

LONG-TIME BEHAVIOR FOR DISCRETIZATION SCHEMES OF FOKKER–PLANCK EQUATIONS VIA COUPLINGS

ANSGAR JÜNGEL AND KATHARINA SCHUH

ABSTRACT. Continuous-time Markov chains associated to finite-volume discretization schemes of Fokker–Planck equations are constructed. Sufficient conditions under which quantitative exponential decay in the ϕ -entropy and Wasserstein distance are established, implying modified logarithmic Sobolev, Poincaré, and discrete Beckner inequalities. The results are not restricted to additive potentials and do not make use of discrete Bochner-type identities. The proof for the ϕ -decay relies on a coupling technique due to Conforti, while the proof for the Wasserstein distance uses the path coupling method. Furthermore, exponential equilibration for discrete-time Markov chains is proved, based on an abstract discrete Bakry–Emery method and a path coupling.

1. INTRODUCTION

We investigate finite-volume discretizations for Fokker–Planck equations, which preserve the exponential decay of the continuous equation with respect to the entropy and Wasserstein distance. Exponential equilibration is often proved by means of the Bakry–Emery method [4]. The computations needed to apply this method use the chain rule, which is not easily available on the discrete level. Possible approaches, which avoid any discrete chain rule, are discrete Bochner-type inequalities [11] or nonlinear summation-by-parts formulas [22]. Recently, Conforti suggested in [14] a new approach that combines the Bakry–Emery method and coupling arguments as a probabilistic alternative to the discrete Bochner identities. In this paper, we extend his approach to Markov chains arising from finite-volume schemes for Fokker–Planck equations. An advantage of this approach is that we can allow for non-additive potentials, which were needed in, e.g., [25].

1.1. The setting. We are interested in the long-time behavior of finite-volume discretizations on rectangular grids of the Fokker–Planck equation

$$(1) \quad \partial_t u = \sigma^2 \Delta u + \operatorname{div}(u \nabla V)$$

on the cube $D = [-K, K]^d$ ($K > 0$) with homogeneous Neumann boundary conditions and the initial condition $u(0, x) = u_0(x)$ for $x \in D$. The diffusion constant $\sigma > 0$ is positive

Date: March 14, 2024.

2000 Mathematics Subject Classification. 60J10, 60J27, 65C40, 65J08, 65M08.

Key words and phrases. Continuous-time Markov chains, Fokker–Planck equations, finite-volume method, coupling method, discrete-time Markov chains.

The authors acknowledge partial support from the Austrian Science Fund (FWF), grants P33010 and F65. This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme, ERC Advanced Grant no. 101018153.

and the potential $V = V(x)$ is assumed to be smooth and strongly convex (see below). Equation (1) is the analytic counterpart of the stochastic differential equation (SDE)

$$(2) \quad dX_t = -\nabla V(X_t)dt + \sqrt{2\sigma^2}dB_t + \sum_{j=1}^d e_j(dL^{j+} - dL^{j-}) \quad \text{on } D$$

with the initial condition X_0 with law u_0 , e_j are the Euclidean basis vectors, and the local time $L^{j\pm}$ is a continuous, nondecreasing process such that $L_0^{j\pm} = 0$, which increases only when the process $(X_t)_{t \geq 0}$ is in the half-spaces associated to D , i.e. $L_t^{j\pm} = \int_0^t 1_{\{X_s \in \partial D_j^\pm\}} dL_s^{j\pm}$, where $\partial D_j^\pm = \{x \in D : \langle x \mp Ke_j, e_j \rangle = 0\}$ and $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^d$. The existence and uniqueness of a strong solution to (2) holds by [30]. The terms involving $L_t^{j\pm}$ ensure that the solution to the SDE is reflected at the boundary ∂D .

The probability density of the solution to the SDE (2) solves the Fokker–Planck equation (1), and the corresponding generator to (2) is of the form

$$(3) \quad \mathcal{L}f(x) = \sigma^2 \Delta f(x) - \nabla V(x) \cdot \nabla f(x) \quad \text{for } f \in \mathcal{A}, x \in D,$$

where the domain of $\mathcal{L}$ is

$$(4) \quad \mathcal{A} = \{f \in C_0^2(\mathbb{R}^d; \mathbb{R}) : \partial_j f(x) = 0 \quad \text{for } x \in \partial D_j^\pm, j = 1, \dots, d\}.$$

Note that its adjoint $\mathcal{L}^*$ is given by the right-hand side of (1).

The finite-volume discretization for (1) is defined as follows. Let $\mathcal{K}_h$ be the centers of the cells of a regular rectangular grid (see Section 2 for details). We study the semi-discretized equation of (1),

$$\partial_t u(t, i) = \mathcal{L}_h^* u(t, i) \quad \text{for } t > 0, i \in \mathcal{K}_h,$$

where $\mathcal{L}_h^*$ is a discretization of $\mathcal{L}^*$. As in the continuous case, this operator is the adjoint of some operator $\mathcal{L}_h$ of some stochastic process, namely a continuous-time Markov chain on the state space $\mathcal{K}_h$, and it can be written as

$$\mathcal{L}_h f(i) = \sum_{\gamma \in G_i} c(i, \gamma) (f(\gamma i) - f(i)),$$

where G_i is the set of all moves $\pm_j$ to the neighboring states $i \mapsto i \pm he_j$ (with the grid size $h > 0$) and $c(i, \gamma)$ denotes the rate that a jump happens from i in the direction γ ,

$$(5) \quad c(i, \pm_j) = \frac{\sigma^2}{h^2} \begin{cases} e^{-(V^h(i \pm he_j) - V^h(i))/(2\sigma^2)} & \text{if } i \pm he_j \in \mathcal{K}_h, \\ 0 & \text{else,} \end{cases}$$

where V^h is the mean value of V over the control volume with center i ,

$$V^h(i) = \frac{1}{h^d} \int_{[-h/2, h/2]^d} V(i + s) ds.$$

Because of the homogeneous Neumann boundary condition, the rate function for states at the boundary vanishes for moves going outside of the domain. We discuss the choice (5) in Remark 2. The Markov chain associated to this scheme has a unique invariant measure m_h on $\mathcal{K}_h$; see Lemma 1.

1.2. Main results and key ideas. We suppose that

- (i) ϕ and Φ , defined by $\Phi(a, b) = (\phi'(a) - \phi'(b))(a - b)$ for $a, b \in \mathbb{R}$, are convex;
- (ii) the potential V is strongly convex in the sense $\langle x - y, \nabla V(x) - \nabla V(y) \rangle \geq \kappa |x - y|^2$ for all $x, y \in \mathbb{R}^d$ and some $\kappa > 0$;
- (iii) the Hessian of the potential is strictly diagonally dominant in a sense specified in (18)–(19) below.

Condition (i) is satisfied for the power-law entropies $\phi_\alpha(x) = x^\alpha$ with $1 < \alpha \leq 2$ and the Boltzmann entropy density $\phi_\alpha(x) = x \log x$ with $\alpha = 1$. For this family of functions, the exponential decay is closely related to convex Sobolev inequalities, namely the modified logarithmic Sobolev inequality for $\alpha = 1$, the Poincaré inequality for $\alpha = 2$, and Beckner-type inequalities for $1 < \alpha < 2$. The convexity of ϕ is natural in this context, while the convexity of Φ is needed in Conforti's coupling approach; see Theorem 15 below. The strong convexity of the potential in Condition (ii) is required to conclude exponential decay. Condition (iii) is satisfied for additive potentials being of the form $V(x) = \sum_{j=1}^d V_j(x_j)$ with $x = (x_1, \dots, x_d) \in \mathbb{R}^d$; see Remark 5.

Our main results include the quantitative decay rates for the continuous-time Markov chain and its discrete-time version (obtained from an explicit Euler scheme) in the relative ϕ -entropy

$$\mathcal{H}^\phi(f|m_h) = \sum_{i \in \mathcal{K}_h} \phi(f(i))m_h(i) - \phi\left(\sum_{i \in \mathcal{K}_h} f(i)m_h(i)\right)$$

and the L^1 Wasserstein distance $\mathcal{W}_1$:

$$(6) \quad \mathcal{H}^\phi(S_t f|m_h) \leq e^{-\kappa_\phi t} \mathcal{H}^\phi(f|m_h) \quad \text{for } f \geq 0, t \geq 0,$$

$$(7) \quad \mathcal{W}_1(\nu p_t, \eta p_t) \leq \sqrt{d} e^{-\kappa_1 t} \mathcal{W}_1(\nu, \eta) \quad \text{for } t \geq 0,$$

where S_t denotes the semigroup of the Markov chain, i.e. $\partial_t(S_t f) = \mathcal{L}_h(S_t f)$, ν and η are two probability measures, and $(p_t)_{t \geq 0}$ is the transition function of the Markov chain. The constants $\kappa_\phi > 0$ and $\kappa_1 > 0$ depend on the gap of the strictly diagonally dominance for the Hessian of V from condition (iii).

The first idea of the proof of inequality (6) is based on the Bakry–Emery method. Indeed, if the convex Sobolev inequality

$$(8) \quad \kappa_\phi \mathcal{H}^\phi(f|m_h) \leq \mathcal{E}(\phi'(f), f)$$

holds for all functions $f \geq 0$, where $\mathcal{E}(f, g)$ is the Dirichlet form associated to $\mathcal{L}_h$ (see (16)), a formal computation shows that

$$\begin{aligned} \frac{d}{dt} \mathcal{H}^\phi(S_t f|m_h) &= \sum_{i \in \mathcal{K}_h} \phi'(S_t f(i)) \partial_t(S_t f)(i) m_h(i) = \sum_{i \in \mathcal{K}_h} \phi'(S_t f(i)) (\mathcal{L}_h S_t f)(i) m_h(i) \\ &= -\mathcal{E}(\phi'(S_t f), S_t f) \leq -\kappa_\phi \mathcal{H}^\phi(S_t f|m_h), \end{aligned}$$

and Grönwall's inequality implies (6). To prove the convex Sobolev inequality (8), Bakry and Emery [4] have shown that the second derivative $d^2 \mathcal{H}^\phi / dt^2$ is bounded from below by

the Dirichlet form, which implies (8). However, this task is very delicate in the discrete setting.

Therefore, the second idea is to use the coupling method of Conforti [14]. Roughly speaking, the idea is to compare two probability measures ν and η and to construct a joint probability spaces with marginals ν and η . Conforti proved (6) for Markov chains with transition rates (5) but with a different value of κ_ϕ . We recalculate κ_ϕ and interpret its value in the context of finite-volume schemes with strongly convex potentials. In the proof, we use a synchronous coupling for the contraction rates with maximal probability. When a synchronous move is not possible, Otherwise, we couple the rates of the neighboring states such that the states after a jump are identical with maximal probability or are at least neighboring states.

Inequality (7) is proved by using the path coupling method of [9]. It holds for two continuous-time Markov chains Y_t^1 and Y_t^2 that

$$\frac{d}{dt}\mathbb{E}[d(Y_t^1, Y_t^2)] = \mathbb{E}[\mathcal{L}_h^2(d(Y_t^1, Y_t^2))] \leq -\kappa_1\mathbb{E}[d(Y_t^1, Y_t^2)],$$

where d is the graph distance, $\mathcal{L}_h^2$ is the operator $\mathcal{L}_h$ on the product space $\mathcal{K}_h \times \mathcal{K}_h$, and $\kappa_1 > 0$ is related to the gap of the strictly diagonally dominance of the Hessian of V . We apply Grönwall's inequality, take the minimum over all couplings, and use the fact that the graph and Euclidean distance are equivalent (with constant $\sqrt{d}$) to conclude (7) from the previous inequality.

If the potential is additive, we obtain exponential contraction in the L^2 Wasserstein distance and, in one space dimension, in the L^p Wasserstein distance with $p \geq 2$ (up to some numerical error). The contraction rate κ_ϕ in (6) is of the order of the convexity constant κ .

We prove similar results as (6)–(7) for discrete-time Markov chains associated to an explicit Euler scheme for the Fokker–Planck equation. Here, the idea is to apply the discrete Bakry–Emery method of [22], which is based on an estimate between the Fisher information $\mathcal{E}(\phi'(f), f)$ and the entropy production $-(d\mathcal{H}^\phi/dt)(S_t f|m_h)$ (which are the same in the continuous setting). We do not obtain contraction but exponential decay with a prefactor that depends on the initial datum f .

1.3. Comparison to the literature. Based on the Bakry–Emery approach, ϕ -entropy bounds for solutions to the Fokker–Planck equation are provided in [6] from a probabilistic viewpoint and in [3] from a PDE viewpoint. For discrete settings, the analysis becomes more involved because of the lack of a general (nonlinear) chain rule. We are aware of the following approaches.

Caputo et al. used in [11] a new Bochner-type inequality, which replaces the Bochner identity of the continuous case, to prove a modified logarithmic Sobolev inequality for certain continuous-time Markov chains, including zero-range processes and Bernoulli–Laplace models. The idea to employ such an inequality was first presented in [8]. The Bochner–Bakry–Emery method was extended by Fathi and Maas in [18] in the context of Ricci

curvature bounds and in [23] to prove discrete Beckner inequalities, which were also derived in [5] using an iteration method.

Mielke investigated in [26] geodesic convexity properties of nonlocal transportation distances on probability spaces such that continuous-time Markov chains can be formulated as gradient flows. It is known that the geodesic convexity property implies exponential decay [2]. This idea was extended to nonlinear Fokker–Planck equations in [12].

A discrete Bakry–Emery method was suggested in [22]. Unlike in the continuous case, the discrete entropy production and the Fisher information are distinguished and compared to each other. The Bakry–Emery method relies on an estimate of the production of the Fisher information, which requires discrete versions of integrations by parts and suitable chain rules. The nonlinear integration-by-parts formulas are “translated” to the discrete case by using the systematic integration-by-parts method of [21]. This method allows for the treatment of nonlinear equations, but it seems to be restricted to numerical three-point schemes and hence to one-dimensional equations only.

Conforti established in [14] a new probabilistic approach to prove convex Sobolev inequalities of the type (8) and to quantify the exponential decay in terms of the entropy for continuous-time Markov chains. His approach uses coupling rates to treat the second time derivative of the entropy, thus avoiding the use of discrete Bochner-type inequalities. Recently, Petrotti [29] has analyzed how contractive coupling rates can be used to prove stronger inequalities in the form of curvature lower bounds for Markov chains and geodesic convexity of entropy functionals.

Maas and Matthes [25] suggested a finite-volume discretization for a nonlinear parabolic fourth-order equation, which can be written as a gradient flow similarly as the Fokker–Planck equation (1). The long-time asymptotics was shown for additive potentials by taking advantage of the factorization of the problem. Cancès and Venel [10] studied an implicit Euler finite-volume approximations of nonlinear Fokker–Planck equations using a discretization of the flux, which was motivated by probabilistic exclusion processes (and which have some similarity to our rates (5)). Their approach allows for non-additive potentials, but the long-time asymptotics was not investigated. The finite-volume scheme of [13] allows for the proof of exponential equilibration, but the decay rate depends on the size of the domain (because of the use of the Poincaré inequality; also see [19]), while we assume strong convexity of the potential. A link between finite-volume-type discretizations for the Fokker–Planck equation in Voronoï meshes and gradient flows with discrete Wasserstein distances on graphs was established in [1].

In this paper, we use Conforti’s coupling method to quantify the exponential decay of finite-volume schemes allowing for non-additive potentials and with a decay rate only depending on the potential and not on the domain size.

This work is organized as follows. In Section 2, the continuous-time Markov chain is defined and its unique invariant measure is determined. Our main results for continuous-time and discrete-time Markov chain are presented in Sections 3 and 4, respectively. We introduce Conforti’s coupling method in Section 5 and prove the main results in Section 6 (for continuous-time Markov chains) and Section 7 (for discrete-time Markov chains).

2. DEFINITION OF THE CONTINUOUS-TIME MARKOV CHAIN

The state space is defined by $D = [-K, K]^d$ for $K > 0$, $d \geq 1$. Let $h > 0$ be such that $2K/h \in \mathbb{N}$, and let $e_1, \dots, e_d$ be the Euclidean basis vectors of $\mathbb{R}^d$. The rectangular finite-volume mesh is given by a family of control volumes (cubes)

$$N_i = [i_1 - h/2, i_1 + h/2] \times \dots \times [i_d - h/2, i_d + h/2]$$

with center $i = (i_1, \dots, i_d) \in \mathcal{K}_h$, where $i_j = (-K + n_j + h/2)$ for $n_j \in \{1, \dots, 2K/h\}$. Let $\sigma > 0$ be a diffusion constant and $V \in C^2(\mathbb{R}^d)$ be a potential.

To define the continuous-time Markov chain, we recall definition (5) of the transition rates. The rate $c(i, \pm j)$ defines the numerical flux from the cell N_i to $N_{i \pm he_j}$. The definition of $c(i, \pm j)$ implies that the numerical flux at the boundary vanishes. The rates are path-independent in the sense

$$(9) \quad c(i, \pm j)c(i \pm he_j, \pm \ell) = c(i, \pm \ell)c(i \pm he_\ell, \pm j) \quad \text{for } i \in \mathcal{K}_h, j, \ell = 1, \dots, d.$$

This identity means that the rate of jumping first in direction j and then in direction ℓ is the same as moving first in direction ℓ and then j . The discrete generator $\mathcal{L}_h$ can be written as

$$(10) \quad \mathcal{L}_h f(i) = \sum_{j=1}^d [c(i, +j)(f(i + he_j) - f(i)) + c(i, -j)(f(i - he_j) - f(i))].$$

With these preparations, we can define the continuous-time Markov chain $(Y_t^h)_{t \geq 0}$ on $\mathcal{K}_h$ via the discrete generator $\mathcal{L}_h$. Recall that G_i denotes the set of all moves from i to $i \pm he_j$ for $j = 1, \dots, d$, written as $\pm j$. Let $(Z_n^h)_{n \in \mathbb{N}}$ be a discrete-time Markov chain on $\mathcal{K}_h$ with initial distribution ν , given by the transition kernel

$$(11) \quad \pi(i, \bar{i}) = \begin{cases} c(i, \gamma)/\mathcal{T} & \text{if there exists } \gamma \in G_i \text{ such that } \bar{i} = \gamma i, \\ 1 - \sum_{\gamma \in G_i} c(i, \gamma)/\mathcal{T} & \text{if } i = \bar{i}, \\ 0 & \text{else,} \end{cases}$$

where $\mathcal{T} = 2 \max_{i \in \mathcal{K}_h} \sum_{\gamma \in G_i} c(i, \gamma)$. Let T_n for $n \in \mathbb{N}$ be independent exponential random variables with parameter one, which are independent of $(Z_n^h)_{n \in \mathbb{N}}$. Set $S_n = \sum_{k=1}^n T_k/\mathcal{T}$. Then the continuous-time Markov chain is defined by

$$(12) \quad Y_t^h = \begin{cases} Z_n^h & \text{if } S_n \leq t < S_{n+1} \text{ for some } n \in \mathbb{N}, \\ \infty & \text{else.} \end{cases}$$

The corresponding transition function is denoted by $(p_t)_{t \geq 0}$.

The discrete Fokker–Planck equation on $\mathcal{K}_h$ is given by $\partial_t u(t, i) = \mathcal{L}_h^* u(t, i)$, where $t \geq 0$, $i \in \mathcal{K}_h$, and the adjoint generator reads as

$$(13) \quad \mathcal{L}_h^* g(i) = \sum_{j=1}^d (c(i - he_j, +j)g(i - he_j) + c(i + he_j, -j)g(i + he_j) - c(i, +j)g(i) - c(i, -j)g(i)),$$

with homogeneous Neumann boundary conditions. We observe that if V is additive, i.e. $V(i) = \sum_{j=1}^d V_j(i_j)$ for $i \in \mathcal{K}_h$ (which allows for a factorization of the invariant measure), the finite-volume scheme and the generator coincide with the scheme and generator considered in [25].

The following lemma determines the invariant measure associated to the Markov chain.

Lemma 1. *The unique invariant measure m_h of the continuous-time Markov chain $(Y_t^h)_{t \geq 0}$ with the transition rates (5) is given by*

$$(14) \quad m_h(i) = Z^{-1} \exp(-V^h(i)/\sigma^2) \quad \text{for } i \in \mathcal{K}_h$$

with the normalization constant $Z = \sum_{i \in \mathcal{K}_h} \exp(-V^h(i)/\sigma^2)$.

Proof. As the state space is finite and the Markov chain is irreducible and aperiodic, there exists a unique invariant measure m_h . To show that (14) is that invariant measure, we verify that $\sum_{i \in \mathcal{K}_h} \mathcal{L}_h f(i) m_h(i) = 0$ for all positive functions f on $\mathcal{K}_h$. We use summation by parts to factor out $f(i)$ and then insert definition (14):

$$\begin{aligned} & \sum_{i \in \mathcal{K}_h} \mathcal{L}_h f(i) m_h(i) \\ &= \sum_{i \in \mathcal{K}_h} \sum_{j=1}^d [c(i, +_j)(f(i + he_j) - f(i)) + c(i, -_j)(f(i - he_j) - f(i))] m_h(i) \\ &= \sum_{i \in \mathcal{K}_h} \sum_{j=1}^d f(i) [c(i - he_j, +_j) m_h(i - he_j) + c(i + he_j, -_j) m_h(i + he_j) \\ & \quad - c(i, +_j) m_h(i) - c(i, -_j) m_h(i)] \\ &= \sum_{i \in \mathcal{K}_h} \sum_{j=1}^d \frac{f(i)}{Z} \left(c(i - he_j, +_j) e^{-V^h(i - he_j)/\sigma^2} 1_{\{i - he_j \in \mathcal{K}_h\}} - c(i, +_j) e^{-V^h(i)/\sigma^2} \right. \\ & \quad \left. + c(i + he_j, -_j) e^{-V^h(i + he_j)/\sigma^2} 1_{\{i + he_j \in \mathcal{K}_h\}} - c(i, -_j) e^{-V^h(i)/\sigma^2} \right). \end{aligned}$$

Finally, we insert definition (5) of the rates c and rearrange the terms:

$$\begin{aligned} \sum_{i \in \mathcal{K}_h} \mathcal{L}_h f(i) m_h(i) &= \sum_{i \in \mathcal{K}_h} \sum_{j=1}^d \frac{\sigma^2}{h^2 Z} f(i) \left(\left[e^{-(V^h(i) - V^h(i - he_j))/(2\sigma^2)} e^{-V^h(i - he_j)/\sigma^2} \right. \right. \\ & \quad \left. - e^{-(V^h(i - he_j) - V^h(i))/(2\sigma^2)} e^{-V^h(i)/\sigma^2} \right] 1_{\{i - he_j \in \mathcal{K}_h\}} \\ & \quad + \left[e^{-(V^h(i) - V^h(i + he_j))/(2\sigma^2)} e^{-V^h(i + he_j)/\sigma^2} \right. \\ & \quad \left. - e^{-(V^h(i + he_j) - V^h(i))/(2\sigma^2)} e^{-V^h(i)/\sigma^2} \right] 1_{\{i + he_j \in \mathcal{K}_h\}} \Big) = 0, \end{aligned}$$

since the expressions in the brackets $[\dots]$ vanish. $\square$

We remark that the continuous-time Markov chain $(Y_n^h)_{n \in \mathbb{N}}$ satisfies the detailed-balance condition $c(i, +_j) m_h(i) = c(i + he_j, -_j) m_h(i + he_j)$ for all $i \in \mathcal{K}_h$ and $j = 1, \dots, d$, which

implies the reversibility of the continuous-time Markov chain starting with initial measure m_h .

Remark 2 (Choice of $c(i, \gamma)$). By Definition (14) of the invariant measure, we can write the discrete generator as

$$\mathcal{L}_h f(i) = \frac{\sigma^2}{h^2 Z} \sum_{j=1}^d \left(\frac{M_h(i, +j)}{m_h(i)} (f(i + he_j) - f(i)) + \frac{M_h(i, -j)}{m_h(i)} (f(i - he_j) - f(i)) \right),$$

$$\text{where } M_h(i, \pm j) = \sqrt{m_h(i \pm he_j) m_h(i)}.$$

The mean value $M_h(i, \pm j)$ corresponds to the geometric mean used in [26, Cor. 5.5] (also see [23, Sec. 4]). Other choices for $M_h(i, \pm j)$ can be found in [20, Sec. 3.2] and [13, Sec. 2.1.3]. We observe that the difference of the flows to and from the neighboring cells satisfies

$$c(i, +j) - c(i + he_j, -j) = 2 \frac{\sigma^2}{h^2} \sinh \left(\frac{V^h(i) - V^h(i + he_j)}{2\sigma^2} \right).$$

This is related to the stochastic jump process in [7, Sec. 1.2] and to the cosh structure of the dissipation potential of generalized gradient flows from [24]. In [7], the transition rates of the one-dimensional Markov jump process depend only on the energy decay of the current state and do not take into account the energy difference between the current and the next state. Therefore, carrying these transition rates to multiple space dimensions, the path-independence property (9), which is essential in our analysis, does not hold in general. $\square$

3. MAIN RESULTS FOR CONTINUOUS-TIME MARKOV CHAINS

We present the main results for continuous-time Markov chains. The proofs will be given in Appendix A and Section 6.

3.1. Convergence of the Markov chain approximation to the SDE. Let $(Y_h^t)_{t \geq 0}$ be the continuous-time Markov chain solving the martingale problem with operator $\mathcal{L}_h$ defined in (10), and let $(X_t)_{t \geq 0}$ be the solution to the SDE (2), which solves the martingale problem with operator $\mathcal{L}$ given by (3). A Taylor expansion shows that $\mathcal{L}f(i) - \mathcal{L}_h f(i) = O(h)$ for $i \in \mathcal{K}_h$ and functions $f \in \mathcal{A}$, defined in (4). This estimate is crucial in the proof that the Markov chain converges in distribution to the solution to (2) as the grid size converges to zero, $h \rightarrow 0$. The following theorem is proved in Appendix A.

Theorem 3 (Convergence to SDE). *Let $(Y_t^h)_{t \geq 0}$ be the continuous-time Markov chain with generator $\mathcal{L}_h$ on the grid $D \cap (h\mathbb{Z})^d$ such that the laws of Y_0^h converge in distribution to the measure μ_0 as $h \rightarrow 0$. Then $(Y_t^h)_{t \geq 0}$ converges in distribution to the solution $(X_t)_{t \geq 0}$ to (2) with initial datum μ_0 .*

To construct a continuous-time Markov chain that converges to the solution to the SDE $dX_t = -\nabla V(X_t)dt + \sqrt{2\sigma}dB_t$ in the whole space, with $(B_t)_{t \geq 0}$ being a Brownian motion, we need to modify the cube $D = [-K, K]^d$. In fact, we replace K by K_h for $h > 0$ such

that $K_h \rightarrow \infty$ as $h \rightarrow 0$. Then Theorem 3 holds similarly after adapting the diffusion approximation of [17, Theorem 7.4.1].

3.2. Exponential decay in ϕ -entropy. Consider the Markov chain $(Y_t^h)_{t \geq 0}$ on $\mathcal{K}_h$, characterized by the generator (10). Let $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and $f, g : \mathcal{K}_h \rightarrow \mathbb{R}_+$ be functions, where $\mathbb{R}_+ := [0, \infty)$. We recall the definition of the ϕ -entropy, relative to the invariant measure m_h :

$$(15) \quad \mathcal{H}^\phi(f|m_h) = \sum_{i \in \mathcal{K}_h} \phi(f(i))m_h(i) - \phi\left(\sum_{i \in \mathcal{K}_h} f(i)m_h(i)\right).$$

We also introduce the Dirichlet form

$$(16) \quad \mathcal{E}(f, g) = - \sum_{i \in \mathcal{K}_h} f(i)(\mathcal{L}_h g)(i)m_h(i) = \frac{1}{2} \sum_{i \in \mathcal{K}_h} \sum_{\gamma \in G} c(i, \gamma) \nabla_\gamma f(i) \nabla_\gamma g(i) m_h(i),$$

where we have set $\nabla_\gamma f(i) = f(\gamma i) - f(i)$. The last identity follows from the reversibility of the Markov chain (see, e.g., [15, (2.12)]).

Our aim is to find the optimal constant κ_ϕ such that the exponential decay (6) in the ϕ -entropy holds for all functions $f : \mathcal{K}_h \rightarrow \mathbb{R}_+$ and $t > 0$, where S_t is the Markovian semigroup generated by $\mathcal{L}_h$. For continuous-time processes, this problem is known to be equivalent to find the optimal constant for the convex Sobolev inequality (8). Introduce the family of functions

$$(17) \quad \phi_\alpha(x) = \begin{cases} (\alpha - 1)^{-1}(x^\alpha - x) - x + 1 & \text{if } 1 < \alpha \leq 2, \\ x \log x - x + 1 & \text{if } \alpha = 1. \end{cases}$$

Then the convex Sobolev inequality becomes the modified logarithmic Sobolev inequality for $\alpha = 1$, the Poincaré inequality for $\alpha = 2$, and the Beckner inequality for $1 < \alpha < 2$.

We need the following assumptions.

- (A1) Convexity of entropy: Let the functions $\phi \in C^1(\mathbb{R}_+; \mathbb{R}_+)$ and $\Phi : (0, \infty)^2 \rightarrow \mathbb{R}_+$, defined by $\Phi(a, b) = (\phi'(a) - \phi'(b))(a - b)$, be convex.
- (A2) Strong κ -convexity: Let $V \in C^1(D; \mathbb{R})$ be such that there exists $\kappa > 0$ such that $\langle x - y, \nabla V(x) - \nabla V(y) \rangle \geq \kappa |x - y|^2$ for all $x, y \in D$, where $\langle \cdot, \cdot \rangle$ is the inner product on $\mathbb{R}^d$.
- (A3) It holds that for all $i, i + he_j \in \mathcal{K}_h$ and $j = 1, \dots, d$,

$$(18) \quad \kappa_+(i, j) := c(i, +_j) - c(i + he_j, +_j) - \sum_{\gamma \in G \setminus \{\pm_j\}} \max\{c(i + he_j, \gamma) - c(i, \gamma), 0\} > 0,$$

and for all $i, i - he_j \in \mathcal{K}_h$ and $j = 1, \dots, d$,

$$(19) \quad \kappa_-(i, j) := c(i, -_j) - c(i - he_j, -_j) - \sum_{\gamma \in G \setminus \{\pm_j\}} \max\{c(i - he_j, \gamma) - c(i, \gamma), 0\} > 0.$$

Theorem 4 (Exponential decay). *Let Assumptions (A1)–(A3) hold. Consider the continuous-time Markov chain $(Y_t^h)_{t \geq 0}$ on $\mathcal{K}_h$ with transition rates (5), which satisfy (9). Then the convex Sobolev inequality (8) holds with constant*

$$\kappa_\phi = 2 \min_{i, i+he_j \in \mathcal{K}_h, j=1, \dots, d} \{\kappa_+(i, j), \kappa_-(i+he_j, j)\},$$

and the exponential decay (6) is valid. Moreover, the modified logarithmic Sobolev inequality holds with $\kappa_1 = 2\kappa_\phi$ and the discrete Beckner inequality holds with $\kappa_\alpha = \alpha\kappa_\phi$ for $\alpha \in (1, 2]$.

Compared to [25], the coupling approach allows us to show exponential decay in the ϕ -entropy for more general (non-additive) potentials. We observe that the constant κ_ϕ is independent of the size of the domain D . Moreover, κ_ϕ scales linearly with the size of the transition rates. In particular, if we multiply the rates by some constant and change the rate when a jump occurs, κ_ϕ changes by this factor.

In the following, we discuss Assumptions (A1)–(A3). The functions ϕ_α from (17) and $\Phi(a, b) = (\phi'_\alpha(a) - \phi'_\alpha(b))(a - b)$ satisfy Assumption (A1).

Remark 5 (Additive potentials). For additive potentials $V(i) = \sum_{j=1}^d V_j(i_j)$ ($i \in \mathcal{K}_h$) that satisfy Assumption (A2), also Assumption (A3) is fulfilled. Indeed, for jumps $\gamma \neq \pm_j$, a computation shows that $c(i \pm he_j, \gamma) - c(i, \gamma) = 0$. Then

$$\begin{aligned} \kappa_+(i, j) &= c(i, +_j) - c(i, +he_j, +_j) \\ &= \frac{\sigma^2}{h^2} \left(e^{-(V_j^h(i_j+h) - V_j^h(i_j))/(2\sigma^2)} - e^{-(V_j^h(i_j+2h) - V_j^h(i_j+h))/(2\sigma^2)} \right) \\ &= \frac{\sigma^2}{h^2} \exp \left(-\frac{1}{2\sigma^2} \int_0^h \partial V_j^h(i_j + s) ds \right) \\ &\quad \times \left[1 - \exp \left(-\frac{1}{2\sigma^2} \int_0^h (\partial V_j^h(i_j + h + s) - \partial V_j^h(i_j + s)) ds \right) \right]. \end{aligned}$$

Assumption (A2) implies for additive potentials that $h(\partial V_j^h(i_j + h + s) - \partial V_j^h(i_j + s)) \geq \kappa h^2$ and hence

$$\kappa_+(i, j) \geq \frac{\sigma^2}{h^2} \exp \left(-\frac{1}{2\sigma^2} \int_0^h \partial V_j^h(i_j + s) ds \right) \left[1 - \exp \left(-\frac{1}{2\sigma^2} \int_0^h \kappa h ds \right) \right] > 0.$$

It follows similarly that $\kappa_-(i, j) > 0$, thus verifying Assumption (A3). We observe that the last expression converges to $\kappa/2$ as $h \rightarrow 0$. Thus, the decay rate κ_ϕ is asymptotically optimal in this situation. $\square$

Remark 6 (Strictly diagonally dominance of D^2V). Assumption (A3) can be interpreted as a modified strictly diagonally dominant condition on the Hessian of V . Indeed, let $V \in C^2(D; \mathbb{R})$ and let the Hessian $D^2V(i) \in \mathbb{R}^{d \times d}$ of V be strictly diagonally dominant, i.e., there exists $c > 0$ such that for all $i \in \mathcal{K}_h$ and $j = 1, \dots, d$,

$$(D^2V)_{jj}(i) - \sum_{\ell \neq j} |(D^2V)_{\ell j}(i)| \geq c.$$

Let $\partial_{\pm j}^h$ be the discrete derivative on $\mathcal{K}_h$ in direction j , defined by $\partial_{\pm j}^h f(i) = \pm(f(i + he_j) - f(i))/h$. A Taylor expansion yields

$$\begin{aligned} c(i, +\ell) - c(i + he_j, +\ell) &= \partial_{+j}^h \partial_{+\ell}^h V^h(i) + O(h), \\ c(i, -\ell) - c(i + he_j, -\ell) &= \partial_{+j}^h \partial_{-\ell}^h V^h(i) + O(h), \text{ etc.} \end{aligned}$$

Then $\kappa_{\pm}(i, j)$ from Assumption (A3) can be approximated by

$$\begin{aligned} \kappa_+(i, j) &= \partial_{+j}^h \partial_{+j}^h V^h(i) - \sum_{\ell \neq j} (\max\{\partial_{+j}^h \partial_{+\ell}^h V^h(i), 0\} + \max\{-\partial_{+j}^h \partial_{-\ell}^h V^h(i), 0\}) + O(h), \\ \kappa_-(i, j) &= \partial_{-j}^h \partial_{-j}^h V^h(i) - \sum_{\ell \neq j} (\max\{\partial_{-j}^h \partial_{-\ell}^h V^h(i), 0\} + \max\{-\partial_{-j}^h \partial_{+\ell}^h V^h(i), 0\}) + O(h). \end{aligned}$$

In the limit $h \rightarrow 0$, the discrete second derivatives of V^h converge to the Hessian of V . Thus, if D^2V is strictly diagonally dominant, Assumption (A3) is satisfied for sufficiently small $h > 0$. $\square$

3.3. Exponential decay in the Wasserstein distance. Let $p \geq 1$ and let ν, η be two probability measures on $\mathcal{K}_h$ with finite p th moment. The L^p Wasserstein distance between ν and η with respect to the Euclidean distance is defined by

$$(20) \quad \mathcal{W}_p(\nu, \eta) = \inf_{\gamma \in \Gamma(\nu, \eta)} \left(\int_{\mathcal{K}_h \times \mathcal{K}_h} |x - y|^p \gamma(dx dy) \right)^{1/p},$$

where $\Gamma(\nu, \eta)$ denotes the set of all couplings between ν and η , i.e. the set of all probability measures on $\mathcal{K}_h \times \mathcal{K}_h$ with marginals ν and η . We first present results for additive potentials.

Theorem 7 (Convergence in Wasserstein distance). *Let the potential $V \in C^2(\mathbb{R})$ be of additive form, $V(i) = \sum_{j=1}^d V_j(i_j)$ for $i \in \mathcal{K}_h$, satisfies Assumption (A2), and has a Lipschitz continuous gradient. Furthermore, let ν and η be two probability measures on $\mathcal{K}_h$ and $(Y_t^h)_{t \geq 0}$ be the continuous-time Markov chain with transition function $(p_t)_{t \geq 0}$ associated to the transition rates (5). Then it holds with the constant $\kappa > 0$ from Assumption (A2) that*

$$\mathcal{W}_2(\nu p_t, \eta p_t) \leq e^{-\kappa t} \mathcal{W}_2(\nu, \eta) + O(h^{1/2}) \quad \text{for all } t \geq 0.$$

Moreover, in one space dimension and for $p \geq 2$,

$$\mathcal{W}_p(\nu p_t, \eta p_t) \leq e^{-\kappa t} \mathcal{W}_p(\nu, \eta) + O(h^{1/p}) \quad \text{for all } t \geq 0.$$

The error term $O(h^{1/p})$ can be avoided by relaxing the contraction rate to $\kappa - O(h)$ for sufficiently small $h > 0$. If the potential is not of additive form, we can still conclude long-time guarantees in the L^1 Wasserstein distance.

Theorem 8 (Convergence in L^1 Wasserstein distance). *Let the potential satisfy Assumptions (A2) and (A3). Furthermore, let ν and η be two probability measures on $\mathcal{K}_h$ and $(Y_t^h)_{t \geq 0}$ be the continuous-time Markov chain with transition function $(p_t)_{t \geq 0}$ associated to the transition rates (5). Then it holds that*

$$\mathcal{W}_1(\nu p_t, \eta p_t) \leq \sqrt{d} e^{-\kappa_1 t} \mathcal{W}_1(\nu, \eta) \quad \text{for all } t \geq 0,$$

where d is the dimension of the state space and

$$(21) \quad \kappa_1 = \min_{i \in \mathcal{K}_h, j=1, \dots, d} (\kappa_+(i, j) + \kappa_-(i, j)).$$

Remark 9 (Discussion of Theorem 8). The contraction rate does not equal $\kappa > 0$ from Assumption (A2) but is generally smaller. This is due to the fact that we use a rectangular mesh in the approximation. In the case of additive potentials, the value κ_1 is of order κ . The prefactor $\sqrt{d}$ originates from the change of the Wasserstein distance $\mathcal{W}_{d,1}$ with respect to the graph distance to the Euclidean Wasserstein distance $\mathcal{W}_1$. Indeed, we prove first, using the path coupling method of [9], that

$$(22) \quad \mathcal{W}_{d,1}(\nu p_t, \eta p_t) \leq e^{-\kappa_1 t} \mathcal{W}_{d,1}(\nu, \eta),$$

where $d(x, y) = \sum_{j=1}^d |x_j - y_j|$ for $x, y \in \mathbb{R}^d$ is the graph distance. Our result then follows from the fact that the graph distance is equivalent to the Euclidean distance, resulting in the prefactor $\sqrt{d}$. $\square$

Remark 10 (Finite-difference scheme). The finite-difference discretization of (1) reads as

$$\partial_t u(t, i) = \mathcal{L}_h^* u(t, i) = \sigma^2 \Delta_h u(t, i) + \operatorname{div}_h(u(t, i) \nabla V(t, i)),$$

where the discrete Laplacian and divergence are defined by

$$\begin{aligned} \Delta_h f(i) &= \frac{1}{h^2} \sum_{j=1}^d (f(i + h e_j) - 2f(i) + f(i - h e_j)), \\ \operatorname{div}_h g(i) &= \frac{1}{2h} \sum_{j=1}^d (g_j(i + h e_j) - g_j(i - h e_j)) \end{aligned}$$

for $i \in \mathcal{K}_h$ and functions $f : \mathcal{K}_h \rightarrow \mathbb{R}$ and $g : \mathcal{K}_h \rightarrow \mathbb{R}^d$. The generator of the associated Markov chain reads as (3), but the jump rates are now given by

$$c(i, \pm j) = \begin{cases} h^{-2}(\sigma^2 \mp h \partial_j V(i)/2) & \text{if } i + h e_j, i - h e_j \in \mathcal{K}_h, \\ h^{-2}(2\sigma^2 \mp h \partial_j V(i)) & \text{if } i \mp h e_j \notin \mathcal{K}_h, \\ 0 & \text{if } i \pm h e_j \notin \mathcal{K}_h. \end{cases}$$

At the boundary $\partial \mathcal{K}_h$, the process is reflected, which corresponds at the level of the Fokker-Planck equation to homogeneous Neumann boundary conditions. To ensure the positivity of the rates, we need to impose the general assumption

$$(23) \quad \inf_{x \in D, j=1, \dots, d} \left(\sigma^2 - \frac{h}{2} |\partial_j V(x)| \right) \geq 0.$$

We observe that the rates given above are a first-order approximation of the jump rates (5). As for the finite-volume scheme, we define the continuous-time Markov chain $(Y_t^h)_{t \geq 0}$ on $\mathcal{K}_h$ via the discrete generator $\mathcal{L}_h$ by (12). Then $(Y_t^h)_{t \geq 0}$ is non-explosive, irreducible, and recurrent.

For additive potentials $V(x) = \sum_{j=1}^d V_j(x_j)$, a computation shows that the invariant measure of $(Y_t^h)_{t \geq 0}$ equals

$$m_h(i) = \frac{1}{Z} \prod_{j=1}^d \prod_{\ell \in \mathcal{K}_h, \ell < i_j} \frac{\sigma^2 - h \partial V_j(\ell)/2}{\sigma^2 + h \partial V_j(\ell + h)/2},$$

where $Z > 0$ is a normalization constant. For non-additive potentials and multiple space dimensions, the structure of the invariant measure becomes more complicated compared to the finite-volume scheme and we leave details to the reader.

The finite-difference scheme has the drawback that the rates are not commutative for more than one space dimension in the sense that (9) does not hold. However, assuming that the potential is additive, it is possible to carry over the results stated for the finite-volume scheme to the finite-difference discretization (supposing condition (23)). Again, we leave the details to the reader. $\square$

4. MAIN RESULTS FOR DISCRETE-TIME MARKOV CHAINS

4.1. Exponential decay in ϕ -entropy. We consider a discrete-time Markov chain as an approximation of the solution to the SDE (2). Instead of exponentially distributed random times, we choose equidistant time steps, which means that for fixed $\tau > 0$, the Fokker–Planck equation is discretized according to the explicit Euler finite-volume scheme

$$\frac{1}{\tau} (u(t_{n+1}, i) - u(t_n, i)) = \mathcal{L}_h^* u(t_n, i) \quad \text{for } i \in \mathcal{K}_h, \quad t_n = n\tau, \quad n \in \mathbb{N},$$

where $\mathcal{L}_h^*$ is defined in (13). The corresponding Markov chain $(Z_n^h)_{n \in \mathbb{N}}$ is defined through the transition kernel (11). Interpreting π as the matrix with entries $\pi(i, k)$, we have

$$\pi f(i) = \sum_{k \in \mathcal{K}_h} \pi(i, k) f(k).$$

With the choice of the rescaling factor $\mathcal{T} = 2 \max_{i \in \mathcal{K}_h} \sum_{\gamma \in G_i} c(i, \gamma)$, we obtain a lazy random walk, where we remain at the current position with probability at least 1/2. Furthermore, we introduce the rescaled jump rates

$$(24) \quad p(i, \gamma) = \frac{c(i, \gamma)}{\mathcal{T}} \quad \text{for } \gamma \in G_i.$$

and we set $\tau := \mathcal{T}^{-1}$.

In the situation of Lemma 1, the unique invariant measure of the discrete-time Markov chain $(Z_n^h)_{n \in \mathbb{N}}$ with transition kernel (11) is still given by (14). Indeed, the laziness of the random walk ensures the aperiodicity of the Markov chain, which yields the uniqueness of an invariant measure and moreover, it holds for all functions $f : \mathcal{K}_h \rightarrow \mathbb{R}$ that

$$(25) \quad \sum_{i \in \mathcal{K}_h} \pi f(i) m_h(i) = \sum_{i \in \mathcal{K}_g} \sum_{\gamma \in G} p(i, \gamma) f(\gamma i) m_h(i) = \sum_{i \in \mathcal{K}_h} f(i) m_h(i).$$

Unfortunately, we cannot use the convex Sobolev inequality (8) to prove the exponential decay in the ϕ -entropy, since the discrete time derivative of the ϕ -entropy does not correspond to the Dirichlet form. We overcome this issue by applying the abstract method of [22], which distinguishes the entropy production and the Fisher information (which both equal the Dirichlet form in the continuous case). Interestingly, while the approach of [22] is based on the implicit Euler scheme, we are able to analyze the explicit Euler scheme.

We introduce for $f \geq 0$ the (discrete) entropy production $\mathcal{P}(f)$ and Fisher information $\mathcal{F}(f)$, respectively, by

$$\mathcal{P}(f) = -\tau^{-1}(\mathcal{H}^\phi(\pi f|m_h) - \mathcal{H}^\phi(f|m_h)), \quad \mathcal{F}(f) = \mathcal{E}(\phi'(f), f).$$

The following results are proved in Section 7.

Proposition 11. *We assume that*

- (i) *There exists $C_P > 0$ such that $0 \leq \mathcal{P}(\pi^n f) \leq C_P \mathcal{F}(\pi^n f)$ for all $f \geq 0$ and $n \in \mathbb{N}$.*
- (ii) *There exists $\lambda > 0$ such that $\mathcal{F}(\pi^{n+1} f) - \mathcal{F}(\pi^n f) \leq -\tau \lambda \mathcal{F}(\pi^n f)$ for all $f \geq 0$ and $n \in \mathbb{N}$.*
- (iii) $\lim_{n \rightarrow \infty} \mathcal{H}^\phi(\pi^n f|m_h) = 0$.

Then for all $f \geq 0$, $n \in \mathbb{N}$, and $\tau < 1/\lambda$,

$$\mathcal{H}^\phi(\pi^n f|m_h) \leq C_f e^{-\lambda n \tau} \mathcal{H}^\phi(f|m_h),$$

where $C_f = C_P \mathcal{F}(f)/(\lambda \mathcal{H}^\phi(f|m_h))$.

We impose a weaker condition compared to [22, Prop. 1], but we obtain a weaker result. In particular, the prefactor C_f depends on f and may be very large. This prefactor equals one in [22], thus providing a contraction result, but assuming the lower bound $c_P \mathcal{F}(\pi^n f) \leq \mathcal{P}(\pi^n f)$ for some $c_P > 0$. Unfortunately, this bound does not hold in our situation. It is not surprising that our result is weaker than in [22], since we consider an explicit scheme, while an implicit scheme is studied in [22].

Using the same coupling approach as for the continuous-time Markov chain, we conclude the exponential decay for the discrete-time Markov chain.

Theorem 12 (Exponential decay). *Let Assumptions (A2) and (A3) hold and let the discrete-time Markov chain with transition matrix (11) be given. Then, for ϕ_α defined in (17), there exists $C_f^\alpha > 0$ such that for $f \geq 0$, $n \in \mathbb{N}$, and $\tau < 1/\kappa_\phi$,*

$$\mathcal{H}^{\phi_\alpha}(\pi^n f|m_h) \leq C_f^\alpha e^{-\kappa_\phi n \tau} \mathcal{H}^{\phi_\alpha}(f|m_h),$$

where $C_f^\alpha > 0$ is as in Proposition 11 with ϕ replaced by ϕ_α , and κ_ϕ is as in Theorem 4.

4.2. Exponential decay in the Wasserstein distance. Analogously to the continuous-time situation, we obtain contraction in the Wasserstein distance up to some error term.

Theorem 13 (Convergence in Wasserstein distance). *Let the potential $V \in C^1(\mathbb{R})$ be of additive form, $V(i) = \sum_{j=1}^d V_j(i_j)$ for $i \in \mathcal{K}_h$, satisfies Assumption (A2), and has a Lipschitz continuous gradient. Furthermore, let ν and η be two probability measures on $\mathcal{K}_h$,*

and consider the discrete-time Markov chain with transition matrix (11). Then it holds with the constant $\kappa > 0$ from Assumption (A2) that

$$\mathcal{W}_2(\nu\pi^n, \eta\pi^n) \leq e^{-\kappa n\tau} \mathcal{W}_2(\nu, \eta) + O(h^{1/2}) \quad \text{for all } n \in \mathbb{N}.$$

Moreover, if the potential satisfies additionally Assumption (A3),

$$(26) \quad \mathcal{W}_{d,1}(\nu\pi^n, \eta\pi^n) \leq e^{-\kappa_1 n\tau} \mathcal{W}_{d,1}(\nu, \eta) \quad \text{for all } n \in \mathbb{N},$$

and

$$\mathcal{W}_1(\nu\pi^n, \eta\pi^n) \leq \sqrt{d} e^{-\kappa_1 n\tau} \mathcal{W}_1(\nu, \eta) \quad \text{for all } n \in \mathbb{N},$$

where $\kappa_1 > 0$ is defined in (21).

The theorem is proved in the same way as Theorems 7 and 8.

Remark 14 (Coarse Ricci curvature). If the transition kernel π on $\mathcal{K}_h$ satisfies (26), we say that the triple $(\mathcal{K}_h, \pi, d)$ has a coarse Ricci curvature at least $\kappa_1\tau$ in the sense of Ollivier [28]. Furthermore, according to [16], the decay (22) implies that

$$\mathcal{W}_{d,1}(\nu, m_h) \leq \left(\frac{2\mathcal{H}^{\phi_1}(\nu|m_h)}{\kappa_1\tau(2 - \kappa_1\tau)} \right)^{1/2},$$

where $\mathcal{H}^{\phi_1}(\nu|m_h)$ denotes the relative entropy between two probability measures, which equals $\mathcal{H}^{\phi_1}(\rho|m_h)$ with $\rho = d\nu/dm_h$ if ν is absolutely continuous with respect to m_h , and $+\infty$ else. Theorem 4 guarantees the validity of the modified logarithmic Sobolev inequality with constant κ_ϕ . Peres and Tetaly have conjectured for Markov chains on discrete spaces, for which a coarse Ricci curvature in the sense of Ollivier holds, that a modified logarithmic Sobolev inequality with the same constant up to a constant factor holds; see [16, Conjecture 3.1]. $\square$

5. COUPLING METHOD

We introduce the coupling rates for Markov chains and present sufficient conditions to conclude the convex Sobolev inequality. These conditions were established by Conforti [14], and we recall them for the convenience of the reader. More precisely, we define a coupling of two Markov chains by coupling their rates. Recall that G_i is the set of all possible moves of the Markov chain starting in $i \in \mathcal{K}_h$. We set $G_i^* = G_i \cup \{e\}$ for the set of all moves starting from i and the null element e (no move).

Definition 1 (Coupling rate). Let $\mathcal{L}_h$ be the generator (10) with transition rates $c(i, \gamma)$. Given $i, \bar{i} \in \mathcal{K}_h$, we call the function $\mathbf{c}(i, \bar{i}, \cdot, \cdot) : G_i^* \times G_{\bar{i}}^* \rightarrow \mathbb{R}_+$ a coupling rate for $(i, \bar{i})$ if and only if

$$\begin{aligned} \sum_{\bar{\gamma} \in G_{\bar{i}}^*} \mathbf{c}(i, \bar{i}, \gamma, \bar{\gamma}) &= c(i, \gamma) \quad \text{for all } \gamma \in G_i, \\ \sum_{\gamma \in G_i^*} \mathbf{c}(i, \bar{i}, \gamma, \bar{\gamma}) &= c(\bar{i}, \bar{\gamma}) \quad \text{for all } \bar{\gamma} \in G_{\bar{i}}. \end{aligned}$$

As in [14], we introduce the nonnegative function

$$f^\phi(i, \delta i) = \Phi(f(i), f(\delta i)) = (\phi'(f(i)) - \phi'(f(\delta i)))(f(i) - f(\delta i))$$

for $i \in \mathcal{K}_h$, $\delta \in G_i$, and $f \geq 0$. Furthermore, let $S = \{(i, \delta) \in \mathcal{K}_h \times G : c(i, \delta) > 0\}$ be the set of all combinations of states and moves with positive transition rate.

Theorem 15. [14, Prop. 2.1] *Let Assumption (A1) hold. Let $c(i, \gamma i, \cdot, \cdot)$ be coupling rates and m_h be the invariant measure associated to the generator $\mathcal{L}_h$.*

(i) *If there exists $\kappa' \geq 0$ such that*

$$(27) \quad \frac{1}{2} \sum_{\substack{(i, \delta) \in S \\ \gamma \in G_i^*, \bar{\gamma} \in G_{\delta i}^*}} c(i, \gamma) \mathbf{c}(i, \delta i, \gamma, \bar{\gamma}) (f^\phi(\gamma i, \bar{\gamma} \delta i) - f^\phi(i, \delta i)) m_h(i) \leq -\kappa' \mathcal{E}(\phi'(f), f)$$

holds uniformly for $f > 0$, the convex Sobolev inequality holds with constant κ' .

(ii) *If in addition to (27) there exists $\kappa'' > 0$ such that*

$$\inf_{(i, \delta) \in S} \min \{ \mathbf{c}(i, \delta i, \delta, e), \mathbf{c}(i, \delta i, e, \delta^{-1}) \} \geq \kappa'',$$

the modified logarithmic Sobolev inequality holds with $\kappa_1 = \kappa' + 2\kappa''$.

(iii) *If in addition to (27) there exists $\kappa''' > 0$ such that*

$$\inf_{(i, \delta) \in S} \sum_{\substack{\gamma \in G_i^*, \bar{\gamma} \in G_{\delta i}^* \\ \gamma i = \bar{\gamma} \delta i}} \mathbf{c}(i, \delta i, \gamma, \bar{\gamma}) \geq \kappa''',$$

the discrete Beckner inequality holds for $1 < \alpha \leq 2$ with $\kappa_\alpha = \kappa' + (\alpha - 1)\kappa'''$.

The proof of this theorem relies on the following construction of the coupling rates, which is also used in the proof of Theorem 4. We define for $i, i + he_j \in \mathcal{K}_h$ with $j = 1, \dots, n$,

$$(28) \quad \mathbf{c}(i, i + he_j, \gamma, \bar{\gamma}) = \begin{cases} \min\{c(i, \gamma), c(i + he_j, \bar{\gamma})\} & \text{if } \gamma = \bar{\gamma} \in G, \\ \max\{c(i + he_j, \bar{\gamma}) - c(i, \bar{\gamma}), 0\} & \text{if } \gamma = +_j, \bar{\gamma} \in G, \bar{\gamma} \neq +_j, -_j, \\ \max\{-c(i + he_j, \gamma) + c(i, \gamma), 0\} & \text{if } \gamma \in G, \gamma \neq +_j, -_j, \bar{\gamma} = -_j, \\ \kappa_+(i, j) & \text{if } \gamma = +_j, \bar{\gamma} = e, \\ \kappa_-(i + he_j, j) & \text{if } \gamma = e, \bar{\gamma} = -_j, \\ 0 & \text{else,} \end{cases}$$

recalling definitions (18) and (19) of $\kappa_\pm(i, j)$. Assumption (A3) guarantees that $\kappa_\pm(i, j) > 0$. The coupling ensures that the new states γi and $\bar{\gamma}(i + he_j)$ are at most neighboring states on the grid and that they are coupled with maximal probability, i.e. $\gamma i = \bar{\gamma}(i + he_j)$ with rate $\kappa_+(i, j) + \kappa_-(i, j)$. Analogously, we define

$$\mathbf{c}(i, i - he_j, \gamma, \bar{\gamma}) = \mathbf{c}(i - he_j, (i - he_j) + he_j, \bar{\gamma}, \gamma) \quad \text{for } i, i - he_j \in \mathcal{K}_h, j = 1, \dots, d.$$

We remark that if the potential has an additive structure, i.e. $V(i) = \sum_{j=1}^d V_j(i_j)$, coupling (28) simplifies to

$$\mathbf{c}(i, i + he_j, \gamma, \bar{\gamma}) = \begin{cases} \min\{c(i, \gamma), c(i + he_j, \bar{\gamma})\} & \text{if } \gamma = \bar{\gamma} \in G, \\ c(i, +_j) - c(i + he_j, +_j) & \text{if } \gamma = +_j, \bar{\gamma} = e, \\ c(i + he_j, -_j) - c(i, -_j) & \text{if } \gamma = e, \bar{\gamma} = -_j, \\ 0 & \text{else.} \end{cases}$$

6. PROOFS FOR CONTINUOUS-TIME MARKOV CHAINS

6.1. Exponential decay in ϕ -entropy. In this subsection, we prove Theorem 4. It is sufficient to verify the conditions of Theorem 15. The most delicate condition is the first one, whose validity is proved in the following lemma.

Lemma 16 (Condition (i)). *Let Assumptions (A2) and (A3) hold and let $f \geq 0$. Then, for κ_ϕ as given in Theorem 4, the following inequality holds:*

$$(29) \quad \sum_{\substack{(i, \delta) \in S \\ \gamma \in G_i^*, \bar{\gamma} \in G_{\delta i}^*}} c(i, \delta) \mathbf{c}(i, \delta i, \gamma, \bar{\gamma}) (f^\phi(\gamma i, \bar{\gamma} \delta i) - f^\phi(i, \delta i)) m_h(i) \leq -\kappa_\phi \sum_{(i, \delta) \in S} c(i, \delta) f^\phi(i, \delta i) m_h(i).$$

Recall that S is the set of pairs $(i, \delta) \in \mathcal{K}_g \times G$ with positive transition rate $c(i, \delta) > 0$.

Proof. Inequality (29) is proved in [14] for a slightly different model, obtaining another value of κ_ϕ . We adapt the proof to the present situation. The idea is to split the left-hand side of (29) into two parts and to reformulate both parts separately:

$$(30) \quad \sum_{\substack{(i, \delta) \in S \\ \gamma \in G_i^*, \bar{\gamma} \in G_{\delta i}^*}} c(i, \delta) \mathbf{c}(i, \delta i, \gamma, \bar{\gamma}) (f^\phi(\gamma i, \bar{\gamma} \delta i) - f^\phi(i, \delta i)) m_h(i) = I - J, \quad \text{where} \\ I := \sum_{\substack{(i, \delta) \in S \\ \gamma \in G_i^*, \bar{\gamma} \in G_{\delta i}^*}} c(i, \delta) \mathbf{c}(i, \delta i, \gamma, \bar{\gamma}) f^\phi(\gamma i, \bar{\gamma} \delta i) m_h(i), \\ J := \sum_{\substack{(i, \delta) \in S \\ \gamma \in G_i^*, \bar{\gamma} \in G_{\delta i}^*}} c(i, \delta) \mathbf{c}(i, \delta i, \gamma, \bar{\gamma}) f^\phi(i, \delta i) m_h(i).$$

To simplify the notation, we write $\nabla_\delta c(i, \gamma) = c(\delta i, \gamma) - c(i, \gamma)$ in the following.

Step 1: Reformulation of J . We insert the coupling rates (28) into the expression for J and observe that we have always two possibilities to move $(\pm_j$ and $\pm_\ell)$. This leads to

$$J = \sum_{i \in \mathcal{K}_h, j=1, \dots, d} c(i, +_j) A^+(i, j) f^\phi(i, i + he_j) m_h(i) \\ + \sum_{i \in \mathcal{K}_h, j=1, \dots, d} c(i, -_j) A^-(i, j) f^\phi(i, i - he_j) m_h(i), \quad \text{where}$$

$$\begin{aligned}
A^\pm(i, j) &= \sum_{\ell=1}^d \left(\min\{c(i, +_\ell), c(i \pm he_j, +_\ell)\} + \min\{c(i, -_\ell), c(i \pm he_j, -_\ell)\} \right) \\
&\quad + \sum_{\ell=1, \ell \neq j}^d \left(\max\{\nabla_{\pm j} c(i, +_\ell), 0\} + \max\{\nabla_{\pm j} c(i, -_\ell), 0\} \right) \\
&\quad + \sum_{\ell=1, \ell \neq j}^d \left(\max\{-\nabla_{\pm j} c(i, +_\ell), 0\} + \max\{-\nabla_{\pm j} c(i, -_\ell), 0\} \right) \\
&\quad + \kappa_\pm(i, j) + \kappa_\mp(i \pm he_j, j).
\end{aligned}$$

We insert into $A^+(i, j)$ definition (18) for $\kappa_\pm$, namely

$$\begin{aligned}
\kappa_+(i, j) &= c(i, +_j) - c(i + he_j, +_j) \\
&\quad - \sum_{\ell=1, \ell \neq j}^d \left(\max\{\nabla_{+j} c(i, +_\ell), 0\} + \max\{\nabla_{+j} c(i, -_\ell), 0\} \right), \\
\kappa_-(i + he_j, j) &= c(i + he_j, -_j) - c(i, -_j) \\
&\quad - \sum_{\ell=1, \ell \neq j}^d \left(\max\{-\nabla_{+j} c(i, +_\ell), 0\} + \max\{-\nabla_{+j} c(i, -_\ell), 0\} \right),
\end{aligned}$$

which cancels the terms in $A^+(i, j)$ involving the maximum, and use the identity $\min\{a, b\} = a - \max\{a - b, 0\}$ for $a, b \in \mathbb{R}$ to find that

$$\begin{aligned}
A^+(i, j) &= \sum_{\ell=1}^d \left(\min(c(i, +_\ell), c(i + he_j, +_\ell)) + \min(c(i, -_\ell), c(i + he_j, -_\ell)) \right) \\
&\quad + c(i, +_j) - c(i + he_j, +_j) - c(i, -_j) + c(i + he_j, -_j) \\
&= \sum_{\ell=1}^d \left(c(i, +_\ell) - \max\{0, -\nabla_{+j} c(i, +_\ell)\} + c(i, -_\ell) - \max\{0, -\nabla_{+j} c(i, -_\ell)\} \right) \\
&\quad + c(i, +_j) - c(i + he_j, +_j) - c(i, -_j) + c(i + he_j, -_j) \\
&= \sum_{\ell=1}^d (c(i, +_\ell) + c(i, -_\ell)) + \kappa_-(i + he_j, j).
\end{aligned}$$

Proceeding in a similar way, we obtain

$$A^-(i, j) = \sum_{\ell=1}^d (c(i, +_\ell) + c(i, -_\ell)) + \kappa_+(i - he_j, j).$$

This shows that

$$J = \sum_{i \in \mathcal{K}_h, j=1, \dots, d} c(i, +_j) \left(\sum_{\ell=1}^d (c(i, +_\ell) + c(i, -_\ell)) + \kappa_-(i + he_j, j) \right) f^\phi(i, i + he_j) m_h(i)$$

(31)

$$+ \sum_{i \in \mathcal{K}_h, j=1, \dots, d} c(i, -j) \left(\sum_{\ell=1}^d (c(i, +\ell) + c(i, -\ell)) + \kappa_+(i - he_j, j) \right) f^\phi(i, i - he_j) m_h(i).$$

Step 2: Reformulation of I . Setting $B_{(i, \delta, \gamma, \bar{\gamma})} := c(i, \delta) \mathbf{c}(i, \delta i, \gamma, \bar{\gamma}) f^\phi(\gamma i, \bar{\gamma} \delta i) m_h(i)$, we write

$$\begin{aligned} I &= \sum_{\substack{(i, \delta) \in S \\ \gamma, \bar{\gamma} \in G^*}} B_{(i, \delta, \gamma, \bar{\gamma})} = I_1 + I_2 + I_3 + I_4, \quad \text{where} \\ I_1 &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \sum_{\ell=1, \ell \neq j}^d (B_{(i, +j, +\ell, +\ell)} + B_{(i, +j, -\ell, -\ell)} + B_{(i, -j, +\ell, +\ell)} + B_{(i, -j, -\ell, -\ell)}), \\ I_2 &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} (B_{(i, +j, +j, +j)} + B_{(i, -j, +j, +j)} + B_{(i, +j, -j, -j)} + B_{(i, -j, -j, -j)}), \\ I_3 &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \sum_{\ell=1, \ell \neq j}^d (B_{(i, +j, +j, +\ell)} + B_{(i, +j, +j, -\ell)} + B_{(i, +j, +\ell, -j)} + B_{(i, +j, -\ell, -j)}), \\ I_4 &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \sum_{\ell=1, \ell \neq j}^d (B_{(i, -j, +\ell, +j)} + B_{(i, -j, -\ell, +j)} + B_{(i, -j, -j, +\ell)} + B_{(i, -j, -j, -\ell)}). \end{aligned}$$

The term I_1 is reformulated by inserting the coupling rates and performing the shift $i \mapsto i \pm he_j$:

$$\begin{aligned} I_1 &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \sum_{\ell=1, \ell \neq j}^d \left[\left(c(i - he_\ell, +j) \min\{c(i - he_\ell, +\ell), c(i + he_j - he_\ell, +\ell)\} m_h(i - he_\ell) \right. \right. \\ &\quad \left. \left. + c(i + he_\ell, +j) \min\{c(i + he_\ell, -\ell), c(i + he_j + he_\ell, -\ell)\} m_j(i + he_\ell) \right) f^\phi(i, i + he_j) \right. \\ &\quad \left. + \left(c(i - he_\ell, -j) \min\{c(i - he_\ell, +\ell), c(i - he_j - he_\ell, +\ell)\} m_h(i - he_\ell) \right. \right. \\ &\quad \left. \left. + c(i + he_\ell, -j) \min\{c(i + he_\ell, -\ell), c(i - he_j + he_\ell, -\ell)\} m_h(i + he_\ell) \right) f^\phi(i, i - he_j) \right]. \end{aligned}$$

Assumption (A3) implies that $c(i, \pm j) > c(i \pm he_j, \pm j)$. Hence,

$$\begin{aligned} \min\{c(i - he_j, +j), c(i, +j)\} &= c(i, +j), \\ \min\{c(i + he_j, -j), c(i + 2he_j, -j)\} &= c(i + he_j, -j), \\ \min\{c(i - he_j, +j), c(i - 2he_j, +j)\} &= c(i - he_j, +j), \\ \min\{c(i + he_j, -j), c(i, -j)\} &= c(i, -j). \end{aligned}$$

Then, inserting the coupling rates into I_2 , we find that

$$\begin{aligned}
I_2 &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \left[\left(c(i - he_j, +_j) \min\{c(i - he_j, +_j), c(i, +_j)\} m_h(i - he_j) \right. \right. \\
&\quad \left. \left. + c(i + he_j, +_j) \min\{c(i + he_j, -_j), c(i + 2he_j, -_j)\} m_h(i + he_j) \right) f^\phi(i, i + he_j) \right. \\
&\quad \left. + \left(c(i - he_j, -_j) \min\{c(i - he_j, +_j), c(i - 2he_j, +_j)\} m_h(i - he_j) \right. \right. \\
&\quad \left. \left. + c(i + he_j, -_j) \min\{c(i + he_j, -_j), c(i, -_j)\} m_h(i + he_j) \right) f^\phi(i, i - he_j) \right] \\
&= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \left[\left(c(i - he_j, +_j) c(i, +_j) m_h(i - he_j) + c(i + he_j, +_j) c(i + he_j, -_j) m_h(i + he_j) \right) \right. \\
&\quad \times f^\phi(i, i + he_j) \\
&\quad \left. + \left(c(i - he_j, -_j) c(i - he_j, +_j) m_h(i - he_j) + c(i + he_j, -_j) c(i, -_j) m_h(i + he_j) \right) \right. \\
&\quad \times f^\phi(i, i - he_j) \left. \right] \\
&= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \left[\left(c(i - he_j, +_j) c(i, +_j) m_h(i - he_j) + c(i, +_j) c(i + he_j, -_j) m_h(i + he_j) \right) \right. \\
&\quad \times f^\phi(i, i + he_j) \\
&\quad \left. + \left(c(i, -_j) c(i - he_j, +_j) m_h(i - he_j) + c(i + he_j, -_j) c(i, -_j) m_h(i + he_j) \right) \right. \\
&\quad \times f^\phi(i, i - he_j) \\
&\quad \left. - \nabla_{-j} c(i, +_j) c(i, -_j) m_h(i) f^\phi(i, i - he_j) - \nabla_{+j} c(i, -_j) c(i, +_j) m_h(i) f^\phi(i, i + he_j) \right],
\end{aligned}$$

where in the last step, we split the second and third terms and performed the shifts $i \mapsto i - he_j$ and $i \mapsto i + he_j$ to these terms, respectively. Similarly, we insert the coupling rates, perform the shifts $i \mapsto i \pm he_j$, and exchange the variables j and ℓ , leading to

$$\begin{aligned}
I_3 &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \sum_{\substack{\ell=1 \\ \ell \neq j}}^d \left(c(i - he_\ell, +_\ell) \max\{-\nabla_{-\ell} c(i, +_j), 0\} m_h(i - he_\ell) f^\phi(i, i + he_j) \right. \\
&\quad \left. + c(i - he_\ell, +_\ell) \max\{-\nabla_{-\ell} c(i, -_j), 0\} m_h(i - he_\ell) f^\phi(i, i - he_j) \right. \\
&\quad \left. + c(i, +_\ell) \max\{-\nabla_{+\ell} c(i, +_j), 0\} m_h(i) f^\phi(i, i + he_j) \right. \\
&\quad \left. + c(i, +_\ell) \max\{-\nabla_{+\ell} c(i, -_j), 0\} m_h(i) f^\phi(i, i - he_j) \right), \\
I_4 &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \sum_{\substack{\ell=1 \\ \ell \neq j}}^d \left(c(i, -_\ell) \max\{-\nabla_{-\ell} c(i, +_j), 0\} m_h(i) f^\phi(i, i + he_j) \right.
\end{aligned}$$

$$\begin{aligned}
& + c(i, -\ell) \max\{-\nabla_{-\ell} c(i, -j), 0\} m_h(i) f^\phi(i, i - h e_j) \\
& + c(i + h e_\ell, -\ell) \max\{-\nabla_{+\ell} c(i, +j), 0\} m_h(i + h e_\ell) f^\phi(i, i + h e_j) \\
& + c(i + h e_\ell, -\ell) \max\{-\nabla_{+\ell} c(i, -j), 0\} m_h(i + h e_\ell) f^\phi(i, i - h e_j) \Big).
\end{aligned}$$

We collect first all terms with factor $f^\phi(i, i + h e_j)$ appearing in I_1 , I_3 , and I_4 and reformulate them by observing that, by (9), we have

$$c(i, +j) \max\{-\nabla_{+j} c(i, \pm\ell), 0\} = c(i, \pm\ell) \max\{-\nabla_{\pm\ell} c(i, +j), 0\}.$$

This yields

$$\begin{aligned}
(32) \quad & \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \sum_{\substack{\ell=1 \\ \ell \neq j}}^d \Big(c(i - h e_\ell, +j) \min\{c(i - h e_\ell, +\ell), c(i + h e_j - h e_\ell, +\ell)\} m_h(i - h e_\ell) \\
& + c(i + h e_\ell, +j) \min\{c(i + h e_\ell, -\ell), c(i + h e_j + h e_\ell, -\ell)\} m_h(i + h e_\ell) \\
& + c(i - h e_\ell, +\ell) \max\{-\nabla_{-\ell} c(i, +j), 0\} m_h(i - h e_\ell) \\
& + c(i + h e_\ell, -\ell) \max\{-\nabla_{+\ell} c(i, +j), 0\} m_h(i + h e_\ell) \\
& + c(i, +\ell) \max\{-\nabla_{+\ell} c(i, +j), 0\} m_h(i) + c(i, -\ell) \max\{-\nabla_{-\ell} c(i, +j), 0\} m_h(i) \Big) \\
& \times f^\phi(i, i + h e_j) \\
& = \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \sum_{\substack{\ell=1 \\ \ell \neq j}}^d \Big(c(i, +j) c(i - h e_\ell, +\ell) m_h(i - h e_\ell) + c(i + h e_\ell, -\ell) c(i, +j) m_h(i + h e_\ell) \\
& + c(i, +j) \max\{-\nabla_{+j} c(i, +\ell), 0\} m_h(i) + c(i, +j) \max\{-\nabla_{+j} c(i, -\ell), 0\} m_h(i) \Big) \\
& \times f^\phi(i, i + h e_j).
\end{aligned}$$

Similarly, by collecting the terms with factor $f^\phi(i, i - h e_j)$ in I_1 , I_3 , and I_4 , we arrive at

$$\begin{aligned}
(33) \quad & \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \sum_{\substack{\ell=1 \\ \ell \neq j}}^d \Big(c(i, -j) c(i - h e_\ell, +\ell) m_h(i - h e_\ell) + c(i + h e_\ell, -\ell) c(i, -j) m_h(i + h e_\ell) \\
& + c(i, -j) \max\{-\nabla_{-j} c(i, +\ell), 0\} m_h(i) + c(i, -j) \max\{-\nabla_{-j} c(i, -\ell), 0\} m_h(i) \Big) \\
& \times f^\phi(i, i - h e_j).
\end{aligned}$$

We insert I_2 and (32)–(33) into I , giving

$$I = \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \left(c(i, +j) f^\phi(i, i + h e_j) E^+ + c(i, -j) f^\phi(i, i - h e_j) E^- \right), \quad \text{where}$$

$$\begin{aligned}
E^\pm &= \sum_{\ell=1, \ell \neq j}^d [c(i - he_\ell, +_\ell)m_h(i - he_\ell) + c(i + he_\ell, -_\ell)m(i + he_\ell)] \\
&\quad + c(i - he_j, +_j)m_h(i - he_j) + c(i + he_j, -_j)m_h(i + he_j) - \nabla_{\pm j}c(i, \mp_j)m_h(i) \\
&\quad + \sum_{\ell=1, \ell \neq j}^d [\max\{-\nabla_{\pm j}c(i, +_\ell), 0\} + \max\{-\nabla_{\pm j}c(i, -_\ell), 0\}]m_h(i).
\end{aligned}$$

We combine the sum over $\ell \neq j$ and the second and third terms in E^+ to obtain

$$\begin{aligned}
E^+ &= \sum_{\ell=1}^d [c(i - he_\ell, +_\ell)m_h(i - he_\ell) + c(i + he_\ell, -_\ell)m(i + he_\ell)] - \nabla_{+j}c(i, -_j)m_h(i) \\
&\quad + \sum_{\ell=1, \ell \neq j}^d [\max\{-\nabla_{+j}c(i, +_\ell), 0\} + \max\{-\nabla_{+j}c(i, -_\ell), 0\}]m_h(i) \\
&= \sum_{\ell=1}^d (c(i, +_\ell) + c(i, -_\ell)) - \kappa_-(i + he_j, j),
\end{aligned}$$

where the last step follows from definition (18) of κ_+ and the identity

$$\begin{aligned}
&\sum_{\ell=1}^d (c(i - he_\ell, +_\ell)m_h(i - he_\ell) + c(i + he_\ell, -_\ell)m_h(i + he_\ell)) \\
&= \sum_{\ell=1}^d (c(i, +_\ell) + c(i, -_\ell))m_h(i) \quad \text{for } i \in \mathcal{K}_h,
\end{aligned}$$

which is similarly shown as in the proof of Lemma 1. A similar computation is done for E^- , resulting eventually in

$$E^\pm = \sum_{\ell=1}^d (c(i, +_\ell) + c(i, -_\ell)) - \kappa_\mp(i \pm he_j, j).$$

This shows that

$$\begin{aligned}
(34) \quad I &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} \left[c(i, +_j)f^\phi(i, i + he_j) \left(\sum_{\ell=1}^d (c(i, +_\ell) + c(i, -_\ell)) - \kappa_-(i + he_j, j) \right) \right. \\
&\quad \left. + c(i, -_j)f^\phi(i, i - he_j) \left(\sum_{\ell=1}^d (c(i, +_\ell) + c(i, -_\ell)) - \kappa_+(i - he_j, j) \right) \right] m_h(i).
\end{aligned}$$

Step 3: End of the proof. We compute the difference $I - J$ by inserting the expressions derived in Steps 1 and 2 and observing that the terms involving $c(i, \pm_\ell)$ cancel out:

$$\begin{aligned} I - J &= \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} c(i, +_j) f^\phi(i, i + he_j) (-2\kappa_-(i + he_j, j)) m_h(i) \\ &\quad + \sum_{\substack{i \in \mathcal{K}_h \\ j=1, \dots, d}} c(i, -_j) f^\phi(i, i - he_j) (-2\kappa_+(i - he_j, j)) m_h(i) \\ &\leq -2\kappa_\phi \sum_{(i, \delta) \in S} c(i, \delta) f^\phi(i, \delta i) m_h(i), \end{aligned}$$

where $\kappa_\phi = \min_{i, i+he_j \in \mathcal{K}_h, j=1, \dots, d} \{\kappa_-(i + he_j, j), \kappa_+(i, j)\}$ and we have used that $f^\phi(i, i \pm he_j) \geq 0$. Note that the terms involving $\kappa_-(i + he_j, j)$ for $i + he_j \notin \mathcal{K}_h$ and $\kappa_+(i - he_j, j)$ for $i - he_j \notin \mathcal{K}_h$ vanish since the rates $c(i, +_j)$ for $i + he_j \notin \mathcal{K}_h$ and $c(i, -_j)$ for $i - he_j \notin \mathcal{K}_h$ are zero. Inserting this inequality into (30) and dividing the resulting estimate by two finishes the proof. $\square$

Lemma 17 (Conditions (ii) and (iii)). *Let Assumptions (A2)–(A3) hold. Then conditions (ii) and (iii) of Theorem 15 are satisfied with $2\kappa'' = \kappa''' = \kappa_\phi$.*

Proof. By definition (28) of the coupling rates, we have

$$\begin{aligned} &\inf_{(i, \delta) \in S} \min\{\mathbf{c}(i, \delta i, \delta, e), \mathbf{c}(i, \delta i, e, \delta^{-1})\} \\ &= \inf_{\substack{i \in \mathcal{K}_h, j=1, \dots, d \\ i+he_j \in \mathcal{K}_h}} \min\{\mathbf{c}(i, i + he_j, +_j, e), \mathbf{c}(i, i + he_j, e, -_j)\} \\ &= \inf_{\substack{i \in \mathcal{K}_h, j=1, \dots, d \\ i+he_j \in \mathcal{K}_h}} \min\{\kappa_+(i, j), \kappa_-(i + he_j, j)\} = \frac{\kappa_\phi}{2}, \\ &\inf_{(i, \delta) \in S} \sum_{\substack{\gamma \in G_i^*, \bar{\gamma} \in G_{\delta i}^* \\ \gamma i = \bar{\gamma} \delta i}} \mathbf{c}(i, \delta i, \gamma, \bar{\gamma}) = \inf_{\substack{i \in \mathcal{K}_h, j=1, \dots, d \\ i+he_j \in \mathcal{K}_h}} (\kappa_+(i, j) + \kappa_-(i, j)) \geq \kappa_\phi, \end{aligned}$$

since $\kappa_\pm(i, j) \geq \kappa_\phi/2$. $\square$

Theorem 4 now follows directly from Theorem 15.

6.2. Exponential decay in Wasserstein distance. We prove Theorems 7 and 8.

Step 1: Decay in the L^2 Wasserstein distance. Because of the infimum property of the L^2 Wasserstein distance, it is sufficient to find a coupling of two copies $(Y_t^1)_{t \geq 0}$ and $(Y_t^2)_{t \geq 0}$ of the Markov chain (with initial data ν and η , respectively) such that

$$\frac{d}{dt} \mathbb{E}(|Y_t^1 - Y_t^2|^2) \leq -\kappa \mathbb{E}(|Y_t^1 - Y_t^2|^2) + O(h),$$

since by Grönwall's inequality and the property $\mathcal{W}_2(\nu p_t, \eta p_t)^2 \leq \mathbb{E}(|Y_t^1 - Y_t^2|^2)$, the conclusion follows after taking the infimum over all couplings of ν and η and taking the square root.

Consider the transition rates (5). We construct a coupling of two copies of the transition kernel by coupling the transition rates in the following way. For $i, k \in \mathcal{K}_h$ and $\gamma, \bar{\gamma} \in G$, we set

$$(35) \quad \mathbf{c}(i, k, \gamma, \bar{\gamma}) = \begin{cases} \min\{c(i, \gamma), c(k, \bar{\gamma})\} & \text{if } \gamma = \bar{\gamma} \in G, \\ \max\{c(i, \gamma) - c(k, \bar{\gamma}), 0\} & \text{if } \gamma \in G, \bar{\gamma} = e, \\ \min\{c(i, \gamma) - c(k, \bar{\gamma}), 0\} & \text{if } \bar{\gamma} \in G, \gamma = e, \\ 0 & \text{else.} \end{cases}$$

By definition (28), this defines indeed a coupling. We denote the generator for the corresponding Markov chain on $\mathcal{K}_h \times \mathcal{K}_h$ by $\mathcal{L}_h^2$. We infer from the definition of the coupling that

$$\begin{aligned} \mathcal{L}_h^2 |i - k|^2 &= \sum_{\gamma, \bar{\gamma} \in G^*} \mathbf{c}(i, k, \gamma, \bar{\gamma}) (|\gamma i - \bar{\gamma} k|^2 - |i - k|^2) \\ &= \sum_{\gamma \in G} [\max\{c(i, \gamma) - c(k, \gamma), 0\} (|\gamma i - k|^2 - |i - k|^2) \\ &\quad + \max\{c(k, \gamma) - c(i, \gamma), 0\} (|i - \gamma k|^2 - |i - k|^2)]. \end{aligned}$$

As G consists of the moves $\pm_j$ with $j = 1, \dots, d$, this sum becomes

$$(36) \quad \mathcal{L}_h^2 |i - k|^2 = \sum_{j=1}^d [C^+ (|i + h e_j - k|^2 - |i - k|^2) + C^- (|i - h e_j - k|^2 - |i - k|^2)],$$

$$\text{where } C^\pm = \max\{c(i, \pm_j) - c(k, \pm_j), 0\} + \max\{c(k, \mp_j) - c(i, \mp_j), 0\}.$$

Since we have assumed that the potential is additive, $V(i) = \sum_{j=1}^d V_j(i_j)$, we have $V^h(i + h e_j) - V^h(i) = V_j^h(i_j + h) - V_j^h(i_j)$. Then it follows from definition (5) of $c(i, \pm_j)$ that

$$\begin{aligned} c(i, \gamma) - c(k, \gamma) &= \frac{\sigma^2}{h^2} (e^{-(V_j^h(i_j+h) - V_j^h(i_j))/(2\sigma^2)} - e^{-(V_j^h(k_j+h) - V_j^h(k_j))/(2\sigma^2)}) \\ &= \frac{\sigma^2}{h^2} \exp\left(-\frac{1}{2\sigma^2} \int_0^h \partial V_j^h(i_j + s) ds\right) \\ &\quad \times \left[1 - \exp\left(-\frac{1}{2\sigma^2} \int_0^h (\partial V_j^h(k_j + s) - \partial V_j^h(i_j + s)) ds\right)\right], \end{aligned}$$

where ∂V_j^h denotes the derivative of the scalar function V_j^h , interpreted as a function defined on D . This gives

$$\begin{aligned} C^+ &= \frac{\sigma^2}{h^2} \max \left\{ \exp\left(-\frac{1}{2\sigma^2} (V_j^h(i_j + h) - V_j^h(i_j))\right) \right. \\ &\quad \times \left[1 - \exp\left(-\frac{1}{2\sigma^2} \int_0^h (\partial V_j^h(k_j + s) - \partial V_j^h(i_j + s)) ds\right)\right], 0 \Big\} \\ &\quad + \frac{\sigma^2}{h^2} \max \left\{ \exp\left(-\frac{1}{2\sigma^2} (V_j^h(i_j - h) - V_j^h(i_j))\right) \right. \end{aligned}$$

$$\times \left[\exp \left(\frac{1}{2\sigma^2} \int_0^h (\partial V_j^h(k_j - s) - \partial V_j^h(i_j - s)) ds \right) - 1 \right], 0 \Big\}.$$

Since $\max\{a, 0\} = a1_{\{a>0\}}$, we may remove the maximum by introducing the factor $1_{\{k_j > i_j\}}$. Then we deduce from Assumption (A2) that

$$\partial V_j^h(k_j + s) - \partial V_j^h(i_j + s) \geq \kappa(k_j - i_j) \quad \text{on } \{k_j > i_j\},$$

and hence,

$$\begin{aligned} C^+ &\geq \frac{\sigma^2}{h^2} \exp \left(-\frac{1}{2\sigma^2} (V_j^h(i_j + h) - V_j^h(i_j)) \right) \left[1 - \exp \left(-\frac{h\kappa}{2\sigma^2} (k_j - i_j) \right) \right] 1_{\{k_j > i_j\}} \\ &\quad + \frac{\sigma^2}{h^2} \exp \left(-\frac{1}{2\sigma^2} (V_j^h(i_j - h) - V_j^h(i_j)) \right) \left[\exp \left(\frac{h\kappa}{2\sigma^2} (k_j - i_j) \right) - 1 \right] 1_{\{k_j > i_j\}}. \end{aligned}$$

By assumption, ∇V is Lipschitz continuous, and so is ∂V_j^h . Then the Taylor expansions $1 - \exp(-x) = x + O(x^2)$ and $\exp(x) - 1 = x + O(x^2)$ show that

$$C^+ \geq \frac{\sigma^2}{h^2} \left(\frac{h\kappa}{\sigma^2} (k_j - i_j) - O(h^2(k_j - i_j)^2) \right) 1_{\{k_j > i_j\}} = \frac{1}{h} (\kappa - O(h))(k_j - i_j) 1_{\{k_j > i_j\}}.$$

A similar computation leads to $C^- \geq h^{-1}(\kappa - O(h))(i_j - k_j) 1_{\{k_j < i_j\}}$. We insert these estimations into (36) and use $|i \pm h e_j - k|^2 - |i - k|^2 = -2h|i_j - k_j| + h^2$ (which is negative for sufficiently small $h > 0$):

$$\begin{aligned} \mathcal{L}_h^2 |i - k|^2 &\leq (\kappa - O(h)) \sum_{j=1}^d |k_j - i_j| (1_{\{k_j > i_j\}} + 1_{\{k_j < i_j\}}) (-2|i_j - k_j| + h) \\ &= -2\kappa |i - k|^2 + O(h). \end{aligned}$$

We infer for the two copies (Y_t^1) and (Y_t^2) of the Markov chain that

$$\frac{d}{dt} \mathbb{E}(|Y_t^1 - Y_t^2|^2) = \mathbb{E}(\mathcal{L}_h^2 |Y_t^1 - Y_t^2|^2) \leq -2\kappa \mathbb{E}(|Y_t^1 - Y_t^2|^2) + O(h),$$

and the result follows after an application of Grönwall's lemma, as detailed above.

Step 2: Decay in the L^p Wasserstein distance. Let $1 < p < \infty$ with $p \neq 2$. The proof is similar as in Step 1, but we need the condition $d = 1$. Using the coupling (35), we obtain

$$\begin{aligned} \mathcal{L}_h^2 |i - k|^p &= (\mathbf{c}(i, k, +, e) + \mathbf{c}(i, k, e, -)) (|i + h - k|^p - |i - k|^p) \\ &\quad + (\mathbf{c}(i, k, e, +) + \mathbf{c}(i, k, -, e)) (|i - h - k|^p - |i - k|^p). \end{aligned}$$

The assumption $d = 1$ enters in the estimate

$$|i \pm h - k|^p - |i - k|^p = \pm p(i - k)^{p-1} h + O(h^2),$$

which generally does not hold for $d > 1$. Then, after a similar computation as in Step 1,

$$\mathcal{L}_h^2 |i - k|^p \leq -p\kappa |i - k|^p + O(h),$$

and we conclude with Grönwall's inequality. This finishes the proof of Theorem 7.

Step 3: Decay in the L^1 Wasserstein distance. We turn to the proof of Theorem 8. Consider the graph distance $d(i, k) = \sum_{j=1}^d |i_j - k_j|$ for $i, k \in \mathcal{K}_h$. With respect to this distance, the set $\mathcal{K}_h$ forms a geodesic graph in the sense that for every $i, k \in \mathcal{K}_h$, there exists a path $i = i^{(0)}, i^{(1)}, \dots, i^{(n)} = k$ such that $d(i, k) = \sum_{\ell=1}^n d(i^{(\ell)}, i^{(\ell-1)})$. We remark that an edge between $i^{(\ell)}$ and $i^{(\ell+1)}$ exists in the graph if there exists $\gamma \in G$ such that $\gamma i^{(\ell)} = i^{(\ell+1)}$ (and vice versa, there exists $\bar{\gamma} \in G$ such that $\bar{\gamma} i^{(\ell+1)} = i^{(\ell)}$).

By the path coupling method of [9], it is sufficient for obtaining contraction in Wasserstein distance with respect to the distance d to prove contraction with respect to the distance d for every neighbouring states (i, k) . Indeed, let (i, k) be two neighbouring states, i.e., there exists $\gamma \in G$ such that $\gamma i = k$ (and vice versa, there exists $\bar{\gamma} \in G$ such that $\bar{\gamma} k = i$). Note that $d(i, k) = h$. Without loss of generality, we can assume that $i, k \in \mathcal{K}_h$ is such that $k = i + he_j$ for some $j \in \{1, \dots, d\}$. Recalling the coupling rates (28), it holds for the corresponding generator $\mathcal{L}_h^2$ on the product space that

$$\begin{aligned} \mathcal{L}_h^2 d(i, k) &= \sum_{\gamma, \bar{\gamma} \in G} \mathbf{c}(i, k, \gamma, \bar{\gamma}) (d(\gamma i, \bar{\gamma} k) - d(i, k)) \\ &= \kappa_+(i, j) (d(i + he_j, k) - d(i, k)) + \kappa_-(k, j) (d(i, k - he_j) - d(i, k)) \\ &= -(\kappa_+(i, j) + \kappa_-(i + he_j, j)) d(i, k) \leq -\kappa_1 d(i, k), \end{aligned}$$

where $\kappa_1 = \max_{j=1, \dots, d, i, i+he_j \in \mathcal{K}_h} (\kappa_+(i, j) + \kappa_-(i + he_j, j))$. It follows from the path coupling method that for two continuous-time Markov chains driven by the transition rates $\mathbf{c}$ given in (28),

$$\frac{d}{dt} \mathbb{E}[d(Y_t^1, Y_t^2)] \leq -\kappa_1 \mathbb{E}[d(Y_t^1, Y_t^2)],$$

and hence by Grönwall's inequality,

$$\mathcal{W}_{d,1}(\nu p_t, \eta p_t) \leq e^{-\kappa_1 t} \mathbb{E}[d(Y_0^1, Y_0^2)],$$

where $\mathcal{W}_{d,1}$ denotes the L^1 Wasserstein distance with respect to the graph distance d . We take the infimum over all couplings ν and η :

$$\mathcal{W}_{d,1}(\nu p_t, \eta p_t) \leq e^{-\kappa_1 t} \mathcal{W}_{d,1}(\nu, \eta).$$

We deduce from the equivalence of the distance d and the Euclidean distance, $|x - y| \leq d(x, y) \leq \sqrt{d}|x - y|$, that

$$\mathcal{W}_1(\nu p_t, \eta p_t) \leq \sqrt{d} e^{-\kappa_1 t} \mathcal{W}_1(\nu, \eta),$$

which concludes the proof.

7. PROOFS FOR DISCRETE-TIME MARKOV CHAINS

7.1. Exponential decay in ϕ -entropy. We present first the proof of Proposition 11 and then of Theorem 12. We consider a discrete-time Markov chain with transition kernel π (see (11)), transition rates $p(i, \gamma)$ (see (24)), and the corresponding invariant measure m_h (see Lemma 1).

Step 1: Proof of Proposition 11. We show Proposition 11 by following the lines of the proof of [22, Prop. 1]. It follows from assumptions (i) and (ii) that

$$\begin{aligned}\mathcal{F}(\pi^{n+1}f) - \mathcal{F}(\pi^n f) &\leq -\tau\lambda\mathcal{F}(\pi^n f) \leq -\tau\lambda C_P^{-1}\mathcal{P}(\pi^n f) \\ &= \lambda C_P^{-1}(\mathcal{H}^\phi(\pi^{n+1}f|m_h) - \mathcal{H}^\phi(\pi^n f|m_h)).\end{aligned}$$

Iterating this argument leads for $k > n$ to

$$(37) \quad \mathcal{F}(\pi^k f) - \mathcal{F}(\pi^n f) \leq \lambda C_P^{-1}(\mathcal{H}^\phi(\pi^k f|m_h) - \mathcal{H}^\phi(\pi^n f|m_h)).$$

We infer from (ii) that $\mathcal{F}(\pi^n f) \leq (1 - \tau\lambda)^n \mathcal{F}(f)$ and thus $\lim_{k \rightarrow \infty} \mathcal{F}(\pi^k f) = 0$. We perform the limit $k \rightarrow \infty$ in (37), using (iii):

$$\mathcal{F}(\pi^n f) \geq \lambda C_P^{-1} \mathcal{H}^\phi(\pi^n f|m_h).$$

Then the result follows from

$$\begin{aligned}\mathcal{H}^\phi(\pi^n f|m_h) &\leq \lambda^{-1} C_P \mathcal{F}(\pi^n f) \leq \lambda^{-1} C_P (1 - \tau\lambda)^n \mathcal{F}(f) \\ &= \frac{C_P \mathcal{F}(f)}{\lambda \mathcal{H}^\phi(f|m_h)} (1 - \tau\lambda)^n \mathcal{H}^\phi(f|m_h) \leq \frac{C_P \mathcal{F}(f)}{\lambda \mathcal{H}^\phi(f|m_h)} e^{-\lambda n \tau} \mathcal{H}^\phi(f|m_h).\end{aligned}$$

Step 2: Proof of Theorem 12. We need to verify conditions (i)–(iii) of Proposition 11. First, we compute

$$\begin{aligned}\tau \mathcal{P}(f) &= \mathcal{H}^\phi(f|m_h) - \mathcal{H}^\phi(\pi f|m_h) \\ &= \sum_{i \in \mathcal{K}_h} \phi(f(i)) m_h(i) - \phi\left(\sum_{i \in \mathcal{K}_h} f(i) m_h(i)\right) \\ &\quad - \sum_{i \in \mathcal{K}_h} \phi(\pi f(i)) m_h(i) - \phi\left(\sum_{i \in \mathcal{K}_h} \pi f(i) m_h(i)\right) \\ &= \sum_{i \in \mathcal{K}_h} (\phi(f(i)) - \phi(\pi f(i))) m_h(i),\end{aligned}$$

where the last step follows from the invariance property (25). Jensen's inequality (recall that ϕ is convex) implies that

$$\tau \mathcal{P}(f) \geq \sum_{i \in \mathcal{K}_h} (\phi(f(i)) - \pi \phi(f(i))) m_h(i) = \sum_{i \in \mathcal{K}_h} (\phi(f(i)) - \phi(f(i))) m_h(i) = 0.$$

To prove that $\mathcal{P}(f) \leq C_P \mathcal{F}(f)$ for some $C_P > 0$, we use the definition of ϕ_α , for all $x, y \geq 0$,

$$(38) \quad \phi_\alpha(x) - \phi_\alpha(y) = \begin{cases} \frac{1}{\alpha}(x\phi'_\alpha(x) - y\phi'_\alpha(y)) - \frac{\alpha}{\alpha-1}(x-y) & \text{if } 1 < \alpha \leq 2, \\ (x\phi'_\alpha(x) - y\phi'_\alpha(y)) - (x-y) & \text{if } \alpha = 1. \end{cases}$$

We deduce from the definition of the Dirichlet form and the invariance property (25) that, for functions $f \geq 0$,

$$\mathcal{F}(f) = \frac{1}{2} \sum_{i \in \mathcal{K}_h} \left(\sum_{\gamma \in G} p(i, \gamma) (f(\gamma i) - f(i)) (\phi'_\alpha(f(\gamma i)) - \phi'_\alpha(f(i))) \right) m_h(i)$$

$$\begin{aligned}
&= \frac{1}{2} \sum_{i \in \mathcal{K}_h} \left(\sum_{\gamma \in G} p(i, \gamma) (2f(i)\phi'_\alpha(f(i)) - f(i)\phi'_\alpha(f(\gamma i)) - f(\gamma i)\phi'_\alpha(f(i))) \right) m_h(i) \\
&\geq \sum_{i \in \mathcal{K}_h} \left(f(i)\phi'_\alpha(f(i)) - \frac{1}{2}f(i)\phi'_\alpha(\pi f(i)) - \frac{1}{2}\pi f(i)\phi'_\alpha(f(i)) \right) m_h(i),
\end{aligned}$$

where we used Jensen's inequality for the concave function ϕ'_α in the last step. We use definition (38) of ϕ_α and invariance property (25) again to find that

$$\begin{aligned}
\tau \mathcal{P}(f) &= \sum_{i \in \mathcal{K}_h} (\phi_\alpha(f(i)) - \phi_\alpha(\pi f(i))) m_h(i) \\
&= \frac{1}{\alpha} \sum_{i \in \mathcal{K}_h} (f(i)\phi'_\alpha(f(i)) - \pi f(i)\phi'_\alpha(\pi f(i))) m_h(i).
\end{aligned}$$

This shows that

$$\mathcal{F}(f) - \frac{\alpha\tau}{2} \mathcal{P}(f) \geq \frac{1}{2} \sum_{i \in \mathcal{K}_h} (f(i) - \pi f(i)) (\phi'_\alpha(f(i)) - \phi'_\alpha(\pi f(i))) m_h(i) \geq 0,$$

since $(a - b)(\phi'_\alpha(a) - \phi'_\alpha(b)) \geq 0$ for all $a, b \geq 0$ by the convexity of ϕ_α . This verifies condition (i) with constant $C_P = 2/(\alpha\tau)$.

To show condition (ii), we consider the same coupling (28) as for the continuous-time Markov chains, except that the rates c are replaced by the transition rates (24). In particular, the resulting coupling rates $\mathbf{p}(i, \bar{i}, \cdot, \cdot)$ satisfy

$$\sum_{\gamma \in G_i^*} \sum_{\bar{\gamma} \in G_{\bar{i}}^*} \mathbf{p}(i, \bar{i}, \gamma, \bar{\gamma}) = \sum_{\gamma \in G_i^*} p(i, \gamma) = \sum_{\bar{\gamma} \in G_{\bar{i}}^*} p(\bar{i}, \bar{\gamma}) = 1 \quad \text{for } i, \bar{i} \in \mathcal{K}_h.$$

We deduce from Lemma 16 that

$$\begin{aligned}
\mathcal{F}(\pi f) - \mathcal{F}(f) &= \frac{1}{2} \sum_{\substack{(i, \delta) \in S \\ \gamma \in G_i^*, \bar{\gamma} \in G_{\bar{i}}^*}} p(i, \delta) \mathbf{p}(i, \delta i, \gamma, \bar{\gamma}) (f^{\phi_\alpha}(\gamma i, \bar{\gamma} \delta i) - f^{\phi_\alpha}(i, \delta i)) m_h(i) \\
&\leq -\frac{\kappa_\phi}{2\mathcal{T}} \sum_{(i, \delta) \in S} p(i, \delta) f^{\phi_\alpha}(i, \delta i) m_h(i) = -\frac{\kappa_\phi}{\mathcal{T}} \mathcal{F}(f).
\end{aligned}$$

We infer that condition (ii) holds with $\tau = \mathcal{T}^{-1}$.

Finally, we verify condition (iii). We recall that since the Markov chain $(Z_n^h)_{n \geq 0}$ is time-homogeneous, aperiodic, irreducible and defined on a finite state space, its laws converge to the unique invariant measure [27, Theorem 1.8.3], and it holds that $\pi^n f(i) \rightarrow M := \sum_{k \in \mathcal{K}_h} f(k) m_h(k)$ as $n \rightarrow \infty$ for all $f \geq 0$ and $i \in \mathcal{K}_h$. Therefore, since $\sum_{i \in \mathcal{K}_h} m_h(i) = 1$,

$$\begin{aligned}
\mathcal{H}^{\phi_\alpha}(\pi^n f | m_h) &= \sum_{i \in \mathcal{K}_h} \phi_\alpha(\pi^n f(i)) m_h(i) - \phi_\alpha \left(\sum_{i \in \mathcal{K}_h} \pi^n f(i) m_h(i) \right) \\
&\rightarrow \sum_{i \in \mathcal{K}_h} \phi_\alpha(M) m_h(i) - \phi_\alpha \left(\sum_{i \in \mathcal{K}_h} M m_h(i) \right) = 0.
\end{aligned}$$

Hence, we can apply Proposition 11 to conclude that

$$\mathcal{H}^{\phi_\alpha}(\pi^n f | m_h) \leq C_f e^{-\kappa_\phi n \tau} \mathcal{H}^{\phi_\alpha}(f | m_h),$$

which finishes the proof.

APPENDIX A. CONVERGENCE OF THE MARKOV CHAIN TO THE SDE

We prove Theorem 3. The proof relies on the diffusion approximation of [17, Theorem 7.4.1], and it is rather standard. We present it for completeness. For $n \in \mathbb{N}$, let $h = 2^{-n}\xi$, where $\xi > 0$ is such that $2K/h \in \mathbb{N}$ holds. We write (Y_t^n) instead of (Y_t^h) , $\mathcal{L}_n$ instead of $\mathcal{L}_h$, $\mathcal{K}_n$ instead of $\mathcal{K}_h$, and V^n instead of V^h . For each $f \in \mathcal{A}$ with $\mathcal{A}$ defined in (4), we introduce the process

$$B_f^n(t) = f(Y_0^n) + \int_0^t \mathcal{L}_n(f(Y_s^n)) ds,$$

and we define for $f, g \in \mathcal{A}$ the process

$$A_{f,g}^n(t) = \int_0^t \sum_{\gamma \in G} c(Y_s^n, \gamma) (f(\gamma Y_s^n) - f(Y_s^n)) (g(\gamma Y_s^n) - g(Y_s^n)) ds.$$

We notice that $(A_{f,g}^n)$ is symmetric in the sense $A_{f,g}^n = A_{g,f}^n$ and that $A_{f,f}^n(t) - A_{f,f}^n(s)$ is positive for $t > s \geq 0$. We set $\mathcal{F}_t^n = \sigma(Y_s^n, B_f^n(s), A_{f,g}^n(s) : s \leq t)$. The process $(Y_t^n)_{t \geq 0}$ does not explode, since it is restricted to $\mathcal{K}_n$. Then $M_f^n(t) := f(Y_t^n) - B_f^n(t)$ is an $\mathcal{F}_t^n$ -martingale and $M_f^n M_g^n - A_{f,g}^n$ is an $\mathcal{F}_t^n$ -martingale, since, using Definition (10) of $\mathcal{L}_n$ and denoting by $\langle \cdot, \cdot \rangle_t$ the angle-bracket process,

$$\begin{aligned} \langle M_f^n, M_g^n \rangle_t &= \int_0^t (\mathcal{L}_n(f(Y_s^n)g(Y_s^n)) - f(Y_s^n)\mathcal{L}_n(g(Y_s^n)) - g(Y_s^n)\mathcal{L}_n(f(Y_s^n))) ds \\ &= \int_0^t \left(\sum_{\gamma \in G} c(Y_s^n, \gamma) (f(\gamma Y_s^n)g(\gamma Y_s^n) - f(Y_s^n)g(Y_s^n)) \right. \\ &\quad \left. - f(Y_s^n) \sum_{\gamma \in G} c(Y_s^n, \gamma) (g(\gamma Y_s^n) - g(Y_s^n)) \right. \\ &\quad \left. - g(Y_s^n) \sum_{\gamma \in G} c(Y_s^n, \gamma) (f(\gamma Y_s^n) - f(Y_s^n)) \right) ds \\ &= \int_0^t \sum_{\gamma \in G} c(Y_s^n, \gamma) (f(\gamma Y_s^n) - f(Y_s^n)) (g(\gamma Y_s^n) - g(Y_s^n)) ds = A_{f,g}^n(t). \end{aligned}$$

Observe that the integrand of $\langle M_f^n, M_g^n \rangle_t$ corresponds, up to the factor 1/2, to the Carré-du-Champ operator associated to the generator $\mathcal{L}_h$. Since each jump has the height $h = 2^{-n}\xi$, we have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{t \leq T} |Y_t^n - Y_{t-}^n|^2 \right] \leq \lim_{n \rightarrow \infty} \mathbb{E}[2^{-2n}\xi^2] = 0.$$

The processes (B_f^n) and $(A_{f,g}^n)$ are continuous in time, which implies for all $T > 0$ and $f, g \in \mathcal{A}$ that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{t \leq T} |B_f^n(t) - B_f^n(t-)|^2 \right] = \lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{t \leq T} |A_{f,g}^n(t) - A_{f,g}^n(t-)|^2 \right] = 0.$$

This shows that the processes M_f^n are uniformly integrable and that the limit is a time-continuous process.

Next, we claim that

$$(39) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left(\sup_{t \leq T} \left| B_f^n(t) - \int_0^t (\sigma^2 \Delta f(Y_s^n) - \nabla V(Y_s^n) \cdot \nabla f(Y_s^n)) ds - f(Y_0^n) \right| \geq \varepsilon \right) = 0.$$

To this end, we introduce the differences

$$\partial_j^+ f(x) = n(f(x + n^{-1}e_j) - f(x)), \quad \partial_j^- f(x) = n(f(x) - f(x - n^{-1}e_j)).$$

Then, by adding and subtracting some terms involving $\partial_j V(Y_s^n) := (\partial V / \partial x_j)(Y_s^n)$,

$$\begin{aligned} & \left| B_f^n(t) - \int_0^t (\sigma^2 \Delta f(Y_s^n) - \nabla V(Y_s^n) \cdot \nabla f(Y_s^n)) ds - f(Y_0^n) \right| \\ &= \left| \int_0^t [\mathcal{L}_n(f(Y_s^n)) - (\sigma^2 \Delta f(Y_s^n) - \nabla V(Y_s^n) \cdot \nabla f(Y_s^n))] ds \right| \\ &\leq I^+ + I^- + J^+ + J^- + J^0, \end{aligned}$$

where, inserting the definition of $\mathcal{L}_h$ and $c(i, \pm j)$,

$$\begin{aligned} I^\pm &= \sigma^2 n^2 \left| \int_0^t \sum_{j=1}^d \left[\exp \left(\mp \frac{\partial_j^\pm V(Y_s^n)}{2\sigma^2 n} \right) - \exp \left(- \frac{\partial_j V(Y_s^n)}{2\sigma^2 n} \right) \right] \right. \\ &\quad \left. \times (f(Y_s^n \pm n^{-1}e_j) - f(Y_s^n)) ds \right|, \\ J^\pm &= \sigma^2 n^2 \left| \int_0^t \sum_{j=1}^d \left[\exp \left(\mp \frac{\partial_j V(Y_s^n)}{2\sigma^2 n} \right) - \left(1 \mp \frac{\partial_j V(Y_s^n)}{2\sigma^2 n} \right) \right] \right. \\ &\quad \left. \times (f(Y_s^n \pm n^{-1}e_j) - f(Y_s^n)) ds \right|, \\ J^0 &= \left| \int_0^t \left\{ \sum_{j=1}^d \left[\left(\sigma^2 n^2 - \frac{n}{2} \partial_j V(Y_s^n) \right) (f(Y_s^n + n^{-1}e_j) - f(Y_s^n)) \right. \right. \right. \\ &\quad \left. \left. + \left(\sigma^2 n^2 + \frac{n}{2} \partial_j V(Y_s^n) \right) (f(Y_s^n - n^{-1}e_j) - f(Y_s^n)) \right] \right. \right. \\ &\quad \left. \left. + \nabla V(Y_s^n) \cdot \nabla f(Y_s^n) - \sigma^2 \Delta f(Y_s^n) \right\} ds \right|. \end{aligned}$$

The terms $I^\pm$ and $J^\pm$ converge in expectation to zero as $n \rightarrow \infty$, since D is bounded domain, V has a Lipschitz continuous gradient, and $f \in \mathcal{A}$ is smooth. We claim that the

expectation of J^0 also converges to zero. Indeed, let the discrete Laplacian and gradient be given by

$$\begin{aligned}\Delta_n f(i) &= n^2 \sum_{j=1}^d (f(i + n^{-1}e_j) - 2f(i) + f(i - n^{-1}e_j)), \\ \nabla_n f(i) &= \frac{n}{2} (f(i + n^{-1}e_j) - f(i - n^{-1}e_j))_{j=1}^d, \quad i \in \mathcal{K}_h.\end{aligned}$$

This gives

$$\begin{aligned}J^0 &\leq \sigma^2 \left| \int_0^t (\Delta_n f(Y_s^n) - \Delta f(Y_s^n)) ds \right| \\ &\quad + \left| \int_0^t \nabla V(Y_s^n) \cdot (\nabla_n f(Y_s^n) - \nabla f(Y_s^n)) ds \right| \rightarrow 0.\end{aligned}$$

It follows from (39) and the Markov inequality that

$$\begin{aligned}\lim_{n \rightarrow \infty} \mathbb{P} \left(\sup_{t \leq T} \left| B_f^n(t) - \int_0^t (\sigma^2 \Delta f(Y_s^n) - \nabla V(Y_s^n) \cdot \nabla f(Y_s^n)) ds - f(Y_0^n) \right| \geq \varepsilon \right) \\ \leq \frac{1}{\varepsilon} \lim_{n \rightarrow \infty} \mathbb{E}(I^+ + I^- + J^+ + J^- + J^0) = 0.\end{aligned}$$

Similarly, we estimate $A_{f,g}^n$:

$$\begin{aligned}\left| A_{f,g}^n(t) - 2\sigma^2 \int_0^t \nabla f(Y_s^n) \cdot \nabla g(Y_s^n) ds \right| \\ = \sigma^2 \left| \int_0^t \left\{ \sum_{j=1}^d \left[\exp \left(-\frac{1}{2\sigma^2} (V^n(Y_s^n + n^{-1}e_j) - V^n(Y_s^n)) \right) \partial_j^+ f(Y_s^n) \partial_j^+ g(Y_s^n) \right. \right. \right. \\ \left. \left. + \exp \left(\frac{1}{2\sigma^2} (V^n(Y_s^n - n^{-1}e_j) - V^n(Y_s^n)) \right) \partial_j^- f(Y_s^n) \partial_j^- g(Y_s^n) \right] \right. \right. \\ \left. \left. - 2\nabla f(Y_s^n) \cdot \nabla g(Y_s^n) \right\} ds \right| \\ \leq L^+ + L^- + L^0,\end{aligned}$$

where

$$\begin{aligned}L^\pm &= \sigma^2 \left| \int_0^t \sum_{j=1}^d \left[\exp \left(\mp \frac{1}{2\sigma^2} (V^n(Y_s^n \pm n^{-1}e_j) - V^n(Y_s^n)) \right) - 1 \right] \right. \\ &\quad \left. \times \partial_j^\pm f(Y_s^n) \partial_j^\pm g(Y_s^n) ds \right|, \\ L^0 &= \sigma^2 \left| \int_0^t (\partial_j^+ f(Y_s^n) \partial_j^+ g(Y_s^n) + \partial_j^- f(Y_s^n) \partial_j^- g(Y_s^n) - 2\nabla f(Y_s^n) \cdot \nabla g(Y_s^n)) ds \right|.\end{aligned}$$

Similarly as above, the expected values of $L^\pm$ and L^0 converge to zero as $n \rightarrow \infty$ and an application of Markov's inequality shows that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{P} \left[\sup_{t \leq T} \left| A_{f,g}^n(t) - 2\sigma^2 \int_0^t \nabla f(Y_s^n) \cdot \nabla g(Y_s^n) ds \right| \geq \varepsilon \right] \\ & \leq \frac{1}{\varepsilon} \lim_{n \rightarrow \infty} \mathbb{E}[L^+ + L^- + L^0] = 0. \end{aligned}$$

Summarizing the previous results and taking into account [17, Theorem 7.4.1], we conclude that the sequence of processes $(M_f^n)_{n \in \mathbb{N}}$ and also $(B_f^n)_{n \in \mathbb{N}}$ are relatively compact in the space of càdlàg functions $f : \mathbb{R}_+^d \rightarrow D$, i.e. of right-continuous functions for which the left limit exists. Hence, $(f(Y_t^n))_{t \geq 0}$ is relatively compact for all $f \in \mathcal{A}$. Then, for given $f \in \mathcal{A}$, there exists a subsequence $(n_k)_{k \in \mathbb{N}}$ such that $f(Y_t^{n_k})_{t \geq 0}$ converges to some limit $(X_t^f)_{t \geq 0}$.

Set $f_j(x) = \sin(\pi x_j / (2K))$ for $j = 1, \dots, d$; then $f_j \in \mathcal{A}$. There exists a subsequence $(n_k^1)_{k \in \mathbb{N}}$ of $(n_k)_{k \in \mathbb{N}}$ such that $f_1(Y_t^{n_k^1})$ converges to a limit $X_t^{f_1}$. Since $Y_t^{n_k^1} \in [-K, K]$ is bounded for $t \geq 0$ and $f_j^{-1}(y) = (2K/\pi) \sin^{-1}(y)$ for $y \in [-1, 1]$, the processes $Y_t^{n_k^1}$ converges to the limit $\bar{X}_t^1 := f_1^{-1}(X_t^{f_1})$. Repeating this argument for the subsequence (n_k^{j+1}) of (n_k^j) for $j = 1, \dots, d-1$, we infer that there exists a subsequence $(\bar{n}_k)_{k \in \mathbb{N}}$ such that for $j = 1, \dots, d$,

$$f_j(Y_t^{\bar{n}_k}) \text{ converges to } X_t^{f_j}, \quad Y_t^{\bar{n}_k} \text{ converges to } \bar{X}_t^j := f_j^{-1}(X_t^{f_j}).$$

Clearly, $(M_f^{\bar{n}_k})_k$, $(B_f^{\bar{n}_k})_k$, and $f(Y_t^{\bar{n}_k})_k$ are relatively compact for $f \in \mathcal{A}$. As in [17, Theorem 7.4.1], $(M_f^{\bar{n}_k})_k$ is uniformly integrable for $f \in \mathcal{A}$. Consequently,

$$M_f(t) = f(\bar{X}_t) - f(\bar{X}_0) - \int_0^t (\sigma^2 \Delta f(\bar{X}_s) - \nabla V(\bar{X}_s) \cdot \nabla f(\bar{X}_s)) ds$$

is a martingale. Now, if $(X_t)_{t \geq 0}$ is a solution to the martingale problem for $\mathcal{L}$ with initial distribution μ_0 , the uniqueness of this martingale problem implies that $(Y_t^n)_{t \geq 0}$ converges in distribution to $(X_t)_{t \geq 0}$, finishing the proof.

REFERENCES

- [1] F. Al Reda and B. Maury. Interpretation of finite volume discretization schemes for the Fokker–Planck equation as gradient flows for the discrete Wasserstein distance. In: M. Bergounioux et al. (eds.), *Topological Optimization and Optimal Transport*, pp. 400–416. De Gruyter, Berlin, 2017.
- [2] L. Ambrosio, N. Gigli and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lecture Notes Math. Birkhäuser, Basel, 2005.
- [3] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations. *Commun. Partial Differ. Eqs.* 26 (2001), 43–100.
- [4] D. Bakry and M. Emery. Diffusions hypercontractives. In: *Séminaire de Probabilités, XIX, 1983/84*. Lecture Notes Math., pages 177–206. Springer, Berlin, 1985.
- [5] S. Bobov and P. Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.* 19 (2006), 289–336.

- [6] F. Bolley and I. Gentil. Phi-entropy inequalities for diffusion semigroups. *J. Math. Pures Appl.* 93 (2010), 449–473.
- [7] G. Bonaschi and M. Peletier. Quadratic and rate-independent limits for a large-deviations functional. *Contin. Mech. Thermodyn.* 28 (2016), 1191–1219.
- [8] A.-S. Boudou, P. Caputo, P. Dai Pra, and G. Posta. Spectral gap estimates for interacting particle systems via a Bochner-type identity. *J. Funct. Anal.* 232 (2006), 222–258.
- [9] R. Bubley and M. Dyer. Path coupling: A technique for proving rapid mixing in Markov chains. *Proceedings of the 38th Annual Symposium on Foundations of Computer Science* (1997), 223–231.
- [10] C. Cancès and J. Venel. On the square-root approximation finite volume scheme for nonlinear drift-diffusion equations. *Comptes Rendus Math.* 361 (2023), 535–558.
- [11] P. Caputo, P. Dai Pra, and G. Posta. Convex entropy decay via the Bochner–Bakry–Emery approach. *Ann. Inst. Henri Poincaré Probab. Stat.* 45 (2009), 734–753.
- [12] J. A. Carrillo, A. Jüngel, and M. Santos. Displacement convexity for the entropy in semidiscrete nonlinear Fokker–Planck equations. *Europ. J. Appl. Math.* 30 (2019), 1103–1122.
- [13] C. Chainais-Hillairet and M. Herda. Large-time behaviour of a family of finite volume schemes for boundary-driven convection–diffusion equations. *IMA J. Numer. Anal.* 40 (2020), 2473–2504.
- [14] G. Conforti. A probabilistic approach to convex (ϕ)-entropy decay for Markov chains. *Ann. Appl. Probab.* 32 (2022), 932–973.
- [15] P. Dai Pra and G. Posta. Entropy decay for interacting systems via the Bochner–Bakry–Emery approach. *Electron. J. Probab.* 18 (2013), 1–21.
- [16] R. Eldan, J. Lee, and J. Lehec. Transport-entropy inequalities and curvature in discrete-space Markov chains. In: M. Loebl, J. Nešetřil, and R. Thomas (eds.), *A Journey Through Discrete Mathematics*, pp. 391–406. Springer, Cham, 2017.
- [17] S. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. Wiley, New York, 1986.
- [18] M. Fathi and J. Maas. Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.* 26 (2016), 1774–1806.
- [19] F. Filbet and M. Herda. A finite volume scheme for boundary-driven convection–diffusion equations with relative entropy structure. *Numer. Math.* 137 (2017), 535–577.
- [20] A. Glitzky. Uniform exponential decay of the free energy for Voronoi finite volume discretized reaction-diffusion systems. *Math. Nachr.* 284 (2011), 2159–2174.
- [21] A. Jüngel and D. Matthes. An algorithmic construction of entropies in higher-order nonlinear PDEs. *Nonlinearity* 19 (2006), 633–659.
- [22] A. Jüngel and S. Schuchnigg. A discrete Bakry–Emery method and its application to the porous-medium equation. *Discrete Contin. Dyn. Syst.* 37 (2017), 5541–5560.
- [23] A. Jüngel and W. Yue. Discrete Beckner inequalities via the Bochner–Bakry–Emery approach for Markov chains. *Ann. Appl. Probab.* 27 (2017), 2238–2269.
- [24] M. Liero, A. Mielke, M. Peletier, and D. Renger. On microscopic origins of generalized gradient structures. *Discrete Contin. Dyn. Syst. Ser. S* 10 (2017), 1–35.
- [25] J. Maas and D. Matthes. Long-time behavior of a finite volume discretization for a fourth order diffusion equation. *Nonlinearity* 29 (2016), 1992–2023.
- [26] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Diff. Eqs.* 48 (2013), 1–31.
- [27] J. Norris. *Markov Chains*. Cambridge University Press, Cambridge, 1997.
- [28] Y. Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.* 256 (2009), 810–864.
- [29] F. Pedrotti. Contractive coupling rates and curvature lower bounds for Markov chains. Submitted for publication, 2023. arXiv:2308.00516.
- [30] Hiroshi Tanaka. Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Math. J.*, 9(1):163–177, 1979.

INSTITUTE OF ANALYSIS AND SCIENTIFIC COMPUTING, VIENNA UNIVERSITY OF TECHNOLOGY,
WIEDNER HAUPTSTRASSE 8–10, 1040 WIEN, AUSTRIA

Email address: `juengel@tuwien.ac.at`

INSTITUTE OF ANALYSIS AND SCIENTIFIC COMPUTING, VIENNA UNIVERSITY OF TECHNOLOGY,
WIEDNER HAUPTSTRASSE 8–10, 1040 WIEN, AUSTRIA

Email address: `katharina.schuh@tuwien.ac.at`