsilon}$ and $B_{\Upsilon'}$ will refer to the discrete variable.

Recall, that $\mu = \mu_d \otimes \mu_c$ is invariant for $(P_t)_{t \geqslant 0}$ and hence

$$\begin{aligned}
\frac{d}{dt} \mathcal{H}(P_t f) &= \int_E \int_X P_t f (L_c \oplus L_d)(\log P_t f) d\mu \\
&= \int_E \int_X P_t f (L_c \log P_t f + L_d \log P_t f) d\mu \\
&= \int_E \int_X P_t f \Big(\frac{(L_c \oplus L_d)(P_t f)}{P_t f} - (\Gamma \oplus \Psi_\Upsilon)(\log P_t f) \Big) d\mu \\
&= - \int_E \int_X P_t f (\Gamma \oplus \Psi_\Upsilon)(\log P_t f) d\mu,
\end{aligned}$$

where in the second to last equality the diffusion property (20) for the term $L_c \log P_t f$ and Lemma 2.2 for the term $L_d \log P_t f$ were used.

Now, $\frac{d}{dt} \Gamma(\log P_t f) = 2\Gamma(\log P_t f, \frac{d}{dt} \log P_t f)$ and $\frac{d}{dt} \Psi_\Upsilon(\log P_t f) = B_{\Upsilon'}(\log P_t f, \frac{d}{dt} \log P_t f)$ together imply

$$\begin{aligned}
\frac{d^2}{dt^2} \mathcal{H}(P_t f) &= - \int_E \int_X (L_c \oplus L_d)(P_t f)(\Gamma \oplus \Psi_\Upsilon)(\log P_t f) d\mu \\
&\quad - \int_E \int_X P_t f \big(2\Gamma(\log P_t f, \frac{d}{dt} \log P_t f) + B_{\Upsilon'}(\log P_t f, \frac{d}{dt} \log P_t f) \big) d\mu \\
&= - \int_E \int_X P_t f (L_c \oplus L_d)(\Gamma \oplus \Psi_\Upsilon)(\log P_t f) d\mu \\
&\quad - \int_E \int_X P_t f \big(2\Gamma(\log P_t f, (L_c \oplus L_d)(\log P_t f)) + B_{\Upsilon'}(\log P_t f, (L_c \oplus L_d)(\log P_t f)) \big) d\mu \\
&\quad - \int_E \int_X P_t f \big(2\Gamma(\log P_t f, (\Gamma \oplus \Psi_\Upsilon)(\log P_t f)) + B_{\Upsilon'}(\log P_t f, (\Gamma \oplus \Psi_\Upsilon)(\log P_t f)) \big) d\mu,
\end{aligned}$$

where in the latter step (20) and Lemma 2.2 were used. Applying (26) and (25) as explained in Section 3, we observe (100).

If additionally L_c satisfies $CD(\kappa_c, \infty)$ and L_d satisfies $CD_\Upsilon(\kappa_d, \infty)$, with $\kappa_c, \kappa_d > 0$, then we conclude by the aid of Lemma 7.2 and (100)

$$\begin{aligned}\frac{d^2}{dt^2}\mathcal{H}(P_t f) &\geqslant \int_E \int_X P_t f\big(\Gamma_2(\log P_t f) + \Psi_{2,\Upsilon}(\log P_t f)\big)d\mu \\ &\geqslant 2\kappa \int_E \int_X P_t f\big(\Gamma \oplus \Psi_\Upsilon\big)(\log P_t f)d\mu \\ &= -2\kappa \frac{d}{dt}\mathcal{H}(P_t f).\end{aligned}$$

□

Recalling the entropy method, which has been outlined in Section 3, Theorem 7.3 yields the following result.

Corollary 7.4. *Let L_c satisfy $CD(\kappa_c, \infty)$ and L_d satisfy $CD_\Upsilon(\kappa_d, \infty)$, $\kappa_c, \kappa_d > 0$. Then the modified logarithmic Sobolev inequality*

$$\mathcal{H}(f) \leqslant \frac{1}{2\kappa}\mathcal{I}(f),$$

with $\kappa = \min\{\kappa_c, \kappa_d\}$, holds true for sufficiently regular positive functions $f \in \mathcal{P}(E \times X)$.

8. Extensions of the calculus to non-local operators

The aim of this section is to highlight that the calculus developed in Section 2 and 3 not only applies to the discrete setting of Markov chains. In fact, we want to point out on a formal level at least, that the ideas can be transferred to non-local operators of the form

$$\mathcal{L}f(x) = \int_\Omega (f(y) - f(x))k(x, dy), \tag{101}$$

where the integral may have to be understood in the principal value sense. Here $\Omega \subset \mathbb{R}^d$ is a domain, $x \in \Omega$ and the kernel $k : \Omega \times \mathcal{B}(\Omega) \to [0, \infty]$, where $\mathcal{B}(\Omega)$ denotes the Borel σ-algebra on Ω, is such that $k(x, \cdot)$ defines a Borel measure on Ω with $k(x, \{x\}) = 0$ for any $x \in \Omega$ and $k(\cdot, B)$ is a Borel measurable mapping for any $B \in \mathcal{B}(\Omega)$. We impose the reversibility condition $k(x, dy)\mu(dx) = k(y, dx)\mu(dy)$ for a fixed measure $\mu : \mathcal{B}(\Omega) \to [0, \infty]$ in the sense that

$$\int_\Omega \int_\Omega f(x, y)k(x, dy)\mu(dx) = \int_\Omega \int_\Omega f(x, y)k(y, dx)\mu(dy) \tag{102}$$

is valid for a sufficiently large class of measurable functions $f : \Omega \times \Omega \to \mathbb{R}$.

We define

$$\Psi_\Upsilon(f)(x) = \int_\Omega \Upsilon(f(y) - f(x))k(x, dy)$$

and

$$\Psi_{2,\Upsilon}(f)(x) = \frac{1}{2}\big(L\Psi_\Upsilon(f)(x) - B_{\Upsilon'}(f, Lf)(x)\big),$$

with

$$B_{\Upsilon'}(f, g)(x) = \int_\Omega \Upsilon'(f(y) - f(x))(g(y) - g(x))k(x, dy),$$

where again the respective integrals may be given in the principal value sense.

Example 8.1. One of the most prominent non-local operators is given by the fractional Laplace operator. Here we choose $\Omega = \mathbb{R}^d$, $d \in \mathbb{N}$, and

$$k(x, dy) = c_{\beta,d} \frac{dy}{|x-y|^{d+\beta}}, \quad x \in \mathbb{R}^d,$$

with $\beta \in (0,2)$ and normalizing constant

$$c_{\beta,d} = \frac{2^\beta \Gamma(\frac{d+\beta}{2})}{\pi^{\frac{d}{2}} \left|\Gamma(-\frac{\beta}{2})\right|},$$

where Γ now denotes the Gamma function. Hence, in this case (101) reads as

$$\mathcal{L}(f)(x) = -\big(-\Delta\big)^{\frac{\beta}{2}}(f)(x) = c_{\beta,d} \lim_{\varepsilon \to 0} \int_{\mathbb{R}^d \setminus B_\varepsilon(x)} \frac{f(y) - f(x)}{|x-y|^{d+\beta}} dy,$$

where $f \in \mathcal{S}(\mathbb{R}^d)$, the space of Schwartz functions.

Since Υ has a quadratic behaviour near 0, we can write

$$\Psi_\Upsilon(f)(x) = c_{\beta,d} \int_{\mathbb{R}^d} \frac{\Upsilon(f(y) - f(x))}{|x-y|^{d+\beta}} dy, \tag{103}$$

with $f \in \mathcal{S}(\mathbb{R}^d)$, without need to take the principal value. Further, analogously to Proposition 2.16 in the discrete setting, we have

$$\Psi_{2,\Upsilon}(f)(x) = c_{\beta,d}^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{e^{f(x+\sigma)-f(x)}}{2} \frac{\Upsilon\big(f(x+\sigma+h) - f(x+h) - f(x+\sigma) + f(x)\big)}{|h|^{d+\beta}|\sigma|^{d+\beta}} dh d\sigma, \tag{104}$$

for $f \in \mathcal{S}(\mathbb{R}^d)$. Thus, we observe that $-(-\Delta)^{\frac{\beta}{2}}$ satisfies $CD_\Upsilon(0,\infty)$ on $\mathcal{S}(\mathbb{R}^d)$.

In the quite recent work [60], it was shown that the fractional Laplacian fails to satisfy $CD(\kappa, N)$ for any $\kappa \in \mathbb{R}$ and finite $N > 0$. By [60, Lemma 2.1], the corresponding operators read as

$$\Gamma(f)(x) = c_{\beta,d} \int_{\mathbb{R}^d} \frac{(f(x+h) - f(x))^2}{|h|^{d+\beta}} dh \tag{105}$$

and

$$\Gamma_2(f)(x) = c_{\beta,d}^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{[f(x+h+\sigma) - f(x+h) - f(x+\sigma) + f(x)]^2}{|h|^{d+\beta}|\sigma|^{d+\beta}} dh d\sigma, \tag{106}$$

which can also be obtained from the formulas (103) and (104) with the scaling argument that has been used in the proof of Proposition 2.11.

It is not hard to see that the fractional Laplacian does not satisfy $CD(\kappa,\infty)$ if $\kappa > 0$. Indeed, a scaling argument shows that $\Gamma(v)(x) = \lambda^\beta \Gamma(u)(x)$ and $\Gamma_2(v)(x) = \lambda^{2\beta}\Gamma_2(u)(x)$, where $v(x) = u(\lambda x)$ and $u \in \mathcal{S}(\mathbb{R}^d)$ (cf. the proof of Proposition 2.3 in [60]). Suppose that $CD(\kappa,\infty)$ holds true. Then

$$\lambda^\beta \Gamma_2(u)(x) \geqslant \kappa \Gamma(u)(x)$$

for any $\lambda > 0$, which implies that $\kappa \leqslant 0$. This shows in particular, that $CD_\Upsilon(0,\infty)$ is best possible for the fractional Laplacian.

Formally the latter follows as well by applying $\Psi_{2,\Upsilon}$ and Ψ_Υ respectively to a non-trivial linear function $f : \mathbb{R}^d \to \mathbb{R}$. Then $\Psi_{2,\Upsilon}(f) = 0$ while $\Psi_\Upsilon(f) > 0$. This formal argument applies of course to a larger class of kernels that allow for the continuous version of the representation of $\Psi_{2,\Upsilon}$ as in Proposition 2.16.

Now, we aim for the corresponding analogue to Theorem 3.4. For this purpose, we assume that $\mathcal{L}$ is symmetric w.r.t. the $L^2(\mu)$-inner product for a sufficiently large class of functions, where μ is the unique invariant measure on $\big(\Omega, \mathcal{B}(\Omega)\big)$ satisfying condition (102) for the Markov semigroup $(P_t)_{t\geqslant 0}$ generated by $\mathcal{L}$. Then we follow exactly the same strategy as in Section 3, i.e. we differentiate

$$\mathcal{H}(P_t f) = \int_\Omega P_t f \log(P_t f) d\mu,$$

with respect to time, where f is a sufficiently regular probability density with respect to μ. As indicated before, we will remain on a formal level in the subsequent lines and do not consider regularity issues that may particularly come from a possible singularity of the kernel.

Since μ is invariant and $\mathcal{L}$ is symmetric with respect to the $L^2(\mu)$-inner product, we obtain similarly as in Section 3

$$\frac{d}{dt}\mathcal{H}(P_t f) = -\int_\Omega P_t f \Psi_\Upsilon(\log(P_t f)) d\mu$$

for sufficiently regular f.

As to the second derivative, we aim for an analogue of Lemma 2.4. We have

$$\begin{aligned}
&\int_\Omega e^{g(x)} \int_\Omega \big(e^{g(y)-g(x)} - 1\big)\big(h(y) - h(x)\big) k(x, dy)\mu(dx) \\
&\quad = \int_\Omega \int_\Omega \big(e^{g(y)}(h(y) - h(x)) - e^{g(x)}(h(y) - h(x))\big) k(x, dy)\mu(dx) \\
&\quad = \int_\Omega \int_\Omega e^{g(y)}(h(y) - h(x)) k(x, dy)\mu(dx) - \int_\Omega e^{g(x)} \mathcal{L}h(x)\mu(dx),
\end{aligned}$$

where the latter step is valid if $\mathcal{L}(h)(x)$ exists for any $x \in \Omega$ in the principal value sense. Note, that this translates to assume that $\mathcal{L}\Psi_\Upsilon(\log(P_t f))$ exists in the calculation of $\frac{d^2}{dt^2}\mathcal{H}(P_t f)$. Now, by (102) and Fubini's theorem and for $\varepsilon > 0$ small enough

$$\begin{aligned}
&\int_\Omega \int_{\Omega\setminus B_\varepsilon(x)} e^{g(y)}(h(y) - h(x)) k(x, dy)\mu(dx) \\
&\quad = \int_\Omega \int_\Omega \mathbb{1}_{B_\varepsilon(x)^C}(y)\; e^{g(y)}(h(y) - h(x)) k(x, dy)\mu(dx) \\
&\quad = \int_\Omega \int_\Omega \mathbb{1}_{B_\varepsilon(y)^C}(x)\; e^{g(y)}(h(y) - h(x)) k(y, dx)\mu(dy) \\
&\quad = \int_\Omega e^{g(y)} \int_{\Omega\setminus B_\varepsilon(y)} (h(y) - h(x)) k(y, dx)\mu(dy).
\end{aligned}$$

For a kernel without or with a weak singularity we can repeat the above calculation with $\varepsilon = 0$. If the kernel has a non-integrable singularity instead, taking the limit can be quite delicate. We will not discuss this here. On a formal level at least we observe as an continuous analogue to Lemma 2.4 the identity

$$\frac{1}{2}\int_\Omega e^g B_{\Upsilon'}(g, h) d\mu = -\int_\Omega e^g \mathcal{L}h d\mu,$$

for sufficiently regular functions g, h.

With this at hand, we can copy the discrete proof verbatim to the result

$$\frac{d^2}{dt^2}\mathcal{H}(P_t f) = 2\int_\Omega P_t f \Psi_{2,\Upsilon}(\log P_t f) d\mu.$$

Having Section 5 in mind, a continuous version of the operator considered in Example 5.2 on a bounded domain seems like a natural candidate for an operator of the form (101) that satisfies

$CD_\Upsilon(\kappa,\infty)$ for some $\kappa>0$. Surprisingly, the following example shows that this is not case. This emphasizes the difference between the discrete and continuous setting.

Example 8.2. Let $\Omega\subset\mathbb{R}^d$ be a domain and $k(x,dy)=l(y)dy$, where $l:\Omega\to(0,\infty)$ with $l\in L^1(\Omega)$. We consider the invariant and reversible measure given by $\mu(dx)=l(x)dx$, (102) following from a simple application of Fubini's theorem. We have

$$\begin{aligned}\mathcal{L}f(x)&=\int_\Omega (f(y)-f(x))l(y)dy,\\ \Psi_\Upsilon(f)(x)&=\int_\Omega \Upsilon(f(y)-f(x))l(y)dy,\\ B_{\Upsilon'}(f,Lf)(x)&=\int_\Omega \Upsilon'(f(y)-f(x))\big(\mathcal{L}f(y)-\mathcal{L}f(x)\big)l(y)dy,\end{aligned}$$

and using the continuous version of the representation formula (40) we can calculate

$$\begin{aligned}2\Psi_{2,\Upsilon}(f)(x)&=\int_\Omega\int_\Omega\Big(\Upsilon(f(z)-f(y))-\Upsilon'(f(y)-f(x))(f(z)-f(y))\Big)l(z)l(y)\,dz\,dy\\ &\quad+\int_\Omega\int_\Omega\Upsilon'(f(y)-f(x))(f(z)-f(x))l(y)l(z)\,dy\,dz\\ &\quad-\int_\Omega\int_\Omega\Upsilon(f(y)-f(x))l(y)l(z)\,dy\,dz\\ &=\|l\|_1\int_\Omega\big(\Upsilon'(f(y)-f(x))(f(y)-f(x))-\Upsilon(f(y)-f(x))\big)l(y)\,dy\\ &\quad+\int_\Omega\int_\Omega\Upsilon(f(z)-f(y))l(y)l(z)\,dz\,dy.\end{aligned}$$

In particular, $CD_\Upsilon(0,\infty)$ is valid since the mappings $r\mapsto\Upsilon'(r)r-\Upsilon(r)$ and $r\mapsto\Upsilon(r)$ are both non-negative.

Regarding positive curvature bounds, we note that $\Psi_{2,\Upsilon}(f)(x)\geqslant\kappa\Psi_\Upsilon(f)(x)$ with $\kappa>0$ is equivalent to

$$\begin{aligned}0\leqslant\|l\|_1\int_\Omega\Big(\Upsilon'(f(y)-f(x))(f(y)-f(x))-\big(1+\frac{2\kappa}{\|l\|_1}\big)\Upsilon(f(y)-f(x))\Big)l(y)dy\\ +\int_\Omega\int_\Omega\Upsilon(f(z)-f(y))l(y)l(z)dzdy.\end{aligned}\tag{107}$$

Fix some $\alpha<0$. Then we find some $r_\alpha<0$ such that $\Upsilon'(r_\alpha)r_\alpha-(1+\frac{2\kappa}{\|l\|_1})\Upsilon(r_\alpha)\leqslant\alpha<0$.

We define for $\varepsilon>0$ with $B_\varepsilon(x)\subset\Omega$ the mapping $f_\varepsilon:\Omega\to\mathbb{R}$ as $f_\varepsilon(x)=-r_\alpha$ and $f_\varepsilon(y)=0$ for any $y\in\Omega\backslash B_\varepsilon(x)$. Further, let $0\leqslant f_\varepsilon(y)\leqslant -r_\alpha$ for any $y\in\Omega$. The latter ensures that we can find a constant $M_\alpha>0$, which is independent of ε, such that

$$\Upsilon'(f_\varepsilon(y)-f_\varepsilon(x))(f_\varepsilon(y)-f_\varepsilon(x))-\big(1+\frac{2\kappa}{\|l\|_1}\big)\Upsilon(f_\varepsilon(y)-f_\varepsilon(x))\leqslant M_\alpha$$

holds for any $y\in\Omega$. Note that we can even choose $f_\varepsilon\in C_c^\infty(\Omega)$. We observe

$$\begin{aligned}&\|l\|_1\int_\Omega\Big(\Upsilon'(f_\varepsilon(y)-f_\varepsilon(x))(f_\varepsilon(y)-f_\varepsilon(x))-\big(1+\frac{2\kappa}{\|l\|_1}\big)\Upsilon(f_\varepsilon(y)-f_\varepsilon(x))\Big)l(y)dy\\ &=\|l\|_1\Big(\Upsilon'(r_\alpha)r_\alpha-(1+\frac{2\kappa}{\|l\|_1})\Upsilon(r_\alpha)\Big)\int_{\Omega\backslash B_\varepsilon(x)}l(y)dy\\ &\quad+\|l\|_1\int_{B_\varepsilon(x)}\Big(\Upsilon'(f_\varepsilon(y)-f_\varepsilon(x))(f_\varepsilon(y)-f_\varepsilon(x))-\big(1+\frac{2\kappa}{\|l\|_1}\big)\Upsilon(f_\varepsilon(y)-f_\varepsilon(x))\Big)l(y)dy\end{aligned}$$

$$\leqslant \|l\|_1 \alpha \int_{\Omega\setminus B_\varepsilon(x)} l(y)dy + \|l\|_1 M_\alpha \int_{B_\varepsilon(x)} l(y)dy.$$

Besides that, we have

$$\begin{aligned}(108)\quad &\int_\Omega \int_\Omega \big(\Upsilon(f_\varepsilon(z) - f_\varepsilon(y))\big) l(y)l(z)dzdy \\ &= \int_{B_\varepsilon(x)} \int_\Omega \Upsilon(f_\varepsilon(z) - f_\varepsilon(y)) l(y)l(z)dzdy + \int_{\Omega\setminus B_\varepsilon(x)} \int_{B_\varepsilon(x)} \Upsilon(f_\varepsilon(z) - f_\varepsilon(y)) l(y)l(z)dzdy.\end{aligned}$$

By dominated convergence (note that f_ε is uniformly bounded with respect to ε) the right-hand side in (108) tends to 0 as $\varepsilon \to 0$. Further, $\|l\|_1 M_\alpha \int_{B_\varepsilon(x)} l(y)dy \to 0$ as $\varepsilon \to 0$. Since $\alpha < 0$, we conclude that there exists some $\varepsilon_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0]$ (107) with $f = f_\varepsilon$ is not valid.

As $x \in \Omega$ was chosen arbitrary, we deduce that $CD_\Upsilon(\kappa, \infty)$ does not hold at any $x \in \Omega$ with $\kappa > 0$ arbitrary.

9. Miscellanea

9.1. **Link to Münch's Γ^ψ-calculus.** Here, we demonstrate the relation of the CD_Υ condition with Münch's Γ^ψ-calculus of [50], which we already described in the introduction and Remark 2.9(i).

To that aim, we first recall the notation of [50]. The graph Laplacian for an unweighted finite graph is given by

$$\Delta f(x) = \sum_{y\sim x} \big(f(y) - f(x)\big). \tag{109}$$

Here, the notation $y \sim x$, $x, y \in X$, means that x and y are neighbours. Clearly, (109) coincides with the Markov generator L when the transition rates $k(x, y)$ for $x \neq y$ are given by $k(x, y) = 1$ whenever $x \sim y$ and 0 otherwise. Let $\psi \in C^1((0, \infty))$ be a concave function. The ψ-Laplacian is defined for positive functions as

$$\Delta^\psi(f)(x) := \Big(\Delta\Big[\psi\Big(\frac{f}{f(x)}\Big)\Big]\Big)(x).$$

With

$$\bar\psi(y) := \psi'(1)(y-1) - \big(\psi(y) - \psi(1)\big)$$

Münch defines

$$\Gamma^\psi(f) := \Delta^{\bar\psi}(f).$$

It is shown that (cf. [50, Lemma 3.13])

$$\Delta^\psi f = \psi'(1)\frac{\Delta f}{f} - \Gamma^\psi(f).$$

Setting

$$\Omega^\psi(f)(x) := \Big(\Delta\Big[\psi'\Big(\frac{f}{f(x)}\Big)\cdot\frac{f}{f(x)}\Big[\frac{\Delta f}{f} - \frac{(\Delta f)(x)}{f(x)}\Big]\Big]\Big)(x)$$

Münch then defines

$$\Gamma_2^\psi(f) = \frac{1}{2}\Big(\Omega^\psi(f) + \frac{(\Delta f)\Delta^\psi(f)}{f} - \frac{\Delta(f\Delta^\psi(f))}{f}\Big)$$

and introduces the so-called $CD\psi(d, K)$ condition via the inequality

$$\Gamma_2^\psi(f) \geqslant \frac{1}{d}\big(\Delta^\psi f\big)^2 + K\Gamma^\psi(f), \tag{110}$$

which has to hold for all positive functions f on X. The focus in [50] is on the case $K = 0$.

Let us now consider the special case $\psi = \log$. Then we have $\Delta^{\log}(f) = \Delta(\log f)$,

$$\bar{\psi}(y) = y - 1 - \log y = \Upsilon(\log y),$$

in particular $\bar{\psi}(1) = 0$, and thus

$$\begin{aligned}\Gamma^{\log}(f)(x) = \Delta^{\overline{\psi}}(f)(x) &= \Big(\Delta\Big[\overline{\psi}\Big(\frac{f}{f(x)}\Big)\Big]\Big)(x)\\ &= \sum_{y\sim x}\overline{\psi}\Big(\frac{f(y)}{f(x)}\Big) = \sum_{y\sim x}\Upsilon\big(\log(f(y)) - \log(f(x))\big) = \Psi_\Upsilon(\log f)(x).\end{aligned}$$

Moreover, inserting $\psi'(y) = \frac{1}{y}$ into the definition of $\Omega^{\log}$ we find that

$$\Omega^{\log}(f)(x) = \Delta\Big(\frac{\Delta f}{f} - \frac{(\Delta f)(x)}{f(x)}\Big)(x) = \Delta\Big(\frac{\Delta f}{f}\Big)(x).$$

Employing Lemma 2.2 and the identity

$$\Delta(gh) = (\Delta g)h + g\Delta h + 2\Gamma(g,h),$$

we therefore obtain

$$\begin{aligned}2\Gamma_2^{\log}(f) &= \Delta\Big(\frac{\Delta f}{f}\Big) + \frac{(\Delta f)\Delta(\log f)}{f} - \frac{\Delta(f\Delta(\log f))}{f}\\ &= \Delta\big(\Psi_\Upsilon(\log f) + \Delta(\log f)\big) + \frac{(\Delta f)\Delta(\log f)}{f}\\ &\quad - \frac{1}{f}(\Delta f)\Delta(\log f) - \Delta\Delta(\log f) - \frac{2}{f}\Gamma\big(f, \Delta(\log f)\big)\\ &= \Delta\Psi_\Upsilon(\log f) - \frac{2}{f}\,\Gamma\big(f, \Delta(\log f)\big).\end{aligned}$$

Finally,

$$\begin{aligned}\frac{2}{f}\,\Gamma\big(f, \Delta(\log f)\big) &= \sum_{y\sim x}\frac{1}{f(x)}\big((f(y) - f(x)\big)\big((\Delta(\log f))(y) - (\Delta(\log f))(x)\big)\\ &= \sum_{y\sim x}\big(e^{(\log f)(y) - (\log f)(x)} - 1\big)\big((\Delta(\log f))(y) - (\Delta(\log f))(x)\big)\\ &= B_{\Upsilon'}\big(\log f, \Delta(\log f)\big),\end{aligned}$$

and so we conclude that

$$\Gamma_2^{\log}(f) = \Psi_{2,\Upsilon}(\log f),$$

which together with $\Gamma^{\log}(f) = \Psi_\Upsilon(\log f)$ shows that $CD\psi(\infty,\kappa)$ with $\psi = \log$ is equivalent to $CD_\Upsilon(\kappa,\infty)$.

9.2. **Link to entropic Ricci curvature.** For $\rho \in \mathcal{P}(X)$ and $\psi \in \mathbb{R}^X$, as we have indicated in the introduction and Remark 2.9(v), the following objects play a fundamental role in [22]:

$$\mathcal{A}(\rho,\psi) = \frac{1}{2}\sum_{x,y\in X}\big(\psi(x) - \psi(y)\big)^2\theta(\rho(x),\rho(y))k(x,y)\pi(x) \tag{111}$$

and

$$\begin{aligned}\mathcal{B}(\rho,\psi) =& \frac{1}{4}\sum_{x,y\in X}\big(\psi(x)-\psi(y)\big)^2\hat{L}\rho(x,y)k(x,y)\pi(x)\\ &-\frac{1}{2}\sum_{x,y\in X}\big(L\psi(x)-L\psi(y)\big)(\psi(x)-\psi(y))\,\theta(\rho(x),\rho(y))k(x,y)\pi(x),\end{aligned}\tag{112}$$

where $\hat{L}\rho(x,y) = \partial_1\theta(\rho(x),\rho(y))L\rho(x) + \partial_2\theta(\rho(x),\rho(y))L\rho(y)$ and for $s,t>0$ the quantity $\theta(s,t) := \int_0^1 s^{1-p}t^p dp = \frac{s-t}{\log s-\log t}$ denotes the logarithmic mean with the latter identity being valid if $s\neq t$. Note, that by the detailed balance condition and the symmetry of the logarithmic mean, we can rewrite

$$\begin{aligned}&\frac{1}{4}\sum_{x,y\in X}\big(\psi(x)-\psi(y)\big)^2\hat{L}\rho(x,y)k(x,y)\pi(x)\\ &\quad=\frac{1}{2}\sum_{x,y,z\in X}\big(\psi(x)-\psi(y)\big)^2\partial_1\theta(\rho(x),\rho(y))(\rho(z)-\rho(x))k(x,y)k(x,z)\pi(x),\end{aligned}$$

cf. the proof of Proposition 4.3 in [22].

We aim now to show the identities (33) and (34). Choosing $\psi=\log\rho$ in (111), yields

$$\begin{aligned}\mathcal{A}(\rho,\log\rho) =& \frac{1}{2}\sum_{x,y\in X}k(x,y)\big(\rho(y)-\rho(x)\big)\big(\log\rho(y)-\log\rho(x)\big)\pi(x)\\ =&\sum_{x\in X}\rho(x)\Psi_\Upsilon(\log\rho)(x)\pi(x)\\ =&\int_X\rho\Psi_\Upsilon(\log(\rho))d\mu,\end{aligned}$$

as it has been shown in Section 3. Further, we have

$$\begin{aligned}&-\frac{1}{2}\sum_{x,y\in X}\big(L(\log\rho)(x)-L(\log\rho)(y)\big)(\log\rho(x)-\log\rho(y))\theta(\rho(x),\rho(y))k(x,y)\pi(x)\\ &\quad=-\sum_{x\in X}\Gamma(\rho,L(\log\rho))(x)\pi(x)=\sum_{x\in X}L(\log\rho)(x)L\rho(x)\pi(x).\end{aligned}$$

Since $\partial_1\theta(s,t)=\frac{\log s-\log t-\frac{s-t}{s}}{(\log s-\log t)^2}$, we observe

$$\begin{aligned}&\frac{1}{2}\sum_{x,y,z\in X}\big(\log\rho(x)-\log\rho(y)\big)^2\partial_1\theta(\rho(x),\rho(y))(\rho(z)-\rho(x))k(x,y)k(x,z)\pi(x)\\ &\quad=\frac{1}{2}\sum_{x,y,z\in X}\Big(\log\rho(x)-\log\rho(y)-\frac{\rho(x)-\rho(y)}{\rho(x)}\Big)(\rho(z)-\rho(x))k(x,y)k(x,z)\pi(x)\\ &\quad=\frac{1}{2}\sum_{x\in X}\frac{(L\rho(x))^2}{\rho(x)}\pi(x)-\frac{1}{2}\sum_{x\in X}L(\log\rho)(x)L\rho(x)\pi(x).\end{aligned}$$

Consequently, we get

$$\mathcal{B}(\rho,\log\rho)=\frac{1}{2}\sum_{x\in X}L(\log\rho)(x)L\rho(x)\pi(x)+\frac{1}{2}\sum_{x\in X}\frac{(L\rho(x))^2}{\rho(x)}\pi(x).\tag{113}$$

As it is well known (see e.g. [14]), the right hand side of (113) equals $\frac{1}{2}\frac{d^2}{dt^2}\mathcal{H}(P_t\rho)\Big|_{t=0}$ and hence, by using Theorem 3.4, it follows that $\mathcal{B}(\rho,\log\rho)=\int_X\rho\,\Psi_{2,\Upsilon}(\log\rho)d\mu$.

Appendix A. Properties of $\nu_{c,d}$

In this section, we collect some properties of the function $\nu_{c,d} : \mathbb{R} \to \mathbb{R}$, given by

$$\nu_{c,d}(r) = c\Upsilon'(r)r + \Upsilon(-r) - d\Upsilon(r), \quad c, d \in \mathbb{R},$$

which was of fundamental importance in Section 5. We have

$$\begin{aligned}
\nu'_{c,d}(r) &= c\Upsilon''(r)r + \Upsilon'(r)(c-d) - \Upsilon'(-r),\\
\nu''_{c,d}(r) &= e^r(cr + 2c - d) + e^{-r},\\
\nu'''_{c,d}(r) &= e^r(cr + 3c - d) - e^{-r},\\
\nu''''_{c,d}(r) &= e^r(cr + 4c - d) + e^{-r}.
\end{aligned}$$

Note that $\nu_{c,d}(0) = 0$ and $r = 0$ is a critical point for $\nu_{c,d}$.

Lemma A.1. *The following assertions hold:*

(i) *If $c \geqslant d$, then $\nu_{c,d}(r) \geqslant 0$ for any $r \in \mathbb{R}$,*
(ii) *Let $c, d, r \in \mathbb{R}$ such that $\nu_{c,d}(r) \geqslant 0$. Then we have $\nu_{c+h,d+h}(r) \geqslant 0$ for any $h \geqslant 0$.*
(iii) *We have $\nu_{c',d'}(r) \geqslant \nu_{c,d}(r)$ for any $r \in \mathbb{R}$ if $c' \geqslant c$ and $d' \leqslant d$.*

Proof. The first two properties follow immediately from $\Upsilon'(r)r \geqslant \Upsilon(r)$, $r \in \mathbb{R}$. The third property then follows from the fact that $r \mapsto \Upsilon(r)$ is non-negative. □

Lemma A.2. *Let $\lambda \geqslant 1$. Then we have*

$$\nu_{\lambda,h(\lambda)}(r) \geqslant 0$$

for any $r \geqslant 0$, where

$$h(\lambda) = \begin{cases} 2\lambda + 1, & \lambda \geqslant 2,\\ 3\lambda - 1, & 1 \leqslant \lambda < 2. \end{cases}$$

Proof. We have

$$\nu''_{\lambda,h(\lambda)}(0) = \begin{cases} 0, & \lambda \geqslant 2,\\ 2 - \lambda & 1 \leqslant \lambda < 2 \end{cases}$$

and

$$\nu'''_{\lambda,h(\lambda)}(r) = \begin{cases} e^r\big(\lambda r + \lambda - 1\big) - e^{-r}, & \lambda \geqslant 2,\\ e^r\big(\lambda r + 1\big) - e^{-r} & 1 \leqslant \lambda < 2. \end{cases}$$

Note that $\nu'''_{\lambda,h(\lambda)}(r) \geqslant 0$ is equivalent to $e^{2r}\big(\lambda r + \lambda - 1\big) \geqslant 1$ in the case of $\lambda \geqslant 2$ and, for $1 \leqslant \lambda < 2$, to $e^{2r}\big(\lambda r + 1\big) \geqslant 1$. Both conditions are apparently satisfied for any $r \geqslant 0$. Hence, we observe $\nu'''_{\lambda,h(\lambda)}(r) \geqslant 0$ for any $\lambda \geqslant 1$ and any $r \geqslant 0$. This suffices to establish the claim, since we can deduce $\nu''_{\lambda,h(\lambda)}(r) \geqslant 0$ from $\nu''_{\lambda,h(\lambda)}(0) \geqslant 0$ and, by the same argument, $\nu'_{\lambda,h(\lambda)}(r) \geqslant 0$ and $\nu_{\lambda,h(\lambda)}(r) \geqslant 0$. □

Lemma A.3. *Let $\lambda \geqslant 2$. Then $\nu_{\lambda,2\lambda+1}(r) \geqslant 0$ for any $r \in \mathbb{R}$ if and only if $\lambda = 2$.*

Proof. We have

$$\begin{aligned}
\nu'_{\lambda,2\lambda+1}(r) &= e^r(\lambda r - (\lambda + 1)) + \lambda + 2 - e^{-r},\\
\nu''_{\lambda,2\lambda+1}(r) &= e^r(\lambda r - 1) + e^{-r},\\
\nu'''_{\lambda,2\lambda+1}(r) &= e^r(\lambda r + \lambda - 1) - e^{-r}.
\end{aligned}$$

In particular, $\nu_{\lambda,2\lambda+1}(0) = \nu'_{\lambda,2\lambda+1}(0) = \nu''_{\lambda,2\lambda+1}(0) = 0$. Moreover, $\nu'''_{\lambda,2\lambda+1}(0) = \lambda - 2$ and thus $x = 0$ is a saddle point whenever $\lambda > 2$. This implies that there exist some $\delta_\lambda > 0$ such that $\nu_{\lambda,2\lambda+1}(r) < 0$ for any $r \in (-\delta_\lambda, 0)$.

On the other hand, we have $\nu_{2,5}'''(r) = 0$ if and only if $e^{2r}(2r+1) = 1$ and a straightforward calculation shows that this only holds for $r = 0$, i.e. $\nu_{2,5}'''(0) = 0$ and 0 is the only critical point for $\nu_{2,5}''$. As $\nu_{2,5}''''(0) = 4$, $\nu_{2,5}$ has a local minimum at 0 and is convex on $\mathbb{R}$, which yields the claim. □

Lemma A.4. *The following assertions hold true:*

(i) *Let $\alpha > 1$ and $\beta \in \mathbb{R}$. Then there exists some $r_\alpha < 0$ and $\overline{\lambda} > 0$ such that*

$$\nu_{\lambda,\alpha\lambda+\beta}(r_\alpha) < 0$$

for any $\lambda \geqslant \overline{\lambda}$.

(ii) *We have for any $r \in \mathbb{R}$ and $\lambda \geqslant 2$*

$$\nu_{\lambda,\lambda+\tau(\lambda)}(r) \geqslant 0,$$

with $\tau(\lambda) = 2^{\frac{3}{2}}\sqrt{\lambda} - 1$.
Moreover, for $\lambda \in [1,2)$, we have $\nu_{\lambda,3\lambda-1}(r) \geqslant 0$ for any $r \in \mathbb{R}$.

Proof. (i): We have

$$\nu_{\lambda,\alpha\lambda+\beta}(r) = \lambda\big(\Upsilon'(r)r - \alpha\Upsilon(r)\big) + \Upsilon(-r) - \beta\Upsilon(r). \tag{114}$$

Since the mappings $r \mapsto \Upsilon'(r)r$ and $r \mapsto \Upsilon(r)$ have the same asymptotic behaviour as $r \to -\infty$, we find some $r_\alpha < 0$, such that $\Upsilon'(r_\alpha)r_\alpha - \alpha\Upsilon(r_\alpha) < 0$. Hence, the claim follows from (114).

(ii): We have $\nu_{\lambda,\lambda+\tau(\lambda)}(r) = \lambda\big(\Upsilon'(r)r - \Upsilon(r)\big) + \Upsilon(-r) - \tau(\lambda)\Upsilon(r)$ and hence

$$\nu'_{\lambda,\lambda+\tau(\lambda)}(r) = \lambda\Upsilon''(r)r - \Upsilon'(-r) - \tau(\lambda)\Upsilon'(r). \tag{115}$$

We aim to show that $\nu'_{\lambda,\lambda+\tau(\lambda)}(r) \leqslant 0$ for any $r \leqslant 0$ and any $\lambda \geqslant 1$. This suffices to deduce the claim, since Lemma A.3 reduces the claim in the case of $\lambda \geqslant 2$ to negative arguments because of Lemma A.1(iii) and the relation

$$\lambda + \tau(\lambda) \leqslant 2\lambda + 1. \tag{116}$$

To see the latter, we define $p(\lambda) = \lambda - 2^{\frac{3}{2}}\sqrt{\lambda}$, $\lambda \in (0,\infty)$. We have $p'(\lambda) = 1 - \frac{\sqrt{2}}{\sqrt{\lambda}}$ and we obtain, that p attains the global minimum, when $\sqrt{\lambda} = \sqrt{2}$. Further, $p(2) = -2$ and hence

$$\lambda - \tau(\lambda) = p(\lambda) + 1 \geqslant -1,$$

which implies (116). Further, one checks that $\lambda + \tau(\lambda) \geqslant 3\lambda - 1$ holds true if $1 \leqslant \lambda < 2$. Consequently, if $\nu'_{\lambda,\lambda+\tau(\lambda)}(r) \leqslant 0$ for any $r \leqslant 0$ is valid, one deduces the claim from Lemma A.2 combined with Lemma A.1(iii).

Thus, replacing r by $-y$ in (115), it remains to show

$$\varphi(y) := \lambda e^{-y}y + e^y - 1 + \tau(\lambda)(e^{-y} - 1) \geqslant 0 \tag{117}$$

for any $y \geqslant 0$. Due to the asymptotic behaviour of φ and the fact that $\varphi(0) = 0$, it suffices to investigate the critical points of φ in $(0,\infty)$. We have

$$\varphi'(y) = e^{-y}\big(\lambda - \lambda y - \tau(\lambda)\big) + e^y$$

and hence, as a necessary condition, $e^{2y} = \lambda y - \lambda + \tau(\lambda)$, i.e.

$$y_* = \frac{e^{2y_*} + \lambda - \tau(\lambda)}{\lambda} \tag{118}$$

must hold at a critical point of φ. Plugging (118) into (117) yields

$$0 \leqslant e^{-y_*}\big(e^{2y_*} + \lambda - \tau(\lambda)\big) + e^{y_*} - 1 + \tau(\lambda)(e^{-y_*} - 1) = 2e^{y_*} + \lambda e^{-y_*} - (\tau(\lambda) + 1).$$

We define $g(y) = 2e^y + \lambda e^{-y}$. If $\lambda \in [1,2)$ we have $g'(y) > 0$ for any $y \geqslant 0$ and the relation $g(0) \geqslant \tau(\lambda)+1$ can be readily checked. In case of $\lambda \geqslant 2$, $g'(y) = 0$ is equivalent to $y = \log\big(\sqrt{\frac{\lambda}{2}}\big)$. Clearly, this is the global minimum point of g in $\mathbb{R}$. We have

$$g\big(\log\big(\sqrt{\frac{\lambda}{2}}\big)\big) = 2\sqrt{\frac{\lambda}{2}} + \lambda\sqrt{\frac{2}{\lambda}} = 2^{\frac{3}{2}}\sqrt{\lambda}.$$

Hence, by definition of $\tau(\lambda)$, we deduce

$$2e^{y*} + \lambda e^{-y*} \geqslant \tau(\lambda) + 1,$$

which establishes the claim. □

References

[1] Ambrosio, L.; Gigli, N.; Savaré, G.: Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. Invent. Math. **195** (2014), 289–391.

[2] Ambrosio, L.; Gigli, N.; Savaré, G.: Metric measure spaces with Riemannian Ricci curvature bounded from below. Duke Math. J. **163** (2014), 1405–1490.

[3] Ambrosio, L.; Gigli, N.; Savaré, G.: Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds. Ann. Probab. **43** (2015), 339–404.

[4] Anderson, W. J.: *Continuous-Time Markov Chains*, Springer Series in Statistics, 1991.

[5] Arnold, A.; Markowich, P.; Toscani, G.; Unterreiter, A.: On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. Comm. Partial Differential Equations **26** (2001), 43–100.

[6] Bakry, D.; Émery, M.: *Diffusions hypercontractives*. Séminaire de probabilités, XIX, 1983/84, 177–206, Lecture Notes in Math., Springer, Berlin, 1985.

[7] Bakry, D.; Gentil, I.; Ledoux, M.: *Analysis and geometry of Markov diffusion operators*. Springer, 2014.

[8] Baudoin, F.: Bakry-Émery meets Villani. J. Funct. Anal. **273** (2017), 2275–2291.

[9] Bauer, F.; Horn, P.; Lin, Y.; Lippner, G.; Mangoubi, D.; Yau, S.-T.: Li-Yau inequality on graphs. J. Differential Geom. **99** (2015), 359–405.

[10] Beckner, W.: A generalized Poincaré inequality for Gaussian measures. Proc. Amer. Math. Soc. **105** (1989), 397–400.

[11] Bobkov, S. G.; Tetali, P.: Modified logarithmic Sobolev inequalities in discrete settings. J. Theoret. Probab. **19** (2006), 289–336.

[12] Bolley, F.; Gentil, I.: Phi-entropy inequalities for diffusion semigroups. J. Math. Pures Appl. (9) **93** (2010), 449–473.

[13] Bonciocat, A.-I.; Sturm, K.-T.: Mass transportation and rough curvature bounds for discrete spaces. J. Funct. Anal. **256** (2009), 2944–2966.

[14] Caputo, P.; Dai Pra, P.; Posta, G.: Convex entropy decay via the Bochner-Bakry-Émery approach. Ann. Inst. Henri Poincaré Probab. Stat. **45** (2009), 734–753.

[15] Chung, F. R. K.; Yau, S.-T. Logarithmic Harnack inequalities. Math. Res. Lett. **3** (1996), 793–812.

[16] Cordero-Erausquin, D.; McCann, R. J.; Schmuckenschläger, M.: A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. Invent. Math. **146** (2001), 219–257.

[17] Cushing, D.; Kamtue, S.; Kangaslampi, R.; Liu, S.; Peyerimhoff, N.: Curvature, Graph Products and Ricci Flatness, Preprint 2019, available at arxiv.

[18] Diaconis, P.; Saloff-Coste, L.: Logarithmic Sobolev inequalities for finite Markov chains. Ann. Appl. Probab. **6** (1996), 695–750.

[19] Dier, D.; Kassmann, M.; Zacher, R.: Discrete versions of the Li-Yau gradient estimate. To appear in Ann. Sc. Norm. Super. Pisa Cl. Sci. (5), DOI: 10.2422/2036-2145.201807-005.

[20] Erbar, M.; Fathi, M.: Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature. J. Funct. Anal. **274** (2018), 3056–3089.

[21] Erbar, M.; Kuwada, K.; Sturm, K.-T.: On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. Invent. Math. **201** (2015), 993–1071.

[22] Erbar, M.; Maas, J.: Ricci curvature of finite Markov chains via convexity of the entropy. Arch. Ration. Mech. Anal. **206** (2012), 997–1038.

[23] Erbar, M.; Maas, J.; Tetali, P.: Discrete Ricci curvature bounds for Bernoulli-Laplace and random transposition models. Ann. Fac. Sci. Toulouse Math. (6) **24** (2015), 781–800.

[24] Fathi, M.; Maas, J.: Entropic Ricci curvature bounds for discrete interacting systems. Ann. Appl. Probab. **26** (2016), 1774–1806.
[25] Freidlin, M.: *Markov processes and differential equations: asymptotic problems.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1996.
[26] Horn, P.; Lin, Y.; Liu, S.; Yau, S.-T.: Volume doubling, Poincaré inequality and Gaussian heat kernel estimate for non-negatively curved graphs. J. Reine Angew. Math. **757** (2019), 89–130.
[27] Hua, B.; Jost, J.; Liu, S.: Geometric analysis aspects of infinite semiplanar graphs with nonnegative curvature. J. Reine Angew. Math. **700** (2015), 1–36.
[28] Hua, B.; Lin, Y.: Graphs with large girth and nonnegative curvature dimension conditions. Comm. Anal. Geom. **27** (2019), 619–638.
[29] Hua, B.; Lin, Y.: Stochastic completeness for graphs with curvature dimension conditions. Advances in Mathematics **306** (2017), 279–302.
[30] Jordan, R.; Kinderlehrer, D.; Otto, F.: The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. **29** (1998), 1–17.
[31] Jost, J.; Liu, S.: Ollivier's Ricci curvature, local clustering and curvature-dimension inequalities on graphs. Discrete Comput. Geom. **51** (2014), 300–322.
[32] Jüngel, A.: *Entropy methods for diffusive partial differential equations.* SpringerBriefs in Mathematics. Springer, 2016.
[33] Jüngel, A.; Yue, W.: Discrete Beckner inequalities via the Bochner-Bakry-Émery approach for Markov chains. Ann. Appl. Probab. **27** (2017), 2238–2269.
[34] Keller, M.; Lenz, D.: Dirichlet forms and stochastic completeness of graphs and subgraphs. J. Reine Angew. Math., **666** (2012), 189-223.
[35] Keller, M.; Münch, F.: Gradient estimates, Bakry-Émery Ricci curvature and ellipticity for unbounded graph Laplacians. arXiv e-prints, page arXiv:1807.10181, July 2018.
[36] Klartag, B.; Kozma, G.; Ralli, P.; Tetali, P.: Discrete curvature and abelian groups. Canad. J. Math. **68** (2016), 655–674.
[37] Li, P.: *Geometric analysis.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2012.
[38] Li, P.; Yau, S.-T.: On the parabolic kernel of the Schrödinger operator. Acta Math. **156** (1986), 153–201.
[39] Lin, Y.; Liu S.: Equivalent properties of CD inequality on graph. In: arXiv preprint arXiv:1512.02677 (2015).
[40] Lin, Y.; Yau, S.-T.: Ricci curvature and eigenvalue estimate on locally finite graphs. Math. Res. Lett. **17** (2010), 343–356.
[41] Liu, S.; Münch, F.; Peyerimhoff, N.: Bakry-Émery curvature and diameter bounds on graphs. Calc. Var. Partial Differential Equations, **57** (2018), no. 2, Art. 67, 9 pp.
[42] Liu, S.; Peyerimhoff, N.: Eigenvalue ratios of non-negatively curved graphs. Combin. Probab. Comput. **27** (2018), 829–850.
[43] Lott, J.; Villani, C.: Ricci curvature for metric-measure spaces via optimal transport. Ann. of Math. (2) **169** (2009), 903–991.
[44] Maas, J.: Entropic Ricci curvature for discrete spaces. Modern approaches to discrete curvature, 159–174, Lecture Notes in Math., 2184, Springer, Cham, 2017.
[45] Maas, J.: Gradient flows of the entropy for finite Markov chains. J. Funct. Anal. **261** (2011), 2250–2292.
[46] McCann, R. J.: A convexity principle for interacting gases. Adv. Math. **128** (1997), 153–179.
[47] Mielke, A.: Geodesic convexity of the relative entropy in reservible Markov chains. Calc. Var. Partial Diff. Equ. **48** (2013), 1–31.
[48] Monmarché, P.: Generalized Γ calculus and application to interacting particles on a graph. Potential Anal. **50** (2019), no. 3, 439–466.
[49] Monmarché, P.: Hypocoercivity in metastable settings and kinetic simulated annealing. Probab. Theory Related Fields **172** (2018), 1215–1248.
[50] Münch, F.: Li-Yau inequality on finite graphs via non-linear curvature dimension conditions. J. Math. Pures Appl. (9) **120** (2018), 130–164.
[51] Münch, F: Remarks on curvature dimension conditions on graphs. Calc. Var. Partial Differential Equations **56** (2017), Paper No. 11, 8 pp.
[52] Münch, F.; Wojciechowski, R. K.: Ollivier Ricci curvature for general graph Laplacians: heat equation, Laplacian comparison, non-explosion and diameter bounds. Adv. Math. **356** (2019), 106759, 45 pp.
[53] Najman, L., Romon, P. (eds.): *Modern approaches to discrete curvature.* Lecture Notes in Mathematics. Springer, Cham, 2017.
[54] Norris, J.: *Markov Chains.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

[55] Ollivier, Y.: Ricci curvature of Markov chains on metric spaces. J. Funct. Anal. **256** (2009), 810–864.
[56] Ollivier, Y.; Villani, C.: A curved Brunn-Minkowski inequality on the discrete hypercube, or: what is the Ricci curvature of the discrete hypercube? SIAM J. Discrete Math. **26** (2012), 983–996.
[57] Otto, F.; Villani, C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. **173** (2000), 361–400.
[58] Soner, H. M.: Stochastic representations for nonlinear parabolic PDEs. Handbook of differential equations: evolutionary equations. Vol. III, 477526, Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2007.
[59] Spener, A.; Weber, F.; Zacher, R.: Curvature-dimension inequalities for non-local operators in the discrete setting. Calc. Var. Partial Diff. Equ. **58** (2019), no. 5, Art. 171, 30 pp.
[60] Spener, A.; Weber, F.; Zacher, R.: The fractional Laplacian has infinite dimension. Comm. Partial Differential Equations **45** (2020), 57–75.
[61] Sturm, K.-T.: On the geometry of metric measure spaces. I. Acta Math. **196** (2006), 65–131.
[62] Sturm, K.-T.: On the geometry of metric measure spaces. II. Acta Math. **196** (2006), 133–177.
[63] Villani, C.: *Hypocoercivity.* Mem. Amer. Math. Soc. **202** (2009), no. 950, 141 pp.
[64] Villani, C.: *Optimal transport old and new.* Springer, 2009.
[65] von Renesse, M.-K.; Sturm, K.-T.: Transport inequalities, gradient estimates, entropy, and Ricci curvature. Comm. Pure Appl. Math. **58** (2005), 923–940.

3. Entropy-information inequalities under curvature-dimension conditions for continuous-time Markov chains

Author: Frederic Weber

Journal: accepted in Electronic Journal of Probability

arXiv: 2010.06309, arXiv preprint, 2020

Copyright: *This is an Accepted Manuscript published by the Institute of Mathematical Statistics and the Bernoulli Society in Electronic Journal of Probability on 23. April 2021, available at: https://doi.org/10.1214/21-EJP627. This article is published with a Creative Commons Attribution 4.0 International License.*

ENTROPY-INFORMATION INEQUALITIES UNDER CURVATURE-DIMENSION CONDITIONS FOR CONTINUOUS-TIME MARKOV CHAINS

FREDERIC WEBER

Abstract. In the setting of reversible continuous-time Markov chains, the CD_Υ condition has been shown recently to be a consistent analogue to the Bakry-Émery condition in the diffusive setting in terms of proving Li-Yau inequalities under a finite dimension term and proving the modified logarithmic Sobolev inequality under a positive curvature bound. In this article we examine the case where both is given, a finite dimension term and a positive curvature bound. For this purpose we introduce the $CD_\Upsilon(\kappa, F)$ condition, where the dimension term is expressed by a so called CD-function F. We derive functional inequalities relating the entropy to the Fisher information, which we will call entropy-information inequalities. Further, we deduce applications of entropy-information inequalities such as ultracontractivity bounds, exponential integrability of Lipschitz functions, finite diameter bounds and a modified version of the celebrated Nash inequality.

Keywords: Markov chain, curvature-dimension inequalities, entropy, Fisher information, ultracontractive bounds, exponential integrability of Lipschitz functions, diameter bounds, modified Nash inequality
MSC(2020): 60J27 (primary), 47D07, 39A12 (secondary).

1. Introduction

1.1. **The curvature-dimension condition of Bakry-Émery.** The origins of the Γ-calculus of Bakry and Émery date back to the seminal work [2]. Meanwhile, this theory, for which the monograph [3] is an excellent source, has been proven itself as a beautiful link between probability theory, geometry and analysis.

For motivational purposes we briefly survey the setting of the Bakry-Émery theory in the sequel. Denoting by L the infinitesimal generator of a Markov semigroup, the carré-du-champ operator Γ and the iterated carré-du-champ operator Γ_2 are defined as

$$\Gamma(f,g) = \frac{1}{2}\big(L(fg) - f\,Lg - g\,Lf\big),$$
$$\Gamma_2(f,g) = \frac{1}{2}\big(L\Gamma(f,g) - \Gamma(f,Lg) - \Gamma(g,Lf)\big)$$

for f and g lying in a suitable algebra of real-valued functions. One defines $\Gamma(f) := \Gamma(f,f)$ and $\Gamma_2(f) := \Gamma_2(f,f)$. Given a fixed invariant and reversible measure μ for the semigroup generated

(Frederic Weber) Institut für Angewandte Analysis, Universität Ulm, Helmholtzstrasse 18, 89081 Ulm, Germany.
E-mail address: frederic.weber@uni-ulm.de.
Date: March 31, 2021.
The author is supported by a PhD-scholarship of the "Studienstiftung des deutschen Volkes", Germany.

by L, the operator L is said to satisfy the curvature-dimension condition $CD(\kappa, n)$ for $\kappa \in \mathbb{R}$ and $n \in [1, \infty]$, if

$$\Gamma_2(f) \geqslant \kappa\Gamma(f) + \frac{1}{n}\big(Lf\big)^2, \ \mu\text{-a.e.}, \tag{1}$$

holds in a sufficiently rich class of functions.

The classical theory is based on the key assumption that the chain rule

$$LH(f) = H'(f)Lf + H''(f)\Gamma(f) \tag{2}$$

is satisfied for every $H \in C^2(\mathbb{R})$ and f lying in a suitable class of functions. Typical examples for operators fitting to the abstract framework of [3], in particular satisfying (2), are given by second-order differential operators with smooth coefficients and without zero-th order term. For instance, consider the Laplace-Beltrami operator $L = \Delta_g$ on a Riemannian manifold (M, g) with invariant and reversible measure given by the canonical Riemannian measure μ_g. In this case one can show by means of the Bochner-Lichnerowicz formula that $CD(\kappa, n)$ is equivalent to $\mathrm{Ric}_g(x) \geqslant \kappa g(x)$ for almost every $x \in M$ and $\dim M \leqslant n$, where Ric_g denotes the Ricci-curvature tensor and $\dim M$ the topological dimension of the manifold M. In this sense, calling (1) a curvature-dimension condition is well motivated.

The curvature-dimension inequality (1) serves as a powerful tool to establish various functional inequalities. In fact, if $CD(\kappa, \infty)$ holds with $\kappa > 0$, then the spectral gap inequality and the logarithmic Sobolev inequality are both satisfied with constant κ (see [3, Chapter 4 and 5]). In the case that $CD(0, n)$ holds for $n < \infty$, one can deduce the Li-Yau inequality, which in turn leads to the parabolic Harnack inequality, cf. [3, Chapter 6]. Our main interest lies in the strongest case of having both, a positive curvature and a finite dimension term. Assuming that $CD(\kappa, n)$ holds with $\kappa > 0$ and $n < \infty$, one obtains Sobolev inequalities or, equivalently, the logarithmic entropy-energy inequality, which reads as

$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{n}{2}\log\Big(1 + \frac{4}{\kappa n}\mathcal{E}(f)\Big), \tag{3}$$

for f being in a sufficiently large class of functions with $\int_X f^2 \mathrm{d}\mu = 1$ (where Ent_μ denotes the Boltzmann entropy and $\mathcal{E}$ the Dirichlet energy on $L^2(\mu)$ associated with L and μ). The functional inequality (3) is an important instance of what is called an entropy-energy inequality, that is

$$\mathrm{Ent}_\mu(f^2) \leqslant \Phi\big(\mathcal{E}(f)\big), \tag{4}$$

where $\Phi : (0, \infty) \to \mathbb{R}$ is a strictly increasing and concave C^1-function. We refer to [3, Chapter 7], where applications of entropy-energy inequalities, such as ultracontractivity or diameter bounds, have been discussed.

1.2. **Existing approaches for finding substitutes for curvature and dimension in the discrete setting.** The issue of finding suitable substitutes of Ricci curvature lower bounds in the discrete setting has been a very vibrant topic of research in the last decade and a half, see e.g. the recent book [24].

Based on the powerful approach of optimal transport, for which we refer to the seminal works [20, 29, 30, 32], Erbar and Maas successfully developed the theory of entropic Ricci curvature in the context of finite Markov chains in [13] and [21]. Another important notion of discrete curvature that relies on ideas from optimal transport is due to Ollivier (see [26]). The latter curvature notion has been studied in a variety of articles concerning the case that the underlying state space is given by a locally finite graph, see e.g. [15, 23].

With regard to the Bakry-Émery approach, it is apparently still possible to define the operators Γ and Γ_2 in the discrete setting, where L now denotes the generator of a Markov chain. However, even though some positive results such as eigenvalue estimates in [17] or diameter bounds in [18]

can be deduced, the Bakry-Émery condition is not as applicable as in the continuous setting, in particular with regard to Li-Yau inequalities and (modified) logarithmic Sobolev inequalities. This is caused especially by the major difficulty that the diffusion property (2) does not hold in the discrete setting. There are several modified versions of curvature-dimension inequalities in the discrete setting which are based on the approach of identifying certain discrete substitutes for the chain rule, e.g. in the context of Li-Yau inequalities we refer to [4, 11, 22]. In particular, in [11] the identity

$$L(\log f) = \frac{Lf}{f} - \Psi_\Upsilon(\log f) \tag{5}$$

has been used as the appropriate replacement for the case of $H = \log$ in (2). Here the operator Ψ_Υ is defined as in (11) below, with $H = \Upsilon$, where $\Upsilon : \mathbb{R} \to \mathbb{R}$ is given by $\Upsilon(r) = e^r - r - 1$, $r \in \mathbb{R}$. We will comment on regularity assumptions for f ensuring the validity of (5) in the next subsection. One of the key ideas of [11] is to make use of so called CD-functions in order to express the dimension term in their CD condition. This in fact leads to significantly improved estimates regarding the corresponding Li-Yau inequalities, which are even sharp in some instances. We will follow the approach of using CD-functions in this article as well.

Regarding positive curvature bounds, based on the identity (5), it has been shown very recently in [31] that the $CD_\Upsilon(\kappa, \infty)$ condition serves as a consistent analogue to the classical curvature-dimension condition with regard to the strategy of proving entropy decay of an exponential rate using the entropy method. The resulting functional inequality, the modified logarithmic Sobolev inequality, holds with constant $\kappa > 0$ provided that $CD_\Upsilon(\kappa, \infty)$ is satisfied (see [31, Corollary 3.5]). In the sense of the relation between curvature-dimension inequalities and related functional inequalities, the modified logarithmic Sobolev inequality with regard to the $CD_\Upsilon(\kappa, \infty)$ condition in the discrete setting serves as the appropriate counterpart to the logarithmic Sobolev inequality with regard to the $CD(\kappa, \infty)$ condition in the diffusive setting.

Moreover, we refer to the discussion in [31, Remark 2.9] that shows that the CD_Υ condition with finite dimension term is strongly related to the articles [22] and [11] (see also Remark 2.4 below) and in particular implies Li-Yau type inequalities. In this sense, the CD_Υ condition serves as a suitable analogue to the Bakry-Émery condition with regard to both, positive curvature in terms of the entropy method and finite dimension in terms of Li-Yau inequalities. The main motivation of this paper is to identify the appropriate discrete counterpart to the logarithmic entropy-energy inequality (3) with regard to the $CD(\kappa, n)$ condition in the diffusive setting, or in other words, to investigate the case where both is satisfied, a positive curvature bound and a finite dimension term.

1.3. **Setting and main results.** We consider a time-homogeneous, continuous-time Markov chain $\big(Z_t\big)_{t\geqslant 0}$ defined on a probability space $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ and with (finite or infinite) countable state space X. The generator L of the Markov chain is defined on a suitable class of functions $f : X \to \mathbb{R}$ by

$$Lf(x) = \sum_{y\in X} k(x,y) f(y) = \sum_{y\in X} k(x,y)\big(f(y) - f(x)\big). \tag{6}$$

Here, we assume $\sum_{y\in X} k(x,y) = 0$, where $k(x,y) \geqslant 0$ denotes the transition rate for jumping from x to y if $x \neq y$. We remark that L determines naturally a graph structure with vertex set X and edge weights given by $k(x,y)$ for $x, y \in X$, $x \neq y$, to which we will refer as the underlying graph to L. If $k(x,y) \in \{0,1\}$ for any $x, y \in X$ with $x \neq y$, then the underlying graph to L is an unweighted graph.

We denote the associated (sub-)Markov semigroup on the space of bounded functions by $\big(P_t\big)_{t\geqslant 0}$, which is given by

$$P_t f(x) = \mathbb{E}\big(f(Z_t)|Z_0 = x\big). \tag{7}$$

Further, we suppose that the Markov chain is irreducible and that a unique (up to positive multiples) invariant measure μ exists such that the detailed balance condition

$$\mu(\{x\})k(x,y) = \mu(\{y\})k(y,x) \tag{8}$$

is valid for any $x, y \in X$. Let $\pi : X \to (0,\infty)$ denote the density for μ with respect to the counting measure on X, i.e. $\mathrm{d}\mu = \pi \mathrm{d}\#$. It is a basic consequence of irreducibility and reversibility that $\pi(x) > 0$ for any $x \in X$.

It is well known that the Markov chain is positive recurrent if and only if μ is finite (and hence can be assumed to be a probability measure) and the Markov chain is non-explosive (see e.g. [25]). In particular, provided that the Markov chain is positive recurrent, stochastic completeness (that is $P_t\mathbb{1} = \mathbb{1}$) and ergodicity (by which we mean what is sometimes called ordinary ergodicity, see e.g. [1]) hold true. In the recent work [31], which is strongly related to this article, positive recurrence has been an important assumption. If one allows for the Markov chain being explosive, then the semigroup given by (7) is only submarkovian, which ensures still that $P_t f$ is bounded provided that f is bounded. For more details on the general theory of continuous-time Markov chains we refer the reader to [1] and [25].

We denote by $\mathbb{R}^X$ the space of real-valued functions on X and by $\ell^p(\mu)$, $1 \leqslant p < \infty$, the elements of $\mathbb{R}^X$ that are p-summable with respect to μ. Further, $\ell^\infty(X)$ denotes the space of bounded real-valued functions on X. Throughout this article the mapping $\|\cdot\|_p : \ell^p(\mu) \to [0,\infty)$, $1 \leqslant p < \infty$, denotes the $\ell^p(\mu)$-norm.

Moreover, we denote by $\mathcal{P}(X)$ the set of probability densities with respect to μ, by $\mathcal{P}_*(X)$ the set of elements in $\mathcal{P}(X)$ that are strictly positive at any $x \in X$ and $P_*^+(X) := \mathcal{P}_*(X) \cap \ell^{\infty,+}(X)$, where

$$\ell^{\infty,+}(X) = \{f \in \ell^\infty(X) : \exists c > 0 \text{ such that } f(x) \geqslant c, \forall x \in X\}.$$

We assume throughout this paper that at any $x \in X$

$$M_1(x) := \sum_{y\in X\setminus\{x\}} k(x,y) < \infty \tag{9}$$

and

$$M_2(x) := \sum_{y\in X\setminus\{x\}} k(x,y) \sum_{z\in X\setminus\{y\}} k(y,z) < \infty. \tag{10}$$

Further, we define $M_{1,\inf} := \inf_{x\in X} M_1(x) \in [0,\infty)$ and $M_{1,\sup} := \sup_{x\in X} M_1(x) \in (0,\infty]$ and introduce

$$N(x) := \sum_{y\in X\setminus\{x\}} k(x,y)k(y,x) \leqslant M_2(x) < \infty.$$

We recall from [31] the definition of the operators Ψ_H and $\Psi_{2,H}$, where $H : \mathbb{R} \to \mathbb{R}$ is a continuous mapping. We define

$$\Psi_H(f)(x) = \sum_{y\in X} k(x,y)H(f(y) - f(x)),\ x \in X, \tag{11}$$

for any $f \in \ell^\infty(X)$ and

$$B_H(f,g)(x) = \sum_{y\in X} k(x,y)H(f(y) - f(x))(g(y) - g(x)),\ x \in X, \tag{12}$$

for suitable functions f and g. In particular, the conditions (9) and (10) ensure that we can choose in (12) $g = Lf$ for $f \in \ell^\infty(X)$. This guarantees that for $f \in \ell^\infty(X)$ and $x \in X$ the operator

$$\Psi_{2,H}(f)(x) = \frac{1}{2}\big(L\Psi_H(f)(x) - B_{H'}(f, Lf)(x)\big)$$

is well defined. In the case of $H(r) = \frac{1}{2}r^2$, $\Psi_H(f)$ coincides with $\Gamma(f)$ and $\Psi_{2,H}(f)$ with $\Gamma_2(f)$. For our purposes, the mapping $\Upsilon(r) = e^r - r - 1$, $r \in \mathbb{R}$, will play a key role. Indeed, the choice of $H(r) = \Upsilon(r)$, which is motivated by the identity (5), leads to the operators $\Psi_\Upsilon(f)$ and $\Psi_{2,\Upsilon}(f)$, which are the central objects of investigation in the recent article [31]. Let us remark that in our setting the identity (5) holds true for any $f \in \mathbb{R}^X$ such that $f, \log f \in \ell^1(k(x,\cdot))$ for any $x \in X$ (cf. [31, Lemma 2.2]), which is for instance the case when $f \in \ell^{\infty,+}(X)$.

We recall a representation formula for the operator $\Psi_{2,\Upsilon}$, which has been used frequently in order to study a large class of examples in [31], and reads as

(13)
$$\begin{aligned} 2\Psi_{2,\Upsilon}(f)(x) = &\sum_{y\in X\setminus\{x\}} k(x,y) \sum_{z\in X} k(y,z)\Big(\Upsilon\big(f(z)-f(y)\big) - \Upsilon'\big(f(y)-f(x)\big)\big(f(z)-f(y)\big)\Big) \\ &+ \sum_{y\in X\setminus\{x\}} k(x,y)\Upsilon'\big(f(y)-f(x)\big) \sum_{z\in X} k(x,z)\big(f(z)-f(x)\big) \\ &- \sum_{y\in X\setminus\{x\}} k(x,y) \sum_{z\in X} k(x,z)\Upsilon\big(f(z)-f(x)\big). \end{aligned}$$

The detailed balance condition (8) ensures that the generator of the Dirichlet form given by

$$\mathcal{E}(f,g) = \frac{1}{2}\sum_{x\in X}\sum_{y\in X} k(x,y)\big(f(y)-f(x)\big)\big(g(y)-g(x)\big)\pi(x) \tag{14}$$

for f, g being suitable functions, coincides with L given by (6) on bounded functions which are contained in the domain of the form generator, see [16]. We will also denote the $\ell^2(\mu)$ operator by L in the sequel. Further, as the corresponding $\ell^2(\mu)$-semigroup generated by L is an extension of the semigroup given by (7) restricted to $\ell^\infty(X) \cap \ell^2(\mu)$, we will also use the notation $(P_t)_{t\geqslant 0}$ for the corresponding $\ell^2(\mu)$-semigroup.

An eminent role will be played by the entropy, being given as

$$\mathrm{Ent}_\mu(f) = \int_X f\log f \mathrm{d}\mu - \int_X f\mathrm{d}\mu \log \int_X f\mathrm{d}\mu,$$

and the Fisher information $\mathcal{I}(f) = \mathcal{E}(f, \log f)$. We refer to the beginning of Section 3 for more details on some elementary properties of the entropy resp. the Fisher information and on corresponding admissible functions.

We say that L satisfies $CD_\Upsilon(\kappa, F)$ if

$$\Psi_{2,\Upsilon}(f) \geqslant \kappa\Psi_\Upsilon(f) + F_0\big(-Lf\big)$$

holds on X for any $f \in \ell^\infty(X)$, where $\kappa \in \mathbb{R}$ and $F_0 : \mathbb{R} \to [0,\infty)$ is the trivial extension of a CD-function $F : [0,\infty) \to [0,\infty)$ (see Definition 2.1 below), i.e. $F_0(r) = 0$ if $r < 0$. Note that the notation of the condition $CD_\Upsilon(\kappa,\infty)$, which states that $\Psi_{2,\Upsilon}(f) \geqslant \kappa\Psi_\Upsilon(f)$ holds on X for any $f \in \ell^\infty(X)$, is a bit missleading since it really means that the dimension term vanishes. This terminology is clearly motivated from the case of the quadratic CD-functions $F(r) = \frac{r^2}{n}$, $n \in [1,\infty)$, to which we will refer as the $CD_\Upsilon(\kappa, n)$ condition (motivated by the classical Bakry-Émery notation). Note that $CD_\Upsilon(\kappa, n)$ has already been mentioned in slightly different form in [31, Remark 2.9] and also in [22] in a rather implicit form (cf. [31, Section 9]).

We now describe our main results. Assuming that the CD-function is convex, continuously differentiable and such that the mapping $r \mapsto \frac{F(r)}{r^{1+\delta}}$ is increasing on $(0,\infty)$ for some $\delta > 0$, we will be able to show in Theorem 3.3 that $CD_\Upsilon(\kappa, F)$ (with $\kappa > 0$) implies the bound

$$\mathrm{Ent}_\mu(f) \leqslant \int_0^\infty G\Big(\frac{\kappa}{e^{2\delta\kappa t}\big(1 + \frac{\kappa\mathcal{I}(f)}{F(\mathcal{I}(f))}\big) - 1}\Big)\mathrm{d}t \tag{15}$$

for any $f \in P_*(X)$ with $\mathrm{Ent}_\mu(f) < \infty$ and $\mathcal{I}(f) \in (0,\infty)$, where $G : (0,\infty) \to (0,\infty)$ denotes the inverse function of the mapping $r \mapsto \frac{F(r)}{r}$, $r > 0$. In particular, in the case of $CD_\Upsilon(\kappa, n)$ with $\kappa > 0$ and $n < \infty$, (15) reads as

$$\mathrm{Ent}_\mu(f) \leqslant \frac{n}{2}\log\Big(1 + \frac{\mathcal{I}(f)}{\kappa n}\Big), \tag{16}$$

see Corollary 3.7. Consequently, (16) with regard to $CD_\Upsilon(\kappa, n)$ is the natural discrete analogue to (3) with regard to $CD(\kappa, n)$ in the diffusive setting. Note that in the diffusive setting the chain rule implies $\mathcal{I}(f^2) = 4\mathcal{E}(f)$ for suitable functions, which yields that in the classical situation the inequalities (3) and (16) coincide. In particular, (16) is an important example of what will be called an entropy-information inequality, i.e. a functional inequality of the form $\mathrm{Ent}_\mu(f) \leqslant \Phi(\mathcal{I}(f))$ for a strictly increasing and concave $C^1\big((0,\infty)\big)$-function Φ, to which we will refer as the growth function.

As the modified logarithmic Sobolev inequality differs from the logarithmic Sobolev inequality in the discrete setting, hypercontractivity of the semigroup, which is equivalent to the latter also in the discrete setting (see [9]), does not characterize the modified logarithmic Sobolev inequality. In [5], Bobkov and Tetali established a hypercontractivity formulation for $e^{P_t f}$ being suitable for the modified logarithmic Sobolev inequality. In this spirit, we show in Theorem 4.2 that certain ultracontractivity bounds for $e^{P_t f}$ hold under corresponding entropy-information inequalities. In particular, in case of the $CD_\Upsilon(\kappa, n)$ condition with $\kappa > 0$ and $n < \infty$, we will be able to derive that

$$\|e^{P_t f}\|_\infty \leqslant \Big(1 + \frac{1}{2\kappa t}\Big)^{\frac{n}{2}}\|e^f\|_1$$

holds for any $f \in \ell^\infty(X)$ and any $t > 0$.

In Theorem 5.4 we prove that μ-integrable 1-Lipschitz functions, by which we mean that $\|\Gamma(f)\|_\infty \leqslant 1$, are exponentially integrable. Moreover, provided that the growth function satisfies $\int_0^\infty \frac{\Phi(s^2)}{s^2}\mathrm{d}s < \infty$, we show that

$$\|f - \int_X f\mathrm{d}\mu\|_\infty \leqslant \int_0^\infty \frac{\Phi(s^2)}{s^2}\mathrm{d}s$$

holds for any 1-Lipschitz function f. This in turn implies a finite diameter bound, which in the special case of $CD_\Upsilon(\kappa, n)$ with $\kappa > 0$ and $n < \infty$ reads as

$$\mathrm{diam}_\varrho \leqslant \pi\sqrt{\frac{n}{\kappa}}, \tag{17}$$

see Corollary 5.5. Interestingly, (17) coincides with the diameter bound that has been obtained by Liu, Münch and Peyerimhoff in [18] for the weaker $CD(\kappa, n)$ condition but under assumptions on the underlying graph to L which can be expected to be more restrictive, see Remark 5.6.

Finally, we also show that a new modified version of the celebrated Nash inequality holds under an entropy-information inequality with logarithmic growth function. In the particular case of the $CD_\Upsilon(\kappa, n)$ condition with $\kappa > 0$ and $n < \infty$, this modified Nash inequality says that

$$\|f\|_2^{n+2} \leqslant \Big(\|f\|_2^2 + \frac{\mathcal{I}(f^2)}{\kappa n}\Big)^{\frac{n}{2}}\|f\|_1^2$$

holds for any non-vanishing $f \in \ell^2(\mu)$, see Theorem 6.1 and Corollary 6.2.

The article is organized as follows. In Section 2 we introduce the curvature-dimension condition $CD_\Upsilon(\kappa, F)$ and discuss several examples. Next, we define the notion of entropy-information inequalities and investigate their relation to the $CD_\Upsilon(\kappa, F)$ condition in the case of power-type CD-functions. In the remaining part of the paper, we discuss applications of entropy-information inequalities and hence also of the $CD_\Upsilon(\kappa, F)$ condition. We derive ultracontractive bounds for $e^{P_t f}$ in Section 4, exponential integrability of Lipschitz functions and diameter bounds in Section 5 and, finally, a modified version of the Nash inequality in Section 6.

2. The CD_Υ condition with finite dimension and some examples

In this section we generalize the $CD_\Upsilon(\kappa, \infty)$ condition from the very recent work [31] by adding a general dimension term involving a CD-function. We first recall the notion of a CD-function that originates from the work of [11] and also has been mentioned in [31, Remark 2.9].

Definition 2.1. *A continuous function $F : [0, \infty) \to [0, \infty)$ is called CD-function, if $F(0) = 0$, $r \mapsto \frac{F(r)}{r}$ is strictly increasing on $(0, \infty)$ and $\frac{1}{F}$ is integrable at ∞.*

For a given CD-function F, we will call the function $F_0 : \mathbb{R} \to [0, \infty)$ given by $F_0(r) = F(r)$ if $r \geqslant 0$ and $F_0(r) = 0$ otherwise the trivial extension of F.

Remark 2.2. If a function $F : [0, \infty) \to [0, \infty)$ with $F(0) = 0$ is strictly convex on $(0, \infty)$, the mapping $r \mapsto \frac{F(r)}{r}$ is strictly increasing on $(0, \infty)$, cf. [11, Remark 3.3]. However, it can not be deduced in general that F is a CD-function as, for instance, the function $r \mapsto \Upsilon(-r)$, $r \in [0, \infty)$, serves as a counterexample.

If we impose instead that $F(0) = 0$ and $r \mapsto \frac{F(r)}{r^{1+\delta}}$ is increasing on $(0, \infty)$ for some $\delta > 0$, as it will be done in Theorem 3.3 (cf. Remark 3.4), then it follows that F is a CD-function. Indeed, it is obvious that $r \mapsto \frac{F(r)}{r}$ is strictly increasing on $(0, \infty)$ and further we have

$$\int_c^\infty \frac{1}{F(r)} \mathrm{d}r \leqslant \frac{c^{1+\delta}}{F(c)} \int_c^\infty r^{-(1+\delta)} \mathrm{d}r < \infty,$$

where $c > 0$.

Definition 2.3. *We say that the Markov generator L satisfies $CD_\Upsilon(\kappa, F)$ at $x \in X$ for $\kappa \in \mathbb{R}$ and a CD-function $F : [0, \infty) \to [0, \infty)$ with trivial extension F_0 if*

$$\Psi_{2,\Upsilon}(f)(x) \geqslant \kappa \Psi_\Upsilon(f)(x) + F_0\big(-Lf(x)\big) \tag{18}$$

holds for all $f \in \ell^\infty(X)$. If L satisfies $CD_\Upsilon(\kappa, F)$ at any $x \in X$, then we say that L satisfies $CD_\Upsilon(\kappa, F)$.
In the special case of $F(r) = \frac{1}{n} r^2$, $r \geqslant 0$, for some $n \in [1, \infty)$, we say that L satisfies $CD_\Upsilon(\kappa, n)$.

According to the ambiguity in the notation, we emphasize that throughout this article a capital letter F in the condition $CD_\Upsilon(\kappa, F)$ always refers to a CD-function, while a small letter n in the condition $CD_\Upsilon(\kappa, n)$ refers to the constant of a quadratic CD-function.

Remark 2.4. (i) It must be pointed out that the condition $CD_\Upsilon(\kappa, n)$ has been introduced in [31, Remark 2.9] in seemingly stricter form, as we only require the dimension term in Definition 2.3 for functions $f \in \ell^\infty(X)$ with $-Lf(x) > 0$. Further, the latter condition is the only difference of $CD_\Upsilon(\kappa, n)$ and the condition $CD\psi(n, \kappa)$ for the specific choice of $\psi = \log$, which has been introduced in the case of finite and unweighted graphs in [22]. We refer to [31, Section 9] for a detailed account on the relation of the operators appearing in (18) and those in [22]. The condition $-Lf(x) > 0$ also appears in the formulation of other curvature-dimension conditions,

such as for instance in the case of the exponential curvature-dimension condition of [4] and the $CD(F;0)$ condition in [11], where F denotes a CD-function. We refer to [31, Remark 2.9(iii)], which shows that the condition $CD_\Upsilon(0,F)$ suffices to deduce Li-Yau inequalities using the results of [11].

(ii) Importantly, the $CD_\Upsilon(\kappa,n)$ condition (or more generally $CD_\Upsilon(\kappa,F)$ with $F(r)\sim\frac{1}{n}r^2$ as $r\to 0^+$) implies the Bakry-Émery condition $CD(\kappa,n)$. This fact relies on the identities

$$\lim_{\lambda\to 0}\frac{\Psi_{2,\Upsilon}(\lambda f)}{\lambda^2}=\Gamma_2(f),\ \lim_{\lambda\to 0}\frac{\Psi_\Upsilon(\lambda f)}{\lambda^2}=\Gamma(f), \tag{19}$$

holding true for any $f\in\ell^\infty(X)$, which have been shown in the proof of [31, Proposition 2.11]. As it has been pointed out in [31, Remark 2.13], the procedure extends easily to a quadratic dimension term. More accurately, in the formulation of Definition 2.3, it first implies $CD(\kappa,n)$ only for $f\in\ell^\infty(X)$ with $-Lf(x)>0$, but then extends to any $f\in\ell^\infty(X)$ by linearity of L, bilinearity of Γ and the definition of Γ_2. The fact that the Bakry-Émery condition is necessary for CD_Υ with a quadratic CD-function also motivates to study the former condition. In particular, we refer to [28], where $CD(0,n)$ has been studied for a large class of operators with long range jumps and state space $\mathbb{Z}$.

(iii) The property (19) has an important consequence for CD-functions that behave like a power-type function near the origin. Indeed, if $CD_\Upsilon(\kappa,F)$ holds with $\kappa\in\mathbb{R}$ and $F(r)\sim r^{1+\delta}$ as $r\to 0^+$ for some $\delta>0$, then we infer from (19) that for any $f\in\ell^\infty(X)$ with $-Lf(x)>0$, $x\in X$, we have at x

$$\begin{aligned}\Gamma_2(f)&=\lim_{\lambda\to 0^+}\frac{\Psi_{2,\Upsilon}(\lambda f)}{\lambda^2}\\&\geqslant\lim_{\lambda\to 0^+}\frac{\kappa\Psi_\Upsilon(\lambda f)+F(-\lambda Lf)}{\lambda^2}\\&=\kappa\Gamma(f)+(-Lf)^{1+\delta}\lim_{\lambda\to 0^+}\frac{F(-\lambda Lf)}{(-\lambda Lf)^{1+\delta}}\lambda^{1+\delta-2}.\end{aligned}$$

Consequently, $\delta\geqslant 1$ must hold, or in other words the best behavior of a CD-function near the origin one can hope for is quadratic.

Several concrete examples have been considered in [31, Section 5] to study the $CD_\Upsilon(\kappa,\infty)$ condition. It has turned out that the functions

$$\nu_{c,d}(r)=c\Upsilon'(r)r+\Upsilon(-r)-d\Upsilon(r),\ r\in\mathbb{R}, \tag{20}$$

with constants $c,d\in\mathbb{R}$, are of eminent importance. We refer the reader to the Appendix of [31], where basic properties of these functions have been collected.

As a warm-up, we begin with the basic example of the two-point space, for which we provide another property of the functions given by (20) in the Appendix below.

Example 2.5. We consider the two-point space $X=\{0,1\}$ with $k(0,1)=a$ and $k(1,0)=b$, where $a,b>0$. Here, the invariant and reversible probability measure μ is given by $\mathrm{d}\mu=\pi\mathrm{d}\#$ with $\pi(0)=\frac{b}{a+b}$ and $\pi(1)=\frac{a}{a+b}$. We write $\tilde{x}=1-x$ for $x\in X$ and $t=f(\tilde{x})-f(x)$. From [31] we know that $CD_\Upsilon(\kappa,\infty)$ holds true for some $\kappa>0$. In order to show that the CD_Υ condition is fulfilled with a positive curvature constant and a finite dimension term it hence suffices to show $CD_\Upsilon(0,F)$, where F is a CD-function. Thus it remains to compare $\Psi_{2,\Upsilon}(f)(x)$ with $F(-Lf(x))=F(-k(x,\tilde{x})t)$, for F being specified below. We have

$$\begin{aligned}2\Psi_{2,\Upsilon}(f)(x)&=L\Psi_\Upsilon(f)(x)-B_{\Upsilon'}(f,Lf)(x)\\&=k(x,\tilde{x})\Big(\big(\Psi_\Upsilon(f)(\tilde{x})-\Psi_\Upsilon(f)(x)\big)-\Upsilon'\big(f(\tilde{x})-f(x)\big)\big(Lf(\tilde{x})-Lf(x)\big)\Big)\end{aligned}$$

$$= k(x,\tilde{x})\,k(\tilde{x},x)\big(\Upsilon(-t)+\Upsilon'(t)t\big) + k(x,\tilde{x})^2\big(\Upsilon'(t)t-\Upsilon(t)\big)$$
$$= k(x,\tilde{x})k(\tilde{x},x)\nu_{1+\frac{k(x,\tilde{x})}{k(\tilde{x},x)},\frac{k(x,\tilde{x})}{k(\tilde{x},x)}}(t).$$

Note that for $c,d\in\mathbb{R}$ and $\eta>0$ we have that $\nu_{c+\eta,d+\eta}(r)\geqslant\nu_{c,d}(r)$ at any $r\in\mathbb{R}$, since $\Upsilon'(r)r\geqslant\Upsilon(r)$, $r\in\mathbb{R}$, holds by convexity. Hence, we can estimate

$$\Psi_{2,\Upsilon}(f)(x)\geqslant\frac{ab}{2}\nu_{1+\lambda,\lambda}(t),$$

where $\lambda:=\min\{\frac{a}{b},\frac{b}{a}\}$. Note that $\nu_{1+\lambda,\lambda}$ is strictly convex by Lemma A.1. Then, we infer from Remark 2.2 and the asymptotic behavior of $\nu_{1+\lambda,\lambda}$ that the mapping $F:[0,\infty)\to[0,\infty)$ defined as

$$F(r)=\frac{ab}{2}\,\nu_{1+\lambda,\lambda}\big(-\frac{r}{\max\{a,b\}}\big),\ r\geqslant 0,$$

is a CD-function and we deduce that $CD_\Upsilon(0,F)$ holds true.

We continue with the following basic observation.

Proposition 2.6. *Let $x\in X$, $\kappa\in\mathbb{R}$, $\gamma:\mathbb{R}\to[0,\infty)$ be convex on $\mathbb{R}$ with $\gamma|_{[0,\infty)}$ being a CD-function and $\alpha:X\to(0,\infty)$ such that*

$$\Psi_{2,\Upsilon}(f)(x)\geqslant\kappa\Psi_\Upsilon(f)(x)+\alpha(x)\sum_{y\in X}k(x,y)\gamma(f(x)-f(y)) \tag{21}$$

holds for any $f\in\ell^\infty(X)$. Then L satisfies $CD_\Upsilon(\kappa,F)$ at x with $F(r)=\alpha(x)M_1(x)\gamma\big(\frac{r}{M_1(x)}\big)$, $r\geqslant 0$. If $0<M_{1,\inf}\leqslant M_{1,\sup}<\infty$, $\alpha_*:=\inf_{x\in X}\alpha(x)>0$ and (21) holds for any $f\in\ell^\infty(X)$ and all $x\in X$, then L satisfies $CD_\Upsilon(\kappa,F)$ with $F(r)=\alpha_*M_{1,\inf}\gamma\big(\frac{r}{M_{1,\sup}}\big)$, $r\geqslant 0$.

Proof. For fixed $x\in X$, we observe by Jensen's inequality

$$\sum_{y\in X\setminus\{x\}}k(x,y)\gamma(f(x)-f(y))\geqslant M_1(x)\gamma\Big(-\frac{Lf(x)}{M_1(x)}\Big),$$

from which the first claim follows by (21). Clearly, γ is increasing on $[0,\infty)$, which implies the second claim. □

Note that the tempting naive approach to deduce a CD_Υ condition with non-negative curvature bound and finite dimension term from $CD_\Upsilon(\kappa,\infty)$ (with $\kappa>0$) alone by using Proposition 2.6 does not work. Indeed, the mapping $r\mapsto\Upsilon(-r)$ can not play the role of the function γ from Proposition 2.6 since it only grows linearly at ∞ and is hence not a CD-function. This is a difference to the Bakry-Émery condition in our setting, where the analogous result to Proposition 2.6 yields at least that $CD(\kappa,\infty)$(with $\kappa>0$) at $x\in X$ implies $CD(\lambda\kappa,(1-\lambda)n)$ at $x\in X$, where $\lambda\in[0,1]$ and $n<\infty$ depends on κ and x. For our purposes however, we need a refined analysis.

In light of the subsequent sections, power-type CD-functions are of particular importance. We note that for any $\delta\geqslant 1$ there exists some optimal $c_\delta>0$ such that the estimate

$$\Upsilon(r)+\Upsilon(-r)\geqslant c_\delta|r|^{1+\delta} \tag{22}$$

holds true for any $r\in\mathbb{R}$, which follows from the asymptotic and monotonic behavior of the mapping $r\mapsto\Upsilon(r)+\Upsilon(-r)$, $r\in\mathbb{R}$, and the fact that $\Upsilon(r)+\Upsilon(-r)\sim r^2$ as $r\to 0$. For instance, it can be easily checked that the optimal constant for $\delta=1$ in (22) is given by $c_1=1$.

We illustrate the practical use of (22) in the following example.

Example 2.7. Let X be an arbitrary countable set with at least two elements and $l : X \to (0, \infty)$ being integrable on X with respect to the counting measure. Further, we set $k(x, y) = l(y)$ for all $x, y \in X$, $x \neq y$. Then μ given by $\mathrm{d}\mu = \pi \mathrm{d}\#$ with $\pi(x) = l(x)$, $x \in X$, is an invariant and reversible measure. In [31, Example 5.2] it has been shown that L satisfies $CD_\Upsilon(\sqrt{2|l|_1 l_*}, \infty)$, where $l_* = \inf_{x \in X} l(x)$ and $|l|_1$ denotes the ℓ^1-norm with respect to the counting measure on X. Clearly, the integrability of l implies that $l_* = 0$ if X is infinite. It has also been shown that $CD_\Upsilon(0, \infty)$ is best possible (concerning the curvature term) in the infinite state space case. Therefore, we will only consider the case of X being finite in the sequel, i.e. $l_* > 0$ holds.

In [31, Example 5.2], the following representation formula has been established for $f \in \mathbb{R}^X$ at $x \in X$ from (13)

$$\begin{aligned} 2\Psi_{2,\Upsilon}(f)(x) = \sum_{y \in X \setminus \{x\}} l(y) \Big(& |l|_1 \Upsilon'(f(y) - f(x))(f(y) - f(x)) + l(x) \Upsilon(f(x) - f(y)) \\ & - \big(|l|_1 - l(x)\big) \Upsilon(f(y) - f(x)) \Big) + \sum_{\substack{y,z \in X \\ y,z \neq x}} l(y) l(z) \Upsilon(f(z) - f(y)). \end{aligned}$$

By positivity of the mapping $r \mapsto \Upsilon(r)$ and (22), we proceed as follows

$$\begin{aligned} 2\Psi_{2,\Upsilon}(f)(x) &\geqslant 2\kappa \Psi_\Upsilon(f)(x) \\ &\quad + \sum_{y \in X \setminus \{x\}} l(y) \Big(|l|_1 \Upsilon'(f(y) - f(x))(f(y) - f(x)) + l(x) \Upsilon(f(x) - f(y)) \\ &\qquad - \big(|l|_1 + 2\kappa - l(x)\big) \Upsilon(f(y) - f(x)) \Big) \\ &\geqslant 2\kappa \Psi_\Upsilon(f)(x) + 2\alpha c_\delta \sum_{y \in X \setminus \{x\}} l(y) \big|f(x) - f(y)\big|^{1+\delta} \\ &\quad + \sum_{y \in X \setminus \{x\}} l(y) \Big(|l|_1 \Upsilon'(f(y) - f(x))(f(y) - f(x)) + (l(x) - 2\alpha) \Upsilon(f(x) - f(y)) \\ &\qquad - (|l|_1 + 2\kappa - (l(x) - 2\alpha)) \Upsilon(f(y) - f(x)) \Big) \\ &= 2\kappa \Psi_\Upsilon(f)(x) + 2\alpha c_\delta \sum_{y \in X \setminus \{x\}} l(y) \big|f(x) - f(y)\big|^{1+\delta} \\ &\quad + (l(x) - 2\alpha) \sum_{y \in X \setminus \{x\}} l(y) \nu_{\frac{|l|_1}{l(x)-2\alpha}, \frac{|l|_1 + 2\kappa}{l(x)-2\alpha} - 1}(f(y) - f(x)), \end{aligned}$$

for $\kappa > 0$, $\delta \geqslant 1$ and $\alpha > 0$ such that $l(x) > 2\alpha$. Now, we aim to apply Proposition 2.6 with $\gamma(r) = |r|^{1+\delta}$. Therefore it is desirable that the last summand of the latter term is non-negative. For this purpose we employ [31, Lemma A.4(ii)], which yields that if $|l|_1 \geqslant 2(l(x) - 2\alpha)$ it suffices that

$$\frac{2\kappa}{l(x) - 2\alpha} - 1 \leqslant 2^{\frac{3}{2}} \sqrt{\frac{|l|_1}{l(x) - 2\alpha}} - 1,$$

which is equivalent to $\kappa \leqslant \sqrt{2|l|_1 (l(x) - 2\alpha)}$. In case that $|l|_1 < 2(l(x) - 2\alpha)$ it suffices by [31, Lemma A.4(ii)] that

$$\frac{2\kappa}{l(x) - 2\alpha} - 1 \leqslant \frac{2|l|_1}{l(x) - 2\alpha} - 1,$$

which is equivalent to $\kappa \leqslant |l|_1$. Clearly, $|l|_1 \geqslant \sqrt{2|l|_1(l(x)-2\alpha)}$ for at least one $x \in X$. Consequently, we obtain by Proposition 2.6 that L satisfies $CD_\Upsilon\big(\sqrt{2|l|_1(l_*-2\alpha)}, F\big)$ with CD-function $F(r) = \frac{\alpha c_\delta}{|l|_1^\delta} r^{1+\delta}$, $r \geqslant 0$, for any $\alpha \in \big(0, \frac{l_*}{2}\big)$ and any $\delta \geqslant 1$.

Example 2.8. We choose $l(x) = 1$ in the setting of Example 2.7 for any $x \in X$ with given finite state space X consisting of n elements, $n \geqslant 2$. Then the underlying graph to L is given by the complete graph K_n and L satisfies $CD_\Upsilon(\sqrt{2n(1-2\alpha)}, \frac{n}{\alpha})$ for any $\alpha \in (0, \frac{1}{2})$ by Example 2.7 in the case of $\delta = 1$ (recall that we have $c_1 = 1$ in (22)). It is natural to ask whether there also exists a dimension bound which is uniform in n while having non-negative curvature as it is the case for the Bakry-Émery condition (see e.g. [15]). Interestingly, in Example 2.14 we are able to give a negative answer to this question.

The procedure described in Example 2.7 can be seen as a guidance for other examples where the mapping $\nu_{c,d}$ plays a similar role, e.g. for weighted 4-cycles, finite birth-death processes and weighted stars as discussed in [31].

The case of (R)-Ricci-flat graphs will be discussed seperatly below. For the reader's convenience we recall the definition of (R)-Ricci flat graphs, which originates from the work of [7].

Definition 2.9. *Let $G = (V, E)$ be an unweighted d-regular graph. We call G (R)-Ricci-flat at $x \in V$ if there exist maps $\eta_i : B_1(x) \to V$ (where $B_1(x)$ denotes the closed ball with radius 1 and center x with respect to the combinatorical graph distance) for $1 \leqslant i \leqslant d$ satisfying the following properties:*

(i) $\eta_i(u) \in B_1(u)\backslash\{u\}$ *for any* $u \in B_1(x)$ *and* $i \in \{1, ..., d\}$,
(ii) $\eta_i(u) \neq \eta_j(u)$, *whenever* $i \neq j$,
(iii) $\bigcup_j \eta_j(\eta_i(x)) = \bigcup_j \eta_i(\eta_j(x))$ *for any* $i \in \{1, ..., d\}$,
(iv) $\eta_i(\eta_i(x)) = x$ *for any* $i \in \{1, ..., d\}$.

We call G (R)-Ricci-flat if G is (R)-Ricci-flat at each $x \in V$.

(R)-Ricci-flat graphs constitute a subclass of Ricci-flat graphs with Bakry-Émery condition $CD(2, \infty)$, see [7]. Important examples are given by complete bipartite graphs and, since (R)-Ricci-flat graphs are invariant under tensorization, by the hypercube, cf. [7]. In [31, Example 5.12] it has been shown that a Markov generator with underlying graph being (R)-Ricci-flat even satisfies $CD_\Upsilon(2, \infty)$.

Example 2.10. Let the transition rates be given such that the underlying graph to L is a (d-regular) (R)-Ricci-flat graph with vertex set X. Further, let $x \in X$ be chosen arbitrary and let $\big(\eta_i\big)_{i=1,...,d}$ denote the corresponding mappings from Definition 2.9. In [31, Example 5.12] it has been shown that for $f \in \mathbb{R}^X$ the estimate

$$2\Psi_{2,\Upsilon}(f)(x) \geqslant 4\Psi_\Upsilon(f)(x) + \sum_{i=1}^{d} \nu_{2,5}(f(\eta_i(x)) - f(x))$$

holds true. Now, one readily checks (see also the proof of [31, Lemma A.3]) that $\nu_{2,5}$ is strictly convex on $\mathbb{R}$. We infer from Proposition 2.6 that L satisfies $CD_\Upsilon(2, F)$ with $F(r) = \frac{d}{2}\nu_{2,5}(-\frac{r}{d})$, $r \geqslant 0$, which is a CD-function due to the asymptotic behavior at ∞ and by Remark 2.2. Interestingly, since the curvature constant in $CD_\Upsilon(2, \infty)$ is optimal in general, which follows, for instance, from the case where the underlying graph to L is given by the hypercube (cf. [31]), we do not need a trade off from the curvature constant in order to achieve the $CD_\Upsilon(2, F)$ condition (which is, for instance, in contrast to the procedure described in Example 2.7). Further, note that $\nu_{2,5}$ behaves only quartic near 0. In fact, there does not exist in general a CD-function $\hat{F}$

behaving quadratically near zero such that L satisfies $CD_\Upsilon(2,\hat{F})$, by combining Remark 2.4(ii) with the fact that the hypercube does not satisfy $CD(2,m)$ for some $m<\infty$ (see [8]).

Example 2.11. Here we consider a birth-death process with infinite state space $X=\mathbb{N}_0$. We use the notation originating from [6], which has also been used in [31], and introduce the functions $a,b:X\to[0,\infty)$ with $a(x)=k(x,x+1)$, $b(x)=k(x,x-1)$, $b(0)=0$, $b(x)>0$ otherwise, and $a(x)>0$ for any $x\in X$. Moreover, we set $k(x,y)=0$ whenever $|x-y|>1$. The detailed balance condition now reads as

$$a(x)\pi(x)=b(x+1)\pi(x+1) \tag{23}$$

for any $x\in X$. Note that the measure μ given by $\mathrm{d}\mu=\pi\mathrm{d}\#$ is a finite measure if and only if

$$\sum_{x=1}^{\infty}\frac{a(x-1)\cdots a(0)}{b(x)\cdots b(1)}<\infty.$$

We assume monotonicity of the rates in the sense that $a(x)\geqslant a(x+1)$ and $b(x+1)\geqslant b(x)$ for any $x\in X$ and moreover that

$$a(x)-a(x+1)+b(x+1)-b(x)\geqslant\kappa \tag{24}$$

holds for any $x\in X$ and some $\kappa>0$. Those assumptions led to modified logarithmic Sobolev inequalities in [6], and in [31] it has been shown that they imply $CD(\frac{\kappa}{2},\infty)$. Apparently they also entail that $a(x)\leqslant a(0)$ for any $x\in X$ and $b(x)\to\infty$ as $x\to\infty$.

We specify for $x\in X$ a function $f_x\in\mathbb{R}^X$ such that $f_x(x+2)=f_x(x+1)+t$ and $f_x(x-2)=f_x(x-1)+s$, where we define $t=f_x(x+1)-f_x(x)$ and $s=f_x(x-1)-f_x(x)$ (this is called minimizing $\Psi_{2,\Upsilon}(f)$ over the second neighborhood throughout [31]) and set $t=0$. Then we observe from (13) (see also the representation formula for $\Psi_{2,\Upsilon}(f)(x)$ that has been established in [31, Example 5.13])

$$2\Psi_{2,\Upsilon}(f_x)(x)=b(x)\Big(\Upsilon(s)\big(b(x-1)-b(x)-a(x)\big)+\Upsilon(-s)a(x-1)+\Upsilon'(s)s\big(a(x-1)+b(x)-b(x-1)\big)\Big).$$

Assuming that $2\Psi_{2,\Upsilon}(f_x)(x)$ is greater than or equal to $\frac{1}{n}b(x)^2s^2$ (which equals $\frac{1}{n}(-Lf_x(x))^2$) for some $n\in[1,\infty)$ implies that

$$\begin{aligned}0&\leqslant b(x)\big(\Upsilon'(s)s-\Upsilon(s)-\frac{1}{n}s^2\big)+b(x-1)\big(\Upsilon(s)-\Upsilon'(s)s\big)\\&\quad+a(x-1)\big(\Upsilon(-s)+\Upsilon'(s)s\big)-a(x)\Upsilon(s)\\&\leqslant a(0)\big(\Upsilon(-s)+\Upsilon'(s)s\big)+b(x)\big(\Upsilon'(s)s-\Upsilon(s)-\frac{1}{n}s^2\big)\end{aligned}$$

Choosing $s<0$ such that $\Upsilon'(s)s-\Upsilon(s)-\frac{1}{n}s^2<0$ and sending $x\to\infty$ yields a contradiction.

Interestingly, in [31] it has been shown that under an assumption which is stronger than (24) the $CD_\Upsilon(\kappa_0,\infty)$ condition holds with some positive constant $\kappa_0>0$. This shows that it is possible to have positive curvature bounds while having no finite dimension bound regarding the $CD_\Upsilon(\kappa,n)$ condition for some $\kappa>0$. Furthermore, note that we will show by means of Corollary 3.10 below that a CD_Υ condition with positive curvature bound and a non-quadratic power-type CD-function does not hold either.

Next, we give a quite simple negative criterion for the existence of a dimension term with regard to the quadratic CD-function in the infinite state space case.

Proposition 2.12. *If there exists a sequence $(x_m)_{m\in\mathbb{N}}\subset X$ such that*

$$\frac{N(x_m)}{(M_1(x_m))^2}\to 0 \tag{25}$$

as $m \to \infty$, then there does not exist some $n < \infty$ such that $CD_\Upsilon(0,n)$ holds true.

Proof. We consider $\big(f_m\big)_{m\in\mathbb{N}} \subset \ell^\infty(X)$ given by $f_m(y) = t$ for any $y \in X\backslash\{x_m\}$ and $f_m(x_m) = 0$, $m \in \mathbb{N}$, with $t < 0$ being specified below. We read from (13) that

$$2\Psi_{2,\Upsilon}(f_m)(x_m) = N(x_m)\big(\Upsilon'(t)t + \Upsilon(-t)\big) + (M_1(x_m))^2\big(\Upsilon'(t)t - \Upsilon(t)\big)$$

holds. Moreover, we have $(-Lf_m(x_m))^2 = t^2(M_1(x_m))^2$. Thus, $2\Psi_{2,\Upsilon}(f_m)(x_m) \geqslant \frac{1}{n}(-Lf_m(x_m))^2$, with $1 \leqslant n < \infty$, is equivalent to

$$\frac{N(x_m)}{(M_1(x_m))^2}\big(\Upsilon'(t)t + \Upsilon(-t)\big) + \Upsilon'(t)t - \Upsilon(t) - \frac{1}{n}t^2 \geqslant 0. \tag{26}$$

Choosing $t < 0$ such that $\Upsilon'(t)t - \Upsilon(t) - \frac{1}{n}t^2 < 0$ and sending $m \to \infty$, yields by (25) a contradiction to (26). □

Clearly, Proposition 2.12 also yields a necessary condition for families of Markov generators satisfying a uniform $CD_\Upsilon(0,n)$ condition. For the sake of clarity we will state this in the following corollary in the case that the underlying graphs to the corresponding Markov generators are unweighted, in which case the mappings N and M_1 are equal. The following corollary follows from the same arguments as in the proof of Proposition 2.12. Despite its simplicity, these findings lead to a remarkable difference between the $CD_\Upsilon(0,n)$ and the $CD(0,n)$ condition, as it will be demonstrated by Example 2.14.

Corollary 2.13. *Let I be an arbitrary index set and $(L_i)_{i\in I}$ a family of Markov generators whose respective underlying graphs are unweighted and with corresponding state space $(X_i)_{i\in I}$. Assume that there exist sequences $(i_m)_{m\in\mathbb{N}} \subset I$ and $(x_m)_{m\in\mathbb{N}} \subset X$, where $X = \bigcup_{i\in I} X_i$, such that $M_1^{(i_m)}(x_m) \to \infty$ as $m \to \infty$. Here the upper index denotes that the function M_1 corresponds to the respective Markov generator. Then there exists no $n < \infty$ such that L_i satisfies $CD_\Upsilon(0,n)$ for all $i \in I$.*

In particular, if the underlying graph to a Markov generator L is given by an unweighted graph with unbounded vertex degree, then there exists no $n < \infty$ such that $CD_\Upsilon(0,n)$ holds true.

Example 2.14. We consider the index set $I = \{n \in \mathbb{N} : n \geqslant 2\}$ (in the sense of Corollary 2.13) and the Markov generator L_n whose underlying graph is given by the complete graph K_n for any $n \in I$. It is known on the one hand that $CD(0,4)$ holds for any L_n (see [15, Proposition 3]), i.e. a dimension-term exists under non-negative curvature with respect to the Bakry-Émery CD-condition that is uniform with regard to n. On the other hand, due to Corollary 2.13 there does not exist a (uniform) $d < \infty$ such that $CD_\Upsilon(0,d)$ holds for any L_n, $n \in I$.

3. Entropy-Information inequalities

From now on we assume that the unique and reversible invariant measure μ is a probability measure.

We consider the entropy

$$\mathrm{Ent}_\mu(f) = \int_X f\log f \mathrm{d}\mu - \int_X f\mathrm{d}\mu \log \int_X f\mathrm{d}\mu \tag{27}$$

for any positive function $f \in \ell^1(\mu)$. It is well known that $\mathrm{Ent}_\mu(f) \geqslant 0$. Note that we also allow for the value of $\mathrm{Ent}_\mu(f) = \infty$.

The Fisher information is given by

$$\mathcal{I}(f) = \frac{1}{2}\sum_{x,y\in X} k(x,y)\big(f(y) - f(x)\big)\big(\log f(y) - \log f(x)\big)\pi(x). \tag{28}$$

If we assume that $M_1 \in \ell^1(\mu)$ then $f \in \ell^{\infty,+}(X)$ ensures that $\mathcal{I}(f) < \infty$. Further, in the latter case we have the representation

$$\mathcal{I}(f) = \int_X f \Psi_\Upsilon(\log f) \mathrm{d}\mu,$$

see [31, Section 3], where the formula has been established for $f \in \ell^{\infty,+}(X)$ being a probability densitiy with respect to μ, although the proof extends verbatim to the general case. Since we sum up in the right-hand side of (28) over non-negative entries, we can extend the functional $\mathcal{I}$ to positive functions $f : X \to (0, \infty)$, where we allow for the value of $\mathcal{I}(f) = \infty$.

Note that the assumption of the Markov chain being irreducible implies that $\mathcal{I}(f)$=0 if and only if f is constant and positive.

Besides that, one readily verifies that the well known scaling behavior

$$\mathrm{Ent}_\mu(cf) = c\,\mathrm{Ent}_\mu(f) \tag{29}$$

and

$$\mathcal{I}(cf) = c\mathcal{I}(f) \tag{30}$$

holds respectively for any constant $c > 0$.

Our main object of investigation in the remaining part of this article will be the following family of functional inequalities.

Definition 3.1. *We say that L satisfies an entropy-information inequality $EI(\Phi)$ with respect to a strictly increasing and concave C^1-function $\Phi : (0, \infty) \to (0, \infty)$, which we refer to as the growth function, if for every $f \in P_*(X)$ with $\mathrm{Ent}_\mu(f) < \infty$ and $\mathcal{I}(f) < \infty$*

$$\mathrm{Ent}_\mu(f) \leqslant \Phi(\mathcal{I}(f)) \tag{31}$$

holds, where we set $\Phi(0) := \lim_{r \to 0^+} \Phi(r)$.

A well known example for an entropy-information inequality is the modified logarithmic Sobolev inequality

$$\mathrm{Ent}_\mu(f) \leqslant \frac{1}{2\kappa}\mathcal{I}(f) \tag{32}$$

with constant $\kappa > 0$. See [5] for an extensive account on modified logarithmic Sobolev inequalities in the discrete setting of Markov chains. Further, the functional inequality (32) was subject of investigation in [6], [12], [13] and [14], as well as in [31] where it has been shown that $CD_\Upsilon(\kappa, \infty)$ (with $\kappa > 0$) together with positive recurrence and the integrability conditions $M_1 \in \ell^2(\mu)$ and $M_2 \in \ell^1(\mu)$ imply (32) with constant κ.

Remark 3.2. (i) The diffusive counterpart to Definition 3.1, so called entropy-energy inequalities, are defined for growth functions mapping to $\mathbb{R}$ instead of $(0, \infty)$, see [3]. This generality in the context of [3] allows to include the quite important special case of the Euclidean logarithmic Sobolev inequality (cf. [3, Proposition 6.2.5]). However, assuming that Φ is non-negative is not a restriction in our setting where we have supposed that μ is a probability measure. Indeed, applying (31) to $f = \mathbb{1}$ shows that $\lim_{r \to 0^+} \Phi(r) < 0$ is impossible to hold.

(ii) It will turn out to be quite useful to write (31) in an equivalent linearized form. More precisely, as Φ is concave we deduce from $\Phi(s) \leqslant \Phi(r) + \Phi'(r)(s - r)$, $s, r \in (0, \infty)$, that (31) implies

$$\mathrm{Ent}_\mu(f) \leqslant \Phi'(r)\,\mathcal{I}(f) + \Theta(r) \tag{33}$$

for any $r \in (0,\infty)$, where $\Theta(r) = \Phi(r) - \Phi'(r)r$. Conversely, specifying $r = \mathcal{I}(f)$, (33) implies (31). Note that concavity also implies that $\Theta(r) \geqslant 0$ for any $r > 0$ since $\lim_{r\to 0^+} \Phi(r) \geqslant 0$. In particular, this implies that (33) also holds for the case of $\mathcal{I}(f) = 0$, or equivalently for $f = \mathbb{1}$.

(iii) Let $f \in \ell^1(\mu)$ be positive with $\mathrm{Ent}_\mu(f) < \infty$ and $\mathcal{I}(f) < \infty$. The entropy-information inequality in the form (33) extends to f by

$$\mathrm{Ent}_\mu(f) \leqslant \Phi'(r)\mathcal{I}(f) + \Theta(r)\int_X f d\mu \tag{34}$$

for any $r \in (0,\infty)$. This is a consequence of the scaling behavior (29) and (30), after having applied (33) to $\frac{f}{\|f\|_1}$.

Now, we come to the main theorem of this section, that links the previous section to the notion of entropy-information inequalities.

Theorem 3.3. *Let $M_1 \in \ell^2(\mu)$, $M_2 \in \ell^1(\mu)$ and the Markov chain generated by L be positive recurrent. Further, let L satisfy $CD_\Upsilon(\kappa, F)$ with $\kappa > 0$ and a convex CD-function $F : [0,\infty) \to [0,\infty)$ such that $F|_{(0,\infty)} \in C^1\big((0,\infty)\big)$ and $\frac{F'(r)r}{F(r)} \geqslant 1 + \delta$ holds for any $r > 0$ and some $\delta > 0$. Let $G : (0,\infty) \to (0,\infty)$ denote the inverse function of $r \mapsto \frac{F(r)}{r}$, $r > 0$. Then*

$$\mathrm{Ent}_\mu(f) \leqslant \int_0^\infty G\Big(\frac{\kappa}{e^{2\delta\kappa t}\big(1 + \frac{\kappa\mathcal{I}(f)}{F(\mathcal{I}(f))}\big) - 1}\Big)\mathrm{d}t \tag{35}$$

holds for any $f \in P_*(X)$ with $\mathrm{Ent}_\mu(f) < \infty$ and $\mathcal{I}(f) \in (0,\infty)$.

Proof. It suffices to deduce the claim for $f \in P_*^+(X)$. The full statement follows then from the same standard truncation argument as presented in [31, Lemma 3.2] and the dominated convergence theorem for approximating the right-hand side of (35). For $f \in P_*^+(X)$ we set $\Lambda(t) = \mathrm{Ent}_\mu(P_t f)$, $t \geqslant 0$. It is well known that

$$\Lambda'(t) = -\mathcal{I}(P_t f) \tag{36}$$

is valid provided that $M_1 \in \ell^1(\mu)$. Further, we infer from [31, Theorem 3.4] that

$$\Lambda''(t) = 2\int_X P_t f \Psi_{2,\Upsilon}(\log P_t f)\mathrm{d}\mu \tag{37}$$

holds true given the assumptions $M_1 \in \ell^2(\mu)$ and $M_2 \in \ell^1(\mu)$. We apply $CD_\Upsilon(\kappa, F)$ to deduce

$$\begin{aligned}\Lambda''(t) &\geqslant -2\kappa\Lambda'(t) + 2\int_X P_t f\, F_0(-L(\log P_t f))\mathrm{d}\mu \\ &\geqslant -2\kappa\Lambda'(t) + 2F_0\Big(-\int_X P_t f L(\log P_t f)\mathrm{d}\mu\Big),\end{aligned}$$

where the latter follows from convexity of the trivial extension F_0 (which follows from convexity of F), the fact that $P_t f$ is a probability density with respect to μ, which follows from μ being invariant for $(P_t)_{t\geqslant 0}$, and Jensen's inequality. Now, by the identity (5) (cf. [31, Lemma 2.2]) and μ being invariant, we have

$$-\int_X P_t f L(\log P_t f)\mathrm{d}\mu = \int_X P_t f\, \Psi_\Upsilon(\log P_t f)\mathrm{d}\mu = \mathcal{I}(P_t f)$$

and hence we end up with the differential inequality

$$\Lambda''(t) \geqslant -2\kappa\Lambda'(t) + 2F(-\Lambda'(t)). \tag{38}$$

Note that in fact $\Lambda'(t) < 0$ holds, since we have $f = \mathbb{1}$ otherwise. Further, we observe that

$$\begin{aligned}\frac{d}{dt}\Big[e^{-2\delta\kappa t}\Big(1-\kappa\frac{\Lambda'(t)}{F(-\Lambda'(t))}\Big)\Big] &= e^{-2\delta\kappa t}\Big(\frac{2\delta\kappa^2\Lambda'(t)}{F(-\Lambda'(t))} - 2\delta\kappa - \kappa\Lambda''(t)\frac{F(-\Lambda'(t)) + F'(-\Lambda'(t))\Lambda'(t)}{F(-\Lambda'(t))^2}\Big)\\ &= \frac{\kappa e^{-2\delta\kappa t}}{F(-\Lambda'(t))}\Big(2\delta\kappa\Lambda'(t) - 2\delta F(-\Lambda'(t)) - \Lambda''(t)\Big(1+\frac{F'(-\Lambda'(t))\Lambda'(t)}{F(-\Lambda'(t))}\Big)\Big)\\ &\geqslant \frac{\delta\kappa e^{-2\delta\kappa t}}{F(-\Lambda'(t))}\Big(2\kappa\Lambda'(t) - 2F(-\Lambda'(t)) + \Lambda''(t)\Big),\end{aligned}$$

where we have applied the condition $\frac{-\Lambda'(t)F'(-\Lambda'(t))}{F(-\Lambda'(t))} \geqslant 1+\delta$ and $\Lambda''(t) \geqslant 0$ in the last step. Hence, (38) yields that the mapping $t \mapsto e^{-2\delta\kappa t}\big(1-\kappa\frac{\Lambda'(t)}{F(-\Lambda'(t))}\big)$ is increasing. In particular, this implies that

$$e^{-2\delta\kappa t}\Big(1-\kappa\frac{\Lambda'(t)}{F(-\Lambda'(t))}\Big) \geqslant 1-\kappa\frac{\Lambda'(0)}{F(-\Lambda'(0))},$$

which can be rearranged to

$$\frac{F(-\Lambda'(t))}{-\Lambda'(t)} \leqslant \frac{\kappa}{e^{2\delta\kappa t}\big(1-\kappa\frac{\Lambda'(0)}{F(-\Lambda'(0))}\big) - 1}.$$

Then, using $\mathcal{I}(f) = -\Lambda'(0)$, we obtain

$$-\Lambda'(t) \leqslant G\Big(\frac{\kappa}{e^{2\delta\kappa t}\big(1+\frac{\kappa\mathcal{I}(f)}{F(\mathcal{I}(f))}\big) - 1}\Big).$$

Consequently, we conclude

$$\Lambda(0) - \Lambda(T) = -\int_0^T \Lambda'(t)\mathrm{d}t \leqslant \int_0^T G\Big(\frac{\kappa}{e^{2\delta\kappa t}\big(1+\frac{\mathcal{I}(f)}{F(\mathcal{I}(f))}\big) - 1}\Big)\mathrm{d}t. \tag{39}$$

The claim follows by sending $T \to \infty$. Indeed, $\Lambda(T) \to 0$ as $T \to \infty$ follows from the dominated convergence theorem, the Markov chain being ergodic and $(P_t)_{t\geqslant 0}$ being a Markov semigroup. □

Remark 3.4. (i) The crucial assumption that

$$F'(r) \geqslant \frac{(1+\delta)F(r)}{r}, \tag{40}$$

holds for some $\delta > 0$ and any $r > 0$ implies by Gronwall's inequality that we have for fixed $a > 0$

$$\frac{F(r)}{r^{1+\delta}} \geqslant \frac{F(a)}{a^{1+\delta}},$$

for any $r > a > 0$, i.e. the mapping $r \mapsto \frac{F(r)}{r^{1+\delta}}$, $r > 0$, is increasing. Conversely, differentiating $r \mapsto \frac{F(r)}{r^{1+\delta}}$, $r > 0$, the property (40) follows provided that $r \mapsto \frac{F(r)}{r^{1+\delta}}$, $r > 0$, is increasing. Hence, both properties are equivalent. In particular, we observe that (40) ensures that F grows at least like $r^{1+\delta}$ as $r \to \infty$. On the other hand, recall that we have seen in Remark 2.4(iii) that F can not behave better than quadratic at 0 provided that $CD_\Upsilon(\kappa, F)$ holds for some $\kappa \in \mathbb{R}$.

(ii) Note that the mapping $G : (0,\infty) \to (0,\infty)$ is in fact well defined, since assuming that $r \mapsto \frac{F(r)}{r^{1+\delta}}$ is increasing on $(0,\infty)$ for some $\delta > 0$ implies that $\frac{F(r)}{r} \to 0$ as $r \to 0$ and $\frac{F(r)}{r} \to \infty$ as $r \to \infty$.

(iii) We interpret the integral on the right-hand side of (35) as ∞ in the case that the integral is divergent. In fact, this situation appears even under the assumptions of Theorem 3.3 as the CD-function $F(r) = r^m e^{-\frac{1}{r^m}}$, $r \geqslant 0$, for some $m > 1$ shows. Indeed, one readily verifies that (40) with $\delta = m - 1$ and convexity of F respectively hold true. The problem results from the

behavior of F in the origin. More precisely, $\frac{F(r)}{r}$ converges faster to 0 than $e^{-\frac{1}{r^m}}$ as $r \to 0^+$, which yields that $G(r)$ tends slower to 0 than $(\log \frac{1}{r})^{-\frac{1}{m}}$ as $r \to 0^+$. Consequently the integrand in the right-hand side of (35) dominates a behavior of $t^{-\frac{1}{m}}$ as $t \to \infty$, which yields that the integral does not converge.

For general CD-functions the integral on the right-hand side of (35) can not be calculated explicitly and, moreover, it is not clear whether the mapping $r \mapsto \int_0^\infty G\big(\frac{\kappa}{e^{2\kappa t}(1+\frac{\kappa r}{F(r)})-1}\big)\mathrm{d}t$ is concave. We will focus in the sequel on the situation where the CD-function is given by some power-type function, in which case the mapping $G : (0,\infty) \to (0,\infty)$ of Theorem 3.3 can be given explicitly. In fact, the following result shows in particular that the functional inequality (35) is compatible with Definition 3.1 for power-type CD-functions.

Proposition 3.5. *Let $M_1 \in \ell^2(\mu)$, $M_2 \in \ell^1(\mu)$ and the Markov chain generated by L be positive recurrent. Further, let L satisfy $CD_\Upsilon(\kappa, F)$ with $\kappa > 0$ and $F(r) = \frac{r^{\delta+1}}{n}$, $r \geqslant 0$, for some $\delta \geqslant 1$ and $n \in (0,\infty)$. Then L satisfies $EI(\Phi)$ with the growth function*

$$\Phi(r) = \frac{\sqrt[\delta]{\kappa n}}{2\kappa} \int_{\frac{\sqrt[\delta]{\kappa n}}{r}}^\infty \frac{v^{\delta-2}}{v^\delta + 1}\mathrm{d}v,\ r > 0. \tag{41}$$

Moreover, the growth function Φ satisfies the following assertions:

(i) *Φ is bounded if and only if $\delta > 1$,*

(ii) $\int_0^\infty \frac{\Phi(s^2)}{s^2}\mathrm{d}s < \infty$.

Proof. Clearly, F is convex and $F'(r)r = (1+\delta)F(r)$ holds for any $r > 0$. The mapping $G : (0,\infty) \to (0,\infty)$ from Theorem 3.3 is given by $G(r) = \sqrt[\delta]{nr}$, $r > 0$. Now, it follows from elementary substitution that

$$\int_0^\infty \sqrt[\delta]{\frac{\kappa n}{e^{2\delta\kappa t}(1+\frac{\kappa n}{r^\delta})-1}}\mathrm{d}t = \frac{\sqrt[\delta]{\kappa n}}{2\delta\kappa} \int_{\frac{\kappa n}{r^\delta}}^\infty \frac{1}{\sqrt[\delta]{u}(u+1)}\mathrm{d}u = \frac{\sqrt[\delta]{\kappa n}}{2\kappa} \int_{\frac{\sqrt[\delta]{\kappa n}}{r}}^\infty \frac{v^{\delta-2}}{v^\delta+1}\mathrm{d}v = \Phi(r).$$

By means of Definition 3.1 and Theorem 3.3 it suffices to prove that Φ is concave in order to deduce that L satisfies $EI(\Phi)$ (note that there is nothing to show for the case of $\mathcal{I}(f) = 0$). To that aim, we differentiate Φ and observe

$$\begin{aligned} \Phi'(r) &= \frac{n}{2(\kappa n + r^\delta)}, \\ \Phi''(r) &= -\frac{n\delta r^{\delta-1}}{2(\kappa n + r^\delta)^2}, \end{aligned} \tag{42}$$

which implies concavity of Φ.

The growth function Φ is bounded if and only if the integral in the right-hand side of (41) converges as $r \to \infty$. The latter property holds true if and only if the integral $\int_{\frac{\sqrt[\delta]{\kappa n}}{r}}^1 v^{\delta-2}\mathrm{d}v$ converges as $r \to \infty$, which happens to be true if and only if $\delta > 1$.

Regarding the remaining assertion, we note that there is nothing to show for the behavior at ∞ by boundedness of Φ in case of $\delta > 1$ and by the explicit formula for the growth function in the special case of $\delta = 1$, which will be deduced in Corollary 3.7 below. As to the behavior at 0, we observe for $\varepsilon > 0$ that

$$\int_0^\varepsilon \frac{1}{s^2} \int_{\frac{\sqrt[\delta]{\kappa n}}{s^2}}^\infty \frac{v^{\delta-2}}{v^\delta+1}\mathrm{d}v\,\mathrm{d}s \leqslant \int_0^\varepsilon \frac{1}{s^2} \int_{\frac{\sqrt[\delta]{\kappa n}}{s^2}}^\infty \frac{1}{v^2}\mathrm{d}v\,\mathrm{d}s = \frac{\varepsilon}{\sqrt[\delta]{\kappa n}}.$$

□

Remark 3.6. Let us emphasize an analogy between the relation of the curvature-dimension conditions $CD_\Upsilon(\kappa, F)$ (with power-type CD-function F) and $CD_\Upsilon(\kappa, \infty)$, and the resulting functional inequalities, $EI(\Phi)$ with Φ given by (41) and the modified logarithmic Sobolev inequality (32) with constant $\kappa > 0$. Clearly, $CD_\Upsilon(\kappa, F)$ implies $CD_\Upsilon(\kappa, \infty)$. On the other hand, concavity of Φ (where Φ is given by (41)) implies that $\Phi(r) \leqslant \Phi(s) + \Phi'(s)(r-s)$, $r, s \in (0, \infty)$. Using identity (42), this estimate yields

$$\Phi(r) \leqslant \frac{r}{2\kappa}$$

when sending $s \to 0$. Hence, $EI(\Phi)$ is stronger than the modified logarithmic Sobolev inequality (32) with constant $\kappa > 0$, where the latter is a consequence of $CD_\Upsilon(\kappa, \infty)$ (cf. [31]).

In the case of the $CD_\Upsilon(\kappa, n)$ condition, Theorem 3.3 yields the following important entropy-information inequality.

Corollary 3.7. *Let $M_1 \in \ell^2(\mu)$, $M_2 \in \ell^1(\mu)$ and the Markov chain generated by L be positive recurrent. Further, let L satisfy $CD_\Upsilon(\kappa, n)$ with $\kappa > 0$ and $n < \infty$. Then $EI(\Phi)$ holds with $\Phi(r) = \frac{n}{2} \log\left(1 + \frac{r}{\kappa n}\right)$, $r > 0$, i.e.*

$$\mathrm{Ent}_\mu(f) \leqslant \frac{n}{2} \log\left(1 + \frac{\mathcal{I}(f)}{\kappa n}\right) \tag{43}$$

holds for any $f \in P_*(X)$ with $\mathrm{Ent}_\mu(f) < \infty$.

Proof. Choosing $\delta = 1$ in (41), the claim follows from elementary calculations. □

As we already highlighted in the introduction, we emphasize that (43) serves as a natural discrete analogue to the logarithmic entropy-energy inequality (3), which plays an important role in the diffusive setting of [3] (in which case it holds true provided that $CD(\kappa, n)$ is valid).

In the following example we consider one of the most important instances of a birth-death process from Example 2.11.

Example 3.8. As a special case of a birth-death process from Example 2.11 (with the notation taken from there), we consider the Poisson case which is given by the choice $a(x) = \lambda$, where $\lambda > 0$ is called the intensity rate, and $b(x) = x$, both for any $x \in \mathbb{N}_0$. The invariant and reversible measure is given by the denisity $\pi_\lambda(x) = \frac{\lambda^x}{x!} e^{-\lambda}$, $x \in \mathbb{N}_0$. In [31, Example 5.13] it has been shown that there does not exists some $\kappa > 0$ such that L satisfies $CD_\Upsilon(\kappa, \infty)$. However, it is known that the Poisson case of the birth-death prosess satisfies the modified logarithmic Sobolev inequality $EI(\Phi)$ with $\Phi(r) = r$, see e.g. [6]. Here we show, that this is the best possible entropy-information inequality for the Poisson case in the sense that if Φ grows slower than linear as $r \to \infty$, then L fails to satisfy $EI(\Phi)$.

To that aim we repeat an argument that has been used in [6] to show sharpness of the corresponding modified logarithmic Sobolev inequality. Indeed, we consider $f_k(x) = \frac{e^{kx}}{e^{\lambda(e^k-1)}}$, which can be readily checked to be an element of $P_*^+(X)$ for any $k \in \mathbb{N}$. We have

$$\begin{aligned}
\mathrm{Ent}_\mu(f_k) &= \frac{1}{e^{\lambda e^k}} \sum_{x \in \mathbb{N}_0} \frac{e^{kx} \lambda^x (kx - \lambda(e^k - 1))}{x!} \\
&= \frac{1}{e^{\lambda e^k}} \Big(k \sum_{x \in \mathbb{N}} \frac{(e^k \lambda)^x}{(x-1)!} - e^{\lambda e^k} \lambda (e^k - 1) \Big) \\
&= \lambda \big(k e^k - e^k + 1 \big).
\end{aligned}$$

Further, it can be easily checked that the detailed balance condition (23) yields that

$$\begin{aligned}\mathcal{E}(f_k, \log f_k) &= \sum_{x\in\mathbb{N}_0} a(x)\big(f_k(x+1) - f_k(x)\big)\big(\log(f_k(x+1)) - \log(f_k(x))\big)\pi(x)\\ &= \frac{\lambda}{e^{\lambda e^k}} \sum_{x\in\mathbb{N}_0} \big(e^{k(x+1)} - e^{kx}\big)\big(k(x+1) - kx\big)\frac{\lambda^x}{x!}\\ &= \lambda k(e^k - 1).\end{aligned}$$

From this, we can see that

$$\frac{\mathrm{Ent}_\mu(f_k)}{\mathcal{I}(f_k)} \to 1, \text{ as } k \to \infty.$$

Consequently, if Φ grows slower than linear at ∞, $EI(\Phi)$ fails for the Poisson case of a birth-death process. In particular, the Poisson case of a birth-death process not only does not satisfy the $CD_\Upsilon(\kappa, n)$ condition (cf. Example 2.11), but also fails on the level of the corresponding entropy-information inequality.

In a somewhat similar fashion, the next result shows quite remarkable consequences of boundedness of the growth function and of the mapping M_1, respectively.

Theorem 3.9. *Let L satisfy $EI(\Phi)$, then the following assertions hold true.*

(i) *If Φ is bounded, then the state space X is finite and the estimate*

$$\mathrm{Ent}_\mu(f) \leqslant \lim_{r\to\infty} \Phi(r)\|f\|_1 \tag{44}$$

holds true for any positive $f \in \mathbb{R}^X$.

(ii) *If $M_{1,\sup} < \infty$ and the state space X is infinite, then Φ grows linearly as $r \to \infty$.*

Proof. We consider for $x \in X$ and some $\varepsilon \in (0,1)$ the function

$$f_x(y) = \begin{cases} \varepsilon, & y \neq x \\ \frac{1-\varepsilon(1-\pi(x))}{\pi(x)}, & y = x \end{cases}.$$

One readily verifies that $f_x \in P_*(X)$ for any $x \in X$. We have

$$\mathrm{Ent}_\mu(f_x) = \int_X f_x \log f_x \,\mathrm{d}\mu = \big(1 - \varepsilon(1-\pi(x))\big)\log\frac{1-\varepsilon(1-\pi(x))}{\pi(x)} + \varepsilon\log\varepsilon\,(1-\pi(x)), \tag{45}$$

for any $x \in X$. Moreover, we observe that

$$\begin{aligned}2\mathcal{I}(f_x) &= \sum_{z,y\in X} k(z,y)\big(f_x(y) - f_x(z)\big)\big(\log f_x(y) - \log f_x(z)\big)\pi(z)\\ &= \sum_{y\in X} k(x,y)\big(f_x(y) - f_x(x)\big)\big(\log f_x(y) - \log f_x(x)\big)\pi(x)\\ &\quad + \sum_{z\in X} k(z,x)\big(f_x(x) - f_x(z)\big)\big(\log f_x(x) - \log f_x(z)\big)\pi(z)\\ &= 2\pi(x)\sum_{y\in X} k(x,y)\big(f_x(y) - f_x(x)\big)\big(\log f_x(y) - \log f_x(x)\big),\end{aligned}$$

where we have applied the detailed balance condition in the last step. Hence, we conclude for any $x \in X$ that

$$\begin{aligned}\mathcal{I}(f_x) &= \pi(x)M_1(x)\Big(\varepsilon - \frac{1-\varepsilon(1-\pi(x))}{\pi(x)}\Big)\log\frac{\varepsilon\pi(x)}{1-\varepsilon(1-\pi(x))}\\ &= M_1(x)(1-\varepsilon)\log\frac{1-\varepsilon(1-\pi(x))}{\varepsilon\pi(x)}.\end{aligned} \tag{46}$$

After this preliminary work we now show the first assertion. The estimate (44) follows from the definition of $EI(\Phi)$ for any $f \in \mathcal{P}_*(X)$ with $\mathrm{Ent}_\mu(f) < \infty$ and $\mathcal{I}(f) < \infty$. We then extend (44) to the more general case of positive $f \in \ell^1(\mu)$ with $\mathrm{Ent}_\mu(f) < \infty$ and $\mathcal{I}(f) < \infty$ by applying (44) to $\frac{f}{\|f\|_1}$. It remains to show that X is finite. Assuming for contradiction that X is infinite, we find a sequence $(x_m)_{m\in\mathbb{N}} \subset X$ such that $\pi(x_m) \to 0$ as $m \to \infty$, since μ is assumed to be a probability measure. We infer from (45) that $\mathrm{Ent}_\mu(f_{x_m}) \to \infty$ as $m \to \infty$, which contradicts what has been shown before.

Let us now turn to the second assertion. Choosing a sequence $(x_m)_{m\in\mathbb{N}} \subset X$ as above and assuming w.l.o.g. that $\pi(x_m) < 1$ for any $m \in \mathbb{N}$, we read from (45) that

$$\mathrm{Ent}_\mu(f_{x_m}) \geqslant \log \frac{1}{\pi(x_m)} + \log(1-\varepsilon) + \varepsilon \log \varepsilon$$

and from (46) that

$$\begin{aligned}\mathcal{I}(f_{x_m}) &\leqslant M_{1,\sup}(1-\varepsilon)\big(\log \frac{1}{\pi(x_m)} - \log \frac{\varepsilon}{1-\varepsilon(1-\pi(x_m))}\big)\\ &\leqslant M_{1,\sup}(1-\varepsilon)\big(\log \frac{1}{\pi(x_m)} - \log \varepsilon\big).\end{aligned}$$

Thus, we have

$$\frac{\mathrm{Ent}_\mu(f_{x_m})}{\mathcal{I}(f_{x_m})} \geqslant \frac{\log \frac{1}{\pi(x_m)} + \log(1-\varepsilon) + \varepsilon \log \varepsilon}{M_{1,\sup}(1-\varepsilon)\big(\log \frac{1}{\pi(x_m)} - \log \varepsilon\big)}.$$

The right hand side of the latter estimate converges to $\frac{1}{(1-\varepsilon)M_{1,\sup}}$ as $m \to \infty$. This yields that there exists a constant $C > 0$ and some $M \in \mathbb{N}$ such that

$$\frac{\mathrm{Ent}_\mu(f_{x_m})}{\mathcal{I}(f_{x_m})} \geqslant C > 0$$

for any $m \geqslant M$. Consequently, it follows from $EI(\Phi)$ that

$$C\,\mathcal{I}(f_{x_m}) \leqslant \Phi(\mathcal{I}(f_{x_m}))$$

for all $m \geqslant M$. Since $\mathcal{I}(f_{x_m}) \to \infty$ as $m \to \infty$ and Φ is concave, we conclude that $\Phi(r)$ must grow linearly as $r \to \infty$. □

While the corresponding growth function in the case of the $CD_\Upsilon(\kappa, n)$ condition (with $\kappa > 0$ and $n < \infty$) grows logarithmically at ∞, we have seen in Proposition 3.5 that for power type CD-functions of higher order the respective growth function is bounded. Combining Proposition 3.5 with Theorem 3.9 leads to the following interesting observation.

Corollary 3.10. *Let $M_1 \in \ell^2(\mu)$, $M_2 \in \ell^1(\mu)$ and the Markov chain generated by L be positive recurrent. Further, let L satisfy $CD_\Upsilon(\kappa, F)$ with $\kappa > 0$ and $F(r) = \frac{r^{1+\delta}}{n}$, $r \geqslant 0$, for some $\delta \geqslant 1$ and $n \in (0, \infty)$. Then the state space X is finite if and only if either $\delta > 1$ or $\delta = 1$ and $M_{1,\sup} < \infty$.*

4. Ultracontractive Bounds under Entropy-Information inequalities

In the classical diffusive setting, entropy-energy inequalities imply under the condition that $r \mapsto \frac{\Phi'(r)}{r}$ is integrable at ∞ ultracontractivity of the semigroup, cf. [3]. But as the entropy-information inequality compares to entropy-energy inequalities like the modified logarithmic Sobolev inequality to logarithmic Sobolev inequalities, it is natural to expect that ultracontractive bounds come in the form of the hypercontractivity bounds from [5] for the modified logarithmic Sobolev inequality, i.e. not with respect to the respective norm of the semigroup, but of $e^{P_t f}$

instead. We recall the following auxiliary result, whose proof is contained in the proof of [5, Theorem 7.1], where the authors have considered an even more general setting. Note in fact that the assumptions (1)–(4) of [5, Section 7] are satisfied provided that $f \in \ell^\infty(X)$. Moreover, $f \in \ell^\infty(X)$ implies that we have $e^{\eta(t)P_t f} \in \ell^{\infty,+}(X)$ for any fixed $t > 0$ (with $\eta(t) \in \mathbb{R}$), which yields that $\mathcal{I}(e^{\eta(t)P_t f}) < \infty$ if we assume in addition that $M_1 \in \ell^1(\mu)$.

Lemma 4.1. *Let $M_1 \in \ell^1(\mu)$, $f \in \ell^\infty(X)$, $t > 0$ and $q : (0, C_0) \to (0, \infty)$ be some differentiable mapping, where $C_0 \in (0, \infty]$. Then we have*

$$q(t)\|e^{P_t f}\|_{q(t)}^{q(t)-1} \frac{d}{dt}\|e^{P_t f}\|_{q(t)} = \frac{q'(t)}{q(t)} \mathrm{Ent}_\mu(e^{q(t)P_t f}) - \mathcal{I}(e^{q(t)P_t f}).$$

Proof. We briefly repeat the calculation of [5] for the reader's convenience and refer for more details to [5, Section 7]. Note that the assumption of $M_1 \in \ell^1(\mu)$ in fact justifies to interchange integration and differentation in the lines below. We have

$$\begin{aligned}
&\frac{d}{dt}\Big(\int_X e^{q(t)P_t f}\mathrm{d}\mu\Big)^{\frac{1}{q(t)}} \\
&\qquad = \|e^{P_t f}\|_{q(t)} \left(\frac{\int_X e^{q(t)P_t f}\big(q'(t)P_t f + q(t)LP_t f\big)\mathrm{d}\mu}{q(t)\int_X e^{q(t)P_t f}\mathrm{d}\mu} - \frac{q'(t)\log\int_X e^{q(t)P_t f}\mathrm{d}\mu}{q(t)^2} \right) \\
&\qquad = \frac{\|e^{P_t f}\|_{q(t)}^{1-q(t)}}{q(t)} \Big(\frac{q'(t)}{q(t)} \mathrm{Ent}_\mu(e^{q(t)P_t f}) - \mathcal{I}(e^{q(t)P_t f}) \Big).
\end{aligned}$$

□

Theorem 4.2. *Let L satisfy $EI(\Phi)$ and $M_1 \in \ell^1(\mu)$. Then for every $1 \leqslant p \leqslant q \leqslant \infty$, every $f \in \ell^\infty(X)$ and every $\varrho > 0$*

$$\|e^{P_{t(\varrho)} f}\|_q \leqslant \|e^f\|_p e^{m(\varrho)}$$

holds true, where

$$t(\varrho) = \int_p^q \frac{\Phi'(\varrho r)}{r}\mathrm{d}r, \ m(\varrho) = \frac{\Phi(\varrho p)}{p} - \frac{\Phi(\varrho q)}{q}. \tag{47}$$

Here the case of $q = \infty$ has to be understood in the limit $q \to \infty$ in both formulas in (47) and can be reached only if $r \mapsto \frac{\Phi'(r)}{r}$ is integrable at ∞.

Proof. We define $\Lambda(t) = \|e^{P_t f}\|_{q(t)}$ for a strictly increasing and differentiable $q : (0, C_0) \to (0, \infty)$, which, together with C_0, will be specified below. By Lemma 4.1 we have

$$q(t)\Lambda(t)^{q(t)-1}\Lambda'(t) = \frac{q'(t)}{q(t)} \mathrm{Ent}_\mu(e^{q(t)P_t f}) - \mathcal{I}(e^{q(t)P_t f}).$$

Applying $EI(\Phi)$ in the form of (34) to $e^{q(t)P_t f}$ yields for any $r > 0$

$$q(t)\Lambda(t)^{q(t)-1}\Lambda'(t) \leqslant \mathcal{I}(e^{q(t)P_t f})\big(\frac{q'(t)}{q(t)}\Phi'(r) - 1\big) + \frac{q'(t)}{q(t)}\Theta(r)\Lambda(t)^{q(t)}.$$

For given $r = r(q)$ (which will be made precise below) we choose $q(t)$ such that the differential equation $q'\Phi'(r(q)) = q$ is satisfied, which in fact can be done by separation of variables. Indeed, let $T : (p, \infty) \to (0, C_0)$ be defined as $T(s) = \int_p^s \frac{\Phi'(r(q))}{q}\mathrm{d}q$, where $C_0 = \int_p^\infty \frac{\Phi'(r(q))}{q}\mathrm{d}q$ (the value $C_0 = \infty$ is allowed). Then T is bijective and $q : (0, C_0) \to (p, \infty)$, given by $q(t) = T^{-1}(t)$, solves the ODE mentioned above on $(0, C_0)$. In particular, $q(0) = p$ extends q continuously onto $[0, C_0)$. We conclude that

$$\Lambda'(t) \leqslant \frac{q'(t)}{q(t)^2}\Theta(r(q(t)))\Lambda(t)$$

holds for any $t \in (0, C_0)$, which is equivalent to the differential inequality

$$(\log \Lambda)' \leqslant \frac{q'}{q^2}\Theta(r(q)). \tag{48}$$

Integrating (48) yields

$$\begin{aligned}\log(\Lambda(t)) &\leqslant \log(\Lambda(0)) + \int_0^t \frac{q'(s)\Theta(r(q(s)))}{q(s)^2}\mathrm{d}s \\ &= \log(\Lambda(0)) + \int_{q(0)}^{q(t)} \frac{\Theta(r(q))}{q^2}\mathrm{d}q.\end{aligned}$$

Note that we can write

$$t = \int_0^t \frac{q'(s)\Phi'(r(q(s)))}{q(s)}\mathrm{d}s = \int_{q(0)}^{q(t)} \frac{\Phi'(r(q))}{q}\mathrm{d}q.$$

Choosing $r(q) = \varrho q$ establishes the formula for $t(\varrho)$. Moreover, recalling that $\Theta(s) = \Phi(s) - \Phi'(s)s$, $s \in (0, \infty)$, we deduce from a simple application of integration by parts that

$$\int_p^q \frac{\Theta(\varrho s)}{s^2}\mathrm{d}s = \int_p^q \frac{\Phi(\varrho s)}{s^2}\mathrm{d}s - \varrho t(\varrho) = \frac{\Phi(\varrho p)}{p} - \frac{\Phi(\varrho q)}{q},$$

which yields the claim. □

We see from Theorem 4.2 that $q = \infty$ can be reached provided that $\frac{\Phi'(r)}{r}$ is integrable at ∞. In particular, in case of the modified logarithmic Sobolev inequality, Theorem 4.2 does not lead to ultracontractive bounds, which is consistent to the role of the logarithmic Sobolev inequality in the diffusive setting. Otherwise, for a growth function Φ that behaves as r^α with $0 < \alpha < 1$ at ∞ we have that $\frac{\Phi'(r)}{r}$ is integrable at ∞. In this sense, the modified logarithmic Sobolev inequality constitutes an extreme case.

The growth function resulting from $CD_\Upsilon(\kappa, F)$, with F being a power-type CD-function from Proposition 3.5, satisfies the integrability condition that we have mentioned throughout the previous lines. This fact can be seen from the identity (42). We close this section with an application of Theorem 4.2 in this particular context.

Corollary 4.3. *Let $M_1 \in \ell^2(\mu)$, $M_2 \in \ell^1(\mu)$ and the Markov chain generated by L be positive recurrent. Further, let L satisfy $CD_\Upsilon(\kappa, F)$ with $\kappa > 0$ and $F(r) = \frac{r^{1+\delta}}{n}$, $r \geqslant 0$, for some $\delta \geqslant 1$ and $n \in (0, \infty)$. Then we have for any $t > 0$ that*

$$\|e^{P_t f}\|_\infty \leqslant e^{\Phi\left(\sqrt[\delta]{\frac{n}{2\delta t}}\right)}\|e^f\|_1 \tag{49}$$

holds for any $f \in \ell^\infty(X)$, where Φ denotes the growth function given by (41). In particular, in case of the $CD_\Upsilon(\kappa, n)$ condition, (49) reads as

$$\|e^{P_t f}\|_\infty \leqslant \left(1 + \frac{1}{2\kappa t}\right)^{\frac{n}{2}}\|e^f\|_1. \tag{50}$$

Proof. Due to Proposition 3.5, L satisfies $EI(\Phi)$ with growth function given by (41). We choose $p = 1$ and $q = \infty$ (in the limit sense) in (47), recall the formula (42) for the derivative of the growth function and observe for $\varrho > 0$

$$t(\varrho) = \frac{n}{2}\int_1^\infty \frac{1}{r(\kappa n + (\varrho r)^\delta)}\mathrm{d}r \leqslant \frac{n}{2\varrho^\delta}\int_1^\infty \frac{1}{r^{1+\delta}}\mathrm{d}r = \frac{n}{2\delta\varrho^\delta}.$$

From this we infer by monotonicity of the growth function that

$$m(\varrho) = \Phi(\varrho) \leqslant \Phi\Big(\sqrt[\delta]{\frac{n}{2\delta t(\varrho)}}\Big).$$

Consequently, by Theorem 4.2 we get that

$$\|e^{P_{t(\varrho)}f}\|_\infty \leqslant e^{\Phi\left(\sqrt[\delta]{\frac{n}{2\delta t(\varrho)}}\right)}\|e^f\|_1$$

holds for any $\varrho > 0$. But as the mapping $\varrho \mapsto t(\varrho)$, $\varrho > 0$, is bijective onto $(0,\infty)$ since it is decreasing with $t(\varrho) \to 0$ as $\varrho \to \infty$ and $t(\varrho) \to \infty$ as $\varrho \to 0$, (49) follows. The special case of (50) now can be established by the explicit formula for the growth function in the case of $\delta = 1$, see Corollary 3.7. □

5. Exponential Integrability of Lipschitz functions and Diameter bounds

Exponential integrability of Lipschitz functions and diameter bounds (see the definitions below) are both important properties to investigate in the classical theory of [3]. In order to reach finite diameter bounds in the diffusive setting, Poincaré inequalities resp. logarithmic Sobolev inequalities are not sufficient. Instead, Sobolev inequalities resp. certain entropy-energy inequalities ensure the validity of a finite diameter. Speaking on the level of CD-inequalities this means that in the diffusive setting positive curvature and finite dimension suffices to deduce finite diameter bounds, while positive curvature alone does not. In this section we will be able to show that the CD_Υ condition behaves consistently in the discrete setting of Markov chains.

Now, we recall the definitions of Lipschitz functions and the diameter.

Definition 5.1. *A function $f \in \mathbb{R}^X$ is called Lipschitz function if $\Gamma(f)(x)$ exists at any $x \in X$ (in the sense that the sum in (11) with $H(r) = \frac{r^2}{2}$ is finite) and $\|f\|_{\mathrm{Lip}} := \sqrt{\|\Gamma(f)\|_\infty} < \infty$. Moreover, we say that f is C-Lipschitz, where $C > 0$, when $\|f\|_{\mathrm{Lip}} \leqslant C$ holds true.*

Definition 5.2. *Considering the mapping $\varrho : X \times X \to [0,\infty)$, given by*

$$\varrho(x,y) = \sup\{f(y) - f(x) : \|f\|_{Lip} \leqslant 1\},$$

we define the diameter with respect to L as $\mathrm{diam}_\varrho = \sup_{x,y\in X} \varrho(x,y)$.

These definitions have been used in the diffusive situation of [3], but also in the discrete setting in [18], where diam_ϱ has been called the resistance diameter. Definition 5.2 is further closely related to the diameter with respect to the combinatorical graph distance on the underlying graph to L. In fact, in [18] it has been shown that the estimate

$$\mathrm{dist}(x,y) \leqslant \sqrt{\frac{M_{1,\sup}}{2}}\varrho(x,y) \tag{51}$$

holds true on locally finite graphs, provided that $M_{1,\sup} < \infty$, where $\mathrm{dist} : X \times X \to [0,\infty)$ denotes the combinatorical graph distance. We emphasize that the bound (51) extends to the case of locally infinite graphs with $M_{1,\sup} < \infty$ as the proof of [18, Lemma 1.4] holds verbatim.

Lemma 5.3. *Let $f \in \mathbb{R}^X$ be a bounded C-Lipschitz function. Then we have that*

$$\mathcal{I}(e^{sf}) \leqslant C^2 s^2 \int_X e^{sf}\mathrm{d}\mu.$$

holds true for any $s \in \mathbb{R}$.

Proof. Clearly, there is nothing to show in the case of $s = 0$. Note that the detailed balance condition implies that

$$k(x,y)s\big(f(y)-f(x)\big)\big(h(sf(y))-h(sf(x))\big)\pi(x) = k(y,x)s\big(f(x)-f(y)\big)\big(h(sf(x))-h(sf(y))\big)\pi(y)$$

holds for any $x, y \in X$ and $s \neq 0$, where $h : \mathbb{R} \to \mathbb{R}$ is some arbitrary function. Defining for $x \in X$ and $s \neq 0$ the set $A_{x,s} := \{y \in X : s(f(y) - f(x)) < 0\}$, we then infer

$$\begin{aligned}
C^2 s^2 \int_X e^{sf}\mathrm{d}\mu &\geqslant s^2 \int_X e^{sf}\Gamma(f)\mathrm{d}\mu \\
&= s^2 \sum_{x\in X} e^{sf(x)} \sum_{y\in A_{x,s}} k(x,y)\big(f(y)-f(x)\big)^2\pi(x) \\
&\geqslant \sum_{x\in X} e^{sf(x)} \sum_{y\in A_{x,s}} k(x,y)|s(f(y)-f(x))|\,|e^{s(f(y)-f(x))}-1|\,\pi(x) \\
&= \sum_{x\in X} e^{sf(x)} \sum_{y\in A_{x,s}} k(x,y)s\big(f(y)-f(x)\big)\big(e^{s(f(y)-f(x))}-1\big)\pi(x) \\
&= \sum_{x\in X}\sum_{y\in A_{x,s}} k(x,y)s\big(f(y)-f(x)\big)\big(e^{sf(y)}-e^{sf(x)}\big)\pi(x) \\
&= \mathcal{I}(e^{sf}).
\end{aligned}$$

Here, we have applied the inequality $|\tau| \geqslant |e^\tau - 1|$, which is valid if $\tau \leqslant 0$. □

We now state the main result of this section.

Theorem 5.4. *Let L satisfy $EI(\Phi)$ and let $f \in \ell^1(\mu)$ be 1-Lipschitz. Further, we assume that $s \mapsto \frac{\Phi(s^2)}{s^2}$ is integrable at 0. Then we have*

$$\int_X e^{tf}\mathrm{d}\mu \leqslant \exp\Big(t\Big(\int_0^t \frac{\Phi(s^2)}{s^2}\mathrm{d}s + \int_X f\mathrm{d}\mu\Big)\Big),\ t \neq 0. \tag{52}$$

If in addition $s \mapsto \frac{\Phi(s^2)}{s^2}$ is also integrable at ∞, then any Lipschitz function is bounded and it holds

$$\|f - \int_X f\mathrm{d}\mu\|_\infty \leqslant \int_0^\infty \frac{\Phi(s^2)}{s^2}\mathrm{d}s \tag{53}$$

for any 1-Lipschitz function f.
In particular, the diameter bound

$$\mathrm{diam}_\varrho \leqslant 2\int_0^\infty \frac{\Phi(s^2)}{s^2}\mathrm{d}s \tag{54}$$

is valid.

Proof. First, let $f \in \mathbb{R}^X$ be 1-Lipschitz and bounded. In particular, by Lemma 5.3 we have that $\mathcal{I}(e^{sf}) < \infty$ for any $s \in \mathbb{R}$. We set $Z(s) = \int_X e^{sf}\mathrm{d}\mu$ and observe that $Z'(s) = \int_X f e^{sf}\mathrm{d}\mu$. Further, we define $\Lambda(s) = \frac{\log(Z(s))}{s}$, $s \neq 0$. Note that we can extend Λ to a continuous function on $\mathbb{R}$ by means of L'Hospital's rule by $\Lambda(0) = \int_X f\mathrm{d}\mu$. For some fixed $s \neq 0$ we obtain

$$\Lambda'(s) = \frac{sZ'(s) - Z(s)\log(Z(s))}{s^2 Z(s)} = \frac{\mathrm{Ent}_\mu(e^{sf})}{s^2 Z(s)}.$$

Now, we apply $EI(\Phi)$ (in the form of (34)) to e^{sf} and observe that

$$\Lambda'(s) \leqslant \frac{\Phi'(r)\mathcal{I}(e^{sf}) + \Theta(r)Z(s)}{s^2 Z(s)} \tag{55}$$

holds for any $r \in (0,\infty)$. Due to f being bounded and 1-Lipschitz, Lemma 5.3 translates to

$$\mathcal{I}(e^{sf}) \leqslant s^2 Z(s).$$

With this at hand, we deduce from (55) that

$$\Lambda'(s) \leqslant \frac{\Phi'(r)s^2 + \Theta(r)}{s^2} = \frac{\Phi(r) + \Phi'(r)(s^2 - r)}{s^2}$$

holds for any $r \in (0,\infty)$. Specifying $r = s^2$, we end up with

$$\Lambda'(s) \leqslant \frac{\Phi(s^2)}{s^2}.$$

Integrating this yields

$$\Lambda(t) \leqslant \Lambda(0) + \int_0^t \frac{\Phi(s^2)}{s^2}\mathrm{d}s,$$

when $t > 0$ and with the reverse inequality in case that $t < 0$. In both situations, (52) follows from the definition of Λ in the case that f is bounded.

Regarding the general case, let $f \in \ell^1(\mu)$ be 1-Lipschitz and consider $f_N \in \ell^\infty(X)$, $N \in \mathbb{N}$, given by $f_N(x) = f(x)$ if $|f(x)| \leqslant N$, $f_N(x) = N$ if $f(x) > N$ and $f_N(x) = -N$ if $f(x) < -N$. Clearly, for any $x, y \in X$, we have $|f_N(x) - f_N(y)| \leqslant |f(x) - f(y)|$ and thus f_N is 1-Lipschitz for any $N \in \mathbb{N}$. As $f \in \ell^1(\mu)$, $\int_X f_N \mathrm{d}\mu$ converges to $\int_X f\mathrm{d}\mu$ as $N \to \infty$ by the dominated convergence theorem. Furthermore, Fatou's lemma implies

$$\begin{aligned}\int_X e^{sf}\mathrm{d}\mu \leqslant \liminf_{N\to\infty} \int_X e^{sf_N}\mathrm{d}\mu &\leqslant \liminf_{N\to\infty} \exp\Big(t\Big(\int_0^t \frac{\Phi(s^2)}{s^2}\mathrm{d}s + \int_X f_N \mathrm{d}\mu\Big)\Big) \\ &= \exp\Big(t\Big(\int_0^t \frac{\Phi(s^2)}{s^2}\mathrm{d}s + \int_X f\mathrm{d}\mu\Big)\Big).\end{aligned}$$

As a first step to establish the second claim, we note that if (53) is valid for bounded 1-Lipschitz functions then any 1-Lipschitz function is an element of $\ell^1(\mu)$. Indeed, let f be a general 1-Lipschitz function and denote by $\big(f_N\big)_{N\in\mathbb{N}} \subset \ell^\infty(X)$ the approximating sequence of 1-Lipschitz functions as defined above. Then, choosing $x \in X$ arbitrary, we observe, since clearly $|f|$ is also 1-Lipschitz, that

$$\big||f(x)| - \int_X |f_N|\mathrm{d}\mu\,\big| \leqslant \big\||f_N| - \int_X |f_N|\mathrm{d}\mu\,\big\|_\infty \leqslant \int_0^\infty \frac{\Phi(s^2)}{s^2}\mathrm{d}s,$$

where $N \geqslant N_0$ with $\mathbb{N}_0 \in \mathbb{N}$ such that $|f(x)| \leqslant N_0$. Sending $N \to \infty$ in the latter estimation, we deduce that $f \in \ell^1(\mu)$.

We will now show that (53) holds for any 1-Lipschitz function that is an element of $\ell^1(\mu)$ (and hence for any 1-Lipschitz function). We set $C = \int_0^\infty \frac{\Phi(s^2)}{s^2}\mathrm{d}s$ and apply (52) to $f - \int_X f\mathrm{d}\mu$, assuming that $f \in \ell^1(\mu)$ is 1-Lipschitz. We obtain for any $t > 0$ that

$$\int_X \exp\Big(t\big(f - \int_X f\mathrm{d}\mu\big)\Big)\mathrm{d}\mu \leqslant \exp\big(tC\big). \tag{56}$$

Now, we assume for contradiction that we can find some $x \in X$ and some $\varepsilon > 0$ such that

$$f(x) - \int_X f\mathrm{d}\mu > C + \varepsilon.$$

Then we have for $t > 0$

$$\int_X \exp\Big(t\big(f - \int_X f\mathrm{d}\mu\big)\Big)\mathrm{d}\mu > e^{t(C+\varepsilon)}\pi(x),$$

which contradicts (56) in the asymptotic behavior of $t \to \infty$. Simultaneously, by considering the asymptotic behavior as $t \to -\infty$, one obtains that $f(x) - \int_X f \mathrm{d}\mu \geqslant -C$ holds for all $x \in X$. This establishes (53) for any 1-Lipschitz function. From that we conclude

$$f(y) - f(x) \leqslant 2\|f - \int_X f \mathrm{d}\mu\|_\infty \leqslant 2 \int_0^\infty \frac{\Phi(s^2)}{s^2} \mathrm{d}s. \tag{57}$$

for any $x, y \in X$ and 1-Lipschitz function f. But (57) implies that f must be bounded. Hence we deduce by scaling that any Lipschitz function is bounded. Furthermore, by (57) and the definition of ϱ, we deduce (54). □

Theorem 5.4 yields finite bounds on diam_ϱ if L satisfies $CD_\Upsilon(\kappa, F)$ with $\kappa > 0$ and a power-type CD-function by means of Proposition 3.5(ii). In the special case of the quadratic CD-function, we get the following bound.

Corollary 5.5. *If $M_1 \in \ell^2(\mu)$, $M_2 \in \ell^1(\mu)$, the Markov chain generated by L is positive recurrent and L satisfies $CD_\Upsilon(\kappa, n)$ with $\kappa > 0$ and $n < \infty$, then the diameter bound*

$$\mathrm{diam}_\varrho \leqslant \pi \sqrt{\frac{n}{\kappa}} \tag{58}$$

holds true.

Proof. By elementary methods one calculates the integral

$$\frac{n}{2} \int_0^\infty \frac{\log\left(1 + \frac{s^2}{\kappa n}\right)}{s^2} \mathrm{d}s = \frac{\pi}{2} \sqrt{\frac{n}{\kappa}}.$$

The claim follows by combining Corollary 3.7 with Theorem 5.4. □

Remark 5.6. By (58) we recover (by different methods) exactly the same diameter bound as in [18], where there it is assumed on the one hand only $CD(\kappa, n)$ but on the other hand that the underlying graph to L is locally finite and satisfies the completeness assumption and non-degeneracy of the vertex measure. Note that by [19, Theorem 2.2] the latter boils down to the case of finite graphs since $\kappa > 0$. Hence, although the curvature-dimension condition of Corollary 5.5 is more restrictive, the setting where it applies can be expected to be more general compared to the one of [18].

With the following example we aim to emphasize that the property that Lipschitz functions are bounded is quite strong in the sense that it fails for a large class of examples that all satisfy corresponding modified logarithmic Sobolev inequalities. In particular, it turns out that $CD_\Upsilon(\kappa, \infty)$, with $\kappa > 0$, is not sufficient for deducing a finite diameter bound.

Example 5.7. We consider a birth-death process on $\mathbb{N}_0$ as in Example 2.11 and employ the notation that has been used therein. In particular, we assume that the rate functions a and b are monotone as in Example 2.11 and that condition (24) holds for some $\kappa > 0$. We set $f(0) = 0$ and define the sequence of partial sums

$$f(n) = \sum_{k=1}^{n} \frac{1}{\sqrt{b(k)}}, \; n \in \mathbb{N}. \tag{59}$$

We claim that Lipschitz functions are bounded if and only if the partial sums given by (59) converge.

First, we assume that the partial sums given by (59) diverge as $n \to \infty$. Then, we have

$$2\Gamma(f)(n) = a(n)(f(n+1) - f(n))^2 + b(n)(f(n-1) - f(n))^2 = \frac{a(n)}{b(n+1)} + 1$$

for any $n \in \mathbb{N}$. From the monotonicity assumption on the rates we infer that $\Gamma(f)$ is bounded, or in other words that f is a Lipschitz function. But apparently, f is unbounded. This yields that the corresponding generator does not satisfy an entropy-information inequality with growth function Φ such that $\int_0^\infty \frac{\Phi(s^2)}{s^2} \mathrm{d}s < \infty$. On the other hand, it is known by [6] that L satisfies a corresponding modified logarithmic Sobolev inequality. Moreover, we emphasize that among those birth-death processes of the present example are also processes that even satisfy the condition $CD_\Upsilon(\kappa, \infty)$, see [31]. This shows that the condition $CD_\Upsilon(\kappa, \infty)$ with $\kappa > 0$ is in general not sufficient to obtain a finite diameter. The latter finding is consistent to the Bakry-Émery condition in the diffusive setting.

Now, let us assume conversely that the partial sums given by (59) converge and let g be C-Lipschitz for some $C > 0$. In particular,

$$b(n)\big(g(n-1) - g(n)\big)^2 \leqslant 2C^2$$

holds for any $n \in \mathbb{N}$. Consequently, we have $|g(n-1) - g(n)| \leqslant \frac{\sqrt{2}C}{\sqrt{b(n)}}$, $n \in \mathbb{N}$, and by the triangle inequality we deduce

$$|g(N) - g(0)| \leqslant \sqrt{2}Cf(N) \leqslant \sqrt{2}C \sum_{n \in \mathbb{N}} \frac{1}{\sqrt{b(n)}}$$

for any $N \in \mathbb{N}$, which yields that g is bounded.

6. Modified Nash inequalities

In the diffusive setting, it is known that logarithmic entropy-energy inequalities imply Nash inequalities (cf. [3, Proposition 6.2.3]). Regarding the discrete setting of Markov chains we refer to [10], and also [27], for an extensive account on Nash inequalities. Clearly, it can not be expected that logarithmic entropy-information inequalities are linked to the classical Nash inequality in the discrete setting as they are in the diffusive setting by the lack of chain rule. We refer to the natural analogue as the modified Nash inequality, which can be induced from corresponding logarithmic entropy-information inequality as will be shown subsequently. We say that a function $f \in \mathbb{R}^X$ is non-vanishing if $f(x) \neq 0$ for any $x \in X$.

Theorem 6.1. *If L satisfies $EI(\Phi)$ with $\Phi(r) = \alpha \log\left(A + \frac{r}{\beta}\right)$ and $\alpha, \beta > 0$, $A \geqslant 1$, then the following modified Nash inequality*

$$\|f\|_2^{2\alpha+2} \leqslant \Big(A\|f\|_2^2 + \frac{\mathcal{I}(f^2)}{\beta}\Big)^\alpha \|f\|_1^2 \tag{60}$$

holds for any non-vanishing $f \in \ell^2(\mu)$.

Proof. Clearly, we can assume that $\mathcal{I}(f^2) < \infty$. Further, it suffices to prove (60) for bounded non-vanishing functions by a standard truncation argument. Indeed, let $(f_N)_{N \in \mathbb{N}}$ denote the sequence of bounded functions that has been considered in the proof of Theorem 5.4. Then it follows readily by means of the monotone convergence theorem that $\|f_N\|_2 \to \|f\|_2$, $\|f_N\|_1 \to \|f\|_1$ and $\mathcal{I}(f_N^2) \to \mathcal{I}(f^2)$ as $N \to \infty$.

It is a well known consequence of Hölder's inequality that the mapping $r \mapsto \|f\|_{\frac{1}{r}}$, $r \in (0, 1]$ is log-convex. Then, for a given non-vanishing $f \in \ell^\infty(X)$ with $\|f\|_2 = 1$, we consider the convex

mapping $\Lambda(r) = \log \|f\|_{\frac{1}{r}}$, $r \in (0,1]$, which is well defined since μ is a probability measure. One readily verifies by a similar calculation as in the proof of Lemma 4.1 that

$$\Lambda'(r) = -\frac{\mathrm{Ent}_\mu(|f|^{\frac{1}{r}})}{\int_X |f|^{\frac{1}{r}} \mathrm{d}\mu},$$

where we use that $|f|$ is bounded in order to interchange differentation and integration. In particular, using $\|f\|_2 = 1$, we observe that $\Lambda'\big(\frac{1}{2}\big) = -\mathrm{Ent}_\mu(f^2)$. By convexity, we thus have

$$2\big(\Lambda(1) - \Lambda\big(\frac{1}{2}\big)\big) \geqslant \Lambda'\big(\frac{1}{2}\big).$$

Consequently, by the entropy-information inequality $EI(\Phi)$ and the fact that $\Lambda\big(\frac{1}{2}\big) = 0$, we observe

$$\log \frac{1}{\|f\|_1^2} \leqslant \mathrm{Ent}_\mu(f^2) \leqslant \log\Big(A + \frac{\mathcal{I}(f^2)}{\beta}\Big)^\alpha,$$

which implies

$$1 \leqslant \big(A + \frac{\mathcal{I}(f^2)}{\beta}\big)^\alpha \|f\|_1^2 \,. \tag{61}$$

Now, for the non-normalized case we apply (61) to $\frac{f}{\|f\|_2}$. By the scaling behavior of the Fisher information (cf. (30)), we deduce

$$\|f\|_2^2 \leqslant \frac{1}{\|f\|_2^{2\alpha}}\Big(A\|f\|_2^2 + \frac{\mathcal{I}(f^2)}{\beta}\Big)^\alpha \|f\|_1^2,$$

from which the claim follows. □

Combining Corollary 3.7 with Theorem 6.1, we observe the following result.

Corollary 6.2. *If $M_1 \in \ell^2(\mu)$, $M_2 \in \ell^1(\mu)$, the Markov chain generated by L is positive recurrent and L satisfies $CD_\Upsilon(\kappa, n)$ with $\kappa > 0$ and $n < \infty$, then L satisfies the following modified Nash inequality*

$$\|f\|_2^{n+2} \leqslant \Big(\|f\|_2^2 + \frac{\mathcal{I}(f^2)}{\kappa n}\Big)^{\frac{n}{2}} \|f\|_1^2$$

for any non-vanishing $f \in \ell^2(\mu)$.

Appendix A. Auxiliary Lemma

In this section we provide an auxiliary result, which has been used to investigate the example of the two-point space in Example 2.5. For further properties of the functions $\nu_{c,d} : \mathbb{R} \to \mathbb{R}$, given by

$$\nu_{c,d}(r) = c\Upsilon'(r)r + \Upsilon(-r) - d\Upsilon(r),\ c, d \in \mathbb{R},$$

which also have been used throughout Section 2, we refer to the Appendix of [31].

Lemma A.1. *$\nu_{1+\lambda,\lambda}$ is strictly convex for any $\lambda \in (0,1]$.*

Proof. We have

$$\nu''_{1+\lambda,\lambda}(r) = e^r\big((1+\lambda)r + 2 + \lambda\big) + e^{-r},$$
$$\nu'''_{1+\lambda,\lambda}(r) = e^r\big((1+\lambda)r + 3 + 2\lambda\big) - e^{-r}.$$

Hence, $\nu'''_{1+\lambda,\lambda}(r) = 0$ if and only if

$$e^{2r}\big((1+\lambda)r + 3 + 2\lambda\big) = 1, \tag{62}$$

which yields by monotonicity that $\nu''_{1+\lambda,\lambda}$ has a unique critical point r_* for which we have $r_* > -\frac{3+2\lambda}{1+\lambda}$. By applying the logarithm we can reformulate (62) into

$$r_* = \log\big(\frac{1}{\sqrt{(1+\lambda)r_* + 3 + 2\lambda}}\big). \tag{63}$$

We infer that

$$\nu''_{1+\lambda,\lambda}(r_*) = \frac{(1+\lambda)r_* + 2 + \lambda}{\sqrt{(1+\lambda)r_* + 3 + 2\lambda}} + \sqrt{(1+\lambda)r_* + 3 + 2\lambda}\,,$$

which is positive if and only if

$$r_* > -\frac{5+3\lambda}{2(1+\lambda)}.$$

Defining $\bar{r} := -\frac{5+3\lambda}{2(1+\lambda)}$, we now observe that

$$e^{2\bar{r}} < 1 \leqslant \frac{2}{\lambda+1} = \frac{1}{(1+\lambda)\bar{r} + 3 + 2\lambda}.$$

In particular, this implies that $\nu'''_{1+\lambda,\lambda}(\bar{r}) < 0$ and hence we have $r_* > \bar{r}$ by monotonicity, which establishes the claim. □

References

[1] Anderson, W. J.: *Continuous-Time Markov Chains*, Springer Series in Statistics, 1991.

[2] Bakry, D.; Émery, M.: *Diffusions hypercontractives.* Séminaire de probabilités, XIX, 1983/84, 177–206, Lecture Notes in Math., Springer, Berlin, 1985.

[3] Bakry, D.; Gentil, I.; Ledoux, M.: *Analysis and geometry of Markov diffusion operators*. Springer, 2014.

[4] Bauer, F.; Horn, P.; Lin, Y.; Lippner, G.; Mangoubi, D.; Yau, S.-T.: Li-Yau inequality on graphs. J. Differential Geom. **99** (2015), 359–405.

[5] Bobkov, S. G.; Tetali, P.: Modified logarithmic Sobolev inequalities in discrete settings. J. Theoret. Probab. **19** (2006), 289–336.

[6] Caputo, P.; Dai Pra, P.; Posta, G.: Convex entropy decay via the Bochner-Bakry-Émery approach. Ann. Inst. Henri Poincaré Probab. Stat. **45** (2009), 734–753.

[7] Cushing, D.; Kamtue, S.; Kangaslampi, R.; Liu, S.; Peyerimhoff, N.: Curvature, Graph Products and Ricci Flatness, Preprint 2019, available at arXiv.

[8] Cushing, D.; Liu, S.; Peyerimhoff, N.: Bakry-Émery curvature functions on graphs. Canad. J. Math. **72** (2020), 89–143.

[9] Diaconis, P.; Saloff-Coste, L.: Logarithmic Sobolev inequalities for finite Markov chains. Ann. Appl. Probab. **6** (1996), 695–750.

[10] Diaconis, P.; Saloff-Coste, L.: Nash inequalities for finite Markov chains. J. Theoret. Probab. **9** (1996), 459–510.

[11] Dier, D.; Kassmann, M.; Zacher, R.: Discrete versions of the Li-Yau gradient estimate. To appear in Ann. Sc. Norm. Super. Pisa Cl. Sci. (5), DOI: 10.2422/2036-2145.201807-005.

[12] Erbar, M.; Fathi, M.: Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature. J. Funct. Anal. **274** (2018), 3056–3089.

[13] Erbar, M.; Maas, J.: Ricci curvature of finite Markov chains via convexity of the entropy. Arch. Ration. Mech. Anal. **206** (2012), 997–1038.

[14] Fathi, M.; Maas, J.: Entropic Ricci curvature bounds for discrete interacting systems. Ann. Appl. Probab. **26** (2016), 1774–1806.

[15] Jost, J.; Liu, S.: Ollivier's Ricci curvature, local clustering and curvature-dimension inequalities on graphs. Discrete Comput. Geom. **51** (2014), 300–322.

[16] Keller, M.; Lenz, D.: Dirichlet forms and stochastic completeness of graphs and subgraphs. J. Reine Angew. Math., **666** (2012), 189-223.

[17] Lin, Y.; Yau, S.-T.: Ricci curvature and eigenvalue estimate on locally finite graphs. Math. Res. Lett. **17** (2010), 343–356.

[18] Liu, S.; Münch, F.; Peyerimhoff, N.: Bakry-Émery curvature and diameter bounds on graphs. Calc. Var. Partial Differential Equations, **57** (2018), no. 2, Art. 67, 9 pp.

[19] Liu, S.: Buser's inequality on infinite graphs. J. Math. Anal. Appl. **475** (2019), no. 2, 1416–1426.

[20] Lott, J.; Villani, C.: Ricci curvature for metric-measure spaces via optimal transport. Ann. of Math. (2) **169** (2009), 903–991.
[21] Maas, J.: Gradient flows of the entropy for finite Markov chains. J. Funct. Anal. **261** (2011), 2250–2292.
[22] Münch, F.: Li-Yau inequality on finite graphs via non-linear curvature dimension conditions. J. Math. Pures Appl. (9) **120** (2018), 130–164.
[23] Münch, F.; Wojciechowski, R. K.: Ollivier Ricci curvature for general graph Laplacians: heat equation, Laplacian comparison, non-explosion and diameter bounds. Adv. Math. **356** (2019), 45 pp.
[24] Najman, L., Romon, P. (eds.): *Modern approaches to discrete curvature*. Lecture Notes in Mathematics. Springer, Cham, 2017.
[25] Norris, J.: *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.
[26] Ollivier, Y.: Ricci curvature of Markov chains on metric spaces. J. Funct. Anal. **256** (2009), 810–864.
[27] Saloff-Coste, L.: Lectures on finite Markov chains. Lectures on probability theory and statistics (Saint-Flour, 1996), 301–413.
[28] Spener, A.; Weber, F.; Zacher, R.: Curvature-dimension inequalities for non-local operators in the discrete setting. Calc. Var. Partial Diff. Equ. **58** (2019), no. 5, Art. 171, 30 pp.
[29] Sturm, K.-T.: On the geometry of metric measure spaces. I. Acta Math. **196** (2006), 65–131.
[30] Sturm, K.-T.: On the geometry of metric measure spaces. II. Acta Math. **196** (2006), 133–177.
[31] Weber, F.; Zacher, R.: The entropy method under curvature-dimension conditions in the spirit of Bakry-Émery in the discrete setting of Markov chains, Preprint 2020, available at arXiv.
[32] Villani, C.: *Optimal transport old and new*. Springer, 2009.

4. Curvature-dimension inequalities for non-local operators in the discrete setting

authors: Adrian Spener, Frederic Weber and Rico Zacher

Journal: Calculus of Variations and Partial Differential Equations **58** (2019), no. 5, Art. 171, 30 pp.

DOI: https://doi.org/10.1007/s00526-019-1616-y

Copyright: *Reprinted by permission from Springer Nature Customer Service Centre GmbH : Springer; Calculus of Variations and Partial Differential Equations; Curvature-dimension inequalities for non-local operators in the discrete setting; Adrian Spener, Frederic Weber, Rico Zacher ©(2019); permission conveyed through Copyright Clearance Center.*

Calc. Var. (2019) 58:171
https://doi.org/10.1007/s00526-019-1616-y

Calculus of Variations

Curvature-dimension inequalities for non-local operators in the discrete setting

Adrian Spener[1] · Frederic Weber[1] · Rico Zacher[1]

Received: 2 April 2019 / Accepted: 31 July 2019 / Published online: 13 September 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract
We study Bakry–Émery curvature-dimension inequalities for non-local operators on the one-dimensional lattice and prove that operators with finite second moment have finite dimension. Moreover, we show that a class of operators related to the fractional Laplacian fails to have finite dimension and establish both positive and negative results for operators with sparsely supported kernels. Furthermore, a large class of operators is shown to have no positive curvature. The results correspond to CD inequalities on locally infinite graphs.

Keywords Gamma calculus · Curvature-dimension inequality · Bakry-Émery inequality · Non-local operator · Fractional Laplacian · Markov chain · Infinite graphs

Mathematics Subject Classification Primary 47D07; Secondary 05C63 · 60G22 · 26A33

1 Introduction and main results

The main purpose of this paper is to study curvature-dimension (CD) inequalities for non-local operators $\mathcal{L}$ on the lattice $\mathbb{Z}$ of the form

$$\mathcal{L}v(x) = \sum_{j\in\mathbb{Z}} k(j)\big(v(x+j) - v(x)\big), \quad x \in \mathbb{Z}, \tag{1}$$

with a (nontrivial) kernel k which is nonnegative, integrable and symmetric, that is

(K1) $k : \mathbb{Z} \to [0,\infty)$, $\sum_{j\in\mathbb{Z}} k(j) < \infty$ and $k(-j) = k(j)$ for all $j \in \mathbb{N}$.

Communicated by J. Jost.

✉ Rico Zacher
rico.zacher@uni-ulm.de

Adrian Spener
adrian.spener@uni-ulm.de

Frederic Weber
frederic.weber@uni-ulm.de

[1] Institut für Angewandte Analysis, Universität Ulm, Helmholtzstraße 18, 89081 Ulm, Germany

Observe that the value of k at 0 does not play a role in the definition of $\mathcal{L}$. It is convenient to assume that

(K2) $k(0) = 0$.

An important example, which will be investigated in this paper, is given by

$$k(j) = \frac{c}{|j|^{1+\beta}}, \quad j \in \mathbb{Z}\backslash\{0\}, \tag{2}$$

where $c, \beta > 0$.

Curvature-dimension inequalities (or conditions) play a central role in the study of functional inequalities associated to Markov semigroups and operators. Important examples of such functional inequalities are the Poincaré or spectral gap inequality, the logarithmic Sobolev inequality and the Sobolev inequality, which, among others, allow to derive various estimates of solutions to related evolution equations, e.g. Harnack inequalities or bounds which imply the exponentially fast trend to an equilibrium. The special feature of CD-inequalities is that they provide a very useful link to the geometric properties (like dimension and curvature) of the underlying structure [3]. For this reason they also constitute an important tool in geometric analysis [15].

There are several different notions of CD-inequalities. Here we use the original one, which goes back to Bakry and Émery and is formulated in terms of the carré du champ operator Γ and the iterated carré du champ operator Γ_2 associated with the infinitesimal generator $\mathcal{L}$ of a Markov semigroup, see [2]. Another powerful approach is based on the theory of optimal transport and displacement convexity inequalities [17,24,25,27]. For Riemannian manifolds it provides, like the Bakry–Émery calculus, an equivalent definition of Ricci curvature lower bounds.

Let $\mathcal{L}$ be the generator of a symmetric Markov semigroup with invariant reversible measure μ on the state space E. The bilinear operators Γ and Γ_2 are defined by

$$\Gamma(u, v) = \frac{1}{2}\left(\mathcal{L}(uv) - u\mathcal{L}v - v\mathcal{L}u\right),$$
$$\Gamma_2(u, v) = \frac{1}{2}\left(\mathcal{L}\Gamma(u, v) - \Gamma(u, \mathcal{L}v) - \Gamma(\mathcal{L}u, v)\right)$$

on a suitable algebra $\mathcal{A}$ of real-valued functions u, v defined on the underlying state space E. Furthermore, one sets

$$\Gamma(u) := \Gamma(u, u) = \frac{1}{2}\mathcal{L}(u^2) - u\mathcal{L}u,$$
$$\Gamma_2(u) := \Gamma_2(u, u) = \frac{1}{2}\mathcal{L}\Gamma(u) - \Gamma(u, \mathcal{L}u).$$

Let $\kappa \in \mathbb{R}$ and $d \in (0, \infty]$. We say that $\mathcal{L}$ satisfies the *Bakry–Émery curvature-dimension inequality* $CD(\kappa, d)$ with dimension d and curvature κ (lower bound) at $x \in E$ if

$$\Gamma_2(u)(x) \geq \frac{1}{d}\big((\mathcal{L}u)(x)\big)^2 + \kappa\Gamma(u)(x) \tag{3}$$

for all functions $u : E \to \mathbb{R}$ in a sufficiently rich class $\mathcal{A}$ of functions. We further say that $\mathcal{L}$ satisfies the $CD(\kappa, d)$-inequality, if (3) holds μ-almost everywhere for all $u \in \mathcal{A}$, cf. [3, Sect. 1.16].

As an illustrating example we let $E = (M, g)$ be a Riemannian manifold with canonical Riemannian measure μ_g and $\mathcal{L} = \Delta_g$ the Laplace–Beltrami operator. Using the Bochner–Lichnerowicz formula one obtains that $CD(\kappa, d)$ is equivalent to $\mathrm{Ric}_g(x) \geq \kappa g(x)$ and $\dim M \leq d$.

In our case, that is, the operator $\mathcal{L}$ is given by (1), we have a countable Markov chain with state space $E = \mathbb{Z}$ and $\mathcal{L}$ is a Markov generator, which can be also written as

$$\mathcal{L}u(x) = \sum_{y\in\mathbb{Z}} l(x, y)u(y), \quad x \in \mathbb{Z},$$

where

$$l(x, y) = |k|_1\big(p(x, y) - \delta(x, y)\big), \quad p(x, y) = \frac{1}{|k|_1}\, k(x - y), \quad x, y \in \mathbb{Z},$$

and $\delta(x, y) = 1$ if $x = y$ and $\delta(x, y) = 0$ otherwise. The infinite matrix $(p(x, y))_{(x,y)\in\mathbb{Z}^2}$ represents the transition probabilities of the Markov chain; $p(x, y)$ is the probability to jump from x to y in the next time step. The counting measure on $\mathbb{Z}$ plays the role of the invariant reversible measure μ. A possible choice for the algebra $\mathcal{A}$ is the space of all bounded functions $l_\infty(\mathbb{Z})$. A straight-forward computation shows that

$$\Gamma(u)(x) = \frac{1}{2}\sum_{j\in\mathbb{Z}} k(j)\big(u(x + j) - u(x)\big)^2, \quad x \in \mathbb{Z}, \tag{4}$$

$$\Gamma_2(u)(x) = \frac{1}{4}\sum_{j,l\in\mathbb{Z}} k(j)k(l)\big(u(x + j + l) - u(x + j) - u(x + l) + u(x)\big)^2, \quad x \in \mathbb{Z}, \tag{5}$$

in particular $\Gamma_2(u) \geq 0$, which implies that the $CD(0, \infty)$-inequality is always true.

The main objective of this paper is to analyse if and under which conditions the operator $\mathcal{L}$ from (1) satisfies $CD(\kappa, d)$, the main focus lying on $CD(0, d)$-conditions with finite dimension $d > 0$. Throughout this paper, we assume that the kernel k in (1) is subject to the conditions (K1) and (K2) from above and that $|k|_1 > 0$. We are especially interested in kernels with *unbounded support* as it is the case, e.g., for the algebraic (or power type) kernel k given by (2). In this situation, arbitrary long jumps are possible. The graphs underlying the Markov chain are *not locally finite*, in contrast to related known results in the literature.

Let us explain the connection to graph theory in more detail. Consider the undirected, weighted graph $\mathcal{G}$ given by the vertices $V = \mathbb{Z}$ with edge weight $\omega(x, y) = k(x - y)$, $x, y \in V$. Setting the weight on the vertices to be constant we obtain (1) for the graph Laplacian on $\mathcal{G}$. Note that $\mathcal{G}$ is locally infinite if and only if k has unbounded support. Curvature-dimension inequalites in the discrete setting, in particular for (locally) finite graphs are studied intensively [14,16,21,22]. Various examples of CD-inequalities in the sense of (3) for finite graphs are given in [1,13]. Some results on locally finite graphs from [12,18] can be adapted to the locally infinite case (see Proposition 2.1). Related notions of curvature-dimension conditions on graphs, e.g. so-called exponential curvature-dimension inequalities, are examined in [4,6,20,23]. Curvature-dimension inequalities in the context of optimal transport in the discrete setting have been studied in [7–11,19].

We now describe the main results of this paper. The first basic observation is that for any kernel with *finite support* the corresponding operator $\mathcal{L}$ satisfies $CD(0, 2N)$, where $2N$ equals the cardinality of the support, see Theorem 2.1. We also prove that the constant $2N$ in the CD-inequality is in general the best one can get.

Turning to the case of unbounded support, we are able to show that for the *power type kernel* $k(j) = c|j|^{-(1+\beta)}$ with $c, \beta > 0$ the value $\beta = 2$ is a critical case. For all $\beta \in (0, 2)$ the $CD(0, d)$-condition fails to hold for all finite $d > 0$ (see Theorem 3.1), whereas for any $\beta \in (2, \infty)$ there exists a finite $d > 0$ such that $CD(0, d)$ is valid, cf. Theorem 4.1. In

the proof of Theorem 3.1, we consider the family of *unbounded* functions $u_\varepsilon(j) = |j|^{\beta-\varepsilon}$ ($\varepsilon > 0$) and prove that $\Gamma_2(u_\varepsilon)(0)/(\mathcal{L}(u_\varepsilon)(0))^2 \to 0$ as $\varepsilon \to 0$. One may ask whether a $CD(0,d)$-condition with $d \in (0,\infty)$ is still true on a smaller class of admissible functions such as $l_\infty(\mathbb{Z})$ or the set of compactly supported functions. This is not the case as we show by means of an appropriate family of functions with compact support, see Theorem 3.2. The basic idea of the construction of these functions is to use the family $(u_\varepsilon)_{\varepsilon>0}$ from before and carefully chosen cut-off functions, a crude cut-off of $(u_\varepsilon)_{\varepsilon>0}$ does not seem to work. The proof is rather technical as the estimate of the Γ_2-term requires distinguishing of several cases due to the double series.

An important consequence of the negative result for $\beta < 2$ is that for all *fractional powers* $\mathcal{L} = -(-\Delta)^{\beta/2}$ of the *discrete Laplacian* Δ on $\mathbb{Z}$ with $\beta \in (0,2)$ the $CD(0,d)$-inequality fails to hold as well for all finite $d > 0$. Here, the discrete Laplacian (which satisfies $CD(0,2)$, see Theorem 2.1) is given by

$$\Delta u(x) = u(x+1) - 2u(x) + u(x-1), \quad x \in \mathbb{Z}, \tag{6}$$

and the fractional powers of the discrete Laplacian can be defined by means of the semigroup method (see [5, (1.3)]) as

$$-(-\Delta)^{\frac{\beta}{2}} u(x) = \frac{1}{\Gamma(-\frac{\beta}{2})} \int_0^\infty \left(u - e^{t\Delta}u\right)(x) \frac{dt}{t^{1+\frac{\beta}{2}}}, \quad x \in \mathbb{Z}.$$

In a very recent work [26], we are able to adapt the discrete estimates and the construction of the counterexamples from the case $\beta < 2$ to the *continuous case*. Employing additional techniques to extend the results to the multi-dimensional case we can show that the *fractional Laplacian on* $\mathbb{R}^N$ with any $N \in \mathbb{N}$ fails to satisfy the $CD(0,d)$-condition for all finite $d > 0$, even in the case where only compactly supported C^∞-functions are admissible.

The main positive result, Theorem 4.1, is not only formulated for purely algebraic kernels with $\beta > 2$ but for a much larger class of sufficiently fast decaying kernels. Besides monotonicity on $\mathbb{N}$ the central assumption is that the kernel k has a *finite second moment*, that is,

$$\sum_{j\in\mathbb{N}} k(j) j^2 < \infty.$$

This class also includes, for example, all kernels of the form

$$k(j) = c\,\frac{e^{-\delta|j|^\alpha}}{|j|^\gamma}, \quad j \in \mathbb{Z}\backslash\{0\}, \tag{7}$$

where $c, \alpha, \delta > 0$ and $\gamma \geq 0$, in particular exponential kernels. We also show that the statement of Theorem 4.1 remains true if the kernel k has a finite second moment and is merely assumed to be non-increasing for all sufficiently large $j \in \mathbb{N}$, see Remark 4.2. This allows to cover also kernels, e.g., of the form (7) where $c, \alpha, \delta > 0$ and $\gamma < 0$.

It is still an interesting open problem whether $CD(0,d)$ holds for some finite $d > 0$ in case of the algebraic kernel with critical value $\beta = 2$, that is, $k(j) = c|j|^{-3}$, $j \in \mathbb{Z}\backslash\{0\}$, with $c > 0$. We can immediately deduce from Theorem 4.1 that this is indeed the case if the kernel has an additional logarithmic factor, more precisely, if k is of the form

$$k(j) = \frac{c}{|j|^3[1+\log(|j|)]^{1+\alpha}}, \quad j \in \mathbb{Z}\backslash\{0\},$$

with $c, \alpha > 0$.

Another striking phenomenon is that all kernels with unbounded support, even those which are bounded above by an exponential kernel, fail the $CD(0,d)$-condition for all finite $d > 0$ if the support enjoys a certain number theoretic property, which in some sense means that *the support is sufficiently thin*. The latter holds in particular if the gaps between the (positive) elements of the support grow faster than the sequence of the powers of three, see Theorem 5.3. Somewhat surprisingly, if the support is precisely the set $\{\pm 3^l : l \in \mathbb{N}\}$ and the kernel decays exponentially then $CD(0,d)$ is still valid for some finite $d > 0$ as we show in Theorem 5.2. Thus the latter situation constitutes an extreme case.

Having summarized the main results on the $CD(0,d)$-condition, we now come to the question whether a *positive curvature* bound is possible, that is $CD(\kappa,\infty)$ for some $\kappa > 0$. In view of the representation formulas (4) and (5) this cannot be expected, since for the function $u(j) = j$ one obtains (at least formally) that $\Gamma_2(u)(0) = 0$ and $\Gamma(u)(0) = \sum_{j\in\mathbb{N}} k(j)j^2 \in (0,\infty]$. For kernels with finite second moment, this formal argument becomes rigorous as $u(j) = j$ is an admissible function in the sense that $\Gamma(u)$ is finite. In the other case, that is, k has no finite second moment, we are able to give a rigorous argument assuming that k is non-increasing on $\mathbb{N}$. By means of suitable approximation, we can here even restrict the class of admissible functions to compactly supported functions u, see Theorem 6.1. Again, the argument is rather technical due to the necessity of distinguishing several cases when estimating the Γ_2-term. The general case (without monotonicity assumption on k) remains open.

Note, however, that $CD(\kappa,\infty)$ with some $\kappa > 0$ implies $CD(0,d)$ for some finite $d > 0$, by Hölder's inequality (see also Lemma 2.2). Consequently, only kernels with $CD(0,d)$ for some $d \in (0,\infty)$ come into consideration with regard to a possible positive curvature bound. Interestingly, all such kernels we know (compare the above mentioned results) do have a finite second moment! This raises the question whether a finite second moment is necessary for $CD(0,d)$ with some $d \in (0,\infty)$. This open question brings us finally back to the open problem about the algebraic kernel in the critical case.

The article is organised as follows. In the next section we present some preliminary results and treat the case of bounded support of k. In Sect. 3 we show that algebraic kernels with $\beta < 2$ do not satisfy $CD(0,d)$ with finite d by constructing a compactly supported counterexample. The positive result for kernels with finite second moment, which in particular applies to algebraic kernels with $\beta > 2$, is shown in the subsequent section. Section 5 is devoted to the combinatorial analysis of kernels with sparse support, and in the last section we show that positive curvature is not possible for monotone kernels.

2 Basic properties and kernels with finite support

In this section we collect some preliminary results and show that finitely supported kernels always have finite dimension, that is $CD(0,d)$ holds for some finite $d > 0$. Moreover, we obtain some $CD(\kappa,d)$-inequalities with finite $d > 0$ but negative curvature κ by adapting previous results on graphs. Recall that we always assume that the kernel k is subject to the conditions (K1) and (K2) with $|k|_1 > 0$.

In order to describe the class of admissible functions, we introduce, for $1 \le p < \infty$ the weighted l_p-spaces

$$l_{p,k}(\mathbb{Z}) := \left\{ v : \mathbb{Z} \to \mathbb{R} : \sum_{j\in\mathbb{Z}} k(j)|v(j)|^p < \infty \right\},$$

endowed with the canonical norm. In what follows we use the following convention with respect to the admissible class of functions for the $CD(\kappa, d)$-inequality at $x \in \mathbb{Z}$. If not stated otherwise, a function $u : \mathbb{Z} \to \mathbb{R}$ is admissible for $CD(\kappa, d)$ at x if $u(\cdot + x) \in l_{1,k}(\mathbb{Z})$ in case $\kappa = 0$ and otherwise $u(\cdot + x) \in l_{2,k}(\mathbb{Z})$. This ensures finiteness of all the terms appearing on the right of (3) when used in the form given by (4) and (5), respectively. Observe that $l_{2,k}(\mathbb{Z}) \hookrightarrow l_{1,k}(\mathbb{Z})$, since $k \in l_1(\mathbb{Z})$ and by Hölder's inequality. Bounded functions are always admissible. Note that, by Lebesgue's theorem, $CD(\kappa, d)$ holds for all bounded functions provided this condition is satisfied for all compactly supported functions, but it seems to be not clear whether the latter implies $CD(\kappa, d)$ for all admissible functions.

The following auxiliary result will be frequently used in the paper.

Lemma 2.1 *When studying the validity of the $CD(\kappa, d)$-condition at $x \in \mathbb{Z}$, we may always assume without loss of generality that $x = 0$ and that the admissible functions satisfy $u(0) = 0$. Moreover, if $CD(-\kappa, d)$ holds for all* symmetric *u, where $\kappa \geq 0$ and $d \in (0, \infty)$, then $CD(-\kappa, d)$ holds for all u.*

Proof. For the first assertion we let $x \in \mathbb{Z}$ and u be an arbitrary admissible function. Setting $\tilde{u}(y) := u(y + x) - u(x)$ we find that $\tilde{u}$ is admissible at 0 and

$$\begin{aligned} 4\Gamma_2(u)(x) &= \sum_{j,l\in\mathbb{Z}} k(j)k(l)\,(u(x+j+l) - u(x+j) - u(x+l) + u(x))^2 \\ &= \sum_{j,l\in\mathbb{Z}} k(j)k(l)\,([u(x+j+l) - u(x)] - [u(x+j) - u(x)] \\ &\quad - [u(x+l) - u(x)])^2 \\ &= 4\Gamma_2(\tilde{u})(0). \end{aligned}$$

Similarly we have $\mathcal{L}u(x) = \mathcal{L}\tilde{u}(0)$ and $\Gamma(u)(x) = \Gamma(\tilde{u})(0)$.

To show the second part we may apply the first part and assume w.l.o.g. that $CD(-\kappa, d)$ holds at 0 for all symmetric functions and u is not symmetric with $u(0) = 0$. We find for the symmetric function $\tilde{u}(x) = \frac{1}{2}(u(x) + u(-x))$ that $\Gamma_2(\tilde{u})(0) \leq \Gamma_2(u)(0)$, since by the inequality $(a + b)^2 \leq 2a^2 + 2b^2$, $a, b \in \mathbb{R}$, we have

$$\begin{aligned} &(\tilde{u}(l + j) - \tilde{u}(l) - \tilde{u}(j))^2 \\ &\quad = \frac{1}{4}((u(l + j) - u(l) - u(j)) + (u(-l - j) - u(-l) - u(-j))]^2 \\ &\quad \leq \frac{1}{2}[((u(l + j) - u(l) - u(j))^2 + (u(-l - j) - u(-l) - u(-j))^2], \end{aligned}$$

and hence

$$\Gamma_2(\tilde{u})(0) \leq \frac{1}{2}\left(\Gamma_2(u)(0) + \Gamma_2(u(-\cdot))(0)\right) = \Gamma_2(u)(0)$$

by symmetry of k. Moreover we have $\mathcal{L}u(0) = \mathcal{L}\tilde{u}(0)$. By the same reasoning as before we have $\Gamma(\tilde{u})(0) \leq \Gamma(u)(0)$, hence $-\kappa\Gamma(\tilde{u})(0) \geq -\kappa\Gamma(u)(0)$ and whence by the assumption

$$\Gamma_2(u)(0) \geq \Gamma_2(\tilde{u})(0) \geq -\kappa\Gamma(\tilde{u})(0) + \frac{1}{d}(\mathcal{L}\tilde{u}(0))^2 \geq -\kappa\Gamma(u)(0) + \frac{1}{d}(\mathcal{L}u(0))^2. \qquad \square$$

Under symmetry of u and assuming $u(0) = 0$ we have (by taking $l = -j$ in formula (5) for Γ_2) the basic estimate

$$\Gamma_2(u)(0) \geq \frac{1}{4}\sum_{j\in\mathbb{Z}} k(j)^2\big(u(0)-u(j)-u(-j)\big)^2 = 2\sum_{j=1}^{\infty} k(j)^2 u(j)^2. \tag{8}$$

Note that (8) is not sufficient to show a $CD(0,d)$-inequality with finite dimension d if k has unbounded support. Indeed, letting $u(j) = \frac{1}{k(j)}$ for $j \in \operatorname{supp} k$ with $|j| \leq N$ and zero otherwise we obtain

$$\sum_{j=1}^{\infty} k(j)^2 u(j)^2 = \#\{j \in \operatorname{supp} k \,:\, 1 \leq j \leq N\}$$

whereas the right hand side satisfies

$$(\mathcal{L}u(0))^2 = \left(2\sum_{j=1}^{\infty} k(j)u(j)\right)^2 = 4(\#\{j \in \operatorname{supp} k \,:\, 1 \leq j \leq N\})^2$$

and grows faster as $N \to \infty$. This observation is a basic ingredient for the results in Sect. 5. In the case of a finitely supported k the lower bound (8) is enough to obtain the following important result.

Theorem 2.1 *If the kernel k has finite support, i.e. $\#\operatorname{supp} k = 2N$ for some $N \in \mathbb{N}$, then the operator $\mathcal{L}$ satisfies $CD(0,2N)$.*

Proof. By Lemma 2.1, it suffices to prove the asserted CD-inequality at the point $x = 0$ for all (admissible) symmetric functions u with $u(0) = 0$. By (K1) and (K2) we may assume that $\operatorname{supp} k = \{-x_N, \ldots, -x_1, x_1, \ldots, x_N\} \subset \mathbb{Z}\backslash\{0\}$. For simplicity of notation we set $x_0 := 0$ and let u be a symmetric admissible function with $u(0) = 0$. Taking $l = -j$ (and hence $x_{-j} = -x_j$) in the sum below we obtain that

$$\begin{aligned}
\Gamma_2(u)(0) &= \frac{1}{4}\sum_{j=-N}^{N}\sum_{l=-N}^{N} k(x_j)k(x_l)\left(u(x_j+x_l)-u(x_j)-u(x_l)\right)^2 \\
&\geq \frac{1}{4}\sum_{j=-N}^{N} k(x_j)k(-x_j)\left(u(x_j+(-x_j))-u(x_j)-u(-x_j)\right)^2 \\
&= 2\sum_{j=1}^{N}(k(x_j)u(x_j))^2 \geq \frac{2}{N}\left(\sum_{j=1}^{N} k(x_j)u(x_j)\right)^2 \\
&= \frac{1}{2N}\left(\sum_{j=-N}^{N} k(x_j)u(x_j)\right)^2 = \frac{1}{2N}(\mathcal{L}u(0))^2
\end{aligned}$$

after employing with $y_i := k(x_i)u(x_i)$ the Cauchy–Schwarz inequality $\left(\sum_{i=1}^{N} y_i\right)^2 \leq N\sum_{i=1}^{N} y_i^2$. □

The following example shows that Theorem 2.1 is optimal in general.

Example 2.1 For $N \in \mathbb{N}$ we define

$$k(x) = \begin{cases} 1, & x = 2j+1 \text{ for some } -N \leq j \leq N-1 \\ 0, & \text{else} \end{cases} \quad \text{and} \quad u(x) = \begin{cases} 1, & x \text{ is odd} \\ 0, & x = 0 \\ 2, & \text{else.} \end{cases}$$

Then k is symmetric and satisfies the assumption of Theorem 2.1, and we calculate

$$\mathcal{L}(u)(0) = \sum_{j=-N}^{N-1} k(2j+1)u(2j+1) = \sum_{j=-N}^{N-1} 1 = 2N.$$

For Γ_2 we find that $u(2j+2l+2) \neq 2$ if and only if $l = -j-1$, whence

$$\begin{aligned}\Gamma_2(u)(0) &= \frac{1}{4}\sum_{j=-N}^{N-1}\sum_{l=-N}^{N-1}\big(u(2l+2j+2)-1-1\big)^2 \\ &= \frac{1}{4}\sum_{j=-N}^{N-1} 4 = 2N = \frac{1}{2N}\big(\mathcal{L}(u)(0)\big)^2.\end{aligned}$$

Clearly, if $CD(\kappa, d)$ is satisfied with some positive curvature κ, then $CD(0, d)$ holds. Similarly we find that $CD(\kappa, \infty)$ with $\kappa > 0$ implies $CD(0, d)$ for some finite d.

Lemma 2.2 *If $CD(\kappa, d)$ holds for some $\kappa \geq 0$ and $d \in (0, \infty]$, then $CD(0, \tilde{d})$ holds, where $\tilde{d} = \frac{2d|k|_1}{d\kappa+2|k|_1}$ for $d < \infty$ and $\tilde{d} = \frac{2|k|_1}{\kappa}$ if $d = \infty$.*

Proof Again, we may restrict ourselves to the described CD-inequalities at $x = 0$ and we may assume that the admissible functions vanish at zero.

First we observe by applying Jensen's inequality that $(\mathcal{L}u(0))^2 \leq 2|k|_1\Gamma(u)(0)$. Indeed, we have

$$(\mathcal{L}u(0))^2 = |k|_1^2\left(\sum_{j\in\mathbb{Z}}\frac{k(j)}{|k|_1}u(j)\right)^2 \leq |k|_1\sum_{j\in\mathbb{Z}}k(j)u(j)^2 = 2|k|_1\Gamma(u)(0),$$

whence

$$\Gamma_2(u)(0) \geq \kappa\Gamma(u)(0) + \frac{1}{d}(\mathcal{L}u(0))^2 \geq \left(\frac{\kappa}{2|k|_1}+\frac{1}{d}\right)(\mathcal{L}u(0))^2 = \frac{1}{\tilde{d}}(\mathcal{L}u(0))^2$$

where $\tilde{d} = \frac{2d|k|_1}{d\kappa+2|k|_1}$. □

Let us present a few adaptions of general results concerning CD-inequalities on graphs.

Proposition 2.1 *For any kernel k we have $CD(-|k|_1, 2)$.*

This proposition is an immediate consequence of the following identity, which can be obtained analogously to [18, Thm 1.3].

Lemma 2.3 [12, (2.9)] *For any kernel we have*

$$\Gamma_2(u)(0) = \frac{1}{4}\sum_{j,l\in\mathbb{Z}}k(l)k(j)[u(j+l)-2u(j)+u(0)]^2 - |k|_1\Gamma(u)(0) + \frac{1}{2}(\mathcal{L}u(0))^2. \quad (9)$$

Corollary 2.1 [18, Thm 1.2] *Assume that $k(j) \geq c$ for $j \in \operatorname{supp} k$. Then $CD(2c-|k|_1, 2)$ holds.*

Proof. Choosing $l = -j$ in (9) we obtain

$$\begin{aligned}\Gamma_2(u)(0) &\geq \frac{4}{4}\sum_{j\in\mathbb{Z}}k(j)^2[u(0)-u(j)]^2 - |k|_1\Gamma(u)(0) + \frac{1}{2}(\mathcal{L}(u)(0))^2 \\ &\geq (2c-|k|_1)\Gamma(u)(0) + \frac{1}{2}(\mathcal{L}(u)(0))^2.\end{aligned}$$

□

Since k is symmetric we have $2c \leq |k|_1$ with equality only for kernels whose support consist of exactly two elements, e.g. for the discrete Laplacian (6).

The next remark shows that for kernels with finite second moment we always have the necessary condition $d \geq 1$ if $CD(0,d)$ holds.

Remark 2.1 Suppose that the kernel is such that $\sum_{j\in\mathbb{N}} j^2 k(j) < \infty$. For $u(j) = j^2$, $j \in \mathbb{Z}$, we find that

$$\begin{aligned}\Gamma_2(u)(0) &= \frac{1}{4}\sum_{j,l\in\mathbb{Z}} k(j)k(l)\big((j+l)^2 - j^2 - l^2\big)^2 = \sum_{j,l\in\mathbb{Z}} k(j)k(l)j^2l^2 \\ &= \big(\mathcal{L}(u)(0)\big)^2.\end{aligned}$$

Hence, if the kernel satisfies $CD(0,d)$, then $d \geq 1$.

3 Kernels of power type with $\beta < 2$

In this section we consider an important class of kernels for which $CD(0,d)$ fails to hold for all finite $d > 0$.

We use the following notation. For functions $f, g : D \to \mathbb{R}$ we write $f(x) \lesssim g(x)$ if there exists some constant $C > 0$ such that $f(x) \leq Cg(x)$ for all $x \in D$. Moreover, if $f(x) \lesssim g(x) \lesssim f(x)$ we write $f(x) \sim g(x)$.

Theorem 3.1 *Let k be a power type kernel as in (2) with $c > 0$ and $\beta \in (0,2)$. Then the corresponding operator $\mathcal{L}$ fails to satisfy $CD(0,d)$ for all finite $d > 0$.*

Proof We may assume without restriction of generality that $c = 1$. We consider the family of functions

$$u_\varepsilon(j) = |j|^{\beta-\varepsilon}, \quad j \in \mathbb{Z},\ \varepsilon \in (0,\beta).$$

We will show that

$$\mathcal{L}(u_\varepsilon)(0) \sim \frac{1}{\varepsilon} \text{ and } \Gamma_2(u_\varepsilon)(0) \sim \frac{1}{\varepsilon} \quad \text{as } \varepsilon \to 0, \tag{10}$$

which implies that for small $\varepsilon > 0$

$$\Gamma_2(u_\varepsilon)(0) \leq C\varepsilon(\mathcal{L}(u_\varepsilon)(0))^2$$

with some constant $C > 0$, contradicting any $CD(0,d)$ inequality with finite $d > 0$. Observe that (10) contains more information than what is actually required for the proof of Theorem 3.1. In fact, concerning Γ_2 it would be enough to show that $\Gamma_2(u_\varepsilon)(0) \lesssim \frac{1}{\varepsilon}$ as $\varepsilon \to 0$.

The first claim in (10) can be easily verified. Indeed,

$$\mathcal{L}(u_\varepsilon)(0) = 2\sum_{j=1}^{\infty} k(j)u_\varepsilon(j) = 2\sum_{j=1}^{\infty} \frac{1}{j^{1+\beta}} j^{\beta-\varepsilon} = 2\sum_{j=1}^{\infty} \frac{1}{j^{1+\varepsilon}} \sim \frac{1}{\varepsilon}$$

as $\varepsilon \to 0$, since by Lemma 3.3 the last sum can be controlled from below and above by a positive constant times the integral

$$\int_1^\infty \frac{dx}{x^{1+\varepsilon}} = \frac{1}{\varepsilon}.$$

Turning to Γ_2, by symmetry it is enough to consider the terms

$$J_\varepsilon := \sum_{j=1}^{\infty}\sum_{l=1}^{j} k(j)k(l)\big(u_\varepsilon(j+l) - u_\varepsilon(j) - u_\varepsilon(l)\big)^2$$

and

$$K_\varepsilon := \sum_{j=1}^{\infty}\sum_{l=1}^{j} k(j)k(l)\big(u_\varepsilon(j-l) - u_\varepsilon(j) - u_\varepsilon(l)\big)^2.$$

The first term can be reformulated as follows.

$$\begin{aligned} J_\varepsilon &= \sum_{j=1}^{\infty}\sum_{l=1}^{j} \frac{1}{j^{1+\beta} l^{1+\beta}} \big((j+l)^{\beta-\varepsilon} - j^{\beta-\varepsilon} - l^{\beta-\varepsilon}\big)^2 \\ &= \sum_{j=1}^{\infty}\sum_{l=1}^{j} \frac{j^{2\beta-2\varepsilon}}{j^{1+\beta} l^{1+\beta}} \left[\Phi_\gamma\left(\frac{l}{j}\right)\right]^2, \end{aligned} \tag{11}$$

where $\gamma = \beta - \varepsilon$ and the function Φ_γ is defined by

$$\Phi_\gamma(x) = (1+x)^\gamma - 1 - x^\gamma, \quad x \in [0,1].$$

Lemma 3.1 *In the case $\gamma \in (1,2)$ there holds*

$$(\gamma - 1)x \le \Phi_\gamma(x) \le 2\gamma x, \quad x \in [0,1].$$

If $\gamma \in (0,1)$, we have

$$-x^\gamma \le \Phi_\gamma(x) \le -(1-\gamma)x^\gamma, \quad x \in [0,1].$$

Proof For $x \in [0,1]$ we have by the mean value theorem that $(1+x)^\gamma - 1 = \gamma\xi^{\gamma-1}x$ for some $\xi \in [1, 1+x]$. Consequently, if $\gamma \in (1,2)$, we see that

$$\gamma x \le (1+x)^\gamma - 1 \le 2\gamma x,$$

which together with $x^\gamma \le x$ implies the first assertion. In the case $\gamma \in (0,1)$, we have for $x \in (0,1]$ that $(1+x)^\gamma - 1 \le \gamma x^{\gamma-1}x = \gamma x^\gamma$, since $x \le \xi$. This shows the second assertion. □

In order to estimate J_ε by the aid of Lemma 3.1 we distinguish two cases.

Case 1: Suppose that $\beta \in (1,2)$. Then $\gamma \in (1,2)$ for sufficiently small ε, and Lemma 3.1 and (11) then show that J_ε can be estimated from below and above as

$$J_\varepsilon \sim C(\beta)\sum_{j=1}^{\infty}\sum_{l=1}^{j} \frac{j^{2\beta-2\varepsilon}}{j^{1+\beta} l^{1+\beta}} \Big(\frac{l}{j}\Big)^2 = C(\beta)\sum_{j=1}^{\infty} \frac{1}{j^{3-\beta+2\varepsilon}} \sum_{l=1}^{j} l^{1-\beta}. \tag{12}$$

By monotonicity of the sequence $(l^{1-\beta})_{l\in\mathbb{N}}$, the sum $\sum_{l=1}^{j} l^{1-\beta}$ can be controlled from below and above by a positive constant times the integral

$$\int_1^{j+1} x^{1-\beta}dx = \frac{1}{2-\beta}\left((j+1)^{2-\beta} - 1\right) \sim j^{2-\beta}.$$

Therefore

$$J_\varepsilon \sim C(\beta)\sum_{j=1}^{\infty}\frac{1}{j^{3-\beta+2\varepsilon}}\,j^{2-\beta} = C(\beta)\sum_{j=1}^{\infty}\frac{1}{j^{1+2\varepsilon}} \sim \frac{1}{\varepsilon}$$

as $\varepsilon \to 0$.

Case 2: Suppose now that $\beta \in (0, 1]$. Then $\gamma \in (0, 1)$, and by Lemma 3.1 it follows that J_ε can be estimated from below and above as

$$J_\varepsilon \sim C(\beta)\sum_{j=1}^{\infty}\sum_{l=1}^{j}\frac{j^{2\beta-2\varepsilon}}{j^{1+\beta}l^{1+\beta}}\big(\frac{l}{j}\big)^{2\beta-2\varepsilon} = C(\beta)\sum_{j=1}^{\infty}\frac{1}{j^{1+\beta}}\sum_{l=1}^{j}l^{\beta-1-2\varepsilon}.$$

Arguing as in the first case, we see that $\sum_{l=1}^{j} l^{\beta-1-2\varepsilon} \sim j^{\beta-2\varepsilon}$ and hence

$$J_\varepsilon \sim C(\beta)\sum_{j=1}^{\infty}\frac{1}{j^{1+\beta}}\,j^{\beta-2\varepsilon} = C(\beta)\sum_{j=1}^{\infty}\frac{1}{j^{1+2\varepsilon}} \sim \frac{1}{\varepsilon}$$

as $\varepsilon \to 0$.

We now come to the term K_ε. We have

$$\begin{aligned} K_\varepsilon &= \sum_{j=1}^{\infty}\sum_{l=1}^{j}\frac{1}{j^{1+\beta}l^{1+\beta}}\big(j^{\beta-\varepsilon}+l^{\beta-\varepsilon}-(j-l)^{\beta-\varepsilon}\big)^2 \\ &= \sum_{j=1}^{\infty}\sum_{l=1}^{j}\frac{j^{2\beta-2\varepsilon}}{j^{1+\beta}l^{1+\beta}}\big[\Psi_\gamma\big(\frac{l}{j}\big)\big]^2, \end{aligned} \tag{13}$$

where again $\gamma = \beta - \varepsilon$ and the function Ψ_γ is defined by

$$\Psi_\gamma(x) = 1 + x^\gamma - (1-x)^\gamma, \quad x \in [0, 1].$$

Lemma 3.2 *In the case $\gamma \in (1, 2)$ there holds*

$$x \le \Psi_\gamma(x) \le (\gamma+1)x, \quad x \in [0, 1].$$

If $\gamma \in (0, 1)$, we have

$$x^\gamma \le \Psi_\gamma(x) \le 3x^\gamma, \quad x \in [0, 1].$$

Proof Suppose first that $\gamma \in (1, 2)$. For $x \in [0, 1]$, we have $1-(1-x)^\gamma = \gamma\xi^{\gamma-1}x$ for some $\xi \in [1-x, 1]$, by the mean value theorem. Since $\gamma - 1 > 0$, this implies $1-(1-x)^\gamma \le \gamma x$, which together with $x^\gamma \le x$ yields the upper bound in the assertion. For $x \in [0, \frac{1}{2}]$ we have $\xi \in [\frac{1}{2}, 1]$, which implies that $\Psi_\gamma(x) \ge 1-(1-x)^\gamma \ge \gamma(1/2)^{\gamma-1}x \ge x$. If $x \in [\frac{1}{2}, 1]$, then clearly $\Psi_\gamma(x) \ge 1 \ge x$. This proves the lower bound in the first claim.

Let now $\gamma \in (0, 1)$. The lower estimate for $\Psi_\gamma(x)$ is evident, since $1-(1-x)^\gamma \ge 0$. For the upper bound we again use the mean value theorem similarly as above to see that for $x \in [0, \frac{1}{2}]$ there holds $1-(1-x)^\gamma \le \gamma(1/2)^{\gamma-1}x$, noting that now $\gamma - 1 < 0$. Since $x = x^{1-\gamma}x^\gamma \le (1/2)^{1-\gamma}x^\gamma$, it follows that $\Psi_\gamma(x) \le 2x^\gamma$. For $x \in [\frac{1}{2}, 1]$ we have $1 \le 2x \le 2x^\gamma$ and hence $\Psi_\gamma(x) \le 1+x^\gamma \le 3x^\gamma$. This shows the upper bound in the second assertion. □

Having Lemma 3.2 at disposal, we may now estimate K_ε appropriately. As before, we distinguish two cases w.r.t. the parameter β.

Case 1: Assume that $\beta \in (1, 2)$. Then (13), the first part of Lemmas 3.2 and 3.3 show that for sufficiently small $\varepsilon > 0$, K_ε can be estimated from below and above as

$$K_\varepsilon \sim C(\beta) \sum_{j=1}^{\infty} \sum_{l=1}^{j} \frac{j^{2\beta-2\varepsilon}}{j^{1+\beta} l^{1+\beta}} \left(\frac{l}{j}\right)^2,$$

which is the same expression as in the estimation of J_ε in case 1. Therefore, $K_\varepsilon \sim \frac{1}{\varepsilon}$ as $\varepsilon \to 0$.

Case 2: Suppose now that $\beta \in (0, 1]$. Then (13), the second part of Lemmas 3.2 and 3.3 show that for sufficiently small $\varepsilon > 0$, K_ε can be controlled from below and above as

$$K_\varepsilon \sim C \sum_{j=1}^{\infty} \sum_{l=1}^{j} \frac{j^{2\beta-2\varepsilon}}{j^{1+\beta} l^{1+\beta}} \left(\frac{l}{j}\right)^{2\beta-2\varepsilon}.$$

This is the same expression as in the estimation of J_ε in case 2. Hence $K_\varepsilon \sim \frac{1}{\varepsilon}$ as $\varepsilon \to 0$. All in all, we see that $\Gamma_2(u_\varepsilon)(0) \sim \frac{1}{\varepsilon}$ as $\varepsilon \to 0$. The proof of Theorem 3.1 is complete. □

Interestingly, in the case $\beta = 2$ the family u_ε considered above does no longer lead to a contradiction of $CD(0, d)$ for finite $d > 0$. This is due to the fact that now $\Gamma_2(u_\varepsilon)(0) \sim \frac{1}{\varepsilon^2}$ as $\varepsilon \to 0$. To see this, we can again use the statements in Lemmas 3.1 and 3.2 where $\gamma \in (1, 2)$. As to the term J_ε, we see from (12) with $\beta = 2$ that

$$J_\varepsilon \sim C \sum_{j=1}^{\infty} \frac{1}{j^{1+2\varepsilon}} \sum_{l=1}^{j} l^{-1}.$$

In contrast to the case $\beta < 2$, the inner sum growths logarithmically in j and thus

$$J_\varepsilon \sim C \sum_{j=1}^{\infty} \frac{1}{j^{1+2\varepsilon}} (1 + \log j) \sim \frac{1}{\varepsilon^2},$$

since $\int_c^\infty \frac{\log x}{x^{1+2\varepsilon}}\, dx \sim \frac{1}{\varepsilon^2}$ as $\varepsilon \to 0$, for all $c \geq 1$. The term K_ε enjoys the same behaviour.

As we have mentioned in the introduction, an important consequence of Theorem 3.1 is the following corollary concerning fractional powers of the discrete Laplacian.

Corollary 3.1 *The operator $-(-\Delta)^{\frac{\beta}{2}}$ fails to satisfy $CD(0, d)$ for all finite $d > 0$.*

Proof Let $\beta \in (0, 2)$ be fixed. From [5, Theorem 1.1] we obtain that $-(-\Delta)^{\frac{\beta}{2}}$ is an operator of the form (1), where the kernel is given by

$$k_\beta(j) = \frac{4^{\frac{\beta}{2}} \Gamma\left(\frac{1+\beta}{2}\right) \Gamma\left(|j| - \frac{\beta}{2}\right)}{\sqrt{\pi} \Gamma\left(-\frac{\beta}{2}\right) \Gamma\left(|j| + 1 + \frac{\beta}{2}\right)}, \quad j \in \mathbb{Z}\backslash\{0\}, \tag{14}$$

and $k_\beta(0) = 0$. Note that in (14), Γ denotes the Gamma function. Throughout this proof we denote by $\mathcal{L}$ and Γ_2 the operators corresponding to the respective power type kernel from (2) and by $\Gamma_2^{(\beta)}$ the iterated carré du champ operator for $-\left(-\Delta\right)^{\frac{\beta}{2}}$. In [5, Theorem 1.1] it is shown that one finds constants $0 < c(\beta) \leq C(\beta)$ such that

$$\frac{c(\beta)}{|j|^{1+\beta}} \leq k_\beta(j) \leq \frac{C(\beta)}{|j|^{1+\beta}}$$

holds for any $j \in \mathbb{Z}\backslash\{0\}$. Thus, choosing the non-negative function u_ε from the proof of Theorem 3.1 we observe $\left(-\left(-\Delta\right)^{\frac{\beta}{2}}(u_\varepsilon)(0)\right)^2 \geq c(\beta)^2\,(\mathcal{L}u_\varepsilon(0))^2$ and $\Gamma_2^{(\beta)}(u_\varepsilon)(0) \leq C(\beta)^2\,\Gamma_2(u_\varepsilon)(0)$. The claim follows from (10). □

A natural question to ask is whether the statement of Theorem 3.1 remains true for a smaller class of functions, e.g. the space of bounded functions, since the above counterexample is obviously unbounded. We can answer this question negatively by an approximation argument.

Theorem 3.2 *Let k be a power type kernel as in (2) with $c > 0$ and $\beta \in (0,2)$. There exists no finite $d > 0$ such that*

$$\Gamma_2(v)(0) \geq \frac{1}{d}\left(\left(\mathcal{L}v\right)(0)\right)^2$$

holds for all compactly supported functions v.

For the proof of this statement we will repeatedly use the following lemma.

Lemma 3.3 *Let $A_1, A_2 \in \mathbb{N}$ with $A_1 < A_2$ and $\gamma \in \mathbb{R}$. Then the following estimates are valid*

$$\sum_{m=A_1}^{A_2} m^\gamma \leq \begin{cases} \frac{1}{\gamma+1}\,(A_2+1)^{1+\gamma}, & 1+\gamma \geq 1 \\ \frac{1}{\gamma+1}A_2^{1+\gamma}, & 1+\gamma \in (0,1), \end{cases} \tag{15}$$

and if $A_1 \geq 2$

$$\sum_{m=A_1}^{A_2} m^\gamma \leq \begin{cases} \log(A_2), & 1+\gamma = 0 \\ \frac{1}{|\gamma+1|}\,(A_1-1)^{1+\gamma}, & 1+\gamma < 0. \end{cases} \tag{16}$$

Proof For $\gamma \geq 0$ we observe

$$\sum_{m=A_1}^{A_2} m^\gamma \leq \sum_{m=A_1}^{A_2} \int_m^{m+1} x^\gamma\,\mathrm{d}x = \frac{1}{\gamma+1}\left((A_2+1)^{\gamma+1} - A_1^{\gamma+1}\right) \leq \frac{1}{\gamma+1}\,(A_2+1)^{\gamma+1}.$$

In case of $\gamma < 0$ we have

$$\sum_{m=A_1}^{A_2} m^\gamma \leq \sum_{m=A_1}^{A_2} \int_{m-1}^{m} x^\gamma\,\mathrm{d}x,$$

since the mapping $x \mapsto x^\gamma$ is now decreasing for $x > 0$. Assuming that $A_1 \geq 2$, we can proceed calculating the integral as above and obtain

$$\sum_{m=A_1}^{A_2} m^\gamma \leq \begin{cases} \frac{1}{\gamma+1}\left(A_2^{\gamma+1} - (A_1-1)^{\gamma+1}\right), & \gamma \in (-\infty,0)\backslash\{-1\} \\ \log(A_2) - \log(A_1-1), & \gamma = -1. \end{cases}$$

From this the claim follows. □

Proof of Theorem 3.2 Let $\beta \in (0,2)$ be fixed and choose $\varepsilon > 0$ such that $\beta - 2\varepsilon > 0$ and $\beta - \varepsilon > 1$ in case of $\beta > 1$. Let u_ε be given as in the proof of Theorem 3.1.

We define for even $N \in 2\mathbb{N}$ the function

$$v_{N,\varepsilon}(j) := \begin{cases} u_\varepsilon(j), & j \in \{0,\dots,N\} \\ -\frac{N^{\beta-\varepsilon}}{N^2-N}\,j + \frac{N^{\beta-\varepsilon+2}}{N^2-N}, & j \in \{N+1,\dots,N^2\} \\ 0, & j > N^2, \end{cases}$$

and extend it symmetrically to $v_{N,\varepsilon} : \mathbb{Z} \to \mathbb{R}$. In the sequel we will denote u_ε by u and $v_{N,\varepsilon}$ by v_N.

Our aim is to prove that $\mathcal{L}(v_N)(0) \to \mathcal{L}(u)(0)$ and $\Gamma_2(v_N)(0) \to \Gamma_2(u)(0)$ as $N \to \infty$, which is sufficient to deduce the claim. Indeed, assuming that there exists some $\Lambda > 0$ such that $\Gamma_2(w) \geq \Lambda(\mathcal{L}w)^2$ at $x = 0$ for all compactly supported functions w, we find from Theorem 3.1 a sufficiently small $\varepsilon > 0$ with u, as given above, satisfying $\Gamma_2(u)(0) \leq \frac{\Lambda}{8}(\mathcal{L}u(0))^2$. Let N be large enough such that $|(\mathcal{L}v_N(0))^2 - (\mathcal{L}u(0))^2| \leq \frac{1}{2}(\mathcal{L}u(0))^2$ and $|\Gamma_2(v_N)(0) - \Gamma_2(u)(0)| \leq \Gamma_2(u)(0)$. Then $(\mathcal{L}u(0))^2 \leq 2(\mathcal{L}v_N(0))^2$ and whence

$$\Gamma_2(v_N)(0) \leq 2\Gamma_2(u)(0) \leq \frac{\Lambda}{4}(\mathcal{L}u(0))^2 \leq \frac{\Lambda}{2}(\mathcal{L}v_N(0))^2,$$

contradicting the assumption and showing the claim.

First we obtain that

$$\begin{aligned}|\mathcal{L}(u)(0) - \mathcal{L}(v_N)(0)| &\leq 2\left(\sum_{j=N+1}^{N^2} \frac{|u(j) - v_N(j)|}{j^{1+\beta}} + \sum_{j=N^2+1}^{\infty} \frac{u(j)}{j^{1+\beta}}\right) \\ &\leq 2\sum_{j=N+1}^{\infty} \frac{u(j)}{j^{1+\beta}} \xrightarrow{N\to\infty} 0.\end{aligned}$$

Next, we fix some $\rho > 0$ and aim to prove $|\Gamma_2(v_N)(0) - \Gamma_2(u)(0)| < \rho$ for each $N \geq N_0(\rho)$ for some sufficiently large $N_0(\rho) \in \mathbb{N}$. For $M \in \mathbb{N}$ we define the kernel $k_M(x) := \frac{1}{|x|^{1+\beta}}\chi_{(\{-M,\dots,M\}\setminus\{0\})}$ and denote the corresponding operator from (1) by $\mathcal{L}^M$. Furthermore, we denote by Γ_2^M the respective iterated carré du champ operator, which can be written as

$$\Gamma_2^M(w)(x) = \frac{1}{4}\sum_{j,l\in\{-M,\dots,M\}\setminus\{0\}} \frac{\big(w(x+j+l) - w(x+j) - w(x+l) + w(x)\big)^2}{j^{1+\beta}l^{1+\beta}}, \quad x \in \mathbb{Z}.$$

From the dominated convergence theorem one obtains that $\Gamma_2^M(u)(0)$ converges to $\Gamma_2(u)(0)$ as M tends to infinity. We fix $M > 0$ large enough such that

$$|\Gamma_2^M(u)(0) - \Gamma_2(u)(0)| < \frac{\rho}{4}.$$

If $N > 2M$ and $|j|, |l| \leq M$ we have $(v_N(j+l) - v_N(j) - v_N(l))^2 = (u(j+l) - u(j) - u(l))^2$ and thus it holds

$$|\Gamma_2^M(v_N)(0) - \Gamma_2^M(u)(0)| = 0.$$

The above observations are beneficial, due to the basic calculation

$$\begin{aligned}&|\Gamma_2(v_N)(0) - \Gamma_2(u)(0)| \\ &\quad\leq |\Gamma_2(v_N)(0) - \Gamma_2^M(v_N)(0)| + |\Gamma_2^M(v_N)(0) - \Gamma_2^M(u)(0)| + |\Gamma_2^M(u)(0) - \Gamma_2(u)(0)|.\end{aligned}$$

Hence, to show the claim it suffices to prove that $|\Gamma_2(v_N)(0) - \Gamma_2^M(v_N)(0)|$ converges to zero as N tends to infinity.

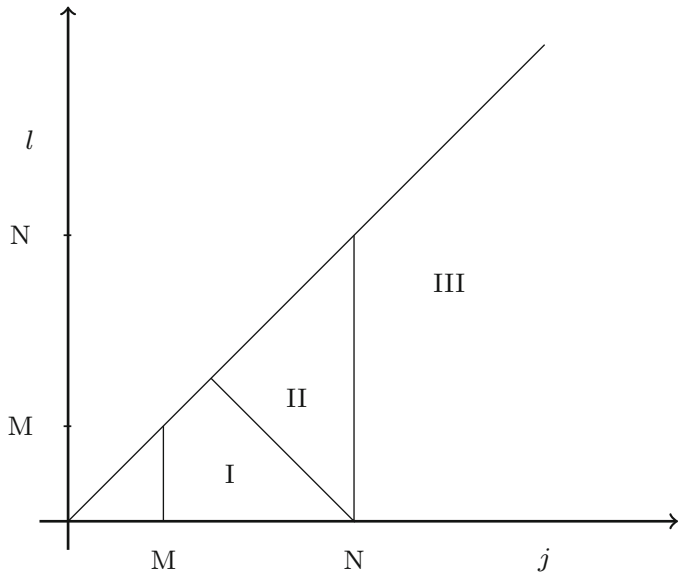


Fig. 1 Splitting of the domain of summation for the term including $j+l$

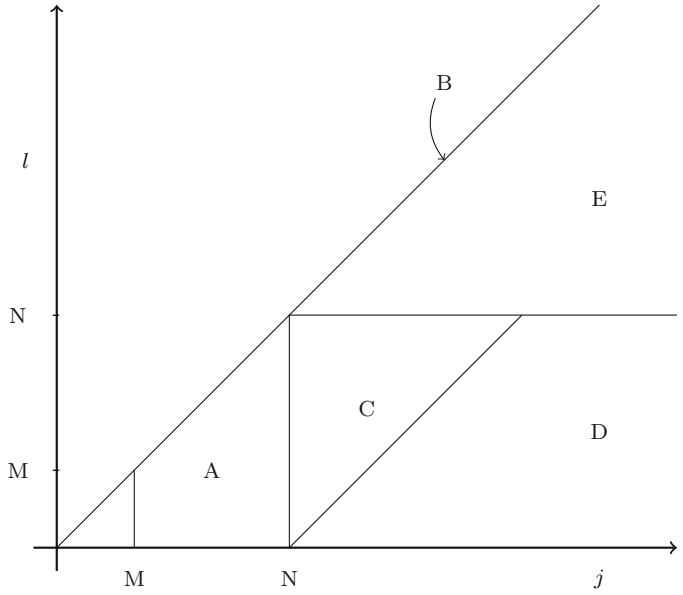


Fig. 2 Splitting of the domain of summation for the term with $j-l$

In order to prove the desired convergence, we have to distinguish several cases. We observe by symmetry of the kernel and v_N that

$$|\Gamma_2(v_N)(0) - \Gamma_2^M(v_N)(0)| \lesssim \sum_{j=M+1}^{\infty} \sum_{l=1}^{j} \frac{(v_N(j+l) - v_N(j) - v_N(l))^2 + (v_N(j-l) - v_N(j) - v_N(l))^2}{j^{1+\beta} l^{1+\beta}}. \tag{17}$$

In order to show that the expression in (17) converges to zero we distinguish several cases (see Figs. 1, 2). First we make the following observation. Given $1 \le x \le y$ one easily obtains from the mean value theorem the upper bound

$$|v_N(x) - v_N(y)| \lesssim \begin{cases} \max\{x^{\beta-\varepsilon-1}, y^{\beta-\varepsilon-1}\}\,(y-x)\,, & y \le N \\ N^{\beta-\varepsilon-2}\,(y-x)\,, & x > N, \end{cases} \tag{18}$$

Here and in the following we write $\lesssim$ if the corresponding constant is independent of N and M. In the remaining situation $x \le N < y$, we apply (18) to obtain

$$\begin{aligned}|v_N(y) - v_N(x)| &\le |v_N(y) - v_N(N)| + |v_N(N) - v_N(x)| \\ &\lesssim N^{\beta-\varepsilon-2}(y-N) + \max\{N^{\beta-\varepsilon-1}, x^{\beta-\varepsilon-1}\}(N-x).\end{aligned}$$

If $\beta > 1$ this leads to

$$|v_N(y) - v_N(x)| \lesssim N^{\beta-\varepsilon-1}(y-x), \tag{19}$$

since $N^{\beta-\varepsilon-2} \le N^{\beta-\varepsilon-1}$. In the other case of $\beta \le 1$, we have

$$|v_N(y) - v_N(x)| \lesssim x^{\beta-\varepsilon-1}(y-x), \tag{20}$$

since $N^{\beta-\varepsilon-2} \le \frac{1}{N}x^{\beta-\varepsilon-1} \le x^{\beta-\varepsilon-1}$.
I : $1 \le l \le j$, $j + l \le N$, $j \ge M+1$: Clearly, if $j + l \le N$ we have $(v_N(j+l) - v_N(j) - v_N(l))^2 = (u(j+l) - u(j) - u(l))^2$. We can control this case by the sum

$$\sum_{j=M+1}^{N-1} \sum_{l=1}^{N-j} \frac{(v_N(j+l) - v_N(j) - v_N(l))^2}{j^{1+\beta} l^{1+\beta}} \le |\Gamma_2^M(u)(0) - \Gamma_2(u)(0)| < \frac{\rho}{4}.$$

II : $1 \le l \le j \le N$, $j + l \ge N+1$: We will make use of the basic estimate

$$(v_N(j \pm l) - v_N(j) - v_N(l))^2 \lesssim v_N(l)^2 + (v_N(j \pm l) - v_N(j))^2 .$$

Observe that

$$\sum_{j=\frac{N}{2}}^{\infty} \frac{1}{j^{1+\beta}} \sum_{l=1}^{j} \frac{v_N(l)^2}{l^{1+\beta}} \le \sum_{j=\frac{N}{2}}^{\infty} \frac{1}{j^{1+\beta}} \sum_{l=1}^{j} \frac{l^{2\beta-2\varepsilon}}{l^{1+\beta}} \lesssim \sum_{j=\frac{N}{2}}^{\infty} \frac{1}{j^{1+2\varepsilon}} \xrightarrow{N\to\infty} 0,$$

where we applied (15) in the second step with $\gamma = \beta - 2\varepsilon - 1$. Note that we will also apply this estimate to the term involving '$j - l$' later. If $\beta > 1$ we have by (19)

$$(v_N(j+l) - v_N(j))^2 \lesssim N^{2\beta-2\varepsilon-2} l^2 \le (j+l)^{2\beta-2\varepsilon-2} l^2 \le (2j)^{2\beta-2\varepsilon-2} l^2.$$

Due to (20), we can thus conclude for any $\beta \in (0, 2)$ it holds

$$(v_N(j+l) - v_N(j))^2 \lesssim j^{2\beta-2\varepsilon-2} l^2.$$

Hence, we obtain with the help of (15) for $\gamma = 1 - \beta$ the upper bound

$$\sum_{j=\frac{N}{2}+1}^{N} \frac{1}{j^{1+\beta}} \sum_{l=N-j+1}^{j} \frac{(v_N(j+l) - v_N(j))^2}{l^{1+\beta}} \lesssim \sum_{j=\frac{N}{2}+1}^{N} \frac{j^{2\beta-2\varepsilon-2}}{j^{1+\beta}} \sum_{l=1}^{j} l^{1-\beta} \lesssim \sum_{j=\frac{N}{2}+1}^{N} \frac{1}{j^{1+2\varepsilon}}$$

which tends to zero as $N \to \infty$.
III : $1 \le l \le j$, $j \ge N+1$: Since $N < j \le j + l$ we observe by (18)

$$\begin{aligned}\sum_{j=N+1}^{\infty} \frac{1}{j^{1+\beta}} \sum_{l=1}^{j} \frac{(v_N(j+l) - v_N(j))^2}{l^{1+\beta}} &\lesssim N^{2\beta-2\varepsilon-4} \sum_{j=N+1}^{N^2-1} \frac{1}{j^{1+\beta}} \sum_{l=1}^{j} l^{1-\beta} \\ &\lesssim N^{2\beta-2\varepsilon-4} \sum_{j=N+1}^{N^2-1} j^{1-2\beta},\end{aligned}$$

where we applied (15) in the last step. According to (15) and (16) we find

$$\sum_{j=N+1}^{N^2-1} j^{1-2\beta} \lesssim \begin{cases} N^{4-4\beta}, & \beta < 1 \\ \log(N^2), & \beta = 1 \\ N^{2-2\beta}, & \beta > 1. \end{cases}$$

Multiplying each of these expressions by $N^{2\beta-2\varepsilon-4}$ yields the desired convergence for any $\beta \in (0,2)$. Hence, the first part of (17) converges to zero as N tends to infinity.

For the estimates of the second part, we need a refined splitting of the domain, sketched in Fig. 2.

A: $1 \le l \le j \le N$, $j \ge M+1$: Since $j - l \le N$ it follows similar to the case I that

$$\sum_{j=M+1}^{N} \sum_{l=1}^{j} \frac{(v_N(j-l) - v_N(j) - v_N(l))^2}{j^{1+\beta} l^{1+\beta}} \le |\Gamma_2^M(u)(0) - \Gamma_2(u)(0)| < \frac{\rho}{4}.$$

B: $l = j \ge N+1$: In this situation we have due to $v_N(j)^2 \le u(j)^2$ the estimate

$$\sum_{j=N+1}^{\infty} \frac{(v_N(0) - 2v_N(j))^2}{j^{2+2\beta}} = 4 \sum_{j=N+1}^{\infty} \frac{1}{j^{2+2\beta}} v_N(j)^2 \le 4 \sum_{j=N+1}^{\infty} \frac{1}{j^{2+2\varepsilon}} \xrightarrow{N\to\infty} 0.$$

C, D, E: $1 \le l \le j-1$, $j \ge N+1$, where $\beta > 1$: We observe in this case by (19) and (15)

$$\begin{aligned} \sum_{j=N+1}^{\infty} \frac{1}{j^{1+\beta}} \sum_{l=1}^{j} \frac{(v_N(j) - v_N(j-l))^2}{l^{1+\beta}} &\lesssim N^{2\beta-2\varepsilon-2} \sum_{j=N+1}^{\infty} \frac{1}{j^{1+\beta}} \sum_{l=1}^{j} l^{1-\beta} \\ &\lesssim N^{2\beta-2\varepsilon-2} \sum_{j=N+1}^{\infty} \frac{1}{j^{2\beta-1}} \lesssim \frac{1}{N^{2\varepsilon}} \to 0 \quad (N \to \infty), \end{aligned}$$

where we applied $\beta > 1$ and (16) in the last step. Recalling the estimates from case II, we have established the claim for $\beta > 1$. We assume from now on $\beta \le 1$.

C: $1 \le l \le N$, $j \ge N+1$, $j - l \le N$, where $\beta \le 1$: Due to (20) we observe

$$\begin{aligned} \sum_{j=N+1}^{2N} \frac{1}{j^{1+\beta}} \sum_{l=j-N}^{N} \frac{(v_N(j) - v_N(j-l))^2}{l^{1+\beta}} &\lesssim \sum_{j=N+1}^{2N} \frac{1}{j^{1+\beta}} \sum_{l=j-N}^{N} l^{1-\beta} (j-l)^{2\beta-2\varepsilon-2} \\ &\lesssim N^{1-\beta} \sum_{j=N+1}^{2N} \frac{1}{j^{1+\beta}} \sum_{m=j-N}^{N} m^{2\beta-2\varepsilon-2}. \end{aligned}$$

In order to treat the latter sum, we proceed with a finer case separation. Note that we can exclude the case of $2\beta - 2\varepsilon - 1 = 0$ by choosing $\varepsilon > 0$ appropriately small. If $2\beta - 2\varepsilon - 1 > 0$ we can apply (15) and prove its convergence to zero similar as before. To conclude convergence in the case of $2\beta - 2\varepsilon - 1 < 0$ we split the expression and use (15) and (16) as follows:

$$\begin{aligned} &N^{1-\beta} \sum_{j=N+1}^{2N} \frac{1}{j^{1+\beta}} \sum_{m=j-N}^{N} m^{2\beta-2\varepsilon-2} \\ &= N^{1-\beta} \left(\sum_{j=N+2}^{2N} \frac{1}{j^{1+\beta}} \sum_{m=j-N}^{N} m^{2\beta-2\varepsilon-2} + \frac{1}{(N+1)^{1+\beta}} \left(1 + \sum_{m=2}^{N} m^{2\beta-2\varepsilon-2} \right) \right) \end{aligned}$$

$$\lesssim N^{1-\beta} \sum_{j=N+2}^{2N} \frac{(j-N-1)^{2\beta-2\varepsilon-1}}{j^{1+\beta}} + \frac{N^{1-\beta}}{N^{1+\beta}}$$
$$\lesssim \frac{1}{N^{2\beta}} \sum_{m=1}^{N-1} m^{2\beta-2\varepsilon-1} + \frac{2}{N^{2\beta}} \lesssim \frac{1}{N^{2\varepsilon}} + \frac{1}{N^{2\beta}} \xrightarrow{N\to\infty} 0.$$

D: $1 \le l \le N < j$, $j - l \ge N + 1$, where $\beta \le 1$: Here we have the comfortable situation that j, $j - l > N$. Therefore, we can apply (18) and observe

$$\sum_{l=1}^{N} \frac{1}{l^{1+\beta}} \sum_{j=N+l+1}^{\infty} \frac{(v_N(j) - v_N(j-l))^2}{j^{1+\beta}} \lesssim N^{2\beta-2\varepsilon-4} \sum_{l=1}^{N} l^{1-\beta} \sum_{j=N+l+1}^{\infty} \frac{1}{j^{1+\beta}}$$
$$\lesssim N^{\beta-2\varepsilon-4} \sum_{l=1}^{N} l^{1-\beta} \lesssim \frac{1}{N^{2+2\varepsilon}} \xrightarrow{N\to\infty} 0,$$

where we applied (16) to the inner sum and (15) to the outer sum. Due to the arguments of the case II, we establish the claim in this situation.
E: $N + 1 \le l \le j - 1$, $j \ge N + 1$, where $\beta \le 1$: Since $\max_{x\in\mathbb{Z}} v_N(x) = N^{\beta-\varepsilon}$ we can estimate

$$\sum_{j=N+1}^{\infty} \frac{1}{j^{1+\beta}} \sum_{l=N+1}^{j-1} \frac{(v_N(j-l) - v_N(j) - v_N(l))^2}{l^{1+\beta}}$$
$$\lesssim N^{2\beta-2\varepsilon} \sum_{j=N+1}^{\infty} \frac{1}{j^{1+\beta}} \sum_{l=N+1}^{j-1} \frac{1}{l^{1+\beta}} \lesssim \frac{1}{N^{2\varepsilon}},$$

where we applied (16) to both sums. Whence, we find $|\Gamma_2(v_N)(0) - \Gamma_2^M(v_N)(0)| < \frac{3\rho}{4}$ for $N \ge N_0(\rho)$ and thus conclude the claim. □

4 Kernels with finite second moment

We have seen in the previous section that for power type kernels with $\beta < 2$ the inequality $CD(0, d)$ fails for all finite $d > 0$. It turns out that a positive result can be obtained if $\beta > 2$. This is a consequence of the following theorem, where an even more general class of kernels is admissible. Recall that we always assume that the kernel k is subject to the conditions (K1) and (K2) with $|k|_1 > 0$.

Theorem 4.1 *Let k be a kernel with finite second moment, that is*

$$\sum_{j\in\mathbb{N}} k(j) j^2 < \infty,$$

and assume that k is non-increasing on $\mathbb{N}$. Then the corresponding operator $\mathcal{L}$ satisfies $CD(0, d)$ for some finite $d > 0$ of the form

$$d = C_0 \cdot \frac{|k|_1 \sum_j k(j) j^2}{k(1)^2},$$

where $C_0 > 0$ is some constant independent of k.

Proof The proof consists of several steps. By Lemma 2.1, we may assume that $u(0) = 0$ as well as $u(-j) = u(j)$ for all $j \in \mathbb{Z}$.

Step 1: Hölder estimate. Using that k has finite second moment, we have by Hölder's inequality that

$$\begin{aligned}\big(\mathcal{L}(u)(0)\big)^2 &= 4\left(\sum_{j=1}^{\infty} k(j)u(j)\right)^2 = 4\left(\sum_{j=1}^{\infty} \sqrt{k(j)}\, j\, \frac{\sqrt{k(j)}u(j)}{j}\right)^2 \\ &\le 4\sum_{j=1}^{\infty} k(j)j^2 \sum_{j=1}^{\infty} \frac{k(j)}{j^2}\, u(j)^2 = C(k)\sum_{j=1}^{\infty} \frac{k(j)}{j^2}\, u(j)^2. \end{aligned} \tag{21}$$

Step 2: Basic lower estimate for Γ_2. Evidently,

$$2\Gamma_2(u)(0) \ge \sum_{j=1}^{\infty} k(1)k(j)\big(u(j+1) - u(j) - u(1)\big)^2,$$

and therefore

$$\begin{aligned}\sum_{j=1}^{\infty} k(j)\big(u(j+1) - u(j)\big)^2 &\le 2\sum_{j=1}^{\infty} k(j)\big(u(j+1) - u(j) - u(1)\big)^2 + 2\sum_{j=1}^{\infty} k(j)u(1)^2 \\ &\le \frac{4}{k(1)}\,\Gamma_2(u)(0) + |k|_1 u(1)^2, \end{aligned} \tag{22}$$

where we use the fact that

$$a^2 = (a - b + b)^2 \le 2(a-b)^2 + 2b^2, \quad a, b \in \mathbb{R}.$$

From the basic estimate (8) we know that $2k(1)^2u(1)^2 \le \Gamma_2(u)(0)$, which together with (22) implies

$$\sum_{j=1}^{\infty} k(j)\big(u(j+1) - u(j)\big)^2 \le \big(\frac{4}{k(1)} + \frac{|k|_1}{2k(1)^2}\big)\Gamma_2(u)(0). \tag{23}$$

Observe that the assumptions on the kernel ensure that $k(1) > 0$.

Step 3: Estimating a squared weighted l_2*-norm of* u *by* Γ_2. Let $N \in \mathbb{N}$ with $N \ge 2$. For any $\delta > 0$ we have by Young's inequality and (23) that

$$\begin{aligned}\sum_{j=1}^{N} \frac{k(j)}{j}\, u(j+1)^2 &= \sum_{j=1}^{N} \frac{k(j)}{j}\, \big(u(j+1) - u(j) + u(j)\big)^2 \\ &\le \sum_{j=1}^{N} \frac{k(j)}{j}\, \left[\left(1 + \frac{j}{\delta}\right)\big(u(j+1) - u(j)\big)^2 + \big(1 + \frac{\delta}{j}\big)u(j)^2\right] \\ &\le \sum_{j=1}^{N} k(j)\, \left(1 + \frac{1}{\delta}\right)\big(u(j+1) - u(j)\big)^2 + k(1)\,(1+\delta)u(1)^2 \\ &\quad + \sum_{j=1}^{N-1} \frac{k(j+1)}{j+1}\, \left(1 + \frac{\delta}{j+1}\right) u(j+1)^2 \\ &\le M_\delta \Gamma_2(u)(0) + \sum_{j=1}^{N} \frac{k(j+1)}{j+1}\, \left(1 + \frac{\delta}{j+1}\right) u(j+1)^2, \end{aligned}$$

with

$$M_\delta = \big(1+\frac{1}{\delta}\big)\big(\frac{4}{k(1)} + \frac{|k|_1}{2k(1)^2}\big) + \frac{1}{2k(1)}\,(1+\delta).$$

This implies that

$$\sum_{j=1}^{N} \Lambda(j,\delta)\, u(j+1)^2 \le M_\delta \Gamma_2(u)(0), \tag{24}$$

where

$$\begin{aligned}\Lambda(j,\delta) &= \frac{k(j)}{j} - \frac{k(j+1)}{j+1} - \delta\,\frac{k(j+1)}{(j+1)^2} \\ &\ge k(j+1)\Big(\frac{1}{j} - \frac{1}{j+1} - \frac{\delta}{(j+1)^2}\Big) \\ &\ge k(j+1)\,\frac{1-\delta}{(j+1)^2},\end{aligned}$$

by monotonicity of the kernel k. Choosing $\delta \in (0,1)$, it follows from (24) that

$$\sum_{j=1}^{N} \frac{k(j+1)}{(j+1)^2}\, u(j+1)^2 \le \frac{M_\delta}{1-\delta}\,\Gamma_2(u)(0).$$

Sending $N \to \infty$ we obtain

$$\sum_{j=1}^{\infty} \frac{k(j+1)}{(j+1)^2}\, u(j+1)^2 \le \frac{M_\delta}{1-\delta}\,\Gamma_2(u)(0). \tag{25}$$

Step 4: Combining the estimates from Step 1 and 3. Using (21), (25) and $2k(1)^2u(1)^2 \le \Gamma_2(u)(0)$ we obtain

$$\begin{aligned}\big(\mathcal{L}(u)(0)\big)^2 &\le C(k)\left(k(1)\,u(1)^2 + \sum_{j=1}^{\infty} \frac{k(j+1)}{(j+1)^2}\, u(j+1)^2\right) \\ &\le C(k)\left(\frac{1}{2k(1)} + \frac{M_\delta}{1-\delta}\right)\Gamma_2(u)(0).\end{aligned}$$

The proof of Theorem 4.1 is complete. □

Theorem 4.1 covers a wide class of non-increasing (on $\mathbb{N}$) kernels which decay sufficiently fast. We illustrate this in the following corollary.

Corollary 4.1 *If the kernel k belongs to one of the subsequent classes of kernels, the corresponding operator $\mathcal{L}$ satisfies the $CD(0,d)$-inequality for some finite $d > 0$.*

(i) Power type kernels:

$$k(j) = \frac{c}{|j|^{1+\beta}}, \quad j \in \mathbb{Z}\backslash\{0\},$$

where $c > 0$ and $\beta > 2$.

(ii) Exponential kernels:

$$k(j) = c\, e^{-\delta |j|^{\alpha}}, \quad j \in \mathbb{Z}\backslash\{0\},$$

where $c, \alpha, \delta > 0$.

(iii) Mixed exponential and power type:

$$k(j) = c\, \frac{e^{-\delta |j|}}{|j|^{\alpha}}, \quad j \in \mathbb{Z}\backslash\{0\},$$

where $c, \alpha, \delta > 0$.

Proof It is easily seen that all the described kernels satisfy the assumptions of Theorem 4.1. □

Remark 4.1 Estimate (25) is one of the key estimates in the proof of Theorem 4.1. In the case of the power type kernel with $\beta > 2$ it can be rephrased as

$$\sum_{j=2}^{\infty} k(j)^{\frac{3+\beta}{1+\beta}} u(j)^2 \leq M \Gamma_2(u)(0)$$

(for symmetric u) with some constant $M > 0$. This is a much stronger estimate than the basic inequality $\sum_{j=2}^{\infty} k(j)^2 u(j)^2 \leq C\Gamma_2(u)(0)$ from (8), which holds for any kernel and symmetric u. The difference becomes more apparent for large values of β, since the exponent $\frac{3+\beta}{1+\beta}$ tends to 1 as $\beta \to \infty$.

Remark 4.2 The statement of Theorem 4.1 (with adapted bound for d) remains true in the more general case where k has a finite second moment but is merely assumed to be non-increasing for all $j \geq j_0$ with some fixed $j_0 \in \mathbb{N}$. The key idea to see this in the case $j_0 > 1$ is to split the sum

$$\sum_{j=1}^{\infty} \frac{k(j+1)}{(j+1)^2} u(j+1)^2 = \sum_{j=1}^{j_0-1} \frac{k(j+1)}{(j+1)^2} u(j+1)^2 + \sum_{j=j_0}^{\infty} \frac{k(j+1)}{(j+1)^2} u(j+1)^2. \tag{26}$$

The second sum on the right-hand side can be estimated from above by $\Gamma_2(u)(0)$ analogously to the argument given in the proof of Theorem 4.1. Instead of the $u(1)^2$ term one has a term involving $u(j_0)^2$, which can be estimated by employing (8). The finite sum on the right of (26) can also be controlled from above by $\Gamma_2(u)(0)$, again by the basic estimate (8).

Example 4.1 Let k be of the form

$$k(j) = c\, |j|^{\alpha} e^{-\delta |j|}, \quad j \in \mathbb{Z},$$

with $c, \alpha, \delta > 0$. Then the associated operator $\mathcal{L}$ satisfies the $CD(0, d)$-inequality for some finite $d > 0$. This is a consequence of Theorem 4.1 and Remark 4.2.

Theorems 3.1 and 4.1 show that for power type kernels the value $\beta = 2$ is a critical case. While $\beta > 2$ ensures $CD(0, d)$ for some finite $d > 0$, the letter fails to be true in the case $\beta < 2$. So what happens in the case $\beta = 2$? This is an interesting question, which we are not able to answer at the moment. However, if we add to a power type kernel with $\beta = 2$ a suitable logarithmic factor, then $CD(0, d)$ is still valid for some finite $d > 0$ as the resulting kernel has a finite second moment and thus Theorem 4.1 applies. We formulate this observation in the following corollary.

Corollary 4.2 *Suppose that the kernel k is of the form*

$$k(j) = \frac{c}{|j|^3[1+\log(|j|)]^{1+\alpha}}, \quad j \in \mathbb{Z}\backslash\{0\},$$

with $c, \alpha > 0$. Then the corresponding operator $\mathcal{L}$ satisfies the $CD(0,d)$-inequality for some finite $d > 0$.

5 Kernels with sparse unbounded support

In this section we consider situations where the support of the kernel k is unbounded but relatively small, in particular k is no longer non-increasing on $\mathbb{N}$. We begin with two positive results. Throughout this section the support of k is denoted by S and $S_+ = S \cap \mathbb{N}$.

Theorem 5.1 *Assume that the support of the kernel k is given by $S = \{\pm 2^l : l \in \mathbb{N}_0\}$ and that k satisfies the condition*

$$\sum_{l \in S_+} \left(\frac{k(2l)}{k(l)}\right)^2 < \infty. \tag{27}$$

Then the corresponding operator $\mathcal{L}$ satisfies $CD(0,d)$ for some finite $d > 0$.

Proof By Lemma 2.1, we may assume that $u(-j) = u(j)$ for all $j \in \mathbb{N}$ as well as $u(0) = 0$. In this situation

$$\begin{aligned}\big(\mathcal{L}(u)(0)\big)^2 &= \left(2\sum_{j\in S_+} k(2j)u(2j) + 2k(1)u(1)\right)^2 \\ &\le 8\left(\sum_{j\in S_+} k(2j)u(2j)\right)^2 + 8k(1)^2u(1)^2 \\ &\le 8\left(\sum_{j\in S_+} k(2j)u(2j)\right)^2 + 4\Gamma_2(u)(0),\end{aligned}$$

by the basic inequality (8). The first term on the right-hand side can be estimated by Hölder's inequality using (27) as follows.

$$\left(\sum_{j\in S_+} k(2j)u(2j)\right)^2 = \left(\sum_{j\in S_+} \frac{k(2j)}{k(j)}\, k(j)u(2j)\right)^2 \le \sum_{j\in S_+}\left(\frac{k(2j)}{k(j)}\right)^2 \sum_{j\in S_+} k(j)^2u(2j)^2.$$

Consequently,

$$\big(\mathcal{L}(u)(0)\big)^2 \le C\left(\sum_{j\in S_+} k(j)^2u(2j)^2 + \Gamma_2(u)(0)\right), \tag{28}$$

with some constant C which only depends on the kernel.

Next, using the inequality $a^2 \le 2(a-b)^2 + 2b^2$, $a, b \in \mathbb{R}$, we have

$$\begin{aligned}\sum_{j\in S_+} k(j)^2 u(2j)^2 &\le 2\sum_{j\in S_+} k(j)^2\big(u(2j)-2u(j)\big)^2 + 8\sum_{j\in S_+} k(j)^2 u(j)^2\\ &= 2\sum_{j\in S_+} k(j)^2\big(u(j+j)-u(j)-u(j)\big)^2 + 8\sum_{j\in S_+} k(j)^2 u(j)^2\\ &\le 4\Gamma_2(u)(0) + 4\Gamma_2(u)(0) = 8\Gamma_2(u)(0),\end{aligned}$$

by the representation formula for Γ_2 and the basic inequality (8). Combining the last estimate and (28) yields the asserted CD-inequality. □

Example 5.1 All kernels k of exponential type or mixed exponential and power type (see Corollary 4.1, (ii) and (iii)) satisfy condition (27), whereas purely algebraic kernels (that is, kernels of power type as described in (2)) do not possess this property.

It turns out that a $CD(0,d)$-inequality with finite $d > 0$ is still possible for larger gaps between the elements of the support. The extreme case is described in the following result.

Theorem 5.2 *Assume that the support of k is given by $S = \{\pm 3^l : l \in \mathbb{N}_0\}$ and that k is subject to the condition*

$$\sum_{l\in S_+} \frac{k(3l)}{k(l)} < \infty. \tag{29}$$

Then the corresponding operator $\mathcal{L}$ satisfies $CD(0,d)$ for some finite $d > 0$.

Proof Assuming w.l.o.g. that $u(0) = 0$ and $u(-j) = u(j)$ for all $j \in \mathbb{N}$ we have analogously to the first part of the proof of Theorem 5.1 that

$$\big(\mathcal{L}(u)(0)\big)^2 \le 8\left(\sum_{j\in S_+} k(3j)u(3j)\right)^2 + 4\Gamma_2(u)(0).$$

Using (29), Hölder's inequality gives

$$\begin{aligned}\left(\sum_{j\in S_+} k(3j)u(3j)\right)^2 &= \left(\sum_{j\in S_+} \sqrt{\frac{k(3j)}{k(j)}}\sqrt{k(3j)k(j)}\,u(3j)\right)^2\\ &\le \sum_{j\in S_+}\frac{k(3j)}{k(j)}\sum_{j\in S_+} k(3j)k(j)u(3j)^2,\end{aligned}$$

and thus there is a constant $C > 0$ depending only on k such that

$$\big(\mathcal{L}(u)(0)\big)^2 \le C\Big(\sum_{j\in S_+} k(3j)k(j)u(3j)^2 + \Gamma_2(u)(0)\Big). \tag{30}$$

The key idea of the proof is the following decomposition and estimate of the term $u(3j)^2$, where we use the fact that $(a+b+c)^2 \le 3(a^2+b^2+c^2)$ for all $a, b, c \in \mathbb{R}$.

$$\begin{aligned}u(3j)^2 &= \Big(\big(u(3j)-u(2j)+u(j)\big) + \big(u(2j)-2u(j)\big) + u(j)\Big)^2\\ &\le 3\big(u(3j)-u(2j)+u(j)\big)^2 + 3\big(u(2j)-2u(j)\big)^2 + 3u(j)^2.\end{aligned}$$

From (29) we obtain some constant $c_0 > 0$ such that $k(3j) \le c_0 k(j)$ for all $j \in S_+$. Using this comparability property and the symmetry of u and k this allows us to estimate

$$\begin{aligned}\sum_{j\in S_+} k(3j)k(j)u(3j)^2 &\le 3\sum_{j\in S_+} k(3j)k(-j)\big(u(3j-j)-u(3j)-u(-j)\big)^2\\ &\quad + 3c_0\sum_{j\in S_+} k(j)^2\big(u(2j)-2u(j)\big)^2 + 3c_0\sum_{j\in S_+} k(j)^2u(j)^2\\ &\le \Big(6+6c_0+\frac{3}{2}c_0\Big)\,\Gamma_2(u)(0).\end{aligned}$$

The last estimate and (30) yield the assertion. □

Example 5.2 All kernels k of exponential type or mixed exponential and power type (see Corollary 4.1, (ii) and (iii)) satisfy condition (29), whereas purely algebraic kernels (that is, kernels of power type as described in (2)) do not enjoy this property.

Writing $S_+ = \{x_l : l \in \mathbb{N}\}$ where the sequence x_l is strictly increasing, we have $x_{l+1} = 3x_l$ for all $l \in \mathbb{N}$ in the situation of Theorem 5.2. This is an extreme case as the next result will imply that for *any* kernel k with

$$x_{l+1} \ge 3x_l + 1, \quad l \in \mathbb{N},$$

the $CD(0,d)$-inequality fails for all finite $d > 0$.

Theorem 5.3 *Let k be a kernel with unbounded support S such that the following conditions are satisfied.*

(i) *There exists a number $N \in \mathbb{N}_0$ such that for all $j, l \in S$ with $\max\{|j|,|l|\} > N$ we have $j + l \notin S$.*
(ii) *For every number $m \in \mathbb{N}$ with $m > 2N$ there exists at most one decomposition $m = j+l$ with $j, l \in S$.*
(iii) *For every number $m \in \mathbb{N}$ with $m \le 2N$ there exist at most finitely many decompositions of the form $m = j + l$ with $j, l \in S$.*

Then the operator $\mathcal{L}$ associated with k fails to satisfy $CD(0,d)$ for all finite $d > 0$.

Proof Letting $n > 2N$ we define the function $u_n : \mathbb{Z} \to \mathbb{R}$ as follows. We set $u_n(j) = 0$ for every $j \in \mathbb{N}_0$ with $j \le 2N$. If $j > 2N$ we distinguish several cases. Suppose first that $j \in S_+$. Then we set $u_n(j) = \frac{1}{k(j)}$ whenever $j \le n$ and otherwise $u_n(j) = 0$. Now suppose that $j \notin S_+$ and $j = l_1 + l_2$ with (unique) $l_1, l_2 \in S$. In this case we set $u_n(j) = u_n(l_1) + u_n(l_2)$. For all remaining $j > 2N$ we set $u_n(j) = 0$. Finally, we put $u_n(-j) = u_n(j)$ for all $j \in \mathbb{N}$. Observe that u_n is well defined, in view of the assumptions (i) and (ii).

We have now

$$\mathcal{L}(u_n)(0) = 2\sum_{j=2N+1}^{n} k(j)u_n(j) = 2\sum_{j=2N+1}^{n} \chi_{S_+}(j),$$

where χ_{S_+} denotes the characteristic function of the set S_+.

On the other hand,

$$\begin{aligned}4\Gamma_2(u_n)(0) &= \sum_{j,l\in S} k(j)k(l)\big(u_n(j+l)-u_n(j)-u_n(l)\big)^2\\ &= 4\sum_{j\in S} k(j)^2u_n(j)^2 + \sum_{j,l\in S, l\ne -j} k(j)k(l)\big(u_n(j+l)-u_n(j)-u_n(l)\big)^2\end{aligned}$$

$$= 8 \sum_{j=2N+1}^{n} \chi_{S_+}(j) + \sum_{j,l\in S, 1\le|j+l|\le 2N} k(j)k(l)\big(u_n(j+l) - u_n(j) - u_n(l)\big)^2 + \sum_{j,l\in S, |j+l|>2N} k(j)k(l)\big(u_n(j+l) - u_n(j) - u_n(l)\big)^2.$$

The last term is equal to zero, by construction of u_n, and the second to the last one is bounded from above by a constant M which is independent of n, by assumption (iii). Since the support of k is unbounded, we have $\xi_n := \sum_{j=2N+1}^{n} \chi_{S_+}(j) \to \infty$ as $n \to \infty$ and thus it follows that

$$\limsup_{n\to\infty} \frac{4\Gamma_2(u_n)(0)}{\big(\mathcal{L}(u_n)(0)\big)^2} \le \limsup_{n\to\infty} \frac{8\xi_n + M}{4\xi_n^2} = 0,$$

which proves the theorem. □

Remark 5.1 In the situation of Theorem 5.3 with $N = 0$, the underlying graph $\mathcal{G}$ (described in the introduction after (5)) has a simple product structure. More precisely, the first and second neighbourhood of any vertex $x \in \mathbb{Z}$ are isomorphic to the first and second neighbourhood, respectively, of $(0, 0, \ldots)$ in the infinite Cartesian product $\times_{j\in S_+} k(j)\mathbb{Z}$. Note that, by definition, $CD(\kappa, d)$ at x only depends on the first and second neighbourhood of x.

6 Is a positive curvature possible?

In this section we show that $CD(\kappa, \infty)$ with $\kappa > 0$ does not hold for arbitrary kernels k provided that k is non-increasing on $\mathbb{N}$.

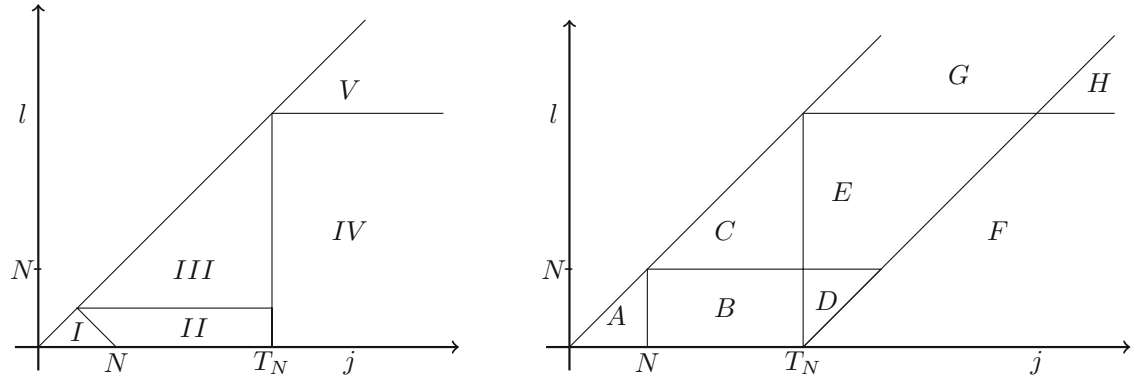
Theorem 6.1 *Suppose that the kernel k is non-increasing on $\mathbb{N}$. Then the corresponding operator $\mathcal{L}$ fails to satisfy the $CD(\kappa, \infty)$-condition for all $\kappa > 0$. This is even true when the class of admissible functions is restricted to compactly supported functions.*

Proof Note that for kernels with finite second moment we immediately obtain the claim for the admissible, but non-compactly supported function $u(j) = j$. To show the desired result for all non-increasing kernels we use the following cut-off procedure. Let $(T_N)_{N\in\mathbb{N}}$ be an increasing sequence of positive real numbers such that $T_N \to \infty$ and $0 < C \le \frac{T_N - N}{T_N}$ holds for some constant C and all $N \ge 1$. We define $u_N : \mathbb{Z} \to \mathbb{R}$ by $u_N(j) = j\varphi_N(j)$, $j \in \mathbb{Z}$, where $\varphi_N : \mathbb{Z} \to \mathbb{R}$ is symmetric and such that $\varphi_N(j) = 1$ for $|j| \le N$, $\varphi_N(j) = \frac{T_N - j}{T_N - N}$ if $N + 1 \le j \le T_N$ and $\varphi_N(j) = 0$ for $j > T_N$. For the sake of convenience we assume that N is even in order to ensure $\frac{N}{2} \in \mathbb{N}$. In what follows we also drop the index N in the notation of the cut-off function, that is we just write φ instead of φ_N. Observe that the definition of φ implies that $|\varphi(j) - \varphi(l)| \le \frac{1}{T_N}|j - l|$, for $j, l \in \mathbb{Z}$.

On one hand we have

$$\Gamma(u_N)(0) = 2\sum_{l=1}^{T_N} k(l)l^2\varphi(l)^2 = 2\left(\sum_{l=1}^{N} k(l)l^2 + \sum_{l=N+1}^{T_N} k(l)l^2\varphi(l)^2\right). \tag{31}$$

On the other hand, due to u_N being anti-symmetric, we have



(A) Summation of terms including $j + l$. **(B)** Summation of terms including $j - l$.

Fig. 3 Splitting the domain of summation

$$\Gamma_2(u_N)(0) \leq 2\sum_{j=1}^{\infty}\sum_{l=1}^{j} k(j)k(l)\left[(u_N(j+l) - u_N(j) - u_N(l))^2 + (u_N(j-l) - u_N(j) + u_N(l))^2\right].$$

In order to estimate the Γ_2-term, we distinguish several cases (see Fig. 3). The strategy in each of the following cases will be the same. In fact, we will always estimate the respective part of the Γ_2-sum by some expression occurring in (31) multiplied by some positive sequence converging to zero as $N \to \infty$. This implies that for any given constant $\kappa > 0$ there exists $N(\kappa) \in \mathbb{N}$ such that

$$0 < \Gamma_2(u_N)(0) \leq \kappa\Gamma(u_N)(0)$$

holds for any $N \geq N(\kappa)$, which will be sufficient to establish the claim. To simplify the following presentation, we use the notation $\delta(N)$ for a positive sequence, which may differ from line to line, such that $\delta(N) \to 0$ as $N \to \infty$.

Note that we can write for $j, l \in \mathbb{Z}$

$$(u_N(j+l) - u_N(j) - u_N(l))^2 = \big(j\,(\varphi(j+l) - \varphi(j)) + l\,(\varphi(j+l) - \varphi(l))\big)^2. \quad (32)$$

I: $1 \leq l \leq j$, $j + l \leq N$: Here we have $\varphi(j+l) = \varphi(j) = \varphi(l) = 1$ and therefore all the corresponding summands are vanishing due to (32).
II: $1 \leq l \leq \frac{N}{2}$, $N + 1 \leq j + l$, $\frac{N}{2} + 1 \leq j \leq T_N$: Due to our choice of φ, we can control the bracket in (32) by $\frac{j^2 l^2}{T_N^2}$ multiplied by some positive constant, and thus the part of $\Gamma_2(u_N)(0)$ that corresponds to the present case can be controlled by the sums

$$\sum_{l=1}^{N} k(l)l^2 \frac{1}{T_N^2} \sum_{j=\frac{N}{2}}^{T_N} k(j)j^2 \leq \sum_{l=1}^{N} k(l)l^2 \sum_{j=\frac{N}{2}}^{T_N} k(j) \leq \delta(N)\Gamma(u_N)(0), \quad (33)$$

since $k \in l_1(\mathbb{Z})$.
III: $\frac{N}{2} + 1 \leq l \leq j \leq T_N$: We use the inequality $(a - b - c)^2 \leq 3\left(a^2 + b^2 + c^2\right)$, $a, b, c \in \mathbb{R}$, and estimate the squares resulting from the left-hand side of (32) and the previous inequality separately. We have

$$\sum_{j=\frac{N}{2}+1}^{T_N} k(j)j^2\varphi(j)^2 \sum_{l=\frac{N}{2}+1}^{j} k(l) \leq \sum_{j=\frac{N}{2}+1}^{T_N} k(j)j^2\varphi(j)^2 \sum_{l=\frac{N}{2}+1}^{\infty} k(l) \leq \delta(N)\Gamma(u_N)(0).$$

Similarly,

$$\sum_{l=\frac{N}{2}+1}^{T_N} k(l)l^2\varphi(l)^2 \sum_{j=l}^{T_N} k(j) \le \sum_{l=\frac{N}{2}+1}^{T_N} k(l)l^2\varphi(l)^2 \sum_{j=\frac{N}{2}+1}^{T_N} k(j) \le \delta(N)\Gamma(u_N)(0).$$

Since $(j+l)^2 \le (2j)^2$ and $\varphi(j+l) \le \varphi(j)$, by using the first estimate, we also obtain the desired estimate for the remaining third term.
IV: $1 \le l \le T_N + 1$, $j \ge T_N$: We have $\varphi(j) = \varphi(j+l) = 0$ and thus it remains to estimate

$$\sum_{l=1}^{T_N} k(l)l^2\varphi(l)^2 \sum_{j=T_N+1}^{\infty} k(j) \le \delta(N)\Gamma(u_N)(0).$$

V: $T_N + 1 \le l \le j$: Now $\varphi(l) = \varphi(j) = \varphi(j+l) = 0$ and therefore the estimate is trivial.
A: $1 \le l \le j \le N$: This case follows for the same reason as I.
B: $1 \le l \le N$, $N+1 \le j \le T_N$: We now write

$$(u_N(j-l) - u_N(j) + u_N(l))^2 = \big(j\,(\varphi(j-l) - \varphi(j)) + l\,(\varphi(l) - \varphi(j-l))\big)^2$$

and note that we can, as in the case II, control this expression by $\frac{j^2l^2}{T_N^2}$ multiplied by some positive constant. This yields the same expression as in (33) with the lower limit of the inner sum replaced by $N+1$. Hence we get the desired estimate.
C: $N+1 \le l \le j$, $j \le T_N$: As in III we estimate the single terms, starting with

$$\begin{aligned}\sum_{l=N+1}^{T_N} k(l) \sum_{j=l}^{T_N} k(j)\,(j-l)^2\varphi(j-l)^2 &= \sum_{l=N+1}^{T_N} k(l) \sum_{j=1}^{T_N-l} k(j+l)\,j^2\varphi(j)^2 \\ &\le \sum_{l=N+1}^{T_N} k(l) \sum_{j=1}^{T_N} k(j)j^2\varphi(j)^2 \le \delta(N)\Gamma(u_N)(0),\end{aligned}$$

where we applied that k is monotone. Next, we estimate

$$\sum_{j=N}^{T_N} k(j)j^2\varphi(j)^2 \sum_{l=N}^{j} k(l) \le \sum_{j=N}^{T_N} k(j)j^2\varphi(j)^2 \sum_{l=N}^{T_N} k(l) \le \delta(N)\Gamma(u_N)(0)$$

and

$$\sum_{l=N}^{T_N} k(l)l^2\varphi(l)^2 \sum_{j=l}^{T_N} k(j) \le \sum_{l=N}^{T_N} k(l)l^2\varphi(l)^2 \sum_{j=N}^{T_N} k(j) \le \delta(N)\Gamma(u_N)(0).$$

D: $T_N + 1 \le j$, $j - T_N \le l \le N$: We have $\varphi(j) = 0$ and hence there are two terms left to consider. First, obtain that

$$\sum_{j=T_N+1}^{T_N+N} k(j) \sum_{l=j-T_N}^{N} k(l)\,(j-l)^2\,\varphi(j-l)^2 = \sum_{j=T_N+1}^{T_N+N} k(j) \sum_{l=-T_N}^{N-j} k(j+l) l^2 \varphi(l)^2$$
$$= \sum_{j=T_N+1}^{T_N+N} k(j) \sum_{l=j-N}^{T_N} k(j-l) l^2 \varphi(l)^2$$
$$\leq \sum_{j=T_N+1}^{T_N+N} k(j) \sum_{l=j-N}^{T_N} k(l) l^2 \varphi(l)^2 \leq \delta(N)\Gamma(u_N)(0),$$

since in this case $j - l \geq l$ holds. In addition, since $\phi(l) = 1$ we have the upper bound

$$\sum_{j=T_N+1}^{T_N+N} k(j) \sum_{l=j-T_N}^{N} k(l) l^2 \leq \sum_{j=T_N+1}^{T_N+N} k(j) \sum_{l=1}^{N} k(l) l^2 \leq \delta(N)\Gamma(u_N)(0).$$

E: $N + 1 \leq l \leq T_N$, $T_N + 1 \leq j \leq T_N + l$: Again, we have $\varphi(j) = 0$. On the one hand it holds

$$\sum_{l=N+1}^{T_N} k(l) \sum_{j=T_N+1}^{T_N+l} k(j)\,(j-l)^2\,\varphi(j-l)^2 = \sum_{l=N+1}^{T_N} k(l) \sum_{j=T_N-l}^{T_N} k(j+l) j^2 \varphi(j)^2$$
$$\leq \sum_{l=N+1}^{T_N} k(l) \sum_{j=1}^{T_N} k(j) j^2 \varphi(j)^2 \leq \delta(N)\Gamma(u_N)(0)$$

and on the other hand

$$\sum_{l=N+1}^{T_N} k(l) l^2 \varphi(l)^2 \sum_{j=T_N+1}^{T_N+l} k(j) \leq \sum_{l=N+1}^{T_N} k(l) l^2 \varphi(l)^2 \sum_{j=T_N+1}^{2\,T_N} k(j) \leq \delta(N)\Gamma(u_N)(0).$$

F: $1 \leq l \leq T_N$, $T_N + l + 1 \leq j$: Here we have $\varphi(j) = \varphi(j - l) = 0$. It remains to consider

$$\sum_{l=1}^{T_N} k(l) l^2 \varphi(l)^2 \sum_{j=T_N+l}^{\infty} k(j) \leq \sum_{l=1}^{T_N} k(l) l^2 \varphi(l)^2 \sum_{j=T_N}^{\infty} k(j) \leq \delta(N)\Gamma(u_N)(0).$$

G: $T_N + 1 \leq l, l \leq j \leq T_N + l$: In this case $\varphi(j)$ and $\varphi(l)$ vanish. Hence, we only need to consider

$$\sum_{l=T_N+1}^{\infty} k(l) \sum_{j=l}^{T_N+l} k(j)\,(j-l)^2\,\varphi(j-l)^2 = \sum_{l=T_N+1}^{\infty} k(l) \sum_{j=1}^{T_N} k(j+l) j^2 \varphi(j)^2$$
$$\leq \sum_{l=T_N+1}^{\infty} k(l) \sum_{j=1}^{T_N} k(j) j^2 \varphi(j)^2 \leq \delta(N)\Gamma(u_N)(0).$$

H: $T_N + 1 \leq l$, $j \leq l + T_N + 1$: This case is obvious, since $\varphi(j - l) = \varphi(l) = \varphi(j) = 0$. □

Remark 6.1 A careful inspection of the proof shows that the monotonicity assumption on k can be weakened by assuming that there exists some $j_0 \in \mathbb{N}$, such that k non-increasing on $\mathbb{N}\backslash\{1, \ldots, j_0\}$.

Acknowledgements Adrian Spener and Rico Zacher are supported by the DFG (Project Number 355354916, GZ ZA 547/4-1). The authors thank the anonymous referee for the valuable comments, in particular the interesting observation described in Remark 5.1.

References

1. Bauer, F., Chung, F., Lin, Y., Liu, Y.: Curvature aspects of graphs. Proc. Am. Math. Soc. **145**(5), 2033–2042 (2017)
2. Bakry, D., Émery, M.: Diffusions hypercontractives. In: Séminaire de probabilités, XIX, 1983/84, volume 1123 of Lecture Notes in Mathematics, pp. 177–206. Springer, Berlin (1985)
3. Bakry, D., Gentil, I., Ledoux, M.: Analysis and geometry of Markov diffusion operators, volume 348 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Cham (2014)
4. Bauer, F., Horn, P., Lin, Y., Lippner, G., Mangoubi, D., Yau, S.-T.: Li–Yau inequality on graphs. J. Differ. Geom. **99**(3), 359–405 (2015)
5. Ciaurri, Ó., Roncal, L., Raúl Stinga, P., Torrea, J.L., Varona, J.L.: Nonlocal discrete diffusion equations and the fractional discrete Laplacian, regularity and applications. Adv. Math. **330**, 688–738 (2018)
6. Dier, D., Kassmann, M., Zacher, R.: Discrete versions of the Li–Yau gradient estimate (2017). ArXiv e-prints arXiv:1701.04807
7. Erbar, M., Fathi, M.: Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature. J. Funct. Anal. **274**(11), 3056–3089 (2018)
8. Erbar, M., Kuwada, K., Sturm, K.-T.: On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. Invent. Math. **201**(3), 993–1071 (2015)
9. Erbar, M., Maas, J.: Ricci curvature of finite Markov chains via convexity of the entropy. Arch. Ration. Mech. Anal. **206**(3), 997–1038 (2012)
10. Erbar, M.: Gradient flows of the entropy for jump processes. Ann. Inst. Henri Poincaré Probab. Stat. **50**(3), 920–945 (2014)
11. Fathi, M., Shu, Y.: Curvature and transport inequalities for Markov chains in discrete spaces. Bernoulli **24**(1), 672–698 (2018)
12. Jost, J., Liu, S.: Ollivier's Ricci curvature, local clustering and curvature-dimension inequalities on graphs. Discrete Comput. Geom. **51**(2), 300–322 (2014)
13. Klartag, B., Kozma, G., Ralli, P., Tetali, P.: Discrete curvature and abelian groups. Can. J. Math. **68**(3), 655–674 (2016)
14. Keller, M., Münch, F.: Gradient estimates, Bakry–Emery Ricci curvature and ellipticity for unbounded graph Laplacians (2018). arXiv e-prints, arXiv:1807.10181
15. Li, P.: Geometric Analysis. Cambridge Studies in Advanced Mathematics, vol. 134. Cambridge University Press, Cambridge (2012)
16. Liu, S., Münch, F., Peyerimhoff, N.: Bakry–Émery curvature and diameter bounds on graphs. Calc. Var. Partial Differ. Equ. **57**(2), 67 (2018)
17. Lott, J., Villani, C.: Ricci curvature for metric-measure spaces via optimal transport. Ann. Math. (2) **169**(3), 903–991 (2009)
18. Lin, Y., Yau, S.-T.: Ricci curvature and eigenvalue estimate on locally finite graphs. Math. Res. Lett. **17**(2), 343–356 (2010)
19. Maas, J.: Gradient flows of the entropy for finite Markov chains. J. Funct. Anal. **261**(8), 2250–2292 (2011)
20. Münch, F.: Li-Yau inequality on finite graphs via non-linear curvature dimension conditions. J. Math. Pures Appl. **120**(9), 130–164 (2018)
21. Münch, F.: Remarks on curvature dimension conditions on graphs. Calc. Var. Partial Differ. Equ. **56**(1), 11 (2017)
22. Najman, L., Romon, P. (eds.): Modern approaches to discrete curvature, volume 2184 of Lecture Notes in Mathematics. Springer, Cham (2017)
23. Ollivier, Y.: Ricci curvature of Markov chains on metric spaces. J. Funct. Anal. **256**(3), 810–864 (2009)
24. Sturm, K.-T.: On the geometry of metric measure spaces. I. Acta Math. **196**(1), 65–131 (2006)
25. Sturm, K.-T.: On the geometry of metric measure spaces. II. Acta Math. **196**(1), 133–177 (2006)

26. Spener, A., Weber, F., Zacher, R.: The fractional Laplacian has infinite dimension. Comm. Partial Diff. Equ. https://doi.org/10.1080/03605302.2019.1663434
27. Villani, C.: Optimal transport, volume 338 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Berlin (2009) (Old and new)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

5. The fractional Laplacian has infinite dimension

Authors: Adrian Spener, Frederic Weber and Rico Zacher

Journal: Communications in Partial Differential Equations **45** (2020), no. 1, 57–75.

DOI: https://doi.org/10.1080/03605302.2019.1663434

Copyright: *This is an Accepted Manuscript of an article published by Taylor & Francis in Communications in Partial Differential Equations on 20. September 2019, available at: http://www.tandfonline.com/10.1080/03605302.2019.1663434.*

The Fractional Laplacian has Infinite Dimension

Adrian Spener[*], Frederic Weber[†], and Rico Zacher[‡]

We show that the fractional Laplacian on $\mathbb{R}^d$ fails to satisfy the Bakry-Émery curvature-dimension inequality $CD(\kappa, N)$ for all curvature bounds $\kappa \in \mathbb{R}$ and all finite dimensions $N > 0$.

Keywords: Gamma Calculus, Bakry-Émery Curvature-Dimension Condition, Fractional Laplacian.
MSC(2010): 35R11 (primary), 47D07, 26D10, 60G22 (secondary).

1 Introduction

The fractional Laplacian is an important example of a non-local operator and has received considerable attention in recent years. It appears in many fields of analysis and probability theory, for example in the theory of stochastic processes as generator of a class of pure jump processes, so-called α-stable Lévy processes [App09, BL02, BBCK09]. It can be further identified with a certain Dirichlet-to-Neumann operator, which is the key property of the so-called harmonic extension method as described by Caffarelli and Silvestre in [CS07]. During the last decade, non-local PDEs involving a fractional Laplacian or other related non-local operators have been studied intensively, see e.g. [Sil06, Sil07, CS07, Kas09, SV09, BBCK09, CS09, CV10, CV11, FK13, BV14, CS14, KMS15, CS15, IS16].
In the present paper we study the fractional Laplacian in $\mathbb{R}^d$ from a geometric analyst's point of view. More precisely, we are interested in the question whether the fractional Laplacian satisfies a curvature-dimension (CD) inequality with *finite* dimension. Curvature-dimension inequalities play a central role in the theory of Markov semigroups and operators and related functional inequalities like e.g. the Poincaré or logarithmic Sobolev inequality (cf. [BGL14]). They also constitute an important tool in geometric analysis as they encode information of the geometric properties (curvature and dimension) of the underlying structure, e.g. a Riemannian manifold [Li12].
There exist several different notions of CD-inequalities. In this paper, we use the classical one, which was introduced by Bakry and Émery and is formulated in terms of the carré du champ operator Γ and its iterated operator Γ_2 associated with the infinitesimal generator $\mathcal{L}$ of a Markov semigroup, see [BE85]. Another widely used approach relies on the theory of optimal transport, see [Vil09]. For Riemannian manifolds, both approaches lead to the same notion of lower bounds for the Ricci curvature.
To describe the notion of CD-inequality in the Bakry-Émery sense, let $\mathcal{L}$ be the generator of a symmetric Markov semigroup with invariant reversible measure μ on the state space E. The bilinear operators Γ and Γ_2 are defined by

$$\Gamma(u,v) = \frac{1}{2}\left(\mathcal{L}(uv) - u\mathcal{L}v - v\mathcal{L}u\right),$$
$$\Gamma_2(u,v) = \frac{1}{2}\left(\mathcal{L}\Gamma(u,v) - \Gamma(u,\mathcal{L}v) - \Gamma(\mathcal{L}u,v)\right)$$

on a suitable algebra $\mathcal{A}$ of real-valued functions u, v defined on the underlying state space E. Furthermore, one sets

$$\Gamma(u) := \Gamma(u,u) = \frac{1}{2}\mathcal{L}(u^2) - u\mathcal{L}u, \tag{1}$$

$$\Gamma_2(u) := \Gamma_2(u,u) = \frac{1}{2}\mathcal{L}\Gamma(u) - \Gamma(u,\mathcal{L}u). \tag{2}$$

[*]Institute of Applied Analysis, University of Ulm, Helmholtzstraße 18, 89081 Ulm, Germany. adrian.spener@uni-ulm.de
[†]Institute of Applied Analysis, University of Ulm, Helmholtzstraße 18, 89081 Ulm, Germany. frederic.weber@uni-ulm.de
[‡]Institute of Applied Analysis, University of Ulm, Helmholtzstraße 18, 89081 Ulm, Germany. rico.zacher@uni-ulm.de

Let $\kappa \in \mathbb{R}$ and $N \in (0, \infty]$. The operator $\mathcal{L}$ is said to satisfy the *Bakry-Émery curvature-dimension inequality* $CD(\kappa, N)$ with dimension N and curvature κ (lower bound) at $x \in E$ if

$$\Gamma_2(u)(x) \geq \frac{1}{N}\big((\mathcal{L}u)(x)\big)^2 + \kappa\Gamma(u)(x) \tag{3}$$

for all functions $u : E \to \mathbb{R}$ in a sufficiently rich class $\mathcal{A}$ of functions. We further say that $\mathcal{L}$ satisfies the $CD(\kappa, N)$-inequality, if (3) holds μ-almost everywhere for all $u \in \mathcal{A}$, cf. [BGL14, Sect. 1.16].
As an illustrating example we let $E = (M, g)$ be a Riemannian manifold with canonical Riemannian measure μ_g and $\mathcal{L} = \Delta_g$ the Laplace-Beltrami operator. Using the Bochner-Lichnerowicz formula one obtains that $CD(\kappa, N)$ is equivalent to $\mathrm{Ric}_g(x) \geq \kappa g(x)$ and $\dim M \leq N$.

In this paper, we take as operator $\mathcal{L}$ the fractional Laplacian of order $\beta \in (0, 2)$ on $E = \mathbb{R}^d$, which can be defined by

$$\mathcal{L}u(x) := -(-\Delta u)^{\frac{\beta}{2}}(x) = c_{\beta,d} \int_{\mathbb{R}^d} \frac{u(x+h) - 2u(x) + u(x-h)}{|h|^{d+\beta}}\,\mathrm{d}h, \quad x \in \mathbb{R}^d, \tag{4}$$

for functions u which are sufficiently regular and do not grow too fast as $|x| \to \infty$, for instance functions in the Schwartz space $\mathcal{S}(\mathbb{R}^d)$. The normalising constant $c_{\beta,d}$ is given by

$$c_{\beta,d} := \frac{2^\beta \Gamma(\frac{d+\beta}{2})}{\pi^{\frac{d}{2}} |\Gamma(-\frac{\beta}{2})|},$$

where Γ now denotes the Gamma function. We refer to [Kwa17, Theorem 1.1. (e)], which collects several other equivalent definitions of the fractional Laplacian. Furthermore, the Lebesgue measure on $\mathbb{R}^d$ plays the role of the invariant reversible measure μ.
Does the fractional Laplacian satisfy the $CD(\kappa, N)$-inequality with some $\kappa \in \mathbb{R}$ and *finite* $N > 0$? This question in the special case $\kappa = 0$ has been recently identified as a key open problem in extending the Gamma calculus of Bakry and Émery to non-local operators such as the fractional Laplacian. It has been posed in the recent survey of Garofalo in [Gar19, (20.14)], who also describes possible applications such as *non-local Li-Yau inequalities*. We remark that the fractional Laplacian satisfies $CD(0, \infty)$ for all $\beta \in (0, 2)$. This is a trivial consequence of the representation formula for the associated Γ_2 operator, which implies that $\Gamma_2(u) \geq 0$ for all u, cf. formula (7) below.
Recall that in the case of the classical Laplacian Δ on $\mathbb{R}^d$ (which corresponds to the limiting case $\beta = 2$) there holds

$$\Gamma_2(u) = |\mathrm{Hess}(u)|^2_{\mathrm{HS}} \geq \frac{1}{d}(\Delta u)^2.$$

That is, the Laplacian satisfies the $CD(0, d)$-inequality. By means of this inequality and the chain rule for the Laplacian one can show that positive solutions u of the heat equation in $\mathbb{R}^d$ satisfy the celebrated Li-Yau gradient estimate or differential Harnack inequality ([LY86])

$$|\nabla(\log u)(t,x)|^2 - \partial_t(\log u)(t,x) \leq \frac{d}{2t}, \quad t > 0,\, x \in \mathbb{R}^d.$$

Integrating this estimate along suitable paths in space-time, one can deduce a sharp version of the parabolic Harnack inequality. More generally, any Markov diffusion operator for which $CD(0, N)$ holds true with a finite $N > 0$ satisfies a Li-Yau inequality by [BGL14, Cor. 6.7.5].

As the main result of this paper, we give a negative answer to Garofalo's question by proving there exists no finite $N > 0$ such that the fractional Laplacian satisfies $CD(0, N)$. As a corollary to this result we obtain by means of a scaling argument that the situation is even worse, namely there exist no pair $(\kappa, N) \in \mathbb{R} \times (0, \infty)$ such that the fractional Laplacian satisfies the $CD(\kappa, N)$-condition. In particular, a negative curvature term does not save the CD-inequality with finite dimension. In this sense, the fractional Laplacian *has infinite dimension*. Our main result reads as follows. Note that here we refer to the representation formulas (6) and (7) for $\Gamma(u, v)$ and $\Gamma_2(u)$, respectively (see Lemma 2.1 below), which hold for sufficiently regular functions which do not grow too fast at infinity, in particular for $\mathcal{C}^\infty_c$-functions.

Theorem 1.1. *Let $\beta \in (0, 2)$, $d \in \mathbb{N}$ and $\mathcal{L}$ be the fractional Laplacian from* (4) *with associated operators Γ and Γ_2 given below by* (6) *and* (7)*, respectively. Then for any $R > 0$, any $\kappa \in \mathbb{R}$ and any finite $N > 0$ there exists some smooth and compactly supported function $u \in \mathcal{C}^\infty_c(\mathbb{R}^d)$ such that*

$$\Gamma_2(u)(x) < \kappa\Gamma(u)(x) + \frac{1}{N}\left(\mathcal{L}(u)(x)\right)^2 \tag{5}$$

for all $x \in \mathbb{R}^d$ with $|x| \leq R$.

Theorem 1.1 follows from Theorem 3.8 and Proposition 2.3.

Our proof of the first part proceeds in several steps. We first consider the case $d = 1$. Given a finite $N > 0$, we construct an *unbounded* function u for which $CD(0, N)$ fails at $x = 0$. The key idea, which is motivated by [SWZ19, Theorem 3.1], is to consider functions which are smooth enough near $x = 0$ and equal to $|x|^{\beta-\varepsilon}$ for $|x| \geq 1$ with small $\varepsilon > 0$. In the very recent work [SWZ19], which deals with CD-inequalities for non-local discrete operators on the lattice $\mathbb{Z}$, the described family of functions is used to prove that $CD(0, N)$ fails for all finite $N > 0$ for a certain class of operators with power type kernel, in particular for all powers $-(-\Delta_{disc})^{\frac{\beta}{2}}$ of the discrete Laplacian Δ_{disc} on $\mathbb{Z}$ with $\beta \in (0, 2)$.
By considering functions which depend only on one real variable, the one-dimensional counterexample can be extended to the multi-dimensional case. To get a counterexample with compactly supported functions, we carry out an approximation argument which involves carefully chosen cut-off functions. The argument is very technical, since the estimates for the Γ_2 operator require distinguishing several cases with respect to the domain of integration. Finally, to see the failure of $CD(0, N)$ on an arbitrary ball around zero, we use a continuity argument to first prove the result for a sufficiently small ball and then use the scaling properties of the fractional Laplacian and its associated Γ_2 operator to get the desired result on any ball.

Acknowledgement

Adrian Spener and Rico Zacher are supported by the DFG (project number 355354916, GZ ZA 547/4-1).

2 Gamma Calculus and the Fractional Laplacian

Let u be smooth enough and growing slowly enough such that the following integrals exist, for instance let u be in the Schwartz space $\mathcal{S}(\mathbb{R}^d)$. Then the quadratic operators from the Γ-calculus are given as follows.

Lemma 2.1. *Let $0 < \beta < 2$ and $\mathcal{L}$ be the fractional Laplacian from (4) with associated Γ from (1) and Γ_2 from (2). Then*

$$\Gamma(u, v)(x) = c_{\beta,d} \int_{\mathbb{R}^d} \frac{(u(x + h) - u(x))(v(x + h) - v(x))}{|h|^{d+\beta}} \, \mathrm{d}h \tag{6}$$

and

$$\Gamma_2(u)(x) = c_{\beta,d}^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{[u(x + h + \sigma) - u(x + h) - u(x + \sigma) + u(x)]^2}{|h|^{d+\beta}|\sigma|^{d+\beta}} \, \mathrm{d}h \, \mathrm{d}\sigma. \tag{7}$$

Proof. The first equality is given for instance in [Gar19, Lemma 20.2], the second one can be obtained from a short calculation, where one uses $a^2 - b^2 - 2b(a - b) = (a - b)^2$ for $a = u(x + h + \sigma) - u(x + \sigma)$, $b = u(x + h) - u(x)$. □

This immediately implies that for all $0 < \beta < 2$ the fractional Laplacian satisfies $CD(0, \infty)$ at all $x \in \mathbb{R}^d$.

Lemma 2.2. *For fixed $x \in \mathbb{R}^d$ in (5) we may assume without loss of generality that $x = 0$ and $u(0) = 0$.*

Proof. Replacing an arbitrary function v with $u(y) := v(x - y) - v(x)$ we obtain a function u satisfying the assertions of Lemma 2.2 as well as $\mathcal{L}u(0) = \mathcal{L}v(x)$, $\Gamma u(0) = \Gamma v(x)$ and $\Gamma_2 u(0) = \Gamma_2 v(x)$. □

Using a scaling argument, we may assume without loss of generality that $\kappa = 0$ in (5), i.e. to show that the fractional Laplacian has infinite dimension we may assume that $\kappa = 0$. This is the content of the next proposition.

Proposition 2.3. *$CD(\kappa, N)$ with some finite $N > 0$ and some $\kappa \in \mathbb{R}$ implies $CD(0, N)$.*

Proof. By means of Lemma 2.2 we may assume that $x = 0$ and $u(0) = 0$. The assertion is evident in the case $\kappa > 0$. Now assume that $CD(-\kappa, N)$ holds for some $\kappa > 0$ and finite $N > 0$. Then we have

$$\frac{1}{N}(\mathcal{L}(u)(0))^2 \leq \Gamma_2(u)(0) + \kappa\Gamma(u)(0) \tag{8}$$

for all u smooth and growing slowly enough. For $\lambda > 0$ we let $v(x) = u(\lambda x)$, and calculate

$$\mathcal{L}(v)(0) = c_{\beta,d}\lambda^{d+\beta} \int_{\mathbb{R}^d} \frac{u(\lambda h) + u(-\lambda h)}{|\lambda h|^{d+\beta}} \, \mathrm{d}h = c_{\beta,d}\lambda^{\beta} \int_{\mathbb{R}^d} \frac{u(x) + u(-x)}{|x|^{d+\beta}} \, \mathrm{d}x = \lambda^{\beta}\mathcal{L}(u)(0).$$

Similarly we find

$$\Gamma(v)(0) = c_{\beta,d}\lambda^{d+\beta}\int_{\mathbb{R}^d}\frac{u(\lambda h)^2}{|\lambda h|^{d+\beta}}\,\mathrm{d}h = \lambda^\beta\Gamma(u)(0)$$

and

$$\Gamma_2(v)(0) = c_{\beta,d}^2\lambda^{2d+2\beta}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{[u(\lambda h+\lambda\sigma)-u(\lambda h)-u(\lambda\sigma)]^2}{|\lambda h|^{d+\beta}|\lambda\sigma|^{d+\beta}}\,\mathrm{d}h\,\mathrm{d}\sigma = \lambda^{2\beta}\Gamma_2(u)(0).$$

Whence, from (8) we find $\frac{1}{N}(\mathcal{L}(v)(0))^2 \le \Gamma_2(v)(0)+\lambda^\beta\kappa\Gamma(v)(0)$ for all $\lambda>0$, which implies the assertion. □

Let us give a small lemma which will simplify the calculation in the third section.

Lemma 2.4. *Let $d=1$ and u be an even function with sufficient smoothness and growing slowly enough which satisfies $u(0)=0$, then*

$$\Gamma_2(u)(0) = 4c_{\beta,1}^2\int_0^\infty\int_0^h\frac{[u(h+\sigma)-u(h)-u(\sigma)]^2+[u(h-\sigma)-u(h)-u(\sigma)]^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h. \tag{9}$$

Proof. Denote the integrand by $A(h,\sigma)=\frac{[u(h+\sigma)-u(h)-u(\sigma)]^2}{h^{1+\beta}\sigma^{1+\beta}}$, then $A(h,\sigma)=A(-h,-\sigma)$ and $A(h,\sigma)=A(\sigma,h)$. The result follows from Fubini's Theorem. □

Lemma 2.5. *Let $u\colon\mathbb{R}^d\to\mathbb{R}$ be smooth and growing slowly enough and assume that u depends only on one variable, i.e. $u(x_1,\dots,x_d)=v(x_1)$ for some $v\colon\mathbb{R}\to\mathbb{R}$. Denote the fractional Laplacian on $\mathbb{R}^d$ temporarily by $\mathcal{L}^{(d)}$, similarly for the corresponding Γ-operators. Then*

$$\mathcal{L}^{(d)}u(0) = A_{d,\beta}\mathcal{L}^{(1)}(v)(0),\quad \Gamma^{(d)}(u)(0)=A_{d,\beta}\Gamma^{(1)}(v)(0),\quad \textit{and}\quad \Gamma_2^{(d)}(u)(0)=A_{d,\beta}^2\Gamma_2^{(1)}(v)(0),$$

where $A_{d,\beta}>0$ is a constant.

Proof. We will use that for $x\neq 0$, $2\alpha>1$ one obtains

$$\int_{\mathbb{R}}\frac{1}{(x^2+y^2)^\alpha}\,\mathrm{d}y = \frac{1}{x^{2\alpha}}\int_{\mathbb{R}}\frac{1}{(1+(\frac{y}{x})^2)^\alpha}\,\mathrm{d}y = \frac{1}{x^{2\alpha-1}}\int_{\mathbb{R}}\frac{1}{(1+z^2)^\alpha}\,\mathrm{d}z = \frac{B_\alpha}{x^{2\alpha-1}},$$

where $B_\alpha=\int_{\mathbb{R}}(1+z^2)^{-\alpha}\,\mathrm{d}z$ is independent of x. Thus, one inductively finds for $2\alpha>n$ that

$$\int_{\mathbb{R}^n}\frac{1}{(x^2+y_1^2+\ldots+y_n^2)^\alpha}\,\mathrm{d}y = \frac{C_{\alpha,n}}{|x|^{2\alpha-n}}$$

with $C_{\alpha,n}=\Pi_{j=0}^{n-1}B_{\alpha-\frac{j}{2}}$.
Whence, we find

$$\begin{aligned}\mathcal{L}^{(d)}(u)(0) &= c_{\beta,d}\int_{\mathbb{R}}(v(h_1)-2v(0)+v(-h_1))\int_{\mathbb{R}^{d-1}}\frac{1}{(h_1^2+h_2^2+\ldots+h_d^2)^{\frac{d+\beta}{2}}}\,\mathrm{d}(h_2,\dots,h_d)\,\mathrm{d}h_1\\ &= c_{\beta,d}C_{\frac{d+\beta}{2},d-1}\int_{\mathbb{R}}\frac{v(h_1)-2v(0)+v(-h_1)}{|h_1|^{1+\beta}}\,\mathrm{d}h_1\\ &= \frac{c_{\beta,d}C_{\frac{d+\beta}{2},d-1}}{c_{\beta,1}}\mathcal{L}^{(1)}(v)(0) = A_{d,\beta}\mathcal{L}^{(1)}(v)(0)\end{aligned}$$

with $A_{d,\beta}=\frac{c_{\beta,d}C_{\frac{d+\beta}{2},d-1}}{c_{\beta,1}}$. The formulas for $\Gamma^{(d)}(u)$ and $\Gamma_2^{(d)}(u)$ follow analogously. □

3 Proof of the Main Result

In this section we gradually construct our counterexample for (3). First, we construct an unbounded but admissible smooth function with unbounded support such that (3) fails at $x=0$. With an elaborate approximation scheme we subsequently construct a smooth and compactly supported function such that (3) fails at $x=0$, which will be used to show (5) using a continuity and a rescaling argument.

3.1 The Unbounded Counterexample with Unbounded Support

In this subsection we construct a class of unbounded functions with noncompact support lying in the domain of both $\mathcal{L}$ and Γ_2 such that (3) fails.

Theorem 3.1. *Let $0 < \beta < 2$, and $u_\varepsilon \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ be an even function defined by*

$$u_\varepsilon(x) = \begin{cases} |x|^{\beta-\varepsilon}, & |x| \geq 1, \\ \phi(x), & |x| < 1, \end{cases}$$

where $0 < \varepsilon < \frac{\beta}{2}$ (and additionally $\varepsilon < \frac{\beta-1}{2}$ if $\beta > 1$), and assume there exist $\Lambda > 0, \delta > 0$ independent of ε with $\beta + \delta > 1$ such that ϕ satisfies

$$0 \leq \phi(x) \leq \Lambda x^{\beta+\delta}, \qquad |\phi'(x)| \leq \Lambda x^{\beta+\delta-1} \quad \text{for } |x| \leq 1.$$

Then there exist constants $C_0 > 0, C_1 > 0$ such that

$$\mathcal{L}(u_\varepsilon)(0) \geq \frac{C_0}{\varepsilon}, \qquad 0 < \Gamma_2(u_\varepsilon)(0) \leq \frac{C_1}{\varepsilon}.$$

In particular, for all $\mu > 0$ there exists some u such that $0 < \Gamma_2(u)(0) < \mu \left(\mathcal{L}(u)(0)\right)^2 < \infty$.

Remark 3.2. *Let $u(x) = |x|^{\beta-\varepsilon}\eta(x)$, where $\eta \geq 0$ is a smooth and even cut-off function satisfying $\eta(x) = 0$ for $|x| \leq \frac{1}{4}$ and $\eta(x) = 1$ for $|x| \geq \frac{3}{4}$. Then u satisfies the assumptions of Theorem 3.1.*

We may extend the previous counterexample to higher spatial dimensions.

Theorem 3.3. *Let $d \geq 2$ and $v(x_1, \ldots, x_d) = u(x_1)$, where $u = u_\varepsilon$ is given as in Theorem 3.1. Then there exist constants $C_2 > 0, C_3 > 0$ such that*

$$\mathcal{L}(v)(0) \geq \frac{C_2}{\varepsilon}, \qquad 0 < \Gamma_2(v)(0) \leq \frac{C_3}{\varepsilon}$$

for all $\varepsilon > 0$ small enough. In particular, for all $\mu > 0$ there exists some v such that $0 < \Gamma_2(v)(0) < \mu \left(\mathcal{L}(v)(0)\right)^2 < \infty$.

Proof. This is a direct consequence of Lemma 2.5 and Theorem 3.1. □

For the proof of Theorem 3.1 we need the following lemmata.

Lemma 3.4. *Let u_ε be as in Theorem 3.1 and $0 \leq x \leq y$. Then we have*

$$\left(u_\varepsilon(y) - u_\varepsilon(x)\right)^2 \leq \begin{cases} \beta^2 \max\{x^{2\beta-2\varepsilon-2}, y^{2\beta-2\varepsilon-2}\}(y-x)^2, & x \geq 1 \\ \Lambda^2 y^{2\beta+2\delta-2}(y-x)^2, & y \leq 1 \\ (\beta+\Lambda)^2 \max\{1, y^{2\beta-2\varepsilon-2}\}(y-x)^2, & \text{else.} \end{cases}$$

Proof. This is a straightforward application of the mean value theorem together with the properties of u_ε. □

Lemma 3.5. *Let u_ε be as in Theorem 3.1 and $1 \leq \sigma \leq h$. Then*

$$\left(u_\varepsilon(h \pm \sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma)\right)^2 \leq \begin{cases} 16h^{2\beta-2\varepsilon-2}\sigma^2, & \beta > 1, \\ 4\sigma^{2\beta-2\varepsilon}, & \beta \leq 1, \end{cases} \tag{10}$$

where we additionally assume $h - \sigma \geq 1$ in the case of '$-$'.

Proof. Introduce $f, g : [0,1] \to \mathbb{R}$ by $f(x) = (1+x)^\gamma - 1 - x^\gamma$ and $g(x) = 1 + x^\gamma - (1-x)^\gamma$, where $0 < \gamma < 2$. Since $f(0) = 0$ we have for any $x \in [0,1]$ some $0 < \xi < x$ such that $f(x) = f'(\xi)x$ by the mean value theorem. Since for $\gamma \geq 1$ we have $f'(\xi) = \gamma\left((1+\xi)^{\gamma-1} - \xi^{\gamma-1}\right) \in (0,4)$ we obtain

$$0 \leq f(x) \leq 4x, \qquad x \in [0,1]. \tag{11}$$

The same argument shows for $\gamma \geq 1$ the bounds

$$0 \leq g(x) \leq 4x, \qquad x \in [0,1]. \tag{12}$$

If $\gamma < 1$, then $f'(\xi) < 0$ and $f(x) \geq -x^\gamma$, thus

$$-x^\gamma \leq f(x) \leq 0, \qquad x \in [0,1]. \tag{13}$$

Similarly $g(x) \geq x^\gamma \geq 0$, and for an upper bound of g we consider $h(x) := x^{-\gamma} g(x), x > 0$. We can extend h continuously by setting $h(0) = 1$. Since $h'(x) = \gamma x^{-1-\gamma}(1-x)^{\gamma-1}(1-(1-x)^{1-\gamma}) > 0$ for $0 < x < 1$ one obtains $h(x) \leq h(1) = 2$, whence

$$0 \leq x^\gamma \leq g(x) \leq 2x^\gamma, \qquad x \in [0,1]. \tag{14}$$

We set $\gamma = \beta - \varepsilon$ and note that $\gamma > 1$ if $\beta > 1$ as well as $\gamma \leq 1$ if $\beta \leq 1$ by the bound of ε in Theorem 3.1. We obtain from the definition of u_ε that $(u_\varepsilon(h+\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma))^2 = h^{2\beta-2\varepsilon} f\left(\frac{\sigma}{h}\right)^2$ and $(u_\varepsilon(h-\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma))^2 = h^{2\beta-2\varepsilon} g\left(\frac{\sigma}{h}\right)^2$, thus (10) follows from (11), (12), (13) and (14). □

We may now show Theorem 3.1.

Proof of Theorem 3.1. Fix some $\varepsilon_0 > 0$ sufficiently small and let $0 < \varepsilon < \varepsilon_0$. Throughout this proof we write $a(\varepsilon) \lesssim b(\varepsilon)$ whenever there exists some constant $C > 0$ independent of ε such that $a(\varepsilon) \leq C b(\varepsilon)$. On the one hand we obtain due to symmetry

$$\mathcal{L}(u_\varepsilon)(0) = 2c_{\beta,1} \int_0^\infty \frac{u_\varepsilon(h)}{h^{1+\beta}}\,\mathrm{d}h = 2c_{\beta,1}\left(\int_0^1 \frac{\phi(h)}{h^{1+\beta}}\,\mathrm{d}h + \int_1^\infty \frac{h^{\beta-\varepsilon}}{h^{1+\beta}}\,\mathrm{d}h\right) \geq 2c_{\beta,1}\int_1^\infty \frac{h^{\beta-\varepsilon}}{h^{1+\beta}}\,\mathrm{d}h = \frac{C_0}{\varepsilon}$$

by our assumptions on ϕ. Clearly, we have $\Gamma_2(u_\varepsilon) > 0$. To show the desired upper bound we will now split the integrals in Γ_2 into multiple parts. From Lemma 2.4 we find

$$\begin{aligned}
\frac{\Gamma_2(u_\varepsilon)(0)}{4c_{\beta,1}^2} &= \int_0^\infty \int_0^h \frac{[u_\varepsilon(h+\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma)]^2 + [u_\varepsilon(h-\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma)]^2}{h^{d+\beta}\sigma^{d+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \\
&= \int_0^1 \int_0^h \frac{[u_\varepsilon(h+\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma)]^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h + \int_0^1 \int_0^h \frac{[u_\varepsilon(h-\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma)]^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \\
&\quad + \int_1^\infty \int_0^1 \frac{[u_\varepsilon(h+\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma)]^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h + \int_1^\infty \int_0^1 \frac{[u_\varepsilon(h-\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma)]^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \\
&\quad + \int_1^\infty \int_1^h \frac{[u_\varepsilon(h+\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma)]^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h + \int_1^\infty \int_1^h \frac{[u_\varepsilon(h-\sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma)]^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \\
&=: A_+ + A_- + B_+ + B_- + C_+ + C_-.
\end{aligned}$$

We will show that $A_+ + A_- + B_+ + B_- \lesssim 1$ and $C_+ + C_- \lesssim \frac{1}{\varepsilon}$. In most of the cases we make use of the basic estimate

$$(u_\varepsilon(h \pm \sigma) - u_\varepsilon(h) - u_\varepsilon(\sigma))^2 \lesssim u_\varepsilon(\sigma)^2 + (u_\varepsilon(h \pm \sigma) - u_\varepsilon(h))^2.$$

Moreover, since $u_\varepsilon(\sigma)^2 = \phi(\sigma)^2$ for $\sigma \leq 1$ we obtain

$$\int_0^1 \frac{1}{h^{1+\beta}} \int_0^h \frac{\phi(\sigma)^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_0^1 \frac{1}{h^{1+\beta}} \int_0^h \sigma^{\beta+2\delta-1}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_0^1 \frac{1}{h^{1-2\delta}}\,\mathrm{d}h \lesssim 1.$$

Hence, for the integrals A_+ and A_- we merely have to estimate the corresponding expressions involving the term $(u(h \pm \sigma) - u(h))^2$. The same observation holds for the integrals B_+ and B_-, since

$$\int_1^\infty \frac{1}{h^{1+\beta}} \int_0^1 \frac{\phi(\sigma)^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_1^\infty \frac{1}{h^{1+\beta}} \int_0^1 \sigma^{\beta+2\delta-1}\,\mathrm{d}\sigma \lesssim 1.$$

$\underline{A_+}$: We split the remaining integral into

$$\int_0^{\frac{1}{2}} \int_\sigma^{1-\sigma} \frac{(u_\varepsilon(h+\sigma) - u_\varepsilon(h)))^2}{\sigma^{1+\beta}h^{1+\beta}}\,\mathrm{d}h\,\mathrm{d}\sigma + \int_{\frac{1}{2}}^1 \int_{1-h}^h \frac{(u_\varepsilon(h+\sigma) - u_\varepsilon(h))^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h.$$

In the first integral we have $h+\sigma \le 1$, thus we obtain $(u_\varepsilon(h+\sigma)-u_\varepsilon(h))^2 \lesssim \sigma^2(h+\sigma)^{2\beta+2\delta-2} \lesssim \sigma^2 h^{2\beta+2\delta-2}$ from Lemma 3.4. Whence

$$\int_0^{\frac{1}{2}} \int_\sigma^{1-\sigma} \frac{(u_\varepsilon(h+\sigma)-u_\varepsilon(h))^2}{\sigma^{1+\beta}h^{1+\beta}}\,\mathrm{d}h\,\mathrm{d}\sigma \lesssim \int_0^{\frac{1}{2}} \sigma^{1-\beta} \int_\sigma^{1-\sigma} h^{\beta+2\delta-3}\,\mathrm{d}h\,\mathrm{d}\sigma \lesssim 1$$

since the last integral can be estimated by $\int_0^{\frac{1}{2}} \sigma^{1-\beta}\,\mathrm{d}\sigma \lesssim 1$ for $\beta+2\delta-2>0$, by $\int_0^{\frac{1}{2}} \log(\frac{1}{\sigma})\sigma^{1-\beta}\,\mathrm{d}\sigma \lesssim 1$ for $\beta+2\delta-2=0$ and if $\beta+2\delta-2<0$ we have the upper bound $\int_0^{\frac{1}{2}} \sigma^{-1+2\delta}\,\mathrm{d}\sigma \lesssim 1$.
In the other case of $h+\sigma \ge 1$, we have $(u_\varepsilon(h+\sigma)-u_\varepsilon(h))^2 \lesssim \max\{1, h^{2\beta-2\varepsilon-2}\}\sigma^2$ by Lemma 3.4, thus

$$\int_{\frac{1}{2}}^1 \int_{1-h}^h \frac{(u_\varepsilon(h+\sigma)-u_\varepsilon(h))^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_{\frac{1}{2}}^1 \frac{\max\{1,h^{2\beta-2\varepsilon-2}\}}{h^{1+\beta}} \int_{1-h}^h \frac{\sigma^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_{\frac{1}{2}}^1 h^{1-2\beta}\max\{1,h^{2\beta-2\varepsilon-2}\}\,\mathrm{d}h,$$

which can be bounded from above by some constant independent of ε.
$\underline{A_-}$: According to Lemma 3.4 we can estimate the remainder by

$$\int_0^1 \int_0^h \frac{(u_\varepsilon(h)-u_\varepsilon(h-\sigma))^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_0^1 h^{\beta+2\delta-3} \int_0^h \sigma^{1-\beta}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_0^1 \frac{1}{h^{1-2\delta}}\,\mathrm{d}h \lesssim 1.$$

$\underline{B_+}$: We apply Lemma 3.4 together with $h \le h+\sigma \le 2h$ to obtain

$$\int_1^\infty \int_0^1 \frac{(u_\varepsilon(h+\sigma)-u_\varepsilon(h))^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_1^\infty \frac{h^{2\beta-2\varepsilon-2}}{h^{1+\beta}} \int_0^1 \sigma^{1-\beta}\,\mathrm{d}\sigma \lesssim \int_1^\infty \frac{1}{h^{3-\beta+2\varepsilon}}\,\mathrm{d}h \le \frac{1}{2-\beta}.$$

$\underline{B_-}$: This part requires the distinction between the cases when $h-\sigma \ge 1$ and $h-\sigma<1$. First, we consider the situation where $h-\sigma<1$ holds. Then we have $h \in [1,2]$, and by Lemma 3.4 we find

$$\int_{h-1}^1 \frac{(u_\varepsilon(h)-u_\varepsilon(h-\sigma))^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma \lesssim \max\{1,h^{2\beta-2\varepsilon-2}\} \int_{h-1}^1 \sigma^{1-\beta}\,\mathrm{d}\sigma \lesssim \max\{1,h^{2\beta-2\varepsilon-2}\}.$$

If $2\beta-2\varepsilon-2>0$, we can calculate $\int_1^2 \frac{h^{2\beta-2\varepsilon-2}}{h^{1+\beta}}\,\mathrm{d}h \le \frac{1}{2-\beta+2\varepsilon} \le \frac{1}{2-\beta}$, while the case of $2\beta-2\varepsilon-2 \le 0$ is clear. Thus

$$\int_1^2 \frac{1}{h^{1+\beta}} \int_{h-1}^1 \frac{(u_\varepsilon(h)-u_\varepsilon(h-\sigma))^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim 1.$$

We will subdivide the case of $h-\sigma \ge 1$ in two integrals. Since $\int_0^{h-1} \frac{\sigma^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma \lesssim (h-1)^{2-\beta}$ and $\int_0^1 \frac{\sigma^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma \lesssim 1$, we can deduce from Lemma 3.4 that

$$\int_1^2 \int_0^{h-1} \frac{(u_\varepsilon(h)-u_\varepsilon(h-\sigma))^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim 1 \quad \text{and} \quad \int_2^\infty \int_0^1 \frac{(u_\varepsilon(h)-u_\varepsilon(h-\sigma))^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim 1.$$

$\underline{C_+}$: Due to Lemma 3.5 we distinguish two cases. If $\beta>1$ one has $(u_\varepsilon(h+\sigma)-u_\varepsilon(h)-u_\varepsilon(\sigma))^2 \lesssim h^{2\beta-2\varepsilon-2}\sigma^2$, whence

$$C_+ \lesssim \int_1^\infty \frac{h^{2\beta-2\varepsilon-2}}{h^{1+\beta}} \int_1^h \frac{\sigma^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_1^\infty \frac{1}{h^{3-\beta+2\varepsilon}} \int_1^h \sigma^{1-\beta}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_1^\infty \frac{1}{h^{1+2\varepsilon}}\,\mathrm{d}h = \frac{1}{2\varepsilon}.$$

If $\beta \le 1$, we deduce

$$C_+ \lesssim \int_1^\infty \frac{1}{h^{1+\beta}} \int_1^h \frac{\sigma^{2\beta-2\varepsilon}}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_1^\infty \frac{1}{h^{1+\beta}} \int_1^h \sigma^{\beta-2\varepsilon-1}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_1^\infty \frac{1}{h^{1+2\varepsilon}}\,\mathrm{d}h = \frac{1}{2\varepsilon}.$$

$\underline{C_-}$: If $h-\sigma<1$ we have the upper bound

$$\int_1^\infty \frac{1}{\sigma^{1+\beta}} \int_\sigma^{\sigma+1} \frac{1}{h^{1+\beta}} h^{2\beta-2\varepsilon}\,\mathrm{d}h\,\mathrm{d}\sigma \lesssim \int_1^\infty \frac{\sigma^{\beta-2\varepsilon-1}}{\sigma^{1+\beta}}\,\mathrm{d}\sigma = \int_1^\infty \frac{1}{\sigma^{2+2\varepsilon}}\,\mathrm{d}\sigma \le \frac{1}{1+2\varepsilon} \le 1,$$

since $\phi(h-\sigma)^2 + u_\varepsilon(\sigma)^2 \lesssim \sigma^{2\beta-2\varepsilon} \le h^{2\beta-2\varepsilon} = u_\varepsilon(h)^2$ holds. If $h-\sigma \ge 1$ holds, we again employ Lemma 3.5 to deduce $C_- \lesssim \frac{1}{2\varepsilon}$ analogously to the estimate of C_+. We conclude $\Gamma_2(u_\varepsilon)(0) \lesssim \frac{1}{\varepsilon}$, which shows the claim. □

3.2 Compactly Supported and Smooth Counterexamples

The aim of this subsection is to give a counterexample from the class of smooth, compactly supported functions.

Theorem 3.6. *Let $0<\beta<2$ and $d\in\mathbb{N}$. Then for all $\mu>0$ there exists some $u\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ satisfying*

$$0 < \Gamma_2(u)(0) < \mu\left(\mathcal{L}(u)(0)\right)^2. \tag{15}$$

Moreover, $u(x)=0$ for all $|x|\le\frac{1}{4}$.

Proof. Let us first show how to reduce Theorem 3.6 to the case of $d=1$. Assume that for given $\mu>0$ there exists a $v\in\mathcal{C}_c^\infty(\mathbb{R})$ with $v(x_1)=0$ for all $|x_1|\le\frac{1}{4}$ and $0<\Gamma_2(v)(0)<\frac{\mu}{4}\left(\mathcal{L}(v)(0)\right)^2$. We will now show (15) for some $u\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ satisfying $u(x)=0$ for all $|x|\le\frac{1}{4}$.
We set $w(x_1,\dots,x_d) := v(x_1)$, then $w\in\mathcal{C}^\infty(\mathbb{R}^d)\cap W^{1,\infty}(\mathbb{R}^d)$. Furthermore, we define $u_n(x) := \chi_n(x)w(x)\in\mathcal{C}_c^\infty(\mathbb{R}^d)$, where $\chi_n\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ is a cutoff function satisfying $0\le\chi_n\le1$, $\chi_n\equiv1$ on $B_n(0)$, $\chi_n\equiv0$ on $\mathbb{R}^d\setminus B_{n+1}(0)$ and $|\nabla\chi_n|_\infty\le C$ for all $n\in\mathbb{N}$. Note that $u_n(x)=0$ for all $|x|\le\frac{1}{4}$ and $(u_n)_n$ is uniformly bounded in $W^{1,\infty}(\mathbb{R}^d)$ by, say, $M>0$. We will now show that

$$\mathcal{L}(u_n)(0)\to\mathcal{L}(w)(0) \quad\text{and}\quad \Gamma_2(u_n)(0)\to\Gamma_2(w)(0) \quad\text{as } n\to\infty \tag{16}$$

using Lebesgue's theorem. Clearly $u_n\to w$ pointwise. For $|h|\le\frac{1}{4}$ we have $u_n(h)=0$, and for $|h|\ge\frac{1}{4}$ we can bound $\frac{|u_n(h)+u_n(-h)|}{|h|^{d+\beta}}\le\frac{2|u_n|_\infty}{|h|^{d+\beta}}\le\frac{2M}{|h|^{d+\beta}}$, which is integrable on $\mathbb{R}^d\setminus B_{\frac{1}{4}}(0)$. The first part of (16) follows. For the latter we define $g_n(h,\sigma) := \frac{[u_n(h+\sigma)-u_n(h)-u_n(\sigma)]^2}{|h|^{d+\beta}|\sigma|^{d+\beta}}$ and $B := B_{\frac{1}{8}}(0)\subset\mathbb{R}^d$, then

$$c_{\beta,d}^{-2}\Gamma_2(u_n)(x) = \int_{\mathbb{R}^d\setminus B}\int_{\mathbb{R}^d\setminus B} g_n(h,\sigma)\,\mathrm{d}\sigma\,\mathrm{d}h + 2\int_{\mathbb{R}^d\setminus B}\int_B g_n(h,\sigma)\,\mathrm{d}\sigma\,\mathrm{d}h.$$

In the first integral we estimate g_n by the integrable function $9M^2|h|^{-d-\beta}|\sigma|^{-d-\beta}$. In the second integral we use $u_n(\sigma)=0$ and apply the mean value theorem to obtain the upper bound $g_n(h,\sigma)\le|\nabla u_n|_\infty^2\sigma^{-d+2-\beta}h^{-d-\beta}\le M^2\sigma^{-d+2-\beta}h^{-d-\beta}$, which is integrable on $(\mathbb{R}^d\setminus B)\times B$ since $\beta<2$. Another application of Lebesgue's theorem yields the second part of (16).
We can now infer the claim for $d>1$ using (16). Pick n large enough such that $u_n\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ satisfies $|(\mathcal{L}(u_n)(0))^2-(\mathcal{L}(w)(0))^2|\le\frac{1}{2}(\mathcal{L}(w)(0))^2$ and $|\Gamma_2(u_n)(0)-\Gamma_2(w)(0)|\le\Gamma_2(w)(0)$. Then $(\mathcal{L}(w)(0))^2\le2(\mathcal{L}(u_n)(0))^2$ and whence, using Lemma 2.5 and our choice of v,

$$0<\Gamma_2(u_n)(0)\le2\Gamma_2(w)(0)=2A_{d,\beta}^2\Gamma_2(v)(0)<2A_{d,\beta}^2\frac{\mu}{4}(\mathcal{L}(v)(0))^2=\frac{\mu}{2}(\mathcal{L}(w)(0))^2\le\mu(\mathcal{L}(u_n)(0))^2.$$

In the remainder we prove the claim for $d=1$. Let u_ε be given as in Theorem 3.1 with a smooth cutoff at zero as explained in Remark 3.2, and define $v_{N,\varepsilon}$ by $v_{N,\varepsilon} := u_\varepsilon\eta_N$, where η_N is a symmetric and smooth cutoff function satisfying $\eta_N\equiv1$ on $[0,N]$, $0<\eta_N<1$ on (N,N^2), $\eta_N\equiv0$ on $[N^2,\infty)$ and a decay behaviour of $|\eta_N'|\lesssim N^{-2}$. Note that we use here and throughout this proof the symbol '$\lesssim$' whenever the respective constant is independent of N. Moreover, $N>0$ is not related to the dimension in the CD-inequality in this subsection.
We have $v_{N,\varepsilon}\in\mathcal{C}_c^\infty(\mathbb{R})$ with the desired vanishing near zero, and $v_{N,\varepsilon}'(x)=0$ for $x\ge N^2$ and

$$|v_{N,\varepsilon}'(x)|\le(\beta-\varepsilon)\eta_N(x)x^{\beta-\varepsilon-1}+|\eta_N'(x)|x^{\beta-\varepsilon}\lesssim x^{\beta-\varepsilon-1},\quad x\in[1,N^2]. \tag{17}$$

Let $\beta\in(0,2)$ be fixed and let $0<\varepsilon<\varepsilon_0$, where $0<4\varepsilon<\beta$ and additionally $4\varepsilon_0<\beta-1$ (if $\beta>1$) and $4\varepsilon_0<\beta-\frac{1}{2}$ if $\beta>\frac{1}{2}$. In the sequel we will denote u_ε by u and $v_{N,\varepsilon}$ by v_N.
First we obtain that

$$|\mathcal{L}(u)(0)-\mathcal{L}(v_N)(0)|\lesssim\left(\int_N^{N^2}\frac{|u(h)-\eta_N(h)u(h)|}{h^{1+\beta}}\,\mathrm{d}h+\int_{N^2}^\infty\frac{u(h)}{h^{1+\beta}}\,\mathrm{d}h\right)\lesssim\int_N^\infty\frac{u(h)}{h^{1+\beta}}\,\mathrm{d}h\to0$$

as $N \to \infty$. If also $\Gamma_2(v_N)(0) \to \Gamma_2(u)(0)$ as $N \to \infty$ then the claim can be shown as above, now using Theorem 3.1. In the remainder of this proof we will thus show that $\Gamma_2(v_N)(0) \to \Gamma_2(u)(0)$.
First, we define for $M > 0$ the kernel $k_M(x) := \frac{1}{|x|^{1+\beta}} \mathbb{1}_{[-M,M]}(x)$. We denote by $\mathcal{L}^M$ the corresponding operator

$$\mathcal{L}^M u(x) := c_{\beta,1} \int_{[-M,M]} \frac{u(x+h) - 2u(x) + u(x-h)}{|h|^{d+\beta}} \, \mathrm{d}h,$$

and by

$$\Gamma_2^M u(x) = c_{\beta,1}^2 \int_{[-M,M]} \int_{[-M,M]} \frac{[u(x+h+\sigma) - u(x+h) - u(x+\sigma) + u(x)]^2}{|h|^{1+\beta}|\sigma|^{1+\beta}} \, \mathrm{d}h \, \mathrm{d}\sigma$$

the corresponding iterated carré du champ operator. Applying the theorem of dominated convergence one finds that $\Gamma_2^M(u)(0)$ converges to $\Gamma_2(u)(0)$ as M tends to infinity. Let $\rho > 0$ and fix $M > 0$ such that

$$|\Gamma_2^M(u)(0) - \Gamma_2(u)(0)| < \rho.$$

If $N > 2M$ and $|h|, |\sigma| \leq M$, then $(v_N(h+\sigma) - v_N(h) - v_N(\sigma))^2 = (u(h+\sigma) - u(h) - u(\sigma))^2$ and thus $|\Gamma_2^M(v_N)(0) - \Gamma_2^M(u)(0)| = 0$. This observation together with

$$|\Gamma_2(v_N)(0) - \Gamma_2(u)(0)| \leq |\Gamma_2(v_N)(0) - \Gamma_2^M(v_N)(0)| + |\Gamma_2^M(v_N)(0) - \Gamma_2^M(u)(0)| + |\Gamma_2^M(u)(0) - \Gamma_2(u)(0)|$$

shows that it suffices to prove that $|\Gamma_2(v_N)(0) - \Gamma_2^M(v_N)(0)|$ converges to zero as N tends to infinity. By symmetry (cf. Lemma 2.4) we have

$$\begin{aligned} &|\Gamma_2(v_N)(0) - \Gamma_2^M(v_N)(0)| \\ &\lesssim \int_M^\infty \int_0^h \frac{(v_N(h+\sigma) - v_N(h) - v_N(\sigma))^2 + (v_N(h-\sigma) - v_N(h) - v_N(\sigma))^2}{h^{1+\beta}\sigma^{1+\beta}} \, \mathrm{d}\sigma \, \mathrm{d}h. \end{aligned} \tag{18}$$

In order to show that this integral converges to zero we distinguish several cases (see Figures 1 and 2). Starting with the '$h+\sigma$'-term we have the following:

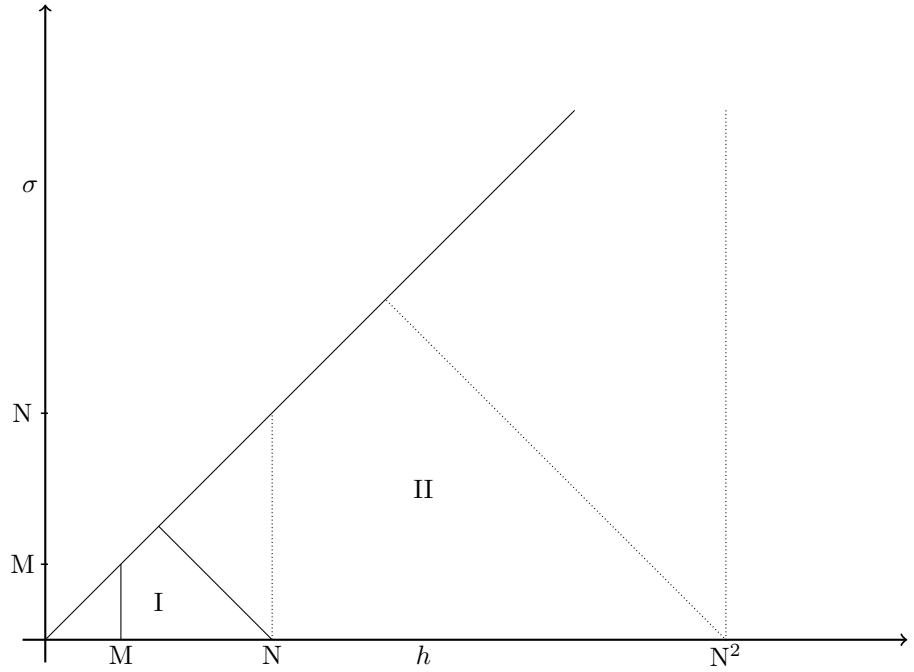


Figure 1: Splitting of the domain of integration for the term with $h + \sigma$.

I: $0 \leq \sigma \leq h$, $h + \sigma \leq N$, $h \geq M$: Clearly, since $h + \sigma \leq N$ one finds $(v_N(h+\sigma) - v_N(h) - v_N(\sigma))^2 = (u(h+\sigma) - u(h) - u(\sigma))^2$ and thus we have the upper bound

$$\int_M^N \int_0^{N-h} \frac{(v_N(h+\sigma) - v_N(h) - v_N(\sigma))^2}{h^{1+\beta}\sigma^{1+\beta}} \, \mathrm{d}\sigma \, \mathrm{d}h \lesssim |\Gamma_2^M(u)(0) - \Gamma_2(u)(0)| < \rho.$$

II: $0 \le \sigma \le h, h+\sigma \ge N$ As in the proof of Theorem 3.1 we will make use of the basic estimate

$$\left(v_N(h\pm\sigma)-v_N(h)-v_N(\sigma)\right)^2 \lesssim v_N(\sigma)^2+\left(v_N(h\pm\sigma)-v_N(h)\right)^2$$

and note that for the $v_N(\sigma)$-term we have

$$\int_{\frac{N}{2}}^{\infty}\frac{1}{h^{1+\beta}}\int_0^1\frac{v_N(\sigma)^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_{\frac{N}{2}}^{\infty}\frac{1}{h^{1+\beta}}\int_0^1\frac{\sigma^{2\beta+2\delta}}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_{\frac{N}{2}}^{\infty}\frac{1}{h^{1+\beta}}\,\mathrm{d}h \lesssim \frac{1}{N^\beta}$$

and

$$\int_{\frac{N}{2}}^{\infty}\frac{1}{h^{1+\beta}}\int_1^h\frac{v_N(\sigma)^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \le \int_{\frac{N}{2}}^{\infty}\frac{1}{h^{1+\beta}}\int_1^h\frac{\sigma^{2\beta-2\varepsilon}}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_{\frac{N}{2}}^{\infty}\frac{1}{h^{1+2\varepsilon}}\,\mathrm{d}h \lesssim \frac{1}{N^{2\varepsilon}}.$$

These estimates will be applied also to the term involving '$h-\sigma$' later. From the mean value theorem we obtain $\left(v_N(h+\sigma)-v_N(h)\right)^2 = v_N'(\xi)^2\sigma^2$ for some $h \le \xi \le h+\sigma \le 2h$. This yields by (17) the upper bound

$$\int_{\frac{N}{2}}^{\infty}\frac{1}{h^{1+\beta}}\int_0^h\frac{\left(v_N(h+\sigma)-v_N(h)\right)^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_{\frac{N}{2}}^{\infty}\frac{h^{2\beta-2\varepsilon-2}}{h^{1+\beta}}\int_0^h\sigma^{1-\beta}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_{\frac{N}{2}}^{\infty}\frac{1}{h^{1+2\varepsilon}}\,\mathrm{d}h \lesssim \frac{1}{N^{2\varepsilon}}$$

which tends to zero as $N\to\infty$. This shows that the first part of (18) tends to zero.
For the estimate of the second part we need a refined splitting of the domain, see Figure 2. Note that we may omit the integrals containing $v_N(\sigma)^2$ by the previous calculation.

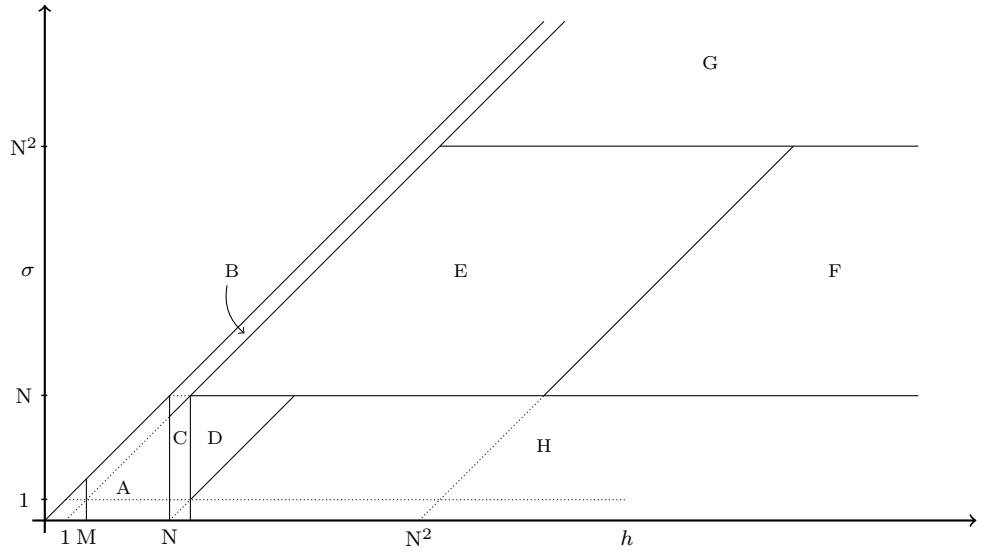


Figure 2: Splitting of the domain of integration for the term with $h-\sigma$.

A: $0 \le \sigma \le h \le N$, $h \ge M$: Since $h-\sigma \le N$ it follows similar to (I) that

$$\int_M^N\int_0^h\frac{\left(v_N(h-\sigma)-v_N(h)-v_N(\sigma)\right)^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h < \rho.$$

B: $h-1 \le \sigma \le h$, $h \ge N$: We can always estimate

$$\left(v_N(h-\sigma)-v_N(h)-v_N(\sigma)\right)^2 \le 9u(h)^2. \tag{19}$$

Hence, we deduce from $(h-1)^{-1-\beta} \lesssim h^{-1-\beta}$ that

$$\int_N^\infty\int_{h-1}^h\frac{\left(v_N(h-\sigma)-v_N(h)-v_N(\sigma)\right)^2}{h^{1+\beta}\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_N^\infty\frac{h^{2\beta-2\varepsilon}}{h^{1+\beta}}\int_{h-1}^h\frac{1}{\sigma^{1+\beta}}\,\mathrm{d}\sigma$$

165

$$\lesssim \int_N^\infty \frac{h^{2\beta-2\varepsilon}}{h^{2+2\beta}}\,\mathrm{d}h \lesssim \frac{1}{N^{1+2\varepsilon}}.$$

C, D, E, F, G, H: $0 \le \sigma \le h-1$, $h \ge N$, where $\beta > \frac{1}{2}$: The mean value theorem implies the existence of some ξ with $1 \le h-\sigma \le \xi \le h$ such that $(v_N(h)-v_N(h-\sigma))^2 = v_N'(\xi)^2\sigma^2$. We consider the case of $\beta > 1$ first. Here we conclude $v_N'(\xi)^2 \lesssim h^{2\beta-2\varepsilon-2}$ from (17). Thus, we can estimate

$$\int_N^\infty \frac{1}{h^{1+\beta}} \int_0^{h-1} \frac{(v_N(h)-v_N(h-\sigma))^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_N^\infty \frac{h^{2\beta-2\varepsilon-2}}{h^{1+\beta}} \int_0^{h-1} \sigma^{1-\beta}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_N^\infty \frac{1}{h^{1+2\varepsilon}}\,\mathrm{d}h \lesssim \frac{1}{N^{2\varepsilon}},$$

establishing the claim for $\beta > 1$. In the case of $\frac{1}{2} < \beta \le 1$ we obtain from $v_N'(\xi)^2 \lesssim (h-\sigma)^{2\beta-2\varepsilon-2}$ and $2\beta-2\varepsilon-1 > 0$ the estimate

$$\begin{aligned}\int_N^\infty \frac{1}{h^{1+\beta}} \int_0^{h-1} \frac{(v_N(h)-v_N(h-\sigma))^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h &\lesssim \int_N^\infty \frac{h^{1-\beta}}{h^{1+\beta}} \int_0^{h-1} (h-\sigma)^{2\beta-2\varepsilon-2}\,\mathrm{d}\sigma\,\mathrm{d}h \\ &= \int_N^\infty \frac{1}{h^{2\beta}} \int_1^h y^{2\beta-2\varepsilon-2}\,\mathrm{d}y\,\mathrm{d}h \lesssim \int_N^\infty \frac{1}{h^{1+2\varepsilon}}\,\mathrm{d}h \lesssim \frac{1}{N^{2\varepsilon}}.\end{aligned}$$

Thus, the claim is established for any $\beta > \frac{1}{2}$. From now on we assume $\beta \le \frac{1}{2}$.
C: $0 \le \sigma \le h-1$, $N \le h \le N+1$, where $\beta \le \frac{1}{2}$: Arguing as before we find

$$\int_N^{N+1} \frac{1}{h^{1+\beta}} \int_0^{h-1} \frac{(v_N(h)-v_N(h-\sigma))^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h \lesssim \int_N^{N+1} \frac{1}{h^{2\beta}} \int_1^h y^{2\beta-2\varepsilon-2}\,\mathrm{d}y\,\mathrm{d}h \lesssim \int_N^{N+1} \frac{1}{h^{2\beta}}\,\mathrm{d}h \lesssim \frac{1}{N^{2\beta}}$$

since $\int_1^h y^{2\beta-2\varepsilon-2}\,\mathrm{d}y$ is bounded, thus this term converges to zero.
D: $N+1 \le h \le 2N$, $\sigma \le N$, $h-\sigma \le N$, where $\beta \le \frac{1}{2}$: Using (19) we obtain the upper bound

$$\begin{aligned}\int_{N+1}^{2N} \frac{u(h)^2}{h^{1+\beta}} \int_{h-N}^{N} \frac{1}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h &\lesssim N^{2\beta-2\varepsilon} \int_{N+1}^{2N} \frac{(h-N)^{-\beta}}{h^{1+\beta}}\,\mathrm{d}h \le \frac{N^{2\beta-2\varepsilon}}{N^{1+\beta}} \int_{N+1}^{2N} (h-N)^{-\beta}\,\mathrm{d}h \\ &= \frac{1}{N^{1+2\varepsilon-\beta}} \int_1^N \frac{1}{y^\beta}\,\mathrm{d}y \lesssim \frac{1}{N^{2\varepsilon}}.\end{aligned}$$

E: $N \le \sigma \le N^2$, $1 \le h-\sigma \le N^2$, where $\beta \le \frac{1}{2}$: In order to estimate the integral

$$\int_N^{N^2} \frac{1}{\sigma^{1+\beta}} \int_{1+\sigma}^{N^2+\sigma} \frac{(v_N(h-\sigma)-v_N(h)-v_N(\sigma))^2}{h^{1+\beta}}\,\mathrm{d}h\,\mathrm{d}\sigma \tag{20}$$

we rewrite

$$\begin{aligned}&(v_N(h-\sigma)-v_N(h)-v_N(\sigma))^2 \\ &\quad= h^{2\beta-2\varepsilon}\left[\eta_N(h-\sigma)\left(1+\left(\frac{\sigma}{h}\right)^{\beta-\varepsilon}-\left(1-\frac{\sigma}{h}\right)^{\beta-\varepsilon}\right)\right. \\ &\qquad\left.+\,\eta_N(h)-\eta_N(h-\sigma)+(\eta_N(\sigma)-\eta_N(h-\sigma))\left(\frac{\sigma}{h}\right)^{\beta-\varepsilon}\right]^2 \\ &\quad\lesssim h^{2\beta-2\varepsilon}\eta_N(h-\sigma)^2\left(1+\left(\frac{\sigma}{h}\right)^{\beta-\varepsilon}-\left(1-\frac{\sigma}{h}\right)^{\beta-\varepsilon}\right)^2 \\ &\qquad+h^{2\beta-2\varepsilon}\left(\eta_N(h)-\eta_N(h-\sigma)\right)^2+h^{2\beta-2\varepsilon}\left(\eta_N(\sigma)-\eta_N(h-\sigma)\right)^2\left(\frac{\sigma}{h}\right)^{2\beta-2\varepsilon}.\end{aligned}$$

Now we want to find suitable estimates for these three terms. First, we obtain from Lemma 3.5, with $\beta \le \frac{1}{2}$, that $h^{2\beta-2\varepsilon}\eta_N(h-\sigma)^2\left(1+\left(\frac{\sigma}{h}\right)^{\beta-\varepsilon}-\left(1-\frac{\sigma}{h}\right)^{\beta-\varepsilon}\right)^2 \lesssim \sigma^{2\beta-2\varepsilon} = u(\sigma)^2$, and therefore we can

166

control this term by the previous calculations in (II). The same holds true for the third term, as $h^{2\beta-2\varepsilon}(\eta_N(\sigma)-\eta_N(h-\sigma))^2\left(\frac{\sigma}{h}\right)^{2\beta-2\varepsilon}\leq\sigma^{2\beta-2\varepsilon}$. The remaining term can be estimated using the mean value theorem and our assumptions on η_N by $h^{2\beta-2\varepsilon}\left(\eta_N(h)-\eta_N(h-\sigma)\right)^2\lesssim\frac{h^{2\beta-2\varepsilon}\sigma^2}{N^4}$. Finally, we obtain

$$\frac{1}{N^4}\int\limits_N^{N^2}\frac{1}{\sigma^{1+\beta}}\int\limits_{1+\sigma}^{N^2+\sigma}\frac{h^{2\beta-2\varepsilon}\sigma^2}{h^{1+\beta}}\,\mathrm{d}h\,\mathrm{d}\sigma\lesssim\frac{N^{2\beta-4\varepsilon}}{N^4}\int\limits_N^{N^2}\sigma^{1-\beta}\,\mathrm{d}\sigma\lesssim\frac{1}{N^{4\varepsilon}}$$

and conclude that (20) converges to 0 as N tends to infinity.
F: $N\leq\sigma\leq N^2$, $N^2\leq h-\sigma$, where $\beta\leq\frac{1}{2}$: Here, $v_N(h-\sigma)=v_N(h)=0$ and hence this integral converges to zero by the previous calculations in (II).
G: $N^2\leq\sigma\leq h-1$, where $\beta\leq\frac{1}{2}$: Since $v_N(x)\leq N^{2\beta-2\varepsilon}$ for all x we have

$$\begin{aligned}\int\limits_{N^2+1}^{\infty}\frac{1}{h^{1+\beta}}\int\limits_{N^2}^{h}\frac{\left(v_N(h-\sigma)-v_N(h)-v_N(\sigma)\right)^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h&=\int\limits_{N^2+1}^{\infty}\frac{1}{h^{1+\beta}}\int\limits_{N^2}^{h}\frac{v_N(h-\sigma)^2}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h\\&\leq N^{4\beta-4\varepsilon}\int\limits_{N^2+1}^{\infty}\frac{1}{h^{1+\beta}}\int\limits_{N^2}^{h}\frac{1}{\sigma^{1+\beta}}\,\mathrm{d}\sigma\,\mathrm{d}h\lesssim\frac{1}{N^{4\varepsilon}}.\end{aligned}$$

H: $0\leq\sigma\leq N, h\geq N+1, h-\sigma\geq N$, where $\beta\leq\frac{1}{2}$: In the remaining case we have $N\leq h-\sigma\leq\xi$ and thus we can estimate

$$\int\limits_0^N\frac{1}{\sigma^{1+\beta}}\int\limits_{N+\sigma}^{\infty}\frac{v_N'(\xi)^2\sigma^2}{h^{1+\beta}}\,\mathrm{d}h\,\mathrm{d}\sigma\lesssim N^{2\beta-2\varepsilon-2}\int\limits_0^N\sigma^{1-\beta}\int\limits_{N+\sigma}^{\infty}\frac{1}{h^{1+\beta}}\,\mathrm{d}h\,\mathrm{d}\sigma\lesssim\frac{1}{N^{2\varepsilon}}.\qquad\square$$

3.3 Failure of $CD(0,N)$ on Arbitrary Balls

In this subsection we finish the proof of Theorem 1.1.

Proposition 3.7. *Let $0<\beta<2$, $d\in\mathbb{N}$. Let $\mu>0$. Then there exists some $\rho>0$ and a smooth and compactly supported function $u\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ such that for all $x\in\mathbb{R}^d$ with $|x|\leq\rho$ one has*

$$0<\Gamma_2(u)(x)<\mu(\mathcal{L}(u)(x))^2.$$

Proof. Let $\mu>0$ and $u\in\mathcal{C}_c^\infty(\mathbb{R}^d)$ be as in Theorem 3.6 satisfying $0<\Gamma_2(u)(0)<\frac{\mu}{4}\left(\mathcal{L}(u)(0)\right)^2$ and $u(y)=0$ for all $|y|\leq\frac{1}{4}$. We will first show continuity of $\mathcal{L}(u)$ and $\Gamma_2(u)$ at zero, i.e.

$$\mathcal{L}(u)(x)\to\mathcal{L}(u)(0)\quad\text{ and }\quad\Gamma_2(u)(x)\to\Gamma_2(u)(0)\quad\text{ as }x\to0.\tag{21}$$

Since u is continuous we can show (21) by finding integrable functions dominating the integrands and applying Lebesgue's theorem, resembling the proof of Theorem 3.6. For the first part we observe

$$\frac{|(u(x+h)-2u(x)+u(x-h)|}{|h|^{d+\beta}}\leq4|u|_\infty\frac{1-\mathbb{1}_{B_{\frac{1}{8}}(0)}(h)}{|h|^{d+\beta}},\quad|x|\leq\frac{1}{8},$$

and the latter is integrable on $\mathbb{R}^d$, which shows $\mathcal{L}(u)(x)\to\mathcal{L}(u)(0)$ as $x\to0$. The continuity of $\Gamma_2(u)$ at zero requires a more detailed analysis , as already indicated in the proof of Theorem 3.6. We define

$$g(x,h,\sigma):=\frac{[u(x+h+\sigma)-u(x+h)-u(x+\sigma)+u(x)]^2}{|h|^{d+\beta}|\sigma|^{d+\beta}},$$

and $B:=B_{\frac{1}{16}}(0)\subset\mathbb{R}^d$, then

$$c_{\beta,d}^{-2}\Gamma_2(u)(x)=\int_B\int_B g(x,h,\sigma)\,\mathrm{d}\sigma\,\mathrm{d}h+\int_{\mathbb{R}^d\setminus B}\int_{\mathbb{R}^d\setminus B}g(x,h,\sigma)\,\mathrm{d}\sigma\,\mathrm{d}h+2\int_{\mathbb{R}^d\setminus B}\int_B g(x,h,\sigma)\,\mathrm{d}\sigma\,\mathrm{d}h.\tag{22}$$

Let $|x|\leq\frac{1}{8}$, then the first integral of (22) vanishes. In the second integral of (22) we estimate g by the integrable function $16|u|_\infty^2|h|^{-d-\beta}|\sigma|^{-d-\beta}$. In the third part of (22) we apply the mean value theorem to obtain the upper bound $g(x,h,\sigma)\leq4|\nabla u|_\infty^2\sigma^{-d+2-\beta}h^{-d-\beta}$, which is integrable on $(\mathbb{R}^d\setminus B)\times B$. Whence

we can apply Lebesgue's theorem to obtain (21).
Having established (21) we can show the claim by choosing $\rho > 0$ small enough such that

$$|\Gamma_2 u(x) - \Gamma_2(u)(0)| \leq \Gamma_2(u)(0) \quad \text{and} \quad |\left(\mathcal{L}(u)(x)\right)^2 - \left(\mathcal{L}(u)(0)\right)^2| \leq \frac{\left(\mathcal{L}(u)(0)\right)^2}{2} \quad \text{for all } |x| \leq \rho.$$

Then $\left(\mathcal{L}(u)(x)\right)^2 \geq \left(\mathcal{L}(u)(0)\right)^2 - |\left(\mathcal{L}(u)(0)\right)^2 - \left(\mathcal{L}(u)(x)\right)^2| \geq \frac{(\mathcal{L}(u)(0))^2}{2}$, whence

$$\Gamma_2(u)(x) \leq 2\Gamma_2(u)(0) < \frac{\mu}{2}\left(\mathcal{L}(u)(0)\right)^2 \leq \mu\left(\mathcal{L}(u)(x)\right)^2, \qquad |x| \leq \rho. \qquad \square$$

Applying a scaling argument we can give a counterexample on arbitrary balls, which is our main result.

Theorem 3.8. *Let $0 < \beta < 2$, $d \in \mathbb{N}$ and $R > 0$. Then (3) with $\kappa = 0$ fails on all of $\overline{B_R(0)} \subset \mathbb{R}^d$, i.e. for all $\mu > 0$ there exists some smooth and compactly supported function $u \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ such that for all $x \in \mathbb{R}^d$ with $|x| \leq R$ one has*

$$0 < \Gamma_2(u)(x) < \mu(\mathcal{L}(u)(x))^2.$$

Proof. Let $R > 0$ and μ be given, and fix $\rho > 0$ and u as in Proposition 3.7. Define $M := \frac{R}{\rho}$ and $v(x) := u\left(\frac{x}{M}\right)$. Then $v \in \mathcal{C}_c^\infty(\mathbb{R}^d)$ satisfies $\mathcal{L}(v)(x) = M^{-\beta}\mathcal{L}(u)\left(\frac{x}{M}\right)$ and $\Gamma_2(v)(x) = M^{-2\beta}\Gamma_2 u\left(\frac{x}{M}\right)$, as a short calculation similar to the proof of Proposition 2.3 shows. Thus we conclude $0 < \Gamma_2(v)(x) < \mu\left(\mathcal{L}(v)(x)\right)^2$ for all $\left|\frac{x}{M}\right| \leq \rho$, i.e. $|x| \leq R$. $\square$

4 References

[App09] David Applebaum. *Lévy processes and stochastic calculus*, volume 116 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2009.

[BBCK09] Martin T. Barlow, Richard F. Bass, Zhen-Qing Chen, and Moritz Kassmann. Non-local Dirichlet forms and symmetric jump processes. *Trans. Amer. Math. Soc.*, 361(4):1963–1999, 2009.

[BE85] Dominique Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[BGL14] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.

[BL02] Richard F. Bass and David A. Levin. Harnack inequalities for jump processes. *Potential Anal.*, 17(4):375–388, 2002.

[BV14] Matteo Bonforte and Juan Luis Vázquez. Quantitative local and global a priori estimates for fractional nonlinear diffusion equations. *Adv. Math.*, 250:242–284, 2014.

[CS07] Luis Caffarelli and Luis Silvestre. An extension problem related to the fractional Laplacian. *Comm. Partial Differential Equations*, 32(7-9):1245–1260, 2007.

[CS09] Luis Caffarelli and Luis Silvestre. Regularity theory for fully nonlinear integro-differential equations. *Comm. Pure Appl. Math.*, 62(5):597–638, 2009.

[CS14] Xavier Cabré and Yannick Sire. Nonlinear equations for fractional Laplacians, I: Regularity, maximum principles, and Hamiltonian estimates. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 31(1):23–53, 2014.

[CS15] Xavier Cabré and Yannick Sire. Nonlinear equations for fractional Laplacians II: Existence, uniqueness, and qualitative properties of solutions. *Trans. Amer. Math. Soc.*, 367(2):911–941, 2015.

[CV10] Luis A. Caffarelli and Alexis Vasseur. Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. *Ann. of Math. (2)*, 171(3):1903–1930, 2010.

[CV11] Luis Caffarelli and Juan Luis Vazquez. Nonlinear porous medium flow with fractional potential pressure. *Arch. Ration. Mech. Anal.*, 202(2):537–565, 2011.

[FK13] Matthieu Felsinger and Moritz Kassmann. Local regularity for parabolic nonlocal operators. *Comm. Partial Differential Equations*, 38(9):1539–1573, 2013.

[Gar19] Nicola Garofalo. Fractional thoughts. In *New developments in the analysis of nonlocal operators*, volume 723 of *Contemp. Math.*, pages 1–135. Amer. Math. Soc., Providence, RI, 2019.

[IS16] Cyril Imbert and Luis Silvestre. The weak Harnack inequality for the Boltzmann equation without cut-off. *arXiv e-prints*, arXiv:1608.07571, Aug 2016.

[Kas09] Moritz Kassmann. A priori estimates for integro-differential operators with measurable kernels. *Calc. Var. Partial Differential Equations*, 34(1):1–21, 2009.

[KMS15] Tuomo Kuusi, Giuseppe Mingione, and Yannick Sire. Nonlocal equations with measure data. *Comm. Math. Phys.*, 337(3):1317–1368, 2015.

[Kwa17] Mateusz Kwaśnicki. Ten equivalent definitions of the fractional Laplace operator. *Fract. Calc. Appl. Anal.*, 20(1):7–51, 2017.

[Li12] Peter Li. *Geometric analysis*, volume 134 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2012.

[LY86] Peter Li and Shing-Tung Yau. On the parabolic kernel of the Schrödinger operator. *Acta Math.*, 156(3-4):153–201, 1986.

[Sil06] Luis Silvestre. Hölder estimates for solutions of integro-differential equations like the fractional Laplace. *Indiana Univ. Math. J.*, 55(3):1155–1174, 2006.

[Sil07] Luis Silvestre. Regularity of the obstacle problem for a fractional power of the Laplace operator. *Comm. Pure Appl. Math.*, 60(1):67–112, 2007.

[SV09] Yannick Sire and Enrico Valdinoci. Fractional Laplacian phase transitions and boundary reactions: a geometric inequality and a symmetry result. *J. Funct. Anal.*, 256(6):1842–1864, 2009.

[SWZ19] Adrian Spener, Frederic Weber, and Rico Zacher. Curvature-dimension inequalities for nonlocal operators in the discrete setting. *to appear in Calc. Var. Partial Differential Equations*, arXiv:1903.00517, 2019.

[Vil09] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

6. Li-Yau inequalities for non-local diffusion equations via reduction to the heat kernel

Authors: Frederic Weber and Rico Zacher

Journal: Mathematische Annalen

arXiv: 2012.12974, arXiv preprint, 2020

Copyright: *This is a Preprint version of an article published by Springer Nature in Mathematische Annalen on 06. January 2022, available at: https://doi.org/10.1007/s00208-021-02350-z. This article is licensed under a Creative Commons Attribution 4.0 International License.*

LI-YAU INEQUALITIES FOR GENERAL NON-LOCAL DIFFUSION EQUATIONS VIA REDUCTION TO THE HEAT KERNEL

FREDERIC WEBER AND RICO ZACHER*

Abstract. We establish a reduction principle to derive Li-Yau inequalities for non-local diffusion problems in a very general framework, which covers both the discrete and continuous setting. Our approach is not based on curvature-dimension inequalities but on heat kernel representations of the solutions and consists in reducing the problem to the heat kernel. As an important application we solve a long-standing open problem by obtaining a Li-Yau inequality for positive solutions u to the fractional (in space) heat equation of the form $(-\Delta)^{\beta/2}(\log u) \leqslant C/t$, where $\beta \in (0,2)$. We also illustrate our general result with an example in the discrete setting by proving a sharp Li-Yau inequality for diffusion on a complete graph.

Keywords: Li-Yau inequality, differential Harnack inequality, non-local diffusion, heat kernel, fractional Laplacian, continuous-time Markov chains, complete graphs
MSC(2020): 35R11, 35K08, 60J27, 39A12

1. Introduction

The celebrated Li-Yau inequality states that on a complete Riemannian manifold M with topological dimension n and non-negative Ricci curvature, we have

$$-\Delta(\log u) \leqslant \frac{n}{2t}, \quad t > 0, \tag{1}$$

for any positive solution u to the heat equation $\partial_t u = \Delta u$ on $(0,\infty) \times M$, where Δ denotes the Laplace-Beltrami operator. This inequality originates from the seminal work [11]. It is optimal in the sense that equality is achieved for the heat kernel in the Euclidean case. As an important application of (1) (sharp) parabolic Harnack inequalities can be deduced. Using the Γ-calculus of Bakry and Émery, the Li-Yau inequality (1) has been generalized to Markov diffusion operators that satisfy the curvature-dimension condition $CD(0,n)$, see [1] and also the extensive monograph [2].

Concerning the non-local situation, the approach via curvature-dimension inequalities has stimulated a lot of research in the *discrete* setting. Based on discrete replacements of certain chain rule identities, in [4, 6, 12] suitable substitutes of curvature-dimension conditions have been introduced in order to prove Li-Yau type inequalities for generalized graph Laplacians. Corresponding Li-Yau inequalities for non-local operators describing diffusion processes with arbitrary long jumps do not seem to be known.

In the continuous setting, the fractional Laplacian on $\mathbb{R}^d$ is one of the most prominent representatives of a non-local operator. It is a natural question to ask in which way properties for

(Frederic Weber, Rico Zacher) Institut für Angewandte Analysis, Universität Ulm, Helmholtzstrasse 18, 89081 Ulm, Germany.
E-mail addresses: frederic.weber@uni-ulm.de, rico.zacher@uni-ulm.de.
Date: April 12, 2021.
*Corresponding author. F. W. is supported by a PhD-scholarship of the "Studienstiftung des deutschen Volkes", Germany. R. Z. is supported by the DFG (project number 355354916, GZ ZA 547/4-2).

the Laplace operator in the Euclidean case carry over to the fractional Laplacian. In particular it is of great interest whether a Li-Yau type inequality is also valid for positive solutions of the fractional heat equation, and if yes, whether the time-dependence is like in (1). The recent survey article [7] of Garofalo dedicates a whole chapter to address these open questions.

The main aim of this article is to establish a general reduction principle to derive Li-Yau inequalities, that applies to the discrete and continuous setting, respectively, and which answers both of the above questions for the fractional heat equation positively. For that purpose, we will not follow the approach of curvature-dimension inequalities. Remarkably, in the quite recent article [14] it has been shown that the fractional Laplacian does not satisfy $CD(\kappa, n)$ for any pair $(\kappa, n) \in \mathbb{R} \times (0, \infty)$, which is in sharp contrast to the Euclidean Laplacian. Instead, we follow the different approach of reducing the problem to the heat kernel. In the case of the classical heat equation on $\mathbb{R}^d$, the heat kernel satisfies (1) with equality and one is able to deduce the Li-Yau inequality (1) by using that positive solutions to the heat equation are given as the convolution of the heat kernel with the respective initial datum, see e.g. [7, Section 21]. We show in a very general framework that this reduction principle also works for positive solutions to the corresponding non-local diffusion equation. Indeed, by Theorem 2.4 we conclude that the heat kernel determines the corresponding Li-Yau inequality for solutions that can be represented via a respective integral representation involving the corresponding heat kernel. As an application, we apply this principle to the fractional heat equation and derive a Li-Yau inequality (see Theorem 3.2), which is optimal with respect to the time-behaviour. To highlight that our method applies to a quite general setting, we also illustrate the case of a discrete state space in Section 4. Here, we consider the example that the underlying graph is given by the complete graph K_n, $n \geqslant 2$, and deduce an optimal Li-Yau type inequality, which improves known results from [6].

The article is organized as follows. In the next section we establish a reduction principle to derive Li-Yau inequalities for non-local diffusion problems in a very general framework. Thereafter, we apply this principle in Section 3 to the fractional heat equation and in Section 4 to the diffusion equation with the Markov generator that corresponds to the complete graph.

We have been informed that Tuhin Ghosh and Moritz Kassmann from Bielefeld (Germany) have recently proved a different version of Li-Yau type inequality for the fractional heat equation by an alternative approach, see [8]. A preprint shall be available soon.

2. A reduction principle for deriving Li-Yau inequalities

Let (M, d) be a metric space and $\mathcal{B}(M)$ denote the Borel σ-algebra on M. We consider a non-local operator of the form

$$Lf(x) = \int_{M\setminus\{x\}} \big(f(y) - f(x)\big) k(x, \mathrm{d}y), \tag{2}$$

where the kernel is such that $k(x, \cdot)$ defines a σ-finite measure on $\mathcal{B}(M\setminus\{x\})$ for any $x \in M$ and $f : M \to \mathbb{R}$ is such that the integral exists. We also include the case where the integral on the right-hand side of (2) is singular. In this situation one replaces $\int_{M\setminus\{x\}}$ by $\lim_{\varepsilon\to 0^+} \int_{M\setminus B_\varepsilon(x)}$ in the right-hand side of (2). This is motivated by the important example of the fractional Laplace operator, where we choose $M = \mathbb{R}^n$ and d as the corresponding Euclidean distance. However, the quite general formulation also covers a variety of other important situations, such as, for instance, the case where L is the generator of a continuous-time Markov chain on a discrete state space endowed with the natural graph structure with M being the set of vertices and edge weights given by the corresponding transition rates between two respective states.

In [6], the following formula serves as a replacement of a classical chain rule. It has been established in the context of locally finite graphs and also for a class of non-local operators on

$\mathbb{R}^n$ that are of the form (2). We state the result here again due to our more general setting. The short proof follows analogously as in [6].

Lemma 2.1. *Let $D \subset \mathbb{R}$ be an open set, $H \in C^1(D;\mathbb{R})$, $f : M \to D$ and $x \in M$ such that $Lf(x)$ and $L(H(f))(x)$ exist. Then we have*

$$L\big(H(f)\big)(x) = H'(f(x))Lf(x) + \int_{M\setminus\{x\}} \Lambda_H\big(f(y), f(x)\big)k(x, \mathrm{d}y),$$

where $\Lambda_H(w, z) := H(w) - H(z) - H'(z)(w - z)$ for any $w, z \in D$.

In the context of this article, we will apply Lemma 2.1 with the specific choice of $H(r) = \log(r)$, $r \in (0, \infty)$. In this case we have

$$\Lambda_{\log}(w, z) = \log w - \log z - \frac{w - z}{z} = -\Upsilon(\log w - \log z),\ w, z \in (0, \infty),$$

where $\Upsilon(r) := e^r - r - 1$, $r \in \mathbb{R}$. Introducing the operator

$$\Psi_\Upsilon(f)(x) = \int_{M\setminus\{x\}} \Upsilon\big(f(y) - f(x)\big)k(x, \mathrm{d}y), \tag{3}$$

we deduce from Lemma 2.1 that

$$L(\log f) = \frac{Lf}{f} - \Psi_\Upsilon(\log f) \tag{4}$$

holds for any positive function f whenever Lf and $L(\log f)$ exist. Note that a possible singularity at $y = x$ in the right-hand side of (3) does not play a role by positivity of Υ in the sense that the integral is either finite or ∞. In the particular case of the fractional Laplacian, a simple Taylor argument shows that the quadratic behavior of Υ near the origin ensures that the singularity of the corresponding integral kernel is compensated provided the function is sufficiently smooth.

Now, we establish the key estimate of this article.

Lemma 2.2. *Let L be an operator of the form (2) and let $H : M \times M \to (0, \infty)$ be such that $H(x, \cdot)$ is $\mathcal{B}(M)$-measurable and the restriction $H|_{M\setminus\{x\}\times M}$ is $\mathcal{B}(M\setminus\{x\}) \otimes \mathcal{B}(M)$-measurable for any $x \in M$, respectively. Further, let $f : M \to (0, \infty)$ be $\mathcal{B}(M)$-measurable. We assume that the integral $Pf(x) = \int_M H(x, y)f(y)\,\mathrm{d}\nu(y)$ and also $\Psi_\Upsilon(\log Pf)$ exist for any $x \in M$ and that for ν-a.e. $y \in M$ the expression $\Psi_\Upsilon(\log H(\cdot, y))(x)$ exists for every $x \in M$. Here $\nu : \mathcal{B}(M) \to [0, \infty]$ is a σ-finite measure. Then we have*

$$\int_M \Psi_\Upsilon(\log H(\cdot, y))(x)H(x, y)f(y)\mathrm{d}\nu(y) \geqslant \Psi_\Upsilon(\log Pf)(x)Pf(x). \tag{5}$$

Proof. Recalling (3), we write

$$\begin{aligned}\Psi_\Upsilon(\log Pf)(x)Pf(x) &= \int_{M\setminus\{x\}} \Upsilon\Big(\log\frac{Pf(h)}{Pf(x)}\Big)Pf(x)k(x, \mathrm{d}h)\\ &= \int_{M\setminus\{x\}} \Upsilon\Big(\log\frac{Pf(h)}{Pf(x)}\Big)\int_M H(x, y)f(y)\mathrm{d}\nu(y)\,k(x, \mathrm{d}h).\end{aligned}$$

Next, by Tonelli's theorem we observe that

$$\begin{aligned}&\int_M \Psi_\Upsilon\big(\log H(\cdot, y)\big)(x)H(x, y)f(y)\mathrm{d}\nu(y)\\ &\qquad = \int_M H(x, y)f(y)\int_{M\setminus\{x\}} \Upsilon\Big(\log\frac{H(h, y)}{H(x, y)}\Big)k(x, \mathrm{d}h)\,\mathrm{d}\nu(y)\end{aligned}$$

$$= \int_{M\setminus\{x\}} \int_M H(x,y) f(y) \Upsilon\Big(\log \frac{H(h,y)}{H(x,y)}\Big) \mathrm{d}\nu(y)\, k(x,\mathrm{d}h).$$

Consequently, we have

(6)

$$\int_M \Psi_\Upsilon\big(\log H(\cdot,y)\big)(x) H(x,y) f(y) \mathrm{d}\nu(y) - \Psi_\Upsilon(\log Pf)(x) Pf(x)$$
$$= \int_{M\setminus\{x\}} \int_M \Big(\Upsilon\Big(\log \frac{H(h,y)}{H(x,y)}\Big) - \Upsilon\Big(\log \frac{Pf(h)}{Pf(x)}\Big)\Big) H(x,y) f(y) \mathrm{d}\nu(y)\, k(x,\mathrm{d}h).$$

Now, it can be readily checked that the mapping $r \mapsto \Upsilon(\log(r))$, $r \in (0,\infty)$, is convex. Since $\frac{\mathrm{d}}{\mathrm{d}r}\Upsilon(\log(r)) = \frac{r-1}{r}$, $r \in (0,\infty)$, we infer from convexity that

$$\Upsilon(\log(r)) - \Upsilon(\log(s)) \geqslant \frac{s-1}{s}(r-s)$$

holds for any $r, s \in (0,\infty)$. With this at hand and using the positivity of f and H, we can now estimate the right-hand side of (6) as follows.

$$\int_{M\setminus\{x\}} \int_M \Big(\Upsilon\Big(\log \frac{H(h,y)}{H(x,y)}\Big) - \Upsilon\Big(\log \frac{Pf(h)}{Pf(x)}\Big)\Big) H(x,y) f(y) \mathrm{d}\nu(y)\, k(x,\mathrm{d}h)$$
$$\geqslant \int_{M\setminus\{x\}} \int_M \frac{Pf(h) - Pf(x)}{Pf(h)} \Big(\frac{H(h,y)}{H(x,y)} - \frac{Pf(h)}{Pf(x)}\Big) H(x,y) f(y) \mathrm{d}\nu(y)\, k(x,\mathrm{d}h).$$

Having a possible singularity in mind, we choose $\varepsilon > 0$ arbitrary (but small enough such that $B_\varepsilon(x) \subsetneq M$) and then use linearity to observe

$$\int_{M\setminus B_\varepsilon(x)} \int_M \frac{Pf(h) - Pf(x)}{Pf(h)} \Big(\frac{H(h,y)}{H(x,y)} - \frac{Pf(h)}{Pf(x)}\Big) H(x,y) f(y) \mathrm{d}\nu(y)\, k(x,\mathrm{d}h)$$
$$= \int_{M\setminus B_\varepsilon(x)} \int_M \frac{Pf(h) - Pf(x)}{Pf(h)} H(h,y) f(y) \mathrm{d}\nu(y)\, k(x,\mathrm{d}h)$$
$$- \int_{M\setminus B_\varepsilon(x)} \int_M \frac{Pf(h) - Pf(x)}{Pf(x)} H(x,y) f(y) \mathrm{d}\nu(y)\, k(x,\mathrm{d}h)$$
$$=: I_{\varepsilon,1} - I_{\varepsilon,2}.$$

Further, we have

$$I_{\varepsilon,1} = \int_{M\setminus B_\varepsilon(x)} \frac{Pf(h) - Pf(x)}{Pf(h)} \int_M H(h,y) f(y) \mathrm{d}\nu(y)\, k(x,\mathrm{d}h)$$
$$= \int_{M\setminus B_\varepsilon(x)} \big(Pf(h) - Pf(x)\big) k(x,\mathrm{d}h)$$

and similarly

$$I_{\varepsilon,2} = \frac{1}{Pf(x)} \int_{M\setminus B_\varepsilon(x)} \big(Pf(h) - Pf(x)\big) \int_M H(x,y) f(y) \mathrm{d}\nu(y)\, k(x,\mathrm{d}h)$$
$$= \int_{M\setminus B_\varepsilon(x)} \big(Pf(h) - Pf(x)\big) k(x,\mathrm{d}h),$$

which yields that $I_{\varepsilon,1} - I_{\varepsilon,2} = 0$ holds. Sending $\varepsilon \to 0$ yields the claim. □

Remark 2.3. Lemma 2.2 is a non-local version of the inequality

$$\int_{\mathbb{R}^n} \big|\nabla_x \log H(x,y)\big|^2 H(x,y) f(y)\, \mathrm{d}y \geqslant \big|\nabla \log Pf(x)\big|^2 Pf(x), \tag{7}$$

where $Pf(x) = \int_{\mathbb{R}^n} H(x,y)f(y)\mathrm{d}y$, for sufficiently regular, positive functions H and f. This shows, as already indicated by the chain rule (4), that the expression $\Psi_\Upsilon(\log f)$ serves as a natural analogue of $|\nabla \log f|^2$ in the non-local setting.

In contrast to inequality (5), the argument for (7) is quite simple. In fact, using Hölder's inequality we have

$$\begin{aligned}(\partial_{x_i} Pf(x))^2 &= \big(\int_{\mathbb{R}^n} \partial_{x_i} H(x,y) f(y)\,\mathrm{d}y\big)^2 \\ &\leqslant \int_{\mathbb{R}^n} \frac{(\partial_{x_i} H(x,y))^2}{H(x,y)} f(y)\,\mathrm{d}y \int_{\mathbb{R}^n} H(x,y) f(y)\,\mathrm{d}y,\end{aligned}$$

which directly leads to (7) by summing up and employing the classical chain rule for the gradient.

We now come to our main result. For $T > 0$ arbitrary, we consider $u : [0,T) \times M \to (0,\infty)$ with $u(\cdot,x) \in C^1\big((0,T)\big)$ for any $x \in M$ and such that

$$\partial_t u(t,x) = L(u(t,\cdot))(x) \tag{8}$$

holds for any $(t,x) \in (0,T) \times M$. We set $u_0(\cdot) = u(0,\cdot)$ and make the following assumptions:

(A1) There exists a mapping $p : (0,T) \times M \times M \to (0,\infty)$ such that $p(t,x,\cdot)$ is $\mathcal{B}(M)$-measurable and the restriction $p(t,\cdot,\cdot)|_{M\setminus\{x\}\times M}$ is $\mathcal{B}(M\setminus\{x\})\otimes\mathcal{B}(M)$-measurable for any $(t,x) \in (0,T)\times M$. Moreover, there exists a σ-finite reference measure $\mu : \mathcal{B}(M) \to [0,\infty]$, such that p is differentiable with respect to time and satisfies for any $(t,x) \in (0,T) \times M$ the equation $\partial_t p(t,x,y) = L(p(t,\cdot,y))(x)$ for $\mu-$a.e. $y \in M$, and, moreover, that the representation formula

$$u(t,x) = \int_M p(t,x,y) u_0(y) \mathrm{d}\mu(y) \tag{9}$$

holds true at any $(t,x) \in (0,T) \times M$.

(A2) We have that

$$\partial_t u(t,x) = \int_M \partial_t p(t,x,y) u_0(y) \mathrm{d}\mu(y)$$

holds at any $(t,x) \in (0,T) \times M$.

Note that assumption (A2) is quite mild in the sense that it can be justified by means of the dominated convergence theorem for a large class of examples.

Combining (4) and (8), we observe that

$$\partial_t \log u = \frac{\partial_t u}{u} = L(\log u) + \Psi_\Upsilon(\log u). \tag{10}$$

Note that due to assumption (A1), in (10), we can also replace u by $p(\cdot,\cdot,y)$ for $\mu-$a.e. $y \in M$.

The above assumptions have a clear motivation from the viewpoint of stochastic processes. Let $\big(X_t\big)_{t\geqslant 0}$ be a Markov process on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ with state space given by a locally compact and separable metric space M. One defines the semigroup

$$P_t f(x) = \mathbb{E}(f(X_t)|X_t = x) \tag{11}$$

for suitable functions f (e.g. bounded and measurable), which is then a solution to the corresponding Cauchy problem. Very often $P_t f$ is given by an integral representation as in (9), where the measure μ commonly plays the role of an invariant measure for the process and the kernel $p(t,x,y)$ describes the transition density for the Markov process. Transition densities play a fundamental role in several aspects of probability theory and analysis. For instance, they form the basis of the parametrix-technique to construct Lévy-type processes, see e.g. [9]. The case

of α-stable isotropic Lévy processes will be discussed in Section 3 below. Of somewhat different flavour is the situation when M is a discrete set, in which case we refer to $(X_t)_{t\geqslant 0}$ as a continuous-time Markov chain. Then the kernel $p(t,x,y)$ coincides with the respective transition probability, i.e. the probability of the process being at time t in the state y provided that it has started in the state x. We illustrate this situation in Section 4 by considering the special case when the underlying graph to the Markov generator is given by a complete graph.

The fundamental estimate of Lemma 2.2 allows to reduce the problem of deriving Li-Yau inequalities for (8) to the kernel function of assumption (A1), as the following result shows.

Theorem 2.4. *Let $u : [0,T) \times M \to (0,\infty)$ solve (8) and satisfy assumptions (A1) and (A2) such that at any $(t,x) \in (0,T) \times M$ the expressions $L(\log p(t,\cdot,y))(x)$ for $\mu-$a.e. $y \in M$ and $L(\log u(t,\cdot))(x)$ exist, respectively. If the estimate*

$$-L\big(\log p(t,\cdot,y)\big)(x) \leqslant \varphi(t,x) \tag{12}$$

holds at any $(t,x) \in (0,T) \times M$ and $\mu-$a.e. $y \in M$, where $\varphi : (0,\infty) \times M \to \mathbb{R}$, then the Li-Yau type inequality

$$-L(\log u(t,\cdot))(x) \leqslant \varphi(t,x) \tag{13}$$

holds true for every $(t,x) \in (0,T) \times M$.

Remark 2.5. Using the evolution equation (10) for $\log u$, the Li-Yau inequality (13) from Theorem 2.4 implies the differential Harnack inequality

$$\partial_t \log u(t,x) \geqslant \Psi_\Upsilon(\log u)(t,x) - \varphi(t,x) \tag{14}$$

for all $(t,x) \in (0,T) \times M$. In the discrete setting, it has been shown that inequalities of the form (14) imply Harnack inequalities for positive solutions of the diffusion equation associated with the operator L, see [6].

Proof. The two key ingredients in the proof of Theorem 2.4 are Lemma 2.2 and the evolution equation (10) for $\log u$.

Using (10), we can rewrite (13) equivalently as

$$u(t,x)\,\Psi_\Upsilon(\log u(t,\cdot))(x) \leqslant \partial_t u(t,x) + \varphi(t,x)\,u(t,x). \tag{15}$$

By the same argument, we know by (A1) that the assumption (12) can be reformulated into

$$p(t,x,y)\,\Psi_\Upsilon(\log p(t,\cdot,y))(x) \leqslant \partial_t p(t,x,y) + \varphi(t,x)\,p(t,x,y) \tag{16}$$

for $\mu-$a.e. $y \in M$. We observe from the assumptions (A1) and (A2) that

$$\begin{aligned}
\partial_t u(t,x) + \varphi(t,x)\,u(t,x) &= \int_M \partial_t p(t,x,y) u_0(y) \mathrm{d}\mu(y) + \varphi(t,x) \int_M p(t,x,y) u_0(y) \mathrm{d}\mu(y) \\
&= \int_M \Big(\partial_t p(t,x,y) + \varphi(t,x) p(t,x,y)\Big) u_0(y) \mathrm{d}\mu(y) \\
&\geqslant \int_M \Psi_\Upsilon\big(\log p(t,\cdot,y)\big)(x)\, p(t,x,y) u_0(y) \mathrm{d}\mu(y) \\
&\geqslant \Psi_\Upsilon\big(\log u(t,\cdot)\big)(x)\, u(t,x),
\end{aligned}$$

where we used (16) in the second to last line and applied Lemma 2.2 in the last line (with $H(x,y) = p(t,x,y)$) at $(t,x) \in (0,T) \times M$. □

Remark 2.6. Note that we could relax the assumptions of Theorem 2.4 in order to get a local version of the above result. More precisely, given some fixed $(t,x) \in (0,T) \times M$, we only need to assume that $L\big(\log p(t,\cdot,y)\big)(x)$ for $\mu-$a.e. $y \in M$ and $L\big(\log u(t,\cdot)\big)(x)$ exist, respectively. If

then $-L\big(\log p(t,\cdot,y)\big)(x) \leqslant C$ holds for $\mu-$a.e. $y \in M$ and a constant $C \in \mathbb{R}$, we infer that $-L\big(\log u(t,\cdot)\big)(x) \leqslant C$ by the same arguments as in the proof of Theorem 2.4.

3. Application to the fractional heat equation

In this section we want to apply Theorem 2.4 to derive Li-Yau type inequalities for positive solutions to the fractional heat equation

$$\partial_t u + \big(-\Delta\big)^{\frac{\beta}{2}} u = 0 \tag{17}$$

on the full space $\mathbb{R}^d$, where $\beta \in (0,2)$.

The fractional Laplacian on $\mathbb{R}^d$ can be defined as

$$-\big(-\Delta\big)^{\frac{\beta}{2}} f(x) = c_{\beta,d}\ \text{p.v.} \int_{\mathbb{R}^d} \frac{f(y)-f(x)}{|x-y|^{d+\beta}}\,\mathrm{d}y,$$

where $\beta \in (0,2)$, the normalizing constant $c_{\beta,d}$ is given by $c_{\beta,d} = \frac{2^\beta \Gamma(\frac{d+\beta}{2})}{\pi^{\frac{d}{2}} |\Gamma(-\frac{\beta}{2})|}$ and f is a suitable function (e.g. in $C^2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$). There is also an equivalent definition in the pointwise sense that reads as

$$-\big(-\Delta\big)^{\frac{\beta}{2}} f(x) = \frac{c_{\beta,d}}{2} \int_{\mathbb{R}^d} \frac{f(x+y)+f(x-y)-2f(x)}{|y|^{d+\beta}}\mathrm{d}y$$

and has the advantage of avoiding the principal value integral.

Using classical Fourier techniques, one can express solutions to (17) as

$$u(t,x) = \big(G^{(\beta)}(t,\cdot) * u_0\big)(x) = \int_{\mathbb{R}^d} G^{(\beta)}(t,x-y)u_0(y)\mathrm{d}y, \tag{18}$$

where we refer to $G^{\beta)}$ as the fractional heat kernel of order $\beta \in (0,2)$. We will discuss below that this representation formula in fact holds true for the class of solutions we are interested in.

But before that, we collect some properties of the fractional heat kernels, which we will use subsequently. We refer to [7] and the references therein for a more detailed account.

In the special case of $\beta = 1$, there is an explicit formula for the fractional heat kernel available, which reads as

$$G^{(1)}(t,x) = \frac{\Gamma\big(\frac{d+1}{2}\big)}{\pi^{\frac{d+1}{2}}} \frac{t}{\big(t^2+|x|^2\big)^{\frac{d+1}{2}}}. \tag{19}$$

Concerning the general situation, it goes back to [5] that the fractional heat kernels behave as follows

$$G^{(\beta)}(t,x) \asymp \frac{t}{\big(t^{\frac{2}{\beta}}+|x|^2\big)^{\frac{d+\beta}{2}}}, \quad (t,x) \in (0,\infty)\times\mathbb{R}^d, \tag{20}$$

where the symbol $\asymp$ means that the ratio is bounded by a constant positive factor from above and below. This yields in particular that we have for any non-negative u_0

$$\int_{\mathbb{R}^d} G^{(\beta)}(t,x-y)u_0(y)\mathrm{d}y \asymp t\int_{\mathbb{R}^d} \frac{u_0(y)}{\big(t^{\frac{2}{\beta}}+|x-y|^2\big)^{\frac{d+\beta}{2}}}\mathrm{d}y, \tag{21}$$

which yields that the integral on the right-hand side of (18) exists if and only if the integral on the right-hand side of (21) is finite.

Further, we have $G^{(\beta)}(t,\cdot) \in C^\infty(\mathbb{R}^d)$ and that the fractional heat kernels can be expressed in the following self-similar form

$$G^{(\beta)}(t,x) = t^{-\frac{d}{\beta}} \Phi_\beta\big(\frac{x}{t^{\frac{1}{\beta}}}\big), \tag{22}$$

where $(t,x) \in (0,\infty) \times \mathbb{R}^d$ and $\Phi_\beta(y) = G^{(\beta)}(1,y)$ for any $y \in \mathbb{R}^d$.

There also exist specific estimates for the derivatives of the fractional heat kernels, for which we refer to [10], where those have been studied extensively in quite great generality. We will only make use of

$$|\nabla \Phi_\beta(x)| \lesssim \frac{1}{|x|^{d+\beta+1}} \tag{23}$$

and

$$\|\nabla^2 \Phi_\beta(x)\| \lesssim \frac{1}{|x|^{d+\beta+2}}, \tag{24}$$

which hold whenever $|x| \geqslant 1$, cf. [10]. Here we have used the symbol '$\lesssim$' to indicate that the corresponding constant is independent of x.

In [3] a Widder-type theorem for the fractional heat equation has been established. More precisely, any non-negative strong solution to (17) is given by (18). Here, by strong solution it is meant that for any fixed $T > 0$ we have $\partial_t u \in C\big((0,T) \times \mathbb{R}^d\big)$, $u \in C\big([0,T) \times \mathbb{R}^d\big)$ and that (17) holds in the pointwise sense for every $(t,x) \in (0,T) \times \mathbb{R}^d$. Consequently, the assumption (A1) is satisfied for any positive strong solution u, where the kernel is given by $p(t,x,y) = G^{(\beta)}(t,x-y)$ and the measure μ is the Lebesgue measure on $\mathbb{R}^d$.

Concerning assumption (A2), we refer to [15], where the bound

$$|\big(-\Delta\big)^{\frac{\beta}{2}}(G^{(\beta)}(t,\cdot))(x)| \leqslant \frac{C}{\big(t^{\frac{2}{\beta}} + |x|^2\big)^{\frac{d+\beta}{2}}}$$

has been established for some constant $C > 0$. This yields, that assumption (A2) follows from the dominated convergence theorem due to the existence of the integral on the right-hand side of (18) (cf. (21)).

Lemma 3.1. *We have that*

$$\big(-\Delta\big)^{\frac{\beta}{2}}(\log G^{(\beta)}(t,\cdot))(x) \leqslant \frac{C_{LY}(\beta,d)}{t}, \tag{25}$$

holds at any $x \in \mathbb{R}^d$, $t > 0$ and $\beta \in (0,2)$, where the finite constant $C_{LY}(\beta,d) > 0$ is given by

$$C_{LY}(\beta,d) = \frac{c_{\beta,d}}{2} \sup_{y \in \mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\log\left(\frac{\Phi_\beta(y)^2}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)}\right)}{|\sigma|^{d+\beta}} \mathrm{d}\sigma. \tag{26}$$

Proof. We consider

$$\begin{aligned}\big(-\Delta\big)^{\frac{\beta}{2}}(\log G^{(\beta)}(t,\cdot))(x) &= \frac{c_{\beta,d}}{2} \int_{\mathbb{R}^d} \frac{2\log G^{(\beta)}(t,x) - \log G^{(\beta)}(t,x+h) - \log G^{(\beta)}(t,x-h)}{|h|^{d+\beta}} \mathrm{d}h \\ &= \frac{c_{\beta,d}}{2} \int_{\mathbb{R}^d} \frac{\log\left(\frac{G^{(\beta)}(t,x)^2}{G^{(\beta)}(t,x+h)\, G^{(\beta)}(t,x-h)}\right)}{|h|^{d+\beta}} \mathrm{d}h.\end{aligned}$$

Using the self-similar form (22) and setting $y = \frac{x}{t^{\frac{1}{\beta}}}$, we get that

$$\int_{\mathbb{R}^d} \frac{\log\left(\frac{G^{(\beta)}(t,x)^2}{G^{(\beta)}(t,x+h)G^{(\beta)}(t,x-h)}\right)}{|h|^{d+\beta}} \mathrm{d}h = \frac{1}{t} \int_{\mathbb{R}^d} \frac{\log\left(\frac{\Phi_\beta(y)^2}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)}\right)}{|\sigma|^{d+\beta}} \mathrm{d}\sigma,$$

where we have substituted $\sigma = \frac{h}{t^{\frac{1}{\beta}}}$. Consequently, it suffices to show that the mapping $J : \mathbb{R}^d \to \mathbb{R}$ given by

$$J(y) = \int_{\mathbb{R}^d} \frac{\log\left(\frac{\Phi_\beta(y)^2}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)}\right)}{|\sigma|^{d+\beta}} \mathrm{d}\sigma$$

is bounded. By continuity we only need to show that J is bounded on $\mathbb{R}^d \backslash B_R(0)$ for some $R > 2$. In the sequel, we assume that $|y| \geqslant R$. We write

$$J(y) = \int_{\mathbb{R}^d \backslash B_1(0)} \frac{\log\left(\frac{\Phi_\beta(y)^2}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)}\right)}{|\sigma|^{d+\beta}} \mathrm{d}\sigma + \int_{B_1(0)} \frac{\log\left(\frac{\Phi_\beta(y)^2}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)}\right)}{|\sigma|^{d+\beta}} \mathrm{d}\sigma =: J_1(y) + J_2(y).$$

First, we consider $J_1(y)$ and use the behaviour described by (20)(with $t = 1$) to find a constant $M > 1$ such that

$$J_1(y) \leqslant \frac{d+\beta}{2} \int_{\mathbb{R}^d \backslash B_1(0)} \frac{\log\left(M \frac{(1+|y+\sigma|^2)(1+|y-\sigma|^2)}{(1+|y|^2)^2}\right)}{|\sigma|^{d+\beta}} \mathrm{d}\sigma.$$

We have that

$$(1 + |y+\sigma|^2)(1 + |y-\sigma|^2) \leqslant \left(1 + 2|y|^2 + 2|\sigma|^2\right)^2.$$

In the case of $|y| > |\sigma|$, this yields that

$$\frac{(1 + |y+\sigma|^2)(1 + |y-\sigma|^2)}{(1+|y|^2)^2} \leqslant \left(\frac{1+4|y|^2}{1+|y|^2}\right)^2 \leqslant M',$$

where $M' > 1$ is some constant that is independent of y. In the other case of $|y| \leqslant |\sigma|$, we have

$$\frac{(1 + |y+\sigma|^2)(1 + |y-\sigma|^2)}{(1+|y|^2)^2} \leqslant \left(1 + 4|\sigma|^2\right)^2.$$

Putting these estimates together and using the monotonicity of the logarithm, we end up with

$$\begin{aligned} J_1(y) &\leqslant \frac{d+\beta}{2} \int_{B_{|y|}(0)\backslash B_1(0)} \frac{\log(MM')}{|\sigma|^{d+\beta}} \mathrm{d}\sigma + \frac{d+\beta}{2} \int_{\mathbb{R}^d \backslash B_{|y|}(0)} \frac{\log(M(1+4|\sigma|^2)^2)}{|\sigma|^{d+\beta}} \mathrm{d}\sigma \\ &\leqslant \frac{d+\beta}{2} \int_{\mathbb{R}^d \backslash B_1(0)} \frac{\log\left(M^2 M'(1+4|\sigma|^2)^2\right)}{|\sigma|^{d+\beta}} \mathrm{d}\sigma \leqslant C_1 < \infty, \end{aligned}$$

where $C_1 > 0$ is some constant which is independent of y.

Now, we turn to $J_2(y)$. Using the inequality $\log r \leqslant r - 1$, $r \in (0, \infty)$, we infer

$$J_2(y) \leqslant \int_{B_1(0)} \frac{\Phi_\beta(y)^2 - \Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)} \frac{\mathrm{d}\sigma}{|\sigma|^{d+\beta}}. \tag{27}$$

Given some $\eta \in B_1(0)$, we have by the Cauchy-Schwarz inequality that

$$1 - \frac{2}{R} \leqslant 1 - \frac{2|\eta|}{|y|} \leqslant 1 - \frac{2|\eta|}{|y|} + \frac{|\eta|^2}{|y|^2} \leqslant \frac{|y+\eta|^2}{|y|^2} \leqslant 1 + \frac{2|\eta|}{|y|} + \frac{|\eta|^2}{|y|^2} \leqslant 1 + \frac{2}{R} + \frac{1}{R^2}.$$

Consequently, since $R > 2$, we observe that

$$|y+\eta|^2 \asymp |y|^2, \tag{28}$$

where the respective constants are in particular independent of η.

Further, by using Taylor's expansion we get that

$$\Phi_\beta(y+\sigma) = \Phi_\beta(y) + \sigma \cdot \nabla\Phi_\beta(y)^T + \frac{1}{2}\sigma \cdot \nabla^2\Phi_\beta(y+\zeta_+\sigma)\sigma,$$

$$\Phi_\beta(y-\sigma) = \Phi_\beta(y) - \sigma\cdot\nabla\Phi_\beta(y)^T + \frac{1}{2}\sigma\cdot\nabla^2\Phi_\beta(y-\zeta_-\sigma)\sigma,$$

where $\zeta_-, \zeta_+ \in [0,1]$. Then we get

$$\begin{aligned}\Phi_\beta(y)^2 - \Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma) = &-\frac{\Phi_\beta(y)}{2}\Big(\sigma\cdot\nabla^2\Phi_\beta(y+\zeta_+\sigma)\sigma + \sigma\cdot\nabla^2\Phi_\beta(y-\zeta_-\sigma)\sigma\Big)\\ &- \frac{\sigma\cdot\nabla\Phi_\beta(y)^T}{2}\Big(\sigma\cdot\nabla^2\Phi_\beta(y-\zeta_-\sigma)\sigma - \sigma\cdot\nabla^2\Phi_\beta(y+\zeta_+\sigma)\sigma\Big)\\ &- \frac{1}{4}\big(\sigma\cdot\nabla^2\Phi_\beta(y+\zeta_+\sigma)\sigma\big)\big(\sigma\cdot\nabla^2\Phi_\beta(y-\zeta_-\sigma)\sigma\big)\\ &+ \big(\sigma\cdot\nabla\Phi_\beta(y)^T\big)^2.\end{aligned}$$

Next, we aim to suitably estimate the integrand on the right hand side of (27). For that purpose, we will subsequently make use of the estimates (20) (with $t=1$), (23), (24), (28) and the Cauchy-Schwarz inequality. Moreover, we will use the symbol '$\lesssim$' whenever the corresponding constant is independent of $|y|$.

We have

$$\begin{aligned}\frac{|\Phi_\beta(y)\,\sigma\cdot\nabla^2\Phi_\beta(y+\zeta_+\sigma)\sigma|}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)} &\lesssim \Big(\frac{(1+|y+\sigma|^2)(1+|y-\sigma|^2)}{1+|y|^2}\Big)^{\frac{d+\beta}{2}}|\sigma|^2\|\nabla^2\Phi_\beta(y+\zeta_+\sigma)\|\\ &\lesssim \Big(\frac{|y+\sigma|^2|y-\sigma|^2}{|y|^2}\Big)^{\frac{d+\beta}{2}}\frac{|\sigma|^2}{|y+\zeta_+\sigma|^{d+\beta+2}} \lesssim \frac{|\sigma|^2}{|y|^2}.\end{aligned}$$

Clearly, the corresponding term with $-\zeta_-$ instead of ζ_+ can be treated analogously. Next, we consider

$$\begin{aligned}&\frac{\big|\big(\sigma\cdot\nabla\Phi_\beta(y)^T\big)\big(\sigma\cdot\nabla^2\Phi_\beta(y+\zeta_+\sigma)\sigma\big)\big|}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)}\\ &\qquad\lesssim |\sigma|^3|\nabla\Phi_\beta(y)|\|\nabla^2\Phi_\beta(y+\zeta_+\sigma)\|\big((1+|y+\sigma|^2)(1+|y-\sigma|^2)\big)^{\frac{d+\beta}{2}}\\ &\qquad\lesssim \frac{|\sigma|^3|y+\sigma|^{d+\beta}|y-\sigma|^{d+\beta}}{|y|^{d+\beta+1}|y+\zeta_+\sigma|^{d+\beta+2}} \lesssim \frac{|\sigma|^3}{|y|^3}\end{aligned}$$

and again, the corresponding term with $-\zeta_-$ instead of ζ_+ can be treated analogously. Furthermore, we have

$$\begin{aligned}&\frac{\big|\big(\sigma\cdot\nabla^2\Phi_\beta(y+\zeta_+\sigma)\sigma\big)\big(\sigma\cdot\nabla^2\Phi_\beta(y-\zeta_-\sigma)\sigma\big)\big|}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)}\\ &\qquad\lesssim |\sigma|^4\|\nabla^2\Phi_\beta(y+\zeta_+\sigma)\|\|\nabla^2\Phi_\beta(y-\zeta_-\sigma)\|\big((1+|y+\sigma|^2)(1+|y-\sigma|^2)\big)^{\frac{d+\beta}{2}}\\ &\qquad\lesssim \frac{|\sigma|^4|y+\sigma|^{d+\beta}|y-\sigma|^{d+\beta}}{|y+\zeta_+\sigma|^{d+\beta+2}|y-\zeta_-\sigma|^{d+\beta+2}} \lesssim \frac{|\sigma|^4}{|y|^4}\end{aligned}$$

and last but not least

$$\begin{aligned}\frac{\big(\sigma\cdot\nabla\Phi_\beta(y)^T\big)^2}{\Phi_\beta(y+\sigma)\Phi_\beta(y-\sigma)} &\lesssim |\sigma|^2|\nabla\Phi_\beta(y)|^2\big((1+|y+\sigma|^2)(1+|y-\sigma|^2)\big)^{\frac{d+\beta}{2}}\\ &\lesssim \frac{|\sigma|^2|y+\sigma|^{d+\beta}|y-\sigma|^{d+\beta}}{|y|^{2d+2\beta+2}} \lesssim \frac{|\sigma|^2}{|y|^2}.\end{aligned}$$

Putting everything together, we infer from (27) that

$$J_2(y) \lesssim \frac{1}{|y|^2}\int_{B_1(0)}|\sigma|^{2-d-\beta}\mathrm{d}\sigma \leqslant C_2 < \infty,$$

for some constant $C_2 > 0$ being independent of y. □

Combining Theorem 2.4 (and Remark 2.6) with Lemma 3.1, we have shown the following Li-Yau type inequality for the fractional heat equation.

Theorem 3.2. *Let $u : [0,T) \times \mathbb{R}^d \to (0,\infty)$ be a strong solution to the fractional heat equation (17). Then the Li-Yau type inequality*

$$\big(-\Delta\big)^{\frac{\beta}{2}} (\log u(t,\cdot))(x) \leqslant \frac{C_{LY}(\beta,d)}{t} \tag{29}$$

holds at any $(t,x) \in (0,T) \times \mathbb{R}^d$ where the left-hand side exists. The constant $C_{LY}(\beta,d) > 0$ is given by (26).

In the special case $\beta = 1$ an analytic expression in closed form is available for the fractional heat kernel, which allows to explicitly compute the constant of the Li-Yau inequality from Theorem 3.2.

Proposition 3.3. *For $\beta = 1$ the optimal constant (26) is given by*

$$C_{LY}(1,d) = \frac{\pi\, d(d+1)\, c_{1,d}\, \omega_d}{2} = \frac{d(d+1)}{2B\big(\frac{d+1}{2}, \frac{1}{2}\big)}, \tag{30}$$

where ω_d denotes the volume of the d-dimensional unit ball in $\mathbb{R}^d$ and B the Beta function.

Proof. First, we show that

$$\sup_{x\in\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\log\Big(\frac{\Phi_1(x)^2}{\Phi_1(x+\sigma)\Phi_1(x-\sigma)}\Big)}{|\sigma|^{d+1}} \mathrm{d}\sigma = \int_{\mathbb{R}^d} \frac{\log\Big(\frac{\Phi_1(0)^2}{\Phi_1(\sigma)\Phi_1(-\sigma)}\Big)}{|\sigma|^{d+1}} \mathrm{d}\sigma. \tag{31}$$

Let $x \in \mathbb{R}^d$ be fixed and $\sigma \in \mathbb{R}^d$ arbitrary. Using the explicit representation (19), we get that

$$\log\Big(\frac{\Phi_1(x)^2}{\Phi_1(x+\sigma)\Phi_1(x-\sigma)}\Big) = \frac{d+1}{2} \log\Big(\frac{(1+|x+\sigma|^2)(1+|x-\sigma|^2)}{(1+|x|^2)^2}\Big).$$

We have

$$\begin{aligned}\frac{(1+|x+\sigma|^2)(1+|x-\sigma|^2)}{(1+|x|^2)^2} &= 1 + \frac{2|\sigma|^2 + |\sigma|^4 + 2|x|^2|\sigma|^2 - 4(x\cdot\sigma)^2}{1+2|x|^2+|x|^4} \\ &\leqslant 1 + \frac{2|\sigma|^2 + |\sigma|^4 + 2|x|^2|\sigma|^2}{1+2|x|^2+|x|^4}.\end{aligned}$$

We set $a := |\sigma|^2$ and consider $h_a : [0,\infty) \to [0,\infty)$ given by $h_a(z) = \frac{2az+2a+a^2}{1+2z+z^2}$. Note that

$$\big(1+h_a(0)\big)^{\frac{d+1}{2}} = \big(1+2|\sigma|^2+|\sigma|^4\big)^{\frac{d+1}{2}} = \frac{\Phi_1(0)^2}{\Phi_1(\sigma)\Phi_1(-\sigma)}.$$

Hence, in order to establish (31) it suffices to show that h_a is decreasing on $[0,\infty)$ for any $a \in [0,\infty)$. We have that

$$h_a'(z) = \frac{(1+2z+z^2)2a - (2az+2a+a^2)(2z+2)}{(1+2z+z^2)^2},$$

which is non-positive if and only if

$$\frac{1+2z+z^2}{2z+2} - z \leqslant 1 + \frac{a}{2}.$$

This inequality holds true since

$$\frac{1+2z+z^2}{2z+2} - z = \frac{1-z^2}{2z+2} = \frac{1-z}{2} \leqslant \frac{1}{2}.$$

Consequently, (31) is established. Due to Φ_1 being radial symmetric, the calculation of the corresponding integral boils down to the one-dimensional case. More precisely, we have

$$\int_{\mathbb{R}^d} \frac{\log\left(\frac{\Phi_1(0)^2}{\Phi_1(\sigma)\Phi_1(-\sigma)}\right)}{|\sigma|^{d+1}} \mathrm{d}\sigma = 2\,\frac{d+1}{2}\,d\,\omega_d \int_0^\infty \frac{\log\big(1+r^2\big)}{r^2}\mathrm{d}r = \pi\, d(d+1)\,\omega_d,$$

where the latter follows from an elementary calculation. The right-hand side in (30) follows from $\omega_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, which yields

$$\pi c_{1,d}\omega_d = \sqrt{\pi}\frac{\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2}+1)} = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2}+1)} = \frac{1}{B\big(\frac{d+1}{2},\frac{1}{2}\big)}.$$

□

4. Illustration of the discrete case

Let us now discuss an application of the general principle developed in Section 2 that is very different from the one of the last section. In the case that M is a finite set, the operator L can be viewed as the generator of a continuous-time Markov chain. More precisely, L induces the generator matrix (which is commonly called Q-matrix) given by $Q = \big(q(x,y)\big)_{x,y\in M}$ where $q(x,\cdot)$ denotes the density of the measure $k(x,\cdot)$ with respect to the counting measure on $M\backslash\{x\}$. Further, one sets $q(x,x) = -\sum_{y\neq x} q(x,y)$. In this setting, the kernel $p(t,x,y)$ in assumption (A1) is given by the transition probabilities and assumption (A2) is obviously satisfied. The transition probabilities are given by the entries of the matrix exponential e^{tQ}. This, and much more information on the theory of continuous-time Markov chains can be found, for instance, in [13].

We illustrate the finite state space case by the example of the (unweighted) complete graph K_n. In this case, the corresponding Q-matrix is given by the entries ($x, y \in \{1,2,\ldots,n\}$)

$$q(x,y) = \begin{cases} 1 & , x \neq y \\ -(n-1) & , x = y. \end{cases} \tag{32}$$

One readily checks that 0 is an eigenvalue of Q of multiplicity 1 with eigenvector $(1,...,1)^T$ and $-n$ is an eigenvalue of Q of multiplicity $n-1$ with eigenvectors $(-1, e_j^{(n-1)})^T$, $j \in \{1,...,n-1\}$, where $e_j^{(n-1)}$ denotes the j-th unit vector in $\mathbb{R}^{n-1}$. From that, one determines the transition probabilities to be given by

$$p(t,x,y) = \begin{cases} \frac{1-e^{-nt}}{n} & , x \neq y, \\ \frac{1+(n-1)e^{-nt}}{n} & , x = y. \end{cases}$$

Now, we have

$$-L\big(\log p(t,\cdot,y)\big)(x) = \sum_{z\neq x} \log\frac{p(t,x,y)}{p(t,z,y)} = \begin{cases} (n-1)\log\big(\frac{1+(n-1)e^{-nt}}{1-e^{-nt}}\big) & , x = y \\ \log\big(\frac{1-e^{-nt}}{1+(n-1)e^{-nt}}\big) & , x \neq y, \end{cases}$$

which yields that $-L\big(\log(p)(t,\cdot,y)\big)(x)$ is maximal if $x = y$ for any fixed $t > 0$.

Applying Theorem 2.4, we obtain the following optimal result.

Theorem 4.1. *Let $u : [0,\infty) \times M \to (0,\infty)$ be a solution to $\partial_t u = Lu$ on $(0,\infty) \times M$, where L denotes the generator of the continuous-time Markov chain such that the complete graph K_n, $n \geqslant 2$, is the corresponding underlying graph. Then the Li-Yau inequality*

$$-L\big(\log u(t,\cdot)\big)(x) \leqslant (n-1)\log\big(\frac{1+(n-1)e^{-nt}}{1-e^{-nt}}\big)$$

holds at any $(t,x) \in (0,\infty) \times M$.

Let us compare this result with what has been obtained in [6] in the case of the complete graph K_n. The method of [6] relies on the curvature-dimension condition $CD(F;0)$ (where F is a so called CD-function). It is shown there, that $CD(F;0)$ implies a Li-Yau inequality of the form

$$-L\big(\log u(t,\cdot)\big)(x) \leqslant \varphi(t),$$

where the so-called relaxation function $\varphi : (0,\infty) \to (0,\infty)$ is defined as the unique positive solution of

$$\varphi'(t) + F(\varphi(t)) = 0, \quad t > 0, \tag{33}$$

such that $\lim_{t\to 0^+} \varphi(t) = \infty$. It is important to notice that for the uniqueness part, there is no need to assume that $\frac{F(r)}{r}$ is increasing on $(0,\infty)$. This can be seen from the proof of [6, Lemma 3.5].

It has been calculated in [6] that the function $F(r) = (n-1)\big(e^{\frac{r}{n-1}} - (n-1)e^{\frac{-r}{n-1}} + n - 2\big)$, $r \geqslant 0$, satisfies the functional inequality that defines the curvature-dimension condition but is not a CD-function in general since $\frac{F(r)}{r}$ is not increasing near 0. Since this property is of eminent importance for the method of [6], one needs to find a suitable lower bound for F which serves as a CD-function. This is when optimality in general gets lost. Remarkably, Theorem 4.1 shows that in this situation this is only an obstacle due to the method of [6], in fact the function F is tailor-made for the Li-Yau inequality of Theorem 4.1. Indeed, we can show that

$$\varphi(t) = (n-1)\log\big(\frac{1+(n-1)e^{-nt}}{1-e^{-nt}}\big)$$

is the relaxation function to F, in the sense that φ is the unique positive solution of (33) with $\lim_{t\to 0^+} \varphi(t) = \infty$.

More precisely, we have

$$\begin{aligned}\varphi'(t) &= \frac{(n-1)ne^{-nt}}{(1+(n-1)e^{-nt})(1-e^{-nt})}\big((1-n)(1-e^{-nt}) - (1+(n-1)e^{-nt})\big)\\ &= -\frac{n^2(n-1)e^{-nt}}{(1+(n-1)e^{-nt})(1-e^{-nt})}\end{aligned}$$

and besides that

$$\begin{aligned}F(\varphi(t)) &= (n-1)\Big(\frac{1+(n-1)e^{-nt}}{1-e^{-nt}} - (n-1)\frac{1-e^{-nt}}{1+(n-1)e^{-nt}} + n - 2\Big)\\ &= \frac{(n-1)\Big((1+(n-1)e^{-nt})^2 - (n-1)(1-e^{-nt})^2 + (n-2)(1-e^{-nt})(1+(n-1)e^{-nt})\Big)}{(1+(n-1)e^{-nt})(1-e^{-nt})}\\ &= \frac{(n-1)\Big(2-n+4(n-1)e^{-nt} - (n-1)(n-2)e^{-2nt} + (n-2)\big(1+(n-2)e^{-nt} - (n-1)e^{-2nt}\big)\Big)}{(1+(n-1)e^{-nt})(1-e^{-nt})}\\ &= \frac{n^2(n-1)e^{-nt}}{(1+(n-1)e^{-nt})(1-e^{-nt})}.\end{aligned}$$

References

[1] Bakry, D.; Bolley, F.; Gentil, I.: The Li-Yau inequality and applications under a curvature-dimension condition. Ann. Inst. Fourier **67** (2017), no. 1, 397–421.

[2] Bakry, D.; Gentil, I.; Ledoux, M.: *Analysis and geometry of Markov diffusion operators*. Springer, 2014.

[3] Barrios, B.; Peral, I.; Soria, F.; Valdinoci, E.: A Widder's type theorem for the heat equation with nonlocal diffusion. Arch. Ration. Mech. Anal. **213** (2014), no. 2, 629–650.

[4] Bauer, F.; Horn, P.; Lin, Y.; Lippner, G.; Mangoubi, D.; Yau, S.-T.: *Li-Yau inequality on graphs*, J. Differential Geom. **99** (2015), 359–405.

[5] Blumenthal, R. M.; Getoor, R. K.: Some theorems on stable processes. Trans. Amer. Math. Soc. **95** (1960), 263–273.

[6] Dier, D.; Kassmann, M.; Zacher, R.: Discrete versions of the Li-Yau gradient estimate. To appear in Ann. Sc. Norm. Super. Pisa Cl. Sci. (5), DOI: 10.2422/2036-2145.201807-005.

[7] Garofalo, N.: Fractional thoughts. New developments in the analysis of nonlocal operators, 1–135, Contemp. Math., **723**, Amer. Math. Soc., Providence, RI, 2019.

[8] Ghosh, T.; Kassmann, M.: Pointwise inequalities for solutions to the fractional heat equation. In preparation.

[9] Jacob, N.: Pseudo differential operators & Markov processes. Vol. II. Generators and their potential theory. Imperial College Press, London, 2002.

[10] Kim, K.-H.; Lim, S.: Asymptotic behaviors of fundamental solution and its derivatives to fractional diffusion-wave equations. J. Korean Math. Soc. **53** (2016), no. 4, 929–967.

[11] Li, P.; Yau, S.-T.: On the parabolic kernel of the Schrödinger operator. Acta. Math. **156** (1986), 153–201.

[12] Münch, F.: Li-Yau inequality on finite graphs via non-linear curvature dimension conditions. J. Math. Pures Appl. **120** (2018), 130–164.

[13] Norris, J.: *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997.

[14] Spener, A.; Weber, F.; Zacher, R.: The fractional Laplacian has infinite dimension. Comm. Partial Differential Equations **45** (2020), no. 1, 57–75.

[15] Vázquez, J.; de Pablo, A.; Quirs, F.; Rodriguez, A.: Classical solutions and higher regularity for nonlinear fractional diffusion equations. J. Eur. Math. Soc. (JEMS) **19** (2017), no. 7, 1949–1975.

A. Formalities

Der Lebenslauf wurde aus datenschutzrechtlichen Gründen entfernt.