

On the optimal convergence rate for the Metropolis algorithm in one dimension

A Patrón^{1,*}, A D Chepelianskii², A Prados¹ and E Trizac^{3,4}

¹ Física Teórica, Universidad de Sevilla, Apartado de Correos 1065, E-41080 Sevilla, Spain

² Laboratoire de Physique des Solides, Université Paris-Saclay, CNRS, 91405 Orsay, France

³ LPTMS, Université Paris-Saclay, CNRS, 91405 Orsay, France

⁴ Ecole Normale Supérieure de Lyon, F-69364 Lyon, France

E-mail: apatron@us.es, alexei.chepelianskii@universite-paris-saclay.fr and prados@us.es

Received 31 October 2024

Accepted for publication 3 January 2025

Published 29 January 2025



Online at stacks.iop.org/JSTAT/2025/013214
<https://doi.org/10.1088/1742-5468/ada7e6>

Abstract. We study the relaxation of the Metropolis Monte Carlo algorithm corresponding to a single particle trapped in a one-dimensional confining potential, with even jump distributions that ensure that the dynamics verifies detailed balance: in particular, how can one minimise the characteristic time for reaching the target equilibrium probability distribution function? To work out the corresponding optimal sampling method, we study the physical mechanisms that affect the dynamics. Previous work suggested that, for smooth jump distributions, the fastest convergence rate is obtained as a result of the competition between diffusive and rejection-dominated dynamics. In this work, we show that a new relevant physical regime comes into play for two-peaked jump distributions, where the relaxation dynamics is dominated neither by diffusion nor by rejection: the eigenmodes adopt an oscillatory form, reminiscent of charge density waves (CDW)—thus we term this new regime the CDW regime. Using a combination of numerical and analytical techniques, the parameter regions corresponding to

* Author to whom any correspondence should be addressed.

diffusion, rejection, and CDW are characterised, as well as the transition lines between them. Therefore, a phase diagram is built, in which there are two order parameters characterising each of the phase transitions: the Inverse Participation Ratio and the Fidelity—reminiscent of the quantum fidelity. The optimal convergence rate is located at the triple point of phase coexistence, where the transition lines (diffusive-rejection, diffusive-CDW, and CDW-rejection) intersect. We then address the fundamental problem of finding the optimal jump distribution, without any assumption on its functional form. Our theoretical framework is checked against the numerical diagonalisation of the master equation for the case of harmonic confinement, which is the paradigmatic example of a convex potential with only one minimum. We also briefly discuss more sophisticated attempts at optimising the convergence rate to equilibrium.

Keywords: classical monte carlo simulations, numerical simulations, stochastic processes

Contents

1. Introduction	3
2. Relaxation dynamics for the Metropolis algorithm	5
2.1. Parametric families of jump distributions	6
2.2. Convergence rate	7
2.3. Localisation transition and eigenvalue crossing	9
3. Ergodicity breaking and CDW	11
3.1. Main assumptions behind the emergence of the CDW eigenfunctions from the master equation	11
3.2. Numerical check of the CDW eigenfunctions	13
4. Phase-diagrams for the diffusion-rejection-CDW picture	15
4.1. Order parameters	15
4.2. Numerical results on the phase behaviour	16
5. Estimating the optimal convergence rate	18
5.1. Estimators for the next-to-largest eigenvalue	20
5.2. Corroborating the physical picture	21
6. On the shape of the optimal jump distribution	23
6.1. More complex jump distributions	24
6.2. Sequence of jumps	26
7. Conclusions	27
Data availability statement	29
Acknowledgments	30

Appendix A. Analytical derivation of the CDW eigenfunctions	30
Appendix B. Robustness of the CDW eigenfunctions	32
Appendix C. Ergodicity breaking and periodic solutions	33
References	35

1. Introduction

Computational algorithms based on Monte Carlo techniques are widely known for their versatility in solving many scientific problems. For instance, in molecular simulations, they are essential for determining the equilibrium properties of physical systems involving many degrees of freedom with a high accuracy [1]. Other branches of knowledge that range from the natural sciences of physics, chemistry, and biology [2, 3] to the fields of economy, machine learning, and archeology [4–6] benefit from these techniques when approaching problems at the deterministic and stochastic levels of description, hence the interest in the study of the convergence of such algorithms, as it provides useful insights on the optimal performance they may reach.

The focus of our work concerns the Metropolis–Hastings algorithm, or simply Metropolis [7], which belongs to the class of Markov Chain Monte Carlo techniques. These are based on creating a sequence of steps from a random walk that makes it possible to sample a desired target probability distribution, which usually corresponds to the equilibrium distribution, provided that the number of steps is large enough—i.e. in the long-time limit, as discussed below. In the Metropolis algorithm, the random displacements η in parameter space giving the new steps of the sequence are drawn from a so-called jump distribution $\omega(\eta)$ —also known as proposal distribution—that is an even function of the displacements. Once a random displacement is chosen, the associated energy change ΔU is calculated and the step is accepted or rejected according to the Metropolis rule: if $\Delta U < 0$, it is accepted with probability $p = 1$; if $\Delta U > 0$, it is accepted with probability $p = \exp(-\beta\Delta U)$, where β is the inverse temperature. This rule verifies detailed balance, and therefore it ensures that the system equilibrates in the long-time limit [8, 9]—provided that the jump distribution allows for the sampling of all configurations, i.e. ergodicity is not broken [10].

A key issue concerns the overall convergence time of the algorithm: it quantifies the time needed to attain the target distribution, which greatly impacts its performance—or equivalently, the convergence rate, which measures how fast the algorithm attains such a target distribution. The longer the convergence time, the larger the error in the computed observables. On the one hand, if the typical size a of the attempted displacements is too small, most of the jumps are accepted but the phase space is not explored sufficiently, thus involving a long convergence time. On the other hand, if the size a of these displacements is too large, most jumps are rejected because they are likely to involve a high energy cost. This discussion suggests that there must exist an optimal amplitude a_{opt} that minimises the convergence time for the Metropolis algorithm. Although there have been many attempts to derive a_{opt} for specific models [11–13], the general

rule of thumb corresponds to choosing a so that the acceptance and rejection rates are comparable, i.e. the probability of accepting or rejecting the new step is close to 0.5 [1, 14, 15].

In this work, we intend to go one step further. Rather than determining what is the optimal size a_{opt} for a given jump distribution $\omega(\eta)$, our work is focused on studying what is the optimal jump distribution $\omega_{\text{opt}}(\eta)$ that minimises the convergence time. This problem can be viewed as engineering the jump distribution to accelerate the dynamics of the system, thus sharing the objectives of the general field of shortcuts, started in the quantum field of shortcuts to adiabaticity [16] and later transposed to classical and stochastic systems as swift state-to-state transformations [17]. Note that engineering the jump distribution could be viewed as analogous to engineering the potential since changing the shape of the jump distribution varies the jump moments—which provide the coefficients in the Kramers–Moyal expansion of the master equation [8, 9].

In order to tackle the aforementioned optimisation problem, it becomes necessary to study and understand the physical mechanisms affecting the relaxation dynamics of the system. In previous work [18], the relaxation dynamics of the Metropolis algorithm was studied for a particle trapped in a one-dimensional confining potential $U(x)$, and the jump distribution that optimises the convergence rate to equilibrium was investigated. For different potentials and choices of the jump distribution, it is shown that a critical jump length a^* appears where a localisation transition occurs: going from $a < a^*$ to $a > a^*$ radically changes the relaxation behaviour. This localisation transition mainly affects the eigenmodes of the master equation characterising the algorithm: for $a < a^*$, the eigenmodes resemble the diffusive modes obtained at the limit $a \rightarrow 0^+$, while, for $a > a^*$, the eigenmodes become spatially localised at discrete points x_0 , where the rejection probability is maximal. Interestingly, it is shown that $a_{\text{opt}} = a^*$ for most of the cases considered, implying that the optimal convergence time is obtained as a balance between diffusion and rejection.

Here, we show that the diffusion-rejection picture described above breaks down when two-peaked jump distributions are considered for one-dimensional potentials. In such a scenario, although the localisation transition still occurs, the optimal convergence rate is attained instead at a point where a crossing between the leading eigenvalues of the dynamics takes place. Eigenvalue crossing constitutes a fundamental phenomenon that plays a critical role in explaining and predicting the behaviour of physical systems across various domains both in quantum [19, 20] and statistical mechanics [21]. We focus our study on such a phenomenon and how it relates to the emergence of a new regime, in which the spatial dependence of the eigenmodes of the Metropolis algorithm varies drastically to an oscillatory behaviour, reminiscent of the *charge density wave* (CDW) transition in correlated electron systems [22–24]. We will refer to this regime, which competes with both the diffusion and rejection ones, as the CDW regime. For the sake of simplicity, we present the derivation of a novel CDW regime for the case of convex potentials with only one minimum, although the final result applies to a broad class of confining potentials in one dimension.

The first of the main goals of this paper is to characterise the novel CDW regime and analyse its implications for the optimisation of the convergence rate of the Metropolis algorithm. To do so, we introduce two order parameters, the fidelity \mathcal{F} and the inverse

participation ratio (IPR), which allow us to characterise the parameter regions dominated by diffusion, rejection, and CDW. A key finding of our work is that the optimal convergence rate is placed at the triple point of coexistence of the diffusive, rejection, and CDW phases. As our second major goal, we employ a variational approach to tackle the functional problem of finding the optimal jump distribution $\omega_{\text{opt}}(\eta)$, and discuss additional strategies on how to further optimise the convergence rate.

The structure of this paper is as follows. Section 2 presents the mathematical framework behind the Metropolis algorithm. It also provides a brief account of the diffusion-rejection picture from [18] and discusses the emergence of the eigenvalue crossing phenomenon. Section 3 is devoted to the derivation and characterisation of the CDW regime. In addition, we numerically check the emergence of its corresponding oscillatory modes. The phase diagram of the Metropolis algorithm is analysed in section 4, for a specific bi-parametric family of jump distributions. Then, in section 5, we introduce the variational method employed for estimating the optimal jump distribution. Also, we employ this variational method to further corroborate numerically that optimality is attained as the competition between the CDW, diffusion, and rejection regimes. Section 6 presents some preliminary ideas for estimating the shape of the optimal jump distribution of the Metropolis algorithm. Finally, we provide some concluding remarks in section 7. The appendix discusses some technicalities that are omitted in the main text.

2. Relaxation dynamics for the Metropolis algorithm

In this work, we study the relaxation dynamics of a particle confined in a one-dimensional confining potential $U(x)$, with x accounting for its position on the real line, and additionally immersed in a thermal bath at temperature T . At each discrete time step n , the particle's position evolves according to the so-called Metropolis rule

$$x_n = \begin{cases} x_{n-1} + \eta_n, & \text{with prob. } p(x_n, x_{n-1}) \equiv \min(1, e^{-\beta \Delta U(x_n, x_{n-1})}), \\ x_{n-1}, & \text{with prob. } 1 - p(x_n, x_{n-1}), \end{cases} \quad (1)$$

where $\Delta U(x_n, x_{n-1}) \equiv U(x_n) - U(x_{n-1})$ and $\beta \equiv (k_B T)^{-1}$ is the inverse temperature. Independent random variables η_n are drawn from a normalised distribution $\omega(\eta)$, which is usually termed the jump distribution. Note that $\Delta x = x_n - x_{n-1} = \eta_n$ for accepted jumps, while $\Delta x = 0$ otherwise. We further assume the jump distribution to be symmetric, i.e. $\omega(-\eta) = \omega(\eta)$, which ensures that detailed balance holds.

At the ensemble level, the system is described via a position probability distribution $P_n[\omega](x)$, which depends on the functional form of the jump distribution $\omega(\eta)$. Its dynamical evolution is governed by the master equation

$$P_n[\omega](x) = \int_{-\infty}^{+\infty} dx' F[\omega](x, x') P_{n-1}[\omega](x'), \quad (2)$$

in which

$$F[\omega](x, x') \equiv \delta(x - x') R[\omega](x) + \omega(x - x') p(x, x'), \quad (3)$$

is the (temperature-dependent [26]) integral kernel of the master equation—i.e. the Markov kernel of the corresponding Markov chain, and

$$R[\omega](x) \equiv \int_{-\infty}^{+\infty} dx' \omega(x' - x) [1 - p(x', x)]. \quad (4)$$

We remark that $R[\omega](x)$ gives the rejection probability at position x , which is a central quantity in this work. Our notation emphasises the (functional) dependence of P_n , F , and R on the jump distribution $\omega(\eta)$, in addition to their spatial dependence.

The master equation (2) has the canonical equilibrium distribution

$$P_{\text{eq}}(x) = \frac{1}{Z} e^{-\beta U(x)}, \quad Z \equiv \int_{-\infty}^{+\infty} dx e^{-\beta U(x)}, \quad (5)$$

as a stationary solution. Metropolis rule (1) ensures that the equilibrium distribution is always attained in the long-time limit $n \rightarrow \infty$, regardless of the shape of the initial probability distribution $P_0(x)$, i.e.

$$P_{\infty}(x) \equiv \lim_{n \rightarrow \infty} P_n[\omega](x) = P_{\text{eq}}(x), \quad (6)$$

provided that the jump distribution allows for reaching any final position x from an arbitrary initial position x' , i.e. the stochastic matrix $F[\omega](x - x')$ is irreducible—in the physics literature, this is usually identified with ergodicity. Note that, since the rejection probability $R[\omega](x) \neq 0$, the stochastic matrix is always aperiodic, so irreducibility—or ergodicity—ensures the convergence to the equilibrium distribution in the long-time limit.

2.1. Parametric families of jump distributions

We stress that our analysis on the convergence rate of the Metropolis algorithm is general, encompassing a wide class of jump distributions $\omega(\eta)$. Still, for illustration purposes, we consider some specific choices—the following bi-parametric families of jump distributions—throughout this work:

- The (a, σ) -family of Gaussian jump distributions

$$\omega^{(a, \sigma)}(\eta) = C^{(a, \sigma)} \exp\left(-\frac{(|\eta| - a)^2}{2\sigma^2}\right), \quad (7)$$

where $C^{(a, \sigma)}$ is a normalisation constant, and $a, \sigma > 0$.

- The (a, α) -family of algebraic jump distributions

$$\omega^{(a, \alpha)}(\eta) = \frac{\alpha + 1}{2a^{\alpha+1}} |\eta|^{\alpha} \theta(a - |\eta|), \quad (8)$$

with $a, \alpha > 0$. In the above, α accounts for the width of the distribution, a stands for the position of its maximum, and $\theta(x)$ corresponds to the Heaviside step-function; $\theta(x) = 1$ for $x \geq 0$ and $\theta(x) = 0$ for $x < 0$.

- The (a, b) -family of box-like jump distributions:

$$\omega^{(a,b)}(\eta) = \frac{1}{4b} \theta(|\eta| - a + b) \theta(a + b - |\eta|), \quad a > b > 0. \quad (9)$$

This distribution has a simple physical interpretation: only jumps with $a - b < |\eta| < a + b$ are allowed, with a flat distribution thereof.

The reasons behind these specific choices of families of jump-distributions are twofold: On the one hand, the parameter a —common for the three families—accounts for the typical size of the attempted displacements, thus allowing to interpolate between the diffusion and rejection regimes previously uncovered in reference [18]. On the other hand, all these families tend to a double Dirac-delta peak

$$\omega_0(\eta) = \frac{1}{2} [\delta(\eta - a) + \delta(\eta + a)], \quad (10)$$

when their respective widths tend to zero, i.e. when (i) $\sigma \rightarrow 0^+$ in equation (7), (ii) $\alpha \rightarrow +\infty$ in equation (8), and (iii) $b \rightarrow 0^+$ in equation (9). The double delta peak does not allow for the sampling of all positions since the length of the jumps is fixed to $\pm a$. This will be important in connection with the eigenvalue crossing discussed at the end of section 2.3. It is worth mentioning that the considered families of jump distributions are distinct enough such that they provide a sense of generality to the results that we present in the forthcoming sections. Specifically, while $\omega^{(a,\sigma)}(\eta)$ and $\omega^{(a,\alpha)}(\eta)$ have well-defined peaks at $\eta = \pm a$, $\omega^{(a,b)}(\eta)$ does not, as it presents a continuous of maxima in between the interval $a - b \leq \eta \leq a + b$. Moreover, while $\omega^{(a,\sigma)}(\eta)$ and $\omega^{(a,b)}(\eta)$ are symmetric with respect to $\eta = \pm a$, $\omega^{(a,\alpha)}(\eta)$ does not, as it presents an algebraic tail for $|\eta| \leq a$ and it vanishes for $|\eta| \geq a$.

2.2. Convergence rate

To study the convergence rate to equilibrium, we put our focus on the spectrum of eigenvalues of the integral kernel $F[\omega](x, x')$ of the master equation (2). For that purpose, it is useful to introduce a new integral kernel $K[\omega](x, x')$ that share the spectrum of eigenvalues with $F[\omega](x, x')$,

$$K[\omega](x, x') \equiv e^{-\beta \Delta U(x, x')/2} F[\omega](x, x') = \omega(x - x') e^{-\beta |\Delta U(x, x')|/2} + \delta(x - x') R[\omega](x). \quad (11)$$

Moreover, $K[\omega](x, x')$ is symmetric under the exchange of $x \leftrightarrow x'$ —note the absolute value in the exponential factor multiplying $\omega(x - x')$, which makes it self-adjoint and simplifies our mathematical description below.

Now we define the eigenvalues λ_ν and eigenfunctions $\phi_\nu(x)$ of the integral operator \hat{K} ,

$$\hat{K}[\omega] \phi_\nu(x) \equiv \int_{-\infty}^{+\infty} dx' K[\omega](x, x') \phi_\nu(x') = \lambda_\nu[\omega] \phi_\nu[\omega](x), \quad \nu = 0, 1, 2, \dots \quad (12)$$

Since $K[\omega](x, x')$ is self-adjoint, all the eigenvalues of equation (12) are real. Moreover, for jump distributions ω such that the Perron–Frobenius theorem [25] applies, this theorem ensures that (i) the maximum eigenvalue is ω -independent, $\lambda_0 = 1$, which corresponds to the stationary solution being non-degenerate, and (ii) the remaining eigenvalues $\lambda_\nu[\omega]$, with $\nu > 0$, verify the strict inequality $|\lambda_\nu| < 1$. The eigenfunction $\phi_0(x)$ for the eigenvalue $\lambda_0 = 1$, corresponding to the equilibrium distribution, is also ω -independent,

$$\phi_0(x) \propto e^{-\beta U(x)/2}. \quad (13)$$

Explicitly, we can write

$$\int_{-\infty}^{+\infty} dx' K[\omega](x, x') \phi_0(x') = \phi_0(x), \quad (14a)$$

$$\int_{-\infty}^{+\infty} dx' K[\omega](x, x') \phi_\nu[\omega](x') = \lambda_\nu[\omega] \phi_\nu[\omega](x), \quad \nu > 0. \quad (14b)$$

Our notation emphasises the fact that, under the conditions of the Perron–Frobenius theorem, both λ_0 and ϕ_0 are independent of the jump distribution ω , whereas the remainder of the eigenvalues and eigenfunctions are functionals thereof. Making use of the explicit form of the kernel $K[\omega](x, x')$ in equation (11), we have that

$$\int_{-\infty}^{+\infty} dx' \omega(x - x') e^{-\beta|\Delta U(x, x')|/2} \phi_\nu(x') + R[\omega](x) \phi_\nu[\omega](x) = \lambda_\nu[\omega] \phi_\nu[\omega](x), \quad \nu > 0. \quad (15)$$

The solution of the master equation can always be written as a linear combination of the eigenfunctions $\phi_\nu(x)$:

$$P_n[\omega](x) = P_{\text{eq}}(x) + e^{-\beta U(x)/2} \sum_{\nu > 0} \mathcal{A}_\nu[\omega] \phi_\nu[\omega](x) (\lambda_\nu[\omega])^n, \quad (16)$$

where $\mathcal{A}_\nu[\omega]$ are constants determined by the initial condition $P_0(x)$. For long times, the system always reaches equilibrium because $|\lambda_\nu[\omega]| < 1$ for $\nu > 0$. The leading convergence rate to equilibrium of the Metropolis algorithm is thus dominated by the contribution corresponding to the next-to-largest eigenvalue $\lambda_1[\omega]$, since

$$P_n[\omega](x) - P_{\text{eq}}(x) \sim e^{-\beta U(x)/2} \mathcal{A}_1[\omega] \phi_1[\omega](x) (\lambda_1[\omega])^n, \quad n \gg 1. \quad (17)$$

The next-to-largest eigenvalue can always be written as

$$\Lambda[\omega] \equiv \lambda_1[\omega] = \max_{\phi \perp \phi_0} \int dx \int dx' \phi(x) K[\omega](x, x') \phi(x') \quad (18)$$

equation (17) entails that $P_n[\omega](x) - P_{\text{eq}}(x) \propto e^{-n|\ln \Lambda[\omega]|}$, so we can define the convergence rate as

$$\gamma[\omega] \equiv |\ln \Lambda[\omega]|, \quad P_n[\omega](x) - P_{\text{eq}}(x) \propto e^{-n\gamma[\omega]}, \quad n \gg 1. \quad (19)$$

The larger $\gamma[\omega]$, the faster the system equilibrates. Note that $\gamma[\omega]$ increases as $\Lambda[\omega]$ decreases, i.e. $\gamma[\omega]$ is an increasing function of the spectral gap $1 - \Lambda[\omega]$ between the largest and next-to-largest eigenvalues. Our final goal is to select the optimal $\omega^*(\eta)$ that minimises $\Lambda[\omega]$, i.e. maximises $\gamma[\omega]$ or the spectral gap $1 - \Lambda[\omega]$, in order to find the optimal convergence rate to equilibrium:

$$\Lambda^* = \min_{\omega} \Lambda[\omega] \implies \gamma^* = |\ln \Lambda^*|. \quad (20)$$

Therefore, we are actually dealing with a min-max problem. The minimisation in equation (20) should ideally be carried out over the whole space of possible jump distributions but, for practical purposes, it may also be performed within a given family of jump distributions, such as those in equations (7)–(9).

2.3. Localisation transition and eigenvalue crossing

In order to get more information about the physics of the problem, a sufficiently narrow jump distribution $\omega(\eta) = \delta_{\sigma}(\eta)$ around $\eta=0$ with the standard deviation σ was considered in [18]. In this way, equation (12) can be expanded around $x' = x$ due to the smallness of σ , giving as a result, a Schrödinger-like equation with an effective potential,

$$\sigma^2 \left\{ -\frac{\phi_n''(x)}{2} + \left[\frac{\beta^2}{8} (U'(x))^2 - \frac{\beta}{4} U''(x) \right] \phi_n(x) \right\} = (1 - \lambda_n) \phi_n(x). \quad (21)$$

The eigenmodes $\phi_n(x)$ are the eigenfunctions of this Schrodinger-like equation, which have a well-defined parity for even potentials, such that $U(x) = U(-x)$. This makes it possible to split this Schrödinger basis into two parts that contain, respectively, the even and odd eigenfunctions. This fact will become useful later on, specifically, in sections 5 and 6.

In [18], it was also argued that the convergence rate is limited by the rejection probability $R[\omega](x)$. The main idea is the following: starting from an initial distribution $\delta(x - x_0)$ at step $n=0$, the probability distribution function at any later time is

$$P_n(x) = R[\omega](x_0)^n \delta(x - x_0) + p_n(x|x_0), \quad (22)$$

where $p_n(x|x_0)$ is a smooth function. The system cannot relax faster than this exponential decay, i.e. the convergence rate should always be slower than the maximum of $|\ln R[\omega](x_0)|$ as a function of x_0 . Equivalently, we can write

$$\Lambda \geq \max_x R[\omega](x). \quad (23)$$

The main result of [18] is that, for many different choices of jump distributions $\omega(\eta)$ and confining potentials $U(x)$ —even at higher dimensions and including interactions, the optimal convergence rate is obtained at the transition point between the diffusion and rejection regimes, which is where the *localisation transition* takes place—and where the inequality (23) saturates. The term localisation stems from the fact that, in addition to the discrete spectrum of well-defined eigenfunctions, a continuum of eigenvalues appears in the rejection regime, with singular localised eigenfunctions—as suggested

On the optimal convergence rate for the Metropolis algorithm in one dimension

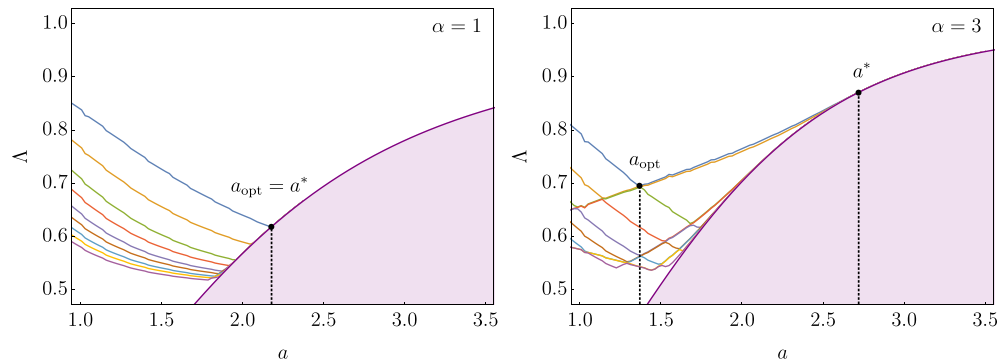


Figure 1. Leading eigenvalues λ_ν , $\nu = 1, \dots, 9$, of the master equation (2) for the Metropolis algorithm, for a particle confined in a harmonic potential $U(x) = \chi x^2/2$. In both panels, we have employed the jump distribution $\omega^{(a,\alpha)}(\eta)$ from equation (8), and we have plotted the eigenvalues against a for two fixed values of α : (left) $\alpha = 1$ and (right) $\alpha = 3$. Purple curves correspond to the maximal rejection probability, with the light purple regions accounting for the continuum of eigenvalues obtained upon the localisation transition. The dashed, vertical lines denote the position of either the optimal value a_{opt} and/or the point a^* at which the localisation transition takes place.

by equation (22), these singular eigenfunctions are peaked at any point x_0 , with eigenvalue $R[\omega](x_0)$. Here we show that, although the above diffusion-rejection picture is quite robust for a wide class of jump distributions, this relatively simple scenario fails when considering two-peaked jump distributions—regardless of the chosen confining potential.

Let us consider the above optimisation problem for a particle confined in a harmonic potential, $U(x) = \chi x^2/2$ —with χ being the stiffness, and the (a, α) -family from equation (8). As α increases, this family of jump distributions varies from a flat distribution in the interval $(-a, a)$, for $\alpha \ll 1$, to a two delta-peak distribution at $\eta = \pm a$, for $\alpha \gg 1$. In figure 1, we plot the leading eigenvalues of the Metropolis algorithm λ_ν , $1 \leq \nu \leq 9$, as functions of a for two specific values of the parameter α , namely $\alpha = 1$ and $\alpha = 3$. For $\alpha = 1$, the left panel agrees with the diffusion-rejection picture developed in reference [18]: the optimal value a_{opt} is obtained at the point a^* where the localisation transition takes place. However, this picture breaks as we increase the value of α , and the jump distribution develops a two-peak structure. For $\alpha = 3$, we show in the right panel that $a_{\text{opt}} \neq a^*$: the optimal convergence rate is not attained at the localisation transition, but at a lower value where two eigenvalues cross and thus Λ takes its minimum possible value. Eigenvalue crossing is related to abrupt changes in the dynamical behaviour of physical systems, thus pointing toward the emergence of new dominant physics—as shown, for instance, in reference [21] in connection with dynamical phase transitions and the Mpemba effect [27–29]. In order to tackle the global problem of maximising the convergence rate of the Metropolis algorithm, it is important to understand all the physical mechanisms that come into play in its relaxation dynamics. The

latter constitutes the first main goal of our work, which we next explore in sections 3 and 4.

3. Ergodicity breaking and CDW

The emergence of a new physical regime for sufficiently peaked jump distributions is intimately related to the loss of ergodicity of the Markov chain for Dirac-delta distributions: for 1D systems, Dirac-delta jump distributions of the form $\delta(|\eta| - a)$ only allows jumps of size $\pm a$, thus leading to discrete points $x_n = x \pm na$, $n = 0, 1, 2, \dots$ in the real line. Therefore, irreducibility is broken and the system is prevented from attaining the equilibrium distribution in the long-time limit. In this section, we show that, for two-peaked jump distributions in 1D, a new regime emerges that solves the above issue. In this regime, the eigenfunctions of the master equation present an oscillatory behaviour, reminiscent of CDWs in electronic systems [22–24]. In condensed matter physics, CDWs constitute a collective transport phenomenon that arises from strong electron–phonon interactions in certain low-dimensional materials. Almost a century ago, Peierls suggested that one-dimensional metals are unstable toward the formation of aperiodic lattice distortions associated with spatially periodic modulations of the electronic charge density [30]—which would be later known as CDWs. The formation of CDWs opens a gap in the Fermi level at low temperatures, which lowers the kinetic energy of the conduction electrons. Due to the similarities between the CDWs found in materials and the explicit form of eigenfunctions reported in this work—i.e. low dimensionality of the system, the lattice periodicity and the inherent relation to the emergence of a phase transition—we refer to this new regime as the CDW regime.

In this section, we provide a rigorous characterisation of this novel CDW regime for the Metropolis algorithm in one dimension, and we corroborate our results by comparing our approximate analytical calculation with the ‘exact’ numerical diagonalisation of the master equation [32].

3.1. Main assumptions behind the emergence of the CDW eigenfunctions from the master equation

We may derive the corresponding CDW eigenfunctions directly from the master equation and study their properties. To do so, we consider here a two-peak family of jump distributions $\omega(\eta; a, \sigma)$, the maxima of which are at $\eta = \pm a$ with a certain small width, measured by a parameter σ . The parameters (a, σ) define the two characteristic lengths of the jump distribution. Let us introduce the auxiliary distribution

$$\delta_\sigma(\eta - a) \equiv 2\omega(\eta)\theta(\eta). \quad (24)$$

By definition, (i) the maximum of $\delta_\sigma(\eta)$ is located at $\eta = 0$ and (ii) $\delta(\eta)$ identically vanishes for $\eta < -a$. The width of the peak of δ_σ is measured by its standard deviation. Using this definition, we write the jump distribution as follows:

$$\omega(\eta; a, \sigma) = \frac{1}{2} [\delta_\sigma(\eta - a) + \delta_\sigma(-\eta - a)]. \quad (25)$$

For $\sigma = 0$, $\delta_0(\eta)$ corresponds to a Dirac-delta distribution and equation (25) reduces to equation (10), which—as discussed above—breaks ergodicity and does not allow the system to equilibrate. Here, we will be mainly interested in the regime $\sigma \ll a$ —but non-zero, where the jump distribution is narrowly peaked around $\pm a$, and thus $\omega(0; a, \sigma) = \delta_\sigma(-a) \approx 0$.

Now we proceed to analyse the eigenfunctions $\phi_\nu(x; a, \sigma)$ of the master equation (2), i.e. the non-trivial solutions of equation (15). It is convenient to introduce the change of variables

$$\phi_\nu(x; a, \sigma) = \phi_0(x) \rho_\nu(x; a, \sigma) \propto e^{-\beta U(x)/2} \rho_\nu(x; a, \sigma), \quad (26)$$

such that the equilibrium distribution corresponds to a constant, $\rho_0(x; a, \sigma) = \text{const.}$ Note that functionals of ω translate here into a parametric dependence on (a, σ) , which we omit henceforth to simplify the notation. For the sake of concreteness and simplicity, we present an analysis for confining potentials $U(x)$ that are convex, i.e. $U''(x) > 0$, and thus have only one minimum—without loss of generality, we assume that the minimum is at $x = 0$. We stress that the convexity condition can be relaxed and our conclusions still hold, but it allows us to simplify the presentation while keeping the main ideas. The analytical derivation of the CDW eigenfunctions is rather technical and cumbersome. Hence, we relegate it to appendix A. Below, we outline the two main assumptions employed in the derivation:

- (1) For jumps where the potential energy increases, we assume that the increase is significant. Consequently, most subsequent jumps to the side of the potential will be rejected, while jumps to the opposite side are predominantly accepted. Thus, the overall rejection probability is approximately 1/2. Mathematically, this applies for sufficiently large values of a and x not too close to the global minimum, such that

$$e^{-\beta \Delta U(x \pm a, x)} \ll 1. \quad (27)$$

- (2) We assume that the CDW ansatz,

$$\rho_\nu(x \pm a) = \rho_\nu(x), \quad \forall x, \quad (28)$$

holds. This assumption follows from the fact that for infinitely sharp jump distributions—i.e. $\sigma = 0$, where ergodicity is broken—the ground state becomes infinitely degenerate. In this case, any periodic solution of the form (28) yields the eigenvalue $\lambda_\nu = 1$. Thus, our theoretical approach can be regarded as a singular perturbation theory within this subspace.

Based on these assumptions, we expand the master equation from equation (14) in powers of σ , resulting in the following equation for $\rho_\nu(x)$:

$$\rho_\nu''(x) + \frac{4(1 - \lambda_\nu)}{\sigma^2} \rho_\nu(x) = 0, \quad \forall x, \quad (29)$$

which provides the solutions

$$\rho_\nu(x) = c_\nu \cos\left(\frac{2\pi\nu}{a}x\right) + d_\nu \sin\left(\frac{2\pi\nu}{a}x\right), \quad \lambda_\nu = 1 - \nu^2 \pi^2 \left(\frac{\sigma}{a}\right)^2. \quad (30)$$

Note that these CDW modes only make sense for $\sigma \neq 0$: as stated above, for $\sigma = 0$ any a -periodic $\rho_\nu(x)$ belongs in the subspace corresponding to $\lambda_0 = 1$, which becomes degenerate due to the breaking of ergodicity. In addition, the CDW modes cease to make sense for very large values of ν : we know that the exact eigenvalues of the master equation cannot become smaller than -1 , whereas $\lim_{\nu \rightarrow \infty} \lambda_\nu \rightarrow -\infty$ [33].

For $\sigma \neq 0$, we have that $\lambda_0 = 1$ is clearly non-degenerate, since the sine mode identically vanishes for $\nu = 0$ and the cosine mode becomes a constant. On the other hand, all the other eigenvalues λ_ν , with $\nu > 0$, seem to be degenerate, since both the sine and the cosine modes are non-zero. Henceforth, we write

$$\phi_\nu^{\text{CDW},e}(x) = c_\nu^e \phi_0(x) \cos(k_\nu x) \propto e^{-\beta U(x)/2} \cos(k_\nu x), \quad (31a)$$

$$\phi_\nu^{\text{CDW},o}(x) = c_\nu^o \phi_0(x) \sin(k_\nu x) \propto e^{-\beta U(x)/2} \sin(k_\nu x), \quad (31b)$$

with

$$k_\nu = \frac{2\pi\nu}{a}, \quad \lambda_\nu = 1 - \frac{\sigma^2}{4} k_\nu^2, \quad (32)$$

for the CDW modes that approximate the exact eigenfunctions $\phi_\nu(x)$ of the master equation for $\sigma \ll a$. In the above, the indexes e and o refer to the even and odd sectors of the CDW basis, respectively. The normalisation constant c_ν is determined by imposing $\|\phi_\nu^{\text{CDW}}\| = 1$, i.e.

$$\int_{-\infty}^{+\infty} dx |\phi_\nu^{\text{CDW}}(x)|^2 = 1. \quad (33)$$

3.2. Numerical check of the CDW eigenfunctions

In order to check the correctness of our theoretical description, we have numerically diagonalised the master equation, i.e. numerically solved the eigenproblem (12), for the harmonic potential with different jump distributions. Again, we have fixed units by taking $k_B T = 1$ and $\chi = 1$, see [32]. For each of the families of jump distributions, equations (7)–(9), we have chosen the parameters so that $\omega(\eta)$ adopts a two-peak structure.

In figure 2, the numerical leading eigenfunction $\phi_1(x)$ is compared with the analytical approximation $\phi_1^{\text{CDW}}(x)$ in equation (31). When it shows oscillatory behaviour, $\phi_1(x)$ is always even: therefore, we have fitted it to a cosine-shaped eigenfunction,

$$\phi_1^{\text{fit}}(x) = c(k_1) \phi_0(x) \cos(k_1^* x), \quad (34)$$

where $c(k_1)$ is a normalisation constant such that $\|\phi_1^{\text{fit}}\| = 1$, by minimising the distance between the two functions,

$$d(\phi_1, \phi_1^{\text{fit}}) \equiv \int_{-\infty}^{+\infty} dx |\phi_1^{\text{fit}}(x) - \phi_1(x)|^2 \quad (35)$$

On the optimal convergence rate for the Metropolis algorithm in one dimension

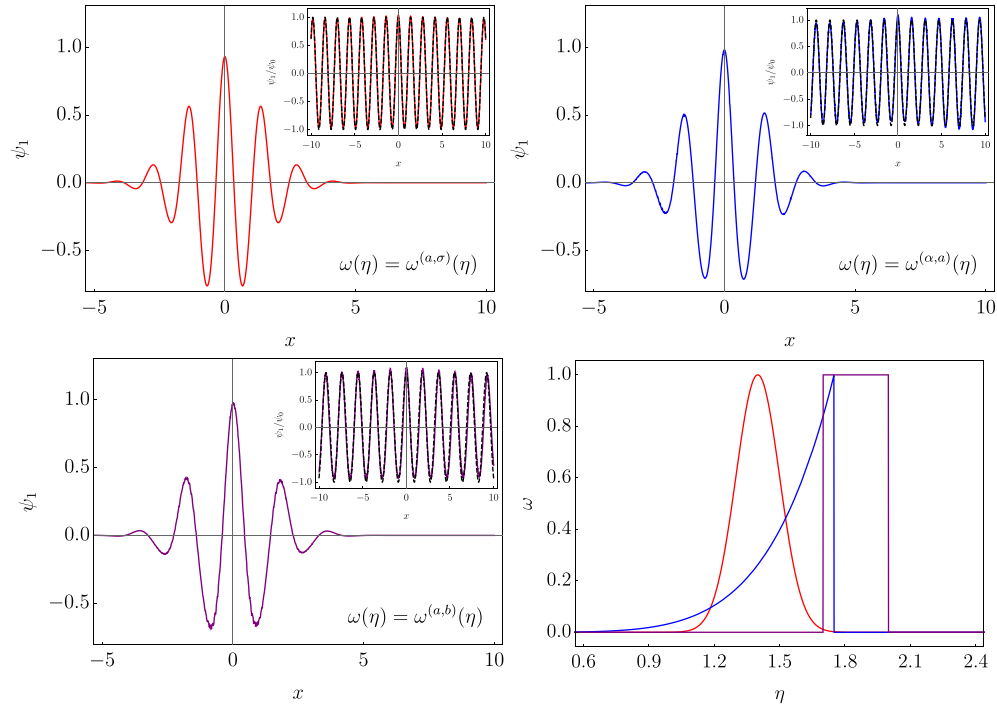


Figure 2. Leading eigenfunction for the Metropolis algorithm. It has been obtained via numerical diagonalisation of the master equation, for: (upper left) $\omega^{(a,\sigma)}(\eta)$, with $(a, \sigma) = (1.4, 0.1)$, (upper right) $\omega^{(a,\alpha)}(\eta)$, with $(a, \alpha) = (1.75, 6)$, and (lower left) $\omega^{(a,b)}(\eta)$ with $(a, b) = (1.85, 0.15)$. The insets present the ratio between the leading eigenfunction and the equilibrium distribution $\phi_0(x)$. The dashed curves correspond to the best fit of such ratios to the function $\cos(k_1^* x)$, with $k_1^* = 2\pi/a^*$, with (upper left) $a^* = 1.40$, (upper right), $a^* = 1.56$, and (lower left) $a^* = 1.85$, respectively. The corresponding jump distributions—only for $\eta > 0$, recall that they are even functions of η —are shown in the lower right panel: $\omega^{(a,\sigma)}(\eta)$ in red, $\omega^{(a,\alpha)}(\eta)$ in blue, and $\omega^{(a,b)}(\eta)$ in purple.

over k_1 . The obtained fits are excellent in all cases, with the largest separation taking place near the origin, where the amplitude of the numerical eigenfunction is slightly larger than that of the theoretical curve. This is reasonable, since the equation for the CDW actually holds for $|x| > a/2$, see Appendix A. On the one hand, the fitted wave vector $k_1^* = 2\pi/a^*$ perfectly matches the theoretical value $k_1 = 2\pi/a$ for the Gaussian and box-like jump distributions in equations (7) and (9), respectively. On the other hand, $k_1^* \neq 2\pi/a$ for the algebraic jump distribution in equation (8). We recall that we described $\delta_\sigma(\eta - a)$ —which is basically the quantity shown in the lower right panel of figure 2—as a peaked distribution around $\eta = a$ with a width proportional to its variance. Although this is a fair description of Gaussian and box-like distributions, it is not so for algebraic jump distributions, which are not symmetric around the point at which the probability accumulates as $\sigma \rightarrow 0^+$. However, a good theoretical prediction for k_1^* is always given by $k_{\text{th}}^* = 2\pi/a_{\text{th}}^*$, with a_{th}^* being the point at which the cumulative distribution equals $1/2$, i.e. $\int_0^{a_{\text{th}}^*} d\eta \delta_\sigma(\eta - a) = 1/2$. For the parameters employed in the figure, $a_{\text{th}}^* = 1.58$ while the fitted value is $a^* = 1.56$ —the difference is approximately

1%. Note that $a_{\text{th}}^* = a$ for both the Gaussian and box-like distributions in equations (7) and (9).

It is interesting that the leading eigenfunction always seems to be even when the CDW oscillatory behaviour emerges. This fact hints at the terms neglected in our derivation of the CDW modes breaking the degeneracy of the sine and cosine eigenfunctions. A more refined approach would be necessary to analytically explain the even parity of the leading eigenfunction.

4. Phase-diagrams for the diffusion-rejection-CDW picture

In this section, we apply our findings in the previous section to investigate the optimisation of the convergence rate for one family of bi-parametric jump distributions, and explore the different phases or regimes that emerge when varying the system parameters. Specifically, we apply our general results to the specific case of harmonic confinement, with the algebraic family of jump distributions of equation (8).

4.1. Order parameters

Our goal is to draw the phase diagram of the system in the (a, α) plane of system parameters. In that way, we will elucidate the interplay between the diffusion, rejection, and CDW phases. The diffusive phase corresponds to the region in the (a, α) plane where the leading eigenvalue Λ is given by the diffusive modes—analogously, we speak of the rejection and CDW phases. As one moves in the (a, α) plane, the jump distribution changes shape. For distributions that allow only small jumps, a diffusive phase is expected; for distributions that allow for large jumps, a rejection phase is expected; for two-peak distributions, a CDW phase is expected. One key finding of our analysis is that the point (a^*, α^*) at which the leading eigenvalue Λ takes its minimum value in the plane is the triple point of phase coexistence.

In order to characterise the different phases, i.e. diffusion-dominated, rejection-dominated, and CDW-dominated, we introduce the following two order parameters, the fidelity \mathcal{F} and the IPR:

- **Fidelity:** Similarly to that employed in the context of coherent states in quantum mechanics, we define the fidelity for the problem of our concern as

$$\mathcal{F} \equiv |\langle \phi_1 | \phi_1^{\text{CDW}} \rangle|^2 = \max_k \left| \int_{-\infty}^{+\infty} dx \, c_1 \phi_0(x) \cos(kx) \phi_1(x) \right|^2, \quad (36)$$

which measures the affinity between the actual leading eigenfunction $\phi_1(x)$ of the Metropolis algorithm with the CDW ansatz ϕ_1^{CDW} from equation (31) with the corresponding optimal wave number k^* . Let us note that the fidelity is closely related to the distance about the involved functions, $d = 2(1 - \mathcal{F})$ —compare equation (36) with equation (35).

The case $\mathcal{F} = 0$ corresponds to null affinity, while $\mathcal{F} = 1$ results in a perfect match. For $k^* \rightarrow 0$, the fidelity tends to zero, since the actual leading eigenfunction $\phi_1(x)$ is orthogonal to the equilibrium solution.

- **IPR:** This quantity, introduced in [18], measures the emergence of the localisation transition in the eigenfunctions of the master equation. In our numerical diagonalisation of the master equation, space is discretised into a lattice with N_d sites. Within this discretisation scheme, the eigenfunction $\phi_\nu(x)$ simplifies into a set of N_d values $\phi_\nu(i)$, $i = 1, \dots, N_d$. The IPR for the leading eigenfunction is

$$\text{IPR} \equiv \sum_{i=1}^{N_d} |\phi_1(i)|^4 / \left(\sum_{i=1}^{N_d} |\phi_1(i)|^2 \right)^2. \quad (37)$$

If $\phi_1(x)$ is completely delocalised, then $\text{IPR} \rightarrow 0$ for large N_d , implying that it has a well-defined continuum limit, with no delta peaks. If part of the eigenfunction gets localised, then $\text{IPR} = O(1)$.

4.2. Numerical results on the phase behaviour

In figures 3 and 4, the heat maps for the fidelity \mathcal{F} , the IPR and the next-to-largest eigenvalue Λ are shown on the (a, α) plane for the bi-parametric family of algebraic jump distributions in equation (8). These diagrams have been obtained by numerically diagonalising the master equation. We recall that the parameter a gives the typical size of the attempted jumps, while the parameter α controls the width of the distribution, which becomes a delta peak at $\eta = a$ in the limit as $\alpha \rightarrow \infty$.

In figure 3, we show the heat map of the order parameters, the fidelity \mathcal{F} (left) and the IPR (right).

- **Behaviour of the fidelity:** For low values of α , the fidelity is always zero. The jump distribution does not have a double-peak structure, and then the CDW modes do not emerge. Instead, for $\alpha > \alpha_c = 1.22$, \mathcal{F} abruptly varies from close-to-zero values to close-to-unity values, when increasing a . For $a < a_1(\alpha)$, the Schrödinger modes dominate and we are in the diffusive regime, for which $\mathcal{F} \approx 0$, while for $a > a_1(\alpha)$, the CDW modes emerge and dominate, and thus $\mathcal{F} \approx 1$. The fidelity order parameter is discontinuous at the line $a = a_1(\alpha)$, which is thus a first-order transition line. As we continue to increase a , the fidelity begins to smoothly decrease back toward zero around a certain value $a_2(\alpha)$ that is less clearly defined. This is related to the change from the CDW regime to the rejection regime—as we will see more clearly when considering the IPR.
- **Behaviour of the IPR:** For $\alpha < \alpha_c$, for which the CDW regime does not emerge, the localisation transition takes place along the line $a = a_3(\alpha)$. Thereat, the IPR abruptly changes from almost null values to $O(1)$ ones, and $a = a_3(\alpha)$ is another first-order transition line—as already reported in [18]. For $\alpha > \alpha_c$, the IPR vanishes both in the regions where the diffusive modes and the CDW modes dominate. As a increases and approaches the region where the fidelity smoothly decreases back from unity to zero, the IPR in turn smoothly increases from zero to unity—and retains non-zero, $O(1)$, values when a is further increased. The transition line is not as clear-cut as $a_1(\alpha)$ and $a_3(\alpha)$, since both \mathcal{F} and the IPR vary smoothly. As a first approach, we could define the second-order transition line $a_2(\alpha)$ as the locus

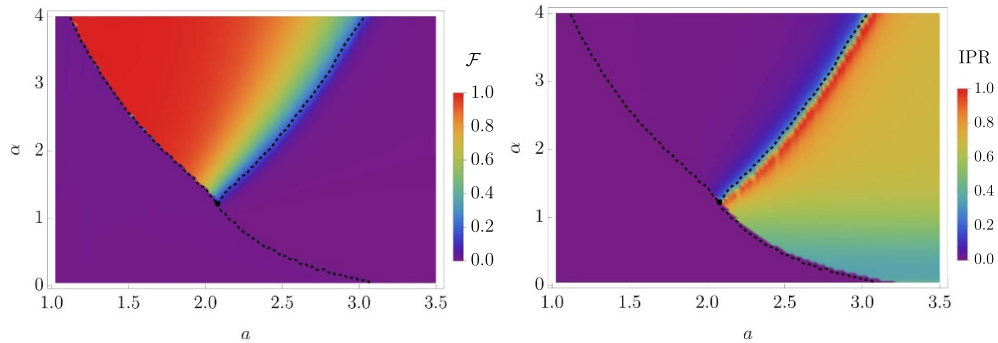


Figure 3. Heat maps of the fidelity \mathcal{F} (left) and the IPR (right). Both are plotted in the (a, α) plane of parameters for the algebraic jump distribution in equation (8). Both order parameters have been evaluated by numerically obtaining the leading eigenfunction of the master equation. The transition lines between the different phases are plotted with dashed lines and intersect at the tricritical point $(a_c, \alpha_c) = (2.08, 1.22)$, which is represented by a black dot. These lines correspond to (i) $a_1(\alpha)$ for $\alpha > \alpha_c$ and $a < a_c$, (ii) $a_2(\alpha)$ for $\alpha > \alpha_c$ and $a > a_c$, and (iii) $a_3(\alpha)$ for $\alpha < \alpha_c$ and $a > a_c$, respectively, as defined in the text.

at which $\mathcal{F} = \text{IPR}$. The three so-defined transition lines cross at the tricritical point $(a_c, \alpha_c) = (2.08, 1.22)$.

Figure 4 shows the corresponding heat map of the next-to-largest eigenvalue Λ , obtained by numerically diagonalising the master equation. We have checked that, up to our numerical precision, Λ attains its global minimum, $\Lambda_{\min} \simeq 0.6170$, at the tricritical point. As is neatly seen, there exists a valley-like region for $\alpha \leq 2$ in which Λ has values that are very close to the minimum. The valley comprises points close to the transition line (i) $a_1(\alpha)$ between the diffusive and CDW phases, for $\alpha_c < \alpha < 2$, and (ii) $a_3(\alpha)$ between the diffusive and rejection phases, for $\alpha < \alpha_c$.

Figure 5 gives us further information about the phase transitions described above. Specifically, we plot the fidelity \mathcal{F} and the IPR for the (a, α) -family of jump-distributions from equation (8), by varying a for fixed $\alpha = 3$ —similarly to right panel of figure 1. For small values of a , we are in the diffusive regime, for which the fidelity and the IPR are close to zero. At the transition point $a_1(\alpha = 3) = 1.37$ between the diffusive and CDW regimes, the eigenvalue crossing takes place—as shown in figure 1, such that the dominant eigenfunction $\phi_1(x)$ changes parity: from odd, for $a < a_1$, to even, for $a > a_1$. This is consistent with the behaviour of the dominant eigenfunctions in the respective phases: in the diffusion regime, the leading eigenfunction $\phi_1(x)$ is provided by the odd sector, whereas, in the CDW regime, the leading eigenfunctions are even—as discussed in section 3.2. Upon further increase in a , the eigenvalues smoothly tend to the rejection curve, marking the onset of the rejection phase. The abrupt changes in the behaviour among the diffusive, CDW and rejection regimes reflect the emergence of different underlying phase transitions in the relaxation dynamics of the Metropolis algorithm. First, the transition between the diffusive and the CDW regimes constitutes a first-order phase transition, characterised by a discontinuous change of the fidelity,

On the optimal convergence rate for the Metropolis algorithm in one dimension

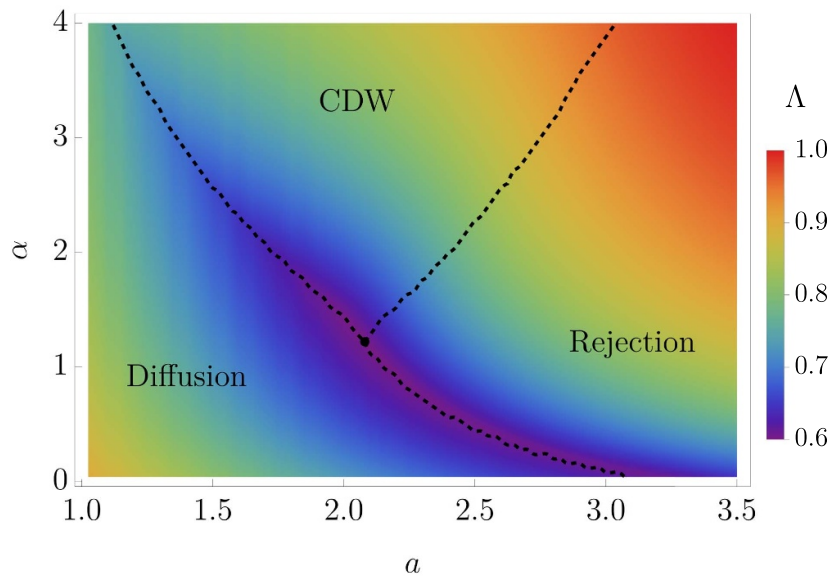


Figure 4. Heat map of the next-to-largest eigenvalue Λ in the (a, α) plane. As in figure 3, the dashed curves correspond to the transition lines between phases. The black point indicates the coordinates of the global minimum, $(a_c, \alpha_c) = (2.08, 1.22)$, which coincides with the tricritical point at which the transition lines intersect.

which is due to the fact that the takeover of the leading relaxation mode involves a change of symmetry in the leading eigenfunction. Second, the transition from the CDW regime to the rejection regime constitutes a higher-order—at least second-order—phase transition, characterised by the smooth changes of both \mathcal{F} and IPR, which is due to the also smooth variation of the leading eigenvalue between the two regimes.

5. Estimating the optimal convergence rate

As we may infer from the approaches followed in the previous sections, analytically determining both $\lambda_1[\omega]$ and its corresponding eigenfunction $\phi_1[\omega](x)$ from the integral equation (12) is generally not possible. Needless to say it is that the same applies to determining the optimal jump distribution $\omega_{\text{opt}}(\eta)$ that maximises the convergence rate. In the remaining sections, we employ a variational approach in order to estimate the convergence rate and the optimal jump distribution for different families of jump distributions. The advantages of this approach are twofold: (i) it is computationally less expensive than the numerical diagonalisation method employed to solve the master equation, and (ii) as we will show, it allows us to further corroborate the diffusion-rejection-CDW picture.

Let us consider a certain known basis of orthogonal functions $\psi_n(x)$, and expand the function $\phi(x)$ in equation (18) in this basis. Truncating the expansion after N_s states,

On the optimal convergence rate for the Metropolis algorithm in one dimension

$\alpha = 3$, IPR (blue) and Fidelity (red)

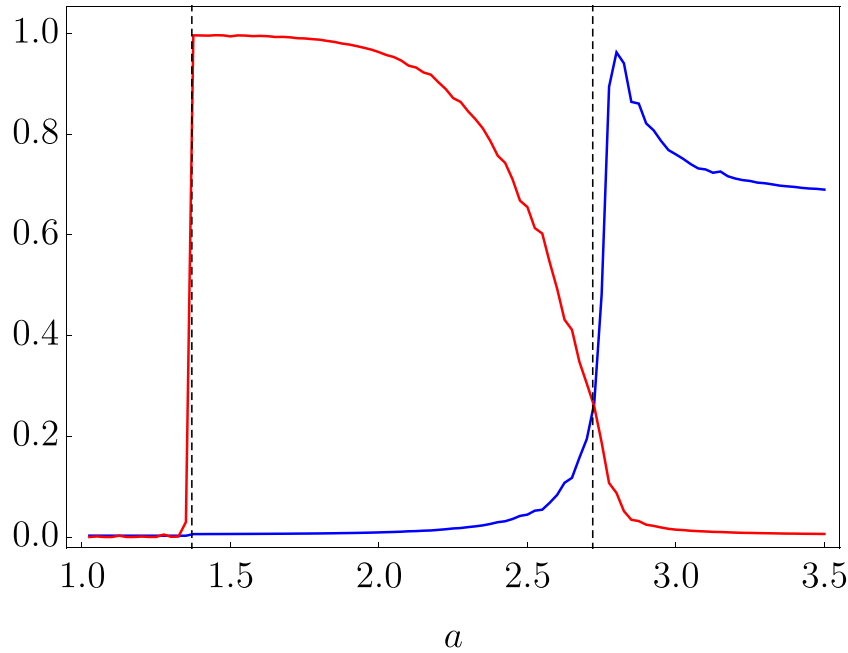


Figure 5. Fidelity \mathcal{F} and IPR as functions of a , for a fixed value of $\alpha = 3 > \alpha_c$ —similar to the right panel of figure 1. Specifically, we have employed the jump distribution $\omega^{(a,\alpha)}(\eta)$ in equation (8). The dashed, vertical lines delimit the three different phases explored in this work: for $a < a_1(\alpha = 3) = 1.37$, the Schrödinger diffusion-like regime, for $a_1(\alpha = 3) = 1.37 < a < a_2(\alpha = 3) = 2.72$, upon eigenvalue crossing, the CDW regime, and for $a > a_2(\alpha = 3) = 2.72$, the rejection regime.

which must be perpendicular to the ground state, we get a lower bound for $\Lambda[\omega]$, i.e.

$$\Lambda[\omega] \geq \tilde{\Lambda}[\omega] \equiv \max_{c_1, c_2, \dots, c_{N_s}} \int dx \int dx' \phi(x) K[\omega](x, x') \phi(x'),$$

$$\phi(x) = \sum_{n=1}^{N_s} c_n \psi_n(x). \quad (38)$$

Let us remark that the above inequality saturates only in the limit $N_s \rightarrow +\infty$. Therein, we have the complete expansion of $\phi(x)$, and therefore the equality is attained regardless of our basis choice. However, certain choices of basis, i.e. certain choices of the functions $\psi_n(x)$ —which we refer to as physical choices—will get us closer to the true next-to-largest eigenvalue than others, for a not so large number of retained eigenmodes N_s .

Now, in order to calculate our estimate $\tilde{\Lambda}[\omega]$ for the next-to-largest eigenvalue, we must find the optimal values of the coefficients $\{c_n\}_{n=1}^{N_s}$ that maximise the integral in equation (38). Such a problem is equivalent to finding the largest eigenvalue of the reduced $N_s \times N_s$ matrix $\tilde{K}[\omega]$ with elements

$$\tilde{K}_{nm}[\omega] = \int dx \int dx' \psi_n(x) K[\omega](x, x') \psi_m(x'), \quad (39)$$

corresponding to the truncation $1 \leq n, m \leq N_s$. It is worth emphasising that for symmetric confining potentials, the eigenbasis of our problem may be split into even and odd eigenfunctions. The integral kernel $K[\omega](x, x')$ inherits these symmetries, implying that the matrix elements $\tilde{K}_{nm}[\omega]$ are zero for those values of n and m for which the corresponding eigenfunctions have different parity. Thus, we may divide the basis into two bases: one containing the even eigenfunctions, and the other one containing the odd eigenfunctions. In this way, the matrix \tilde{K}_{nm} can be written as a direct sum of even and odd sectors.

Using different bases, we can ensure that our lower bound is sufficiently close to the true next-to-largest eigenvalue for every possible jump distribution $\omega(\eta)$. In fact, using M different bases, we may define the functional

$$\tilde{\Lambda}_M[\omega] \equiv \max \left[\tilde{\Lambda}^{(1)}[\omega], \tilde{\Lambda}^{(2)}[\omega], \dots, \tilde{\Lambda}^{(M)}[\omega] \right] \quad (40)$$

to estimate the next-to-largest eigenvalue, taking a sufficiently high number of retained eigenfunctions N_s for each basis. In what follows, we mainly use the estimator above with different values of M , typically with $100 \leq N_s \leq 300$ retained eigenfunctions per basis.

In order to approximately solve the min-max problem presented above, we thus resort to a two-step procedure. First, for each considered jump distribution $\omega(\eta)$, we use the functional (40) to find an accurate estimate for the leading eigenvalue $\lambda_1[\omega]$, i.e. the maximum $\tilde{\Lambda}_M[\omega]$ of its lower bounds $\tilde{\Lambda}^{(i)}[\omega]$, $i = 1, \dots, M$, for the M bases considered. Second, we vary the jump distribution to find the minimum value of such estimates $\tilde{\Lambda}_M^*$, which is attained for a specific jump distribution ω_M^* , i.e. $\tilde{\Lambda}_M^* \equiv \tilde{\Lambda}_M[\omega_M^*]$ [34]. One could use bases without a clear physical meaning and the lower bound obtained in this way might still be quite useful for numerical purposes, but this ‘blind’ procedure would not provide meaningful insights into the behaviour of the eigenfunctions of the master equation and the physics underlying the optimal value for the convergence rate—we recall that the convergence rate is determined by the next-to-largest eigenvalue, as given by equations (19) and (20).

5.1. Estimators for the next-to-largest eigenvalue

Following the Schrodinger basis obtained for sufficiently narrow jump distributions around $\eta = 0$ —briefly reviewed at the beginning of section 2.3, the next-to-largest eigenvalue Λ was estimated in reference [18] by using the 3-bases functional

$$\tilde{\Lambda}_3[\omega] \equiv \max \left[\tilde{\Lambda}_o^{(\text{Sch.})}[\omega], \tilde{\Lambda}_e^{(\text{Sch.})}[\omega], \max_x R[\omega](x) \right], \quad (41)$$

where $\tilde{\Lambda}_o^{(\text{Sch.})}[\omega]$ and $\tilde{\Lambda}_e^{(\text{Sch.})}[\omega]$ are the estimates for Λ obtained from the odd and even sectors of the Schrödinger basis, respectively. Estimate (41) is based on the assumption that the convergence rate is obtained as a balance between the diffusion modes of the Schrödinger equation and the maximum value of the rejection rate. The validity of such an argument was checked for several choices of the confining potential $U(x)$ and different jump distributions $\omega(\eta)$, even for a few number N_s of retained eigenmodes.

Given the results previously presented concerning the characterisation of the CDW regime, we build another estimator for the next-to-largest eigenvalue by adding the CDW basis, as given by equation (31), thereto:

$$\tilde{\Lambda}_4[\omega] \equiv \max \left[\tilde{\Lambda}_0^{(\text{Sch.})}[\omega], \tilde{\Lambda}_e^{(\text{Sch.})}[\omega], \tilde{\Lambda}^{(\text{CDW})}[\omega], \max_x R(x; \omega) \right], \quad (42)$$

where $\tilde{\Lambda}^{(\text{CDW})}[\omega]$ is the estimate given by the CDW basis for the next-to-largest eigenvalue.

Let us also consider the localised Gaussian modes

$$\phi_n(x) = \mathcal{N}_n \left[e^{-\frac{(x-a_n)^2}{2\sigma^2}} + \epsilon e^{-\frac{(x+a_n)^2}{2\sigma^2}} \right], \quad a_n = \frac{L}{N_s} \left(n + \frac{1}{2} \right), \quad \sigma = \frac{L}{N_s}, \quad n = 0, 1, \dots, N_s - 1, \quad (43)$$

where \mathcal{N}_n is a normalisation constant. The parameter ϵ is equal to either $+1$ or -1 and sets the symmetry of the eigenfunctions. Therefore, this basis can also be split into even and odd sectors. By adding the even and odd bases of localised Gaussian modes to $\tilde{\Lambda}_4$, we build a more refined estimator with $M = 6$ bases:

$$\tilde{\Lambda}_6[\omega] \equiv \max \left[\tilde{\Lambda}_0^{(\text{Sch.})}[\omega], \tilde{\Lambda}_e^{(\text{Sch.})}[\omega], \tilde{\Lambda}_e^{(\text{Gau.})}[\omega], \tilde{\Lambda}_o^{(\text{Gau.})}[\omega], \tilde{\Lambda}^{(\text{CDW})}[\omega], \max_x R(x; \omega) \right], \quad (44)$$

where $\tilde{\Lambda}_e^{(\text{Gau.})}[\omega]$ and $\tilde{\Lambda}_o^{(\text{Gau.})}[\omega]$ are the estimates for the next-to-largest eigenvalue provided by the even and odd sectors, respectively, of the localised Gaussian mode basis.

We recall the min-max problem that we solve for estimating the next-to-largest eigenvalue Λ . For each jump distribution ω , the above estimators have a definite value—the maximum over the M bases considered in each estimator, which we denote by $\tilde{\Lambda}_M[\omega]$. Afterwards, by varying the jump distribution, we find out the optimal jump distribution that minimises $\tilde{\Lambda}_M[\omega]$: this is what we denote by ω_M^* , which in turn provides the estimate $\tilde{\Lambda}_M^* \equiv \tilde{\Lambda}_M[\omega_M^*]$ for the considered set of bases. Note that we always have $\tilde{\Lambda}_6^* \geq \tilde{\Lambda}_4^* \geq \tilde{\Lambda}_3^*$; the functional $\tilde{\Lambda}_6[\omega]$ provides the most accurate lower bound for the next-to-largest eigenvalue.

In the following sections, the values of $\tilde{\Lambda}_M[\omega]$ and $\tilde{\Lambda}_M^*$ are given to 10^{-4} precisions. This is a challenging precision, but one that can be obtained by increasing the number of retained modes N_s for each basis. With the typical size employed throughout this work, $100 \leq N_s \leq 300$, we have checked that further increasing N_s do not alter the reported values for our best estimate $\tilde{\Lambda}_6^*$ of the next-to-largest eigenvalue—recall that any basis would perfectly give Λ in the limit as $N_s \rightarrow \infty$.

5.2. Corroborating the physical picture

To start with, let us consider the bi-parametric (a, σ) family of Gaussian jump distributions defined in equation (7) for the Metropolis algorithm. That is, we look for the optimal values of the parameters (a^*, σ^*) that minimise the next-to-largest eigenvalue Λ , following our variational approach. In the numerical scheme for the optimisation, a

On the optimal convergence rate for the Metropolis algorithm in one dimension

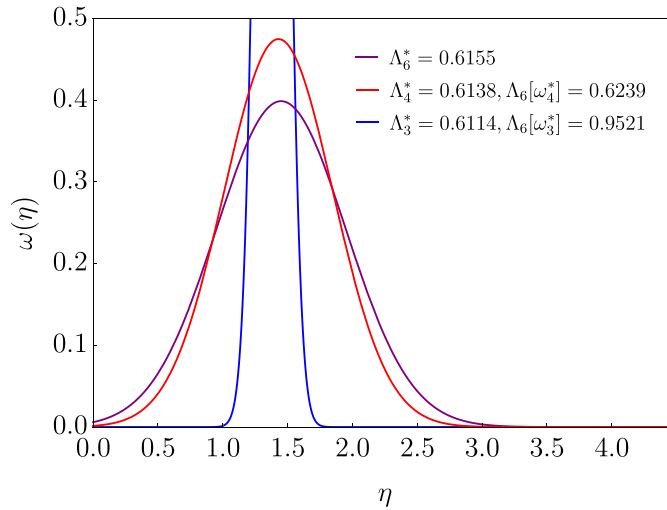


Figure 6. Optimal jump distributions for the Gaussian family in equation (7). These optimal jump distributions have been obtained by employing the three functionals $\tilde{\Lambda}_3[\omega]$, $\tilde{\Lambda}_4[\omega]$, and $\tilde{\Lambda}_6[\omega]$ for the next-to-largest eigenvalue. Specifically, we show ω_6^* (purple), with associated parameters $(a_6^*, \sigma_6^*) = (1.45, 0.50)$ for $\tilde{\Lambda}_6[\omega]$, ω_4^* (red), with $(a_4^*, \sigma_4^*) = (1.43, 0.42)$ for $\tilde{\Lambda}_4[\omega]$, and ω_3^* (blue), with $(a_3^*, \sigma_3^*) = (1.38, 0.10)$ for $\tilde{\Lambda}_3[\omega]$. In the legend, we present the values of the next-to-largest eigenvalue mentioned in the main text.

lower cut-off in the width σ must be imposed, i.e. $\sigma \geq \sigma_{\text{cutoff}}$, since very small values of σ lead to numerical difficulties [35]. If, when decreasing σ_{cutoff} , σ^* consistently reached the cut-off value, this would hint at an incomplete convergence of σ due to the cut-off, i.e. it would hint at actually having $\sigma^* \rightarrow 0$. We refer to this problematic scenario as the *collapse instability*, since the optimal jump distribution seems to converge to a double Dirac-delta peak, which breaks ergodicity as discussed in the preceding sections.

In figure 6, we display the optimal jump distributions obtained via the functionals $\tilde{\Lambda}_3$, $\tilde{\Lambda}_4$ and $\tilde{\Lambda}_6$ for the next-to-largest eigenvalue—note that the latter two functionals include the CDW basis. The cutoff $\sigma_{\text{cutoff}} = 0.1$ has been employed when finding the optimal parameters (a, σ) for each estimator. The failure of the diffusion-rejection picture is illustrated by the fact that $\omega_3^*(\eta)$ reaches the numerical cutoff $\sigma_3^* = \sigma_{\text{cutoff}}$, thus tending toward a Dirac delta-like distribution. It is neatly seen that the variances of $\omega_4^*(\eta)$ and $\omega_6^*(\eta)$ no longer reach the cutoff, so that the collapse instability is fixed by including the CDW basis. Moreover, ω_4^* and ω_6^* are close and quite different from the Dirac-like distribution ω_3^* . On the one hand, all the functionals give estimates for the next-to-largest eigenvalue that are not so different, $\tilde{\Lambda}_3^* = 0.6114$, $\tilde{\Lambda}_4^* = 0.6138$, and $\tilde{\Lambda}_6^* = 0.6155$. On the other hand, the incomplete convergence of ω_3^* can further be verified by inserting it into $\tilde{\Lambda}_4[\omega]$ and $\tilde{\Lambda}_6[\omega]$. This gives $\tilde{\Lambda}_4[\omega_3^*] = 0.9239$ and $\tilde{\Lambda}_6[\omega_3^*] = 0.9521$, which have a relative error of more than 50% with respect to our best estimate of the next-to-largest eigenvalue $\tilde{\Lambda}_6^*$. If we perform a similar check with ω_4^* , by evaluating $\tilde{\Lambda}_6[\omega_4^*] = 0.6239$, the relative error is reduced to 1.36%, thus clearly showing the improvement of the convergence obtained with the incorporation of the CDW basis.

On the optimal convergence rate for the Metropolis algorithm in one dimension

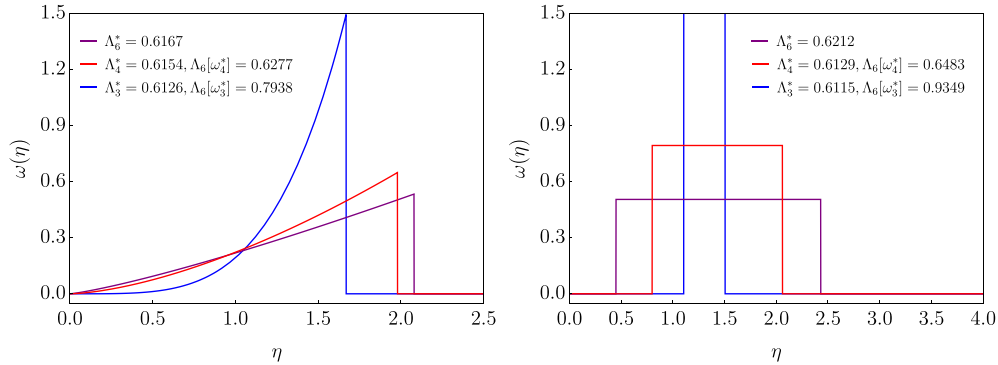


Figure 7. Optimal jump distributions for the algebraic and box-like family of distributions. Specifically, the algebraic (8) and box-like (9) cases are considered in the left and right panels, respectively. Similarly to figure 6, the optimal jump distributions depicted correspond to the functionals $\tilde{\Lambda}_3[\omega]$, $\tilde{\Lambda}_4[\omega]$, and $\tilde{\Lambda}_6[\omega]$. Specifically, we show: ω_6^* (purple), for which $(a_6^*, \alpha_6^*) = (2.08, 1.22)$ and $(a_6^*, b_6^*) = (1.44, 0.99)$, ω_4^* (red), for which $(a_4^*, \alpha_4^*) = (1.98, 1.57)$ and $(a_4^*, b_4^*) = (1.43, 0.63)$ and (blue) ω_3^* , for which $(a_3^*, \alpha_3^*) = (1.67, 4.00)$ and $(a_3^*, b_3^*) = (1.30, 0.20)$. Again, the values of the next-to-largest eigenvalue mentioned in the main text are given in the legend.

We further support the above picture by applying the same methodology to the other two bi-parametric families of jump distributions: (i) the algebraic (a, α) one, equation (8), and (ii) the box-like (a, b) one, equation (9). We recall that the algebraic and the box-like families tend to a double Dirac-delta peak in the limit $\alpha \rightarrow \infty$ and $b \rightarrow 0^+$, respectively. This implies that we have to impose cutoffs in α and b , respectively, to avoid convergence problems: specifically, we have employed the values $\alpha_{\text{cutoff}} = 4$ and $b_{\text{cutoff}} = 0.2$. In figure 7, we show again that the collapse instability arises when considering the diffusion-rejection estimator $\tilde{\Lambda}_3[\omega]$, and that this instability is fixed by incorporating the CDW basis, i.e. by considering, at least, $\tilde{\Lambda}_4[\omega]$ [36]. When using the diffusion-rejection estimator $\tilde{\Lambda}_3[\omega]$, the optimal parameters in the distributions ω_3^* reach the corresponding cutoffs, thus showing the collapse instability. Going to the estimators $\tilde{\Lambda}_4[\omega]$ or $\tilde{\Lambda}_6[\omega]$ solves this issue again. In fact, the global picture is completely analogous to the one found for the family of Gaussian distributions: (i) in all cases, $\tilde{\Lambda}_3^*$, $\tilde{\Lambda}_4^*$ and $\tilde{\Lambda}_6^*$ are quite close—of course, verifying always $\tilde{\Lambda}_3^* < \tilde{\Lambda}_4^* < \tilde{\Lambda}_6^*$, (ii) as compared with $\tilde{\Lambda}_6^*$, $\tilde{\Lambda}_6[\omega_3^*]$ has a large error that is radically diminished by $\tilde{\Lambda}_6[\omega_4^*]$.

6. On the shape of the optimal jump distribution

In previous sections, we have carried out the minimisation of the next-to-largest eigenvalue over several bi-parametric families of jump distributions, all of which tend to a double Dirac-delta peak in a certain limit. Here, we carry out this minimisation over more complex jump distributions. In this way, we try to shed some light on the global optimisation problem, i.e. given a confining potential $U(x)$, look for a jump distribution $\omega^*(\eta)$ that maximises the convergence rate of the dynamics. We highlight once

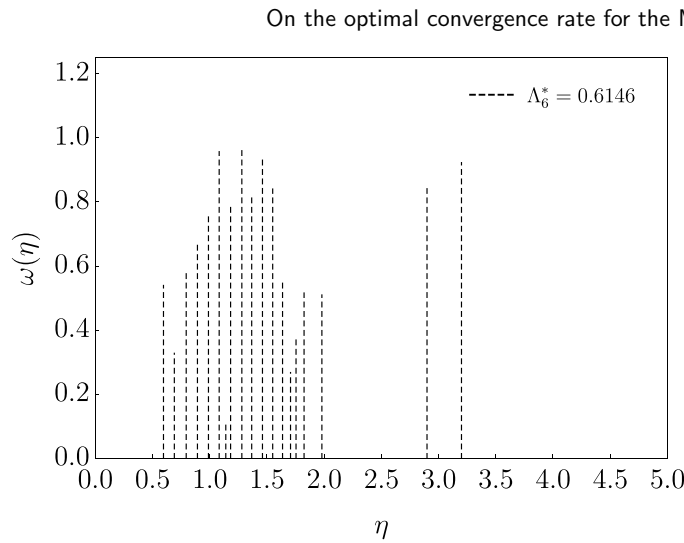


Figure 8. Optimal jump distribution for the discrete-point family of equation (45). Specifically, we have taken $N_P = 20$. The vertical dashed lines give the value of the coefficients a_j of each of the delta peaks at discrete positions η_j . The optimal jump distribution has been obtained using the $\tilde{\Lambda}_6$ functional, and the corresponding estimate $\tilde{\Lambda}_6^*$ for the next-to-largest eigenvalue is shown in the legend.

again that the following analysis is carried out for the case of harmonic confinement $U(x) = \chi x^2/2$. In general, the shape of the optimal jump distribution depends on the form of the confining potential $U(x)$, and the harmonic potential just constitutes the paradigmatic example of a convex potential with a unique global minimum.

6.1. More complex jump distributions

In order to see if we can further improve our estimation of the global minimum value of the next-to-largest eigenvalue, we consider a more complex, discrete-point, family of the form:

$$\omega(\eta) = \sum_{j=1}^{N_P} a_j \delta(|\eta| - \eta_j), \quad (45)$$

where both the weights a_j and the positions η_j of each of the N_P delta peaks are adjustable parameters, over which we carry out the optimisation procedure. Figure 8 shows the optimal jump distribution for the case $N_P = 20$, employing the $\tilde{\Lambda}_6[\omega]$ functional for the next-to-largest eigenvalue. With this approach, we get $\tilde{\Lambda}_6^* = 0.6146$, which is slightly smaller than the previously found values for the two-parametric families in section 5.2. But, more interestingly, although most of the optimal jump distribution is concentrated in a definite interval, $0.5 \lesssim \eta \lesssim 2.0$, there are a couple of delta peaks for larger jumps, around $\eta = 3$. This behaviour hints at the actual optimal jump distribution having a more complicated structure, with at least four peaks—recall that $\omega(\eta)$ is an even function of η , so the two peaks found for $\eta > 0$ are replicated for $\eta < 0$.

On the optimal convergence rate for the Metropolis algorithm in one dimension

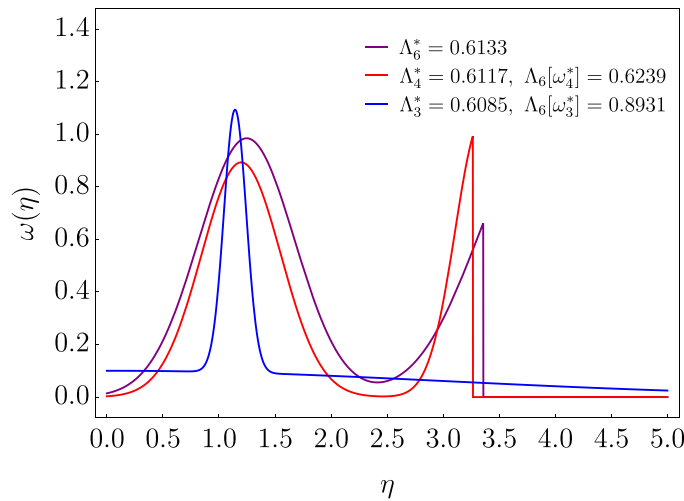


Figure 9. Optimal jump distribution for the 7-parameter family in equation (46). Similarly to the previous figures, the optimal jump distributions have been obtained using three different functionals: $\tilde{\Lambda}_6^*$ (purple), $\tilde{\Lambda}_4^*$ (red), and $\tilde{\Lambda}_3^*$ (blue). In the legend, we present the values of the next-to-largest eigenvalue mentioned in the text. The parameters characterising each jump distribution are shown in table 1.

To check the above insights, we have employed the following two-Gaussian family of jump distributions with a cutoff,

$$\omega(\eta) = C \sum_{n=1}^2 \alpha_n \exp\left(-\frac{(|\eta| - a_n)^2}{2\sigma_n^2}\right) \theta(\eta_{\max} - |\eta|), \quad a_1 < a_2, \quad (46)$$

where C is a normalisation constant. Note that this family has seven parameters: $(a_1, \alpha_1, \sigma_1, a_2, \alpha_2, \sigma_2, \eta_{\max})$. The constants $\alpha_i > 0$ give the relative weights of the four peaks at $\eta = \pm a_1$ and $\eta = \pm a_2$. In addition, this family includes as a particular case the bi-parametric Gaussian family in equation (7), in the limit $\eta_{\max} \rightarrow \infty$ and (e.g.) $\alpha_1 \rightarrow 0^+$.

Figure 9 shows the optimal jump distributions for the 7-parameter family (46), once more employing the functionals $\tilde{\Lambda}_M[\omega]$ with $M = 3, 4$, and 6 . Again, the CDW fixes the collapse instability as shown by the estimate ω_3^* of the diffusion-rejection estimator—note that σ_1 again has reached the cutoff value $\sigma_{\text{cutoff}} = 0.1$. Interestingly, this 7-parameter family slightly improves the optimal next-to-largest eigenvalue: here we have $\tilde{\Lambda}_6^* = 0.6133$, which is smaller than the values for the previous bi-parametric families, displayed in figures 6–7. Apart from the small decrease in the optimal next-to-largest eigenvalue, the optimal jump distributions ω_4^* and ω_6^* have a four-peak structure, analogous to that found in figure 8 for the discrete-point family—recall that we only plot $\omega(\eta)$ for $\eta > 0$, due to its even parity.

The above analysis points out that the global optimal jump distribution may present at least four peaks, corresponding to jumps centered around $\pm a_1$ and $\pm a_2$, instead of the two peaks at $\pm a$ we have considered in the previous sections.

Table 1. Parameters characterising the optimal jump distributions for the 7-parameter family in equation (46). As in previous cases, optimisation is carried out for the functionals $\hat{\Lambda}_M[\omega]$, with $M = 3, 4$ and 6.

	η_{\max}	α_1	a_1	σ_1	α_2	a_2	σ_2
ω_6^*	3.358	0.985	1.247	0.428	0.986	3.845	0.546
ω_4^*	3.265	0.893	1.198	0.354	1.050	3.343	0.235
ω_3^*	5.764	1.000	1.144	0.100	0.100	0.074	2.950

6.2. Sequence of jumps

The fact that four-peak jump distributions improve the convergence rate as compared to two-peak jump distributions, uncovered in section 6.1, also provides insights into the design of other strategies for further optimising the convergence rate. As an example of such a possible strategy, we consider the deterministic alternation between two jump distributions $\omega_1(\eta)$ and $\omega_2(\eta)$ with peaks at $\eta = \pm a_1$ and $\eta = \pm a_2$, respectively.

The corresponding kernel $\hat{K}(\omega_1, \omega_2)$ for the master equation of this modified Metropolis algorithm is

$$\hat{K}(\omega_1, \omega_2) = \hat{K}(\omega_1) \hat{K}(\omega_2) \hat{K}(\omega_2) \hat{K}(\omega_1), \quad (47)$$

with $\hat{K}(\omega_i)$ being the kernel of our original Metropolis algorithm from equation (11), with a fixed choice $\omega_i(\eta)$ for the jump distribution. We employ parenthesis instead of the brackets that appear in equation (11), in order to highlight that ω is no longer a variational parameter. This new kernel describes a modified Metropolis algorithm in which two jumps from ω_1 are attempted, then two from ω_2 are attempted, and the process repeats. The sequence has been chosen so that the kernel form of equation (47) remains symmetric.

In the following, we numerically compare the efficiency of this strategy against the original Metropolis algorithm studied in this work, but with the average jump distribution $\omega = (\omega_1 + \omega_2)/2$. Now, as a consequence of the structure of the sequential kernel in equation (47), the corresponding next-to-largest eigenvalue for the modified Metropolis algorithm is $\Lambda_{\text{seq}} = (\Lambda_1 \Lambda_2)^{1/2}$, with Λ_i being the next-to-largest eigenvalue for the jump distribution ω_i . As shown in figure 10, the sequence of jumps improves the convergence time of the original algorithm, by lowering the next-to-largest eigenvalue from $\Lambda_6^* = 0.6133$ to $\Lambda_{\text{seq}} = 0.5840$. Since the convergence rate is given by $\gamma = |\ln \Lambda|$, as discussed in section 2.2, this entails that the actual speed-up obtained with a sequence of jumps is about 10%. We remark that this improvement is only for a reduced number of degrees of freedom: in this case, one standard deviation and two mean values for the Gaussian ansatz. Increasing the dimension of the parameter space, or even combining different families of jump distributions—such as the ones employed before, given by equations (7)–(9)—should further improve the optimisation of the convergence rate.

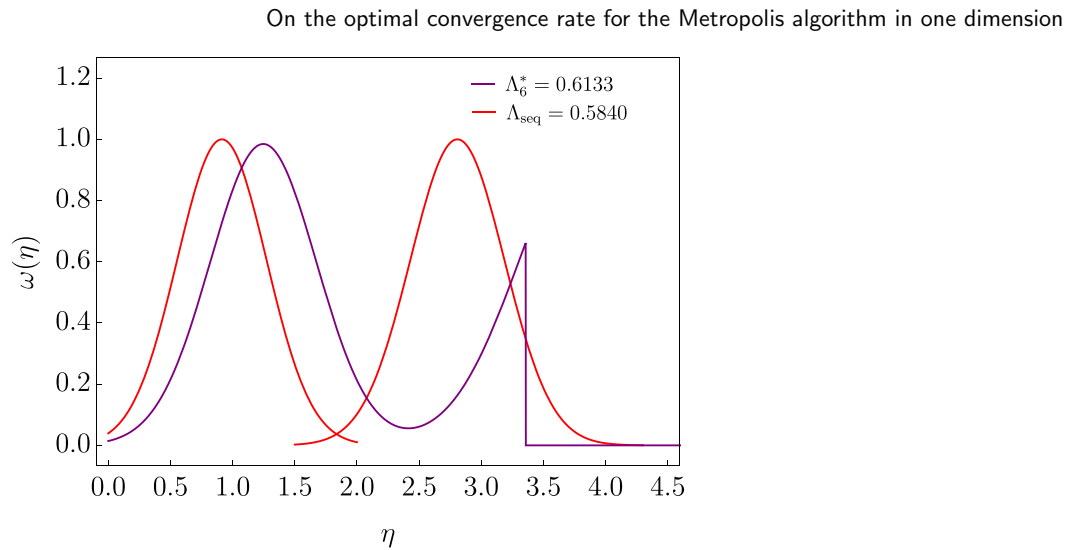


Figure 10. Comparison of optimal jump distributions. Specifically, we show the optimal jump distributions for (i) the 7-parametric jump distribution in equation (46), already shown in figure 9 (purple), and (red) the sequence of jumps described in the text, alternating two Gaussian distributions with the same standard deviation but different mean value. The corresponding optimal next-to-largest eigenvalues mentioned in the text are also plotted in the legend.

We may understand the better performance of the sequence-of-jumps strategy over the original Metropolis scheme with the following intuitive argument. Let us take the characteristic lengths $a_1 < a_2$, with a sufficient gap between the two. Then, for $\omega_1(\eta)$, the first two steps of the algorithm will resemble that of free diffusion, which allows exploration of small regions of the phase space. In order to move further, we take the second two steps, which allow one to jump to further points in space, and thus the rejection rate increases accordingly. Thus, each distribution balances the problems presented for the other.

The previous argument has remarkable implications in terms of choosing between stochastic and deterministic jump strategies. On the one hand, the former would correspond to our original Metropolis algorithm: if we consider the four-peak jump distribution $\omega(\eta)$ as the average of two independent distributions $\omega_1(\eta)$ and $\omega_2(\eta)$ with different characteristic lengths, $\omega = (\omega_1 + \omega_2)/2$, we may understand the algorithm as choosing randomly between these two distributions before attempting any jumps. On the other hand, the deterministic approach would correspond to the sequence of jumps, since we clearly set the order in which we draw the attempted jumps from ω_1 and ω_2 .

7. Conclusions

In this paper, we have analysed the relaxation dynamics of the Metropolis algorithm for a particle confined in a one-dimensional potential, with even jump distributions that ensure that the dynamics verify detailed balance. The central result of this paper is the functional optimisation of the convergence rate of the Metropolis algorithm in one

dimension, which is obtained as a balance of three mechanisms: diffusion, rejection, and the novel CDW regime that we unveil and characterise in this paper. In the CDW regime, the leading eigenfunction $\phi_1(x)$ corresponding to the next-to-largest eigenvalue $\Lambda \equiv \lambda_1$ —we recall that the largest eigenvalue is the equilibrium one $\lambda_0 = 1$ —shows an oscillatory behaviour, characterised by a wave number k . For two-parametric jump distributions, the balance between diffusion, rejection, and CDW occurs at the tricritical point of coexistence—at which the minimum value of the Λ , i.e. the maximum value of the convergence rate $\gamma = |\ln \Lambda|$ is attained.

The first of our core results constitutes completing the physical picture of the dominant mechanisms driving the relaxation dynamics of the Metropolis algorithm. The balance between diffusion and rejection makes it possible to obtain the optimal convergence rate for several choices of confining potentials and jump distributions [18], but it does not explain the eigenvalue crossing taking place for sufficiently narrow jump distributions. Our numerical and analytical approaches show that such crossing points out toward a first-order phase transition between the diffusion and the—newly characterised—CDW regime, which involves a change of symmetry in the dominant eigenmodes.

This interplay between the aforementioned physical factors has been illustrated by the phase diagram for the next-to-largest eigenvalue Λ . Specifically, this phase diagram has been built for the algebraic (a, α) family of jump distributions of equation (8). The different phases are characterised by two order parameters: fidelity \mathcal{F} , which is different from zero when the new CDW regime dominates, and IPR, which is different from zero when the rejection regime dominates. First, for smooth jump distributions without a double-peak structure, we have a localisation transition between the diffusion and rejection regions as the length of the typical jumps is increased—which was characterised in [18]. Then, we have unveiled a new transition between the diffusion and CDW regimes in the parameter range for which the jump distribution approaches a double peak structure, also as the length of the typical jumps is increased. The diffusion-CDW transition is first-order so that (i) \mathcal{F} and the optimal wave number abruptly change, and (ii) the parity of the leading eigenfunction of the master equation changes—from odd to even. Finally, it is only when the length of the typical jumps is further increased that we find a transition between the CDW and rejection regimes. At this transition, both order parameters vary smoothly, so the CDW-rejection transition is at least second order.

As our second major result, we have searched for the functional optimal jump distribution (meaning that no *a priori* assumption is made on the functional shape), so that the relaxation dynamics is the shortest possible. We have found the corresponding jump distribution to present four peaks, rather than two as studied above. Besides, we have shown numerical evidence on the improvement of the convergence rate by combining different jump distributions. The sequence of four jumps exposed in this work is just an example of how the interplay between more than one characteristic length may improve the relaxation to equilibrium. Interestingly, a similar improvement of other figures of merit when considering more than one characteristic length has been observed in other physical systems—e.g. for optimal search strategies and resetting processes [31, 37, 38].

The variational method employed to estimate the optimal convergence rate is quite general and could, in principle, be extended to study optimisation problems in higher

dimensions or in systems with multiple minima. However, it is important to consider the computational challenges associated with such extensions. Specifically, (i) the number of integrals in the master equation increases with each additional dimension, significantly raising computational complexity; (ii) the convergence of the method may be compromised, as a higher number of eigenmodes from the truncated basis would likely be required to achieve accurate estimations in higher-dimensional systems; and (iii) for potentials with multiple minima, a larger and more diverse basis set might be necessary to adequately capture the system's behaviour. Despite these challenges, such generalisations represent a promising avenue for future work.

In this paper, we have completely characterised the CDW regime in the one-dimensional case. Therein, an isotropic jump distribution with a fixed jump size, i.e. the Dirac-delta jump distribution in equation (10), does not allow the system to explore the whole configuration space. It is this ergodicity—or irreducibility—breaking that gives rise to the CDW behaviour. Thus, the natural question is whether or not the CDW regime would play a role in the global optimisation of the Metropolis algorithm in higher dimensions. Intuitively, it seems plausible that the role of the CDW regime in the optimisation problem may be less relevant in higher-dimensional systems. This stems from isotropic jump distributions with a radial Dirac-delta component not inherently breaking ergodicity: the whole configuration space can be explored with a sequence of fixed-size jumps—already in two dimensions. Consequently, the influence of the CDW regime on the convergence properties of the Metropolis algorithm might diminish with dimensionality. Nonetheless, a detailed investigation is required to check the accuracy of the intuitive picture above and to determine whether or not higher-dimensional analogues of the CDW regime can emerge under specific conditions.

Additionally, it is relevant to investigate the optimisation of the convergence rate for other dynamics, different from the Metropolis algorithm—such as the Glauber algorithm for simulating the Ising Model [39], or the Simulated Annealing techniques [40–42], which incorporate the Metropolis algorithm with a time-dependent cooling program for the temperature T in order to attain the global minimum of a function. We would like to stress that our analysis has focused on the leading eigenvalue λ_1 , which as a rule controls the convergence rate to equilibrium. However, there are certain specific situations in which the convergence rate to equilibrium is controlled by the following eigenvalue $\lambda_2 < \lambda_1$, because the coefficient of the first mode vanishes, as in the strong version of the Mpemba effect [27, 29, 43–48] or when specific strategies are employed to tune the dynamical evolution of the system [49, 50]. Therefore, it is also worth investigating whether the general picture found here for λ_1 , i.e. the competition between diffusion, rejection, and CDW regimes, extends to λ_2 or even to the full eigenvalue spectrum of the master equation. Certainly, all these questions open the door to future research.

Data availability statement

The codes employed for generating the data that support the findings of this study, together with the Mathematica notebooks employed for producing the figures presented in the paper, are openly available in the https://github.com/fine-group-us/optimal_convergence_rate_metropolis of University of Sevilla's FINE research group.

Acknowledgments

We thank Satya Majumdar for fruitful discussions. APR and APA acknowledge financial support from Grant PID2021-122588NB-I00 funded by MCIN/AEI/10.13039/501100011033/ and by ‘ERDF A way of making Europe’. APR, APA, and ET acknowledge financial support from Grant ProyExcel_00796 funded by Junta de Andalucía’s PAIDI 2020 programme. APA also acknowledges support from the FPU programme through Grant FPU2019-4110, and additional support from the FPU programme through Grant EST22/00346, which funded his research stay at Univ. Paris-Saclay during autumn 2022. APR also acknowledges the hospitality of LPTMS, which funded his stay at Univ. Paris-Saclay in June 2022. AC acknowledges financial support from Grant ANR-20-CE92-0041 (MARS).

Appendix A. Analytical derivation of the CDW eigenfunctions

We start by analysing the second term on the lhs of equation (15), which involves the rejection probability defined in equation (4). By defining $\ell(x)$ and $h(x)$ as the x' -values at which $U(x') = U(x)$, the Heaviside function restricts the integral over x' to a certain interval $(-\infty, \ell(x)) \cup (h(x), +\infty)$. Depending on the sign of $U'(x)$, either $\ell(x)$ or $h(x)$ equals x —in the following we focus on the latter case, $U'(x) > 0$ ($x > 0$), where $h(x) = x$. With these definitions,

$$R(x; a, \sigma) = \left(\int_{-\infty}^{\ell(x)} + \int_x^{+\infty} \right) dx' \omega(x' - x; a, \sigma) \left[1 - e^{-\beta \Delta U(x', x)} \right]. \quad (\text{A1})$$

The first integral is subdominant, since ω is basically zero over the considered interval, as long as $\ell(x) - x < -a$, i.e. $x - a > \ell(x)$. On the other hand, in the second integral we have that $x' - x > 0$ and therefore only one of the δ_σ ’s contribute:

$$R(x; a, \sigma) \approx \frac{1}{2} \int_x^{+\infty} dx' \delta_\sigma(x' - x - a) \left[1 - e^{-\beta \Delta U(x', x)} \right] \approx \frac{1}{2} \left[1 - e^{-\beta \Delta U(x+a, x)} \right] \approx \frac{1}{2}, \quad (\text{A2})$$

assuming that

$$e^{-\beta \Delta U(x+a, x)} \ll 1, \quad (\text{A3})$$

which is analogous to equation (27) from the main text. This condition implies that a is large enough, see appendix B for more details. In fact, the relationship between equations (A2) and (A3) follows from the intuitive physical arguments presented in section 3.1: If a is sufficiently large and x not too close to zero—i.e. the global minimum, such that the potential energy increases significantly, then most of the jumps to the right will be rejected, while those to the left are basically accepted. Thus, the overall rejection probability is close to $1/2$. Now we move on to analysing the first term on the lhs of equation (15), splitting the integral into two slices: one in which $U(x) > U(x')$, i.e. the

On the optimal convergence rate for the Metropolis algorithm in one dimension

interval $(\ell(x), x)$ and another in which $U(x) < U(x')$, i.e. $(-\infty, \ell(x)) \cup (x, +\infty)$. For the first slice, $U(x) > U(x')$, we have

$$\begin{aligned}
 & \int_{-\infty}^{+\infty} dx' \theta(\Delta U(x, x')) \omega(x - x'; a, \sigma) e^{-\beta|\Delta U(x, x')|/2} \rho_\nu(x') e^{-\beta U(x')/2} \\
 &= \int_{\ell(x)}^x dx' \omega(x - x'; a, \sigma) e^{-\beta(U(x) - U(x'))/2} \rho_\nu(x') e^{-\beta U(x')/2} \\
 &\approx \frac{1}{2} e^{-\beta U(x)/2} \int_{\ell(x)}^x dx' \delta_\sigma(x - x' - a) \rho_\nu(x') \\
 &\approx \frac{1}{2} e^{-\beta U(x)/2} \int_{-\infty}^{+\infty} dx' \delta_\sigma(x - x' - a) \rho_\nu(x') \\
 &= \frac{e^{-\beta U(x)/2}}{2} \int_{-\infty}^{+\infty} dx' \delta_\sigma(x - x') \rho_\nu(x' - a) \\
 &\approx \frac{e^{-\beta U(x)/2}}{2} \int_{-\infty}^{+\infty} dx' \delta_\sigma(x - x') \left[\rho_\nu(x - a) + \frac{1}{2} \rho_\nu''(x - a)(x - x')^2 \right] \\
 &= \frac{1}{2} e^{-\beta U(x)/2} \left[\rho_\nu(x - a) + \frac{\sigma^2}{2} \rho_\nu''(x - a) \right], \tag{A4}
 \end{aligned}$$

where we have taken into account again that (i) only one of the δ_σ 's contribute, and (ii) the interval of integration may be extended from $(\ell(x), x)$ to $(-\infty, +\infty)$, once the peak of the δ_σ is included in the considered interval—again, as long as $x - a > \ell(x)$. For the second slice, $U(x') > U(x)$, we have

$$\begin{aligned}
 & \int_{-\infty}^{+\infty} dx' \theta(\Delta U(x', x)) \omega(x - x'; a, \sigma) e^{-\beta|\Delta U(x, x')|/2} \rho_\nu(x') e^{-\beta U(x')/2} \\
 &= \left(\int_{-\infty}^{\ell(x)} + \int_x^{+\infty} \right) dx' \omega(x - x'; a, \sigma) e^{-\beta \Delta U(x', x)/2} \rho_\nu(x') e^{-\beta U(x')/2} \\
 &\approx \frac{e^{\beta U(x)/2}}{2} \int_x^{+\infty} dx' \delta_\sigma(x + a - x') \rho_\nu(x') e^{-\beta U(x')} \\
 &\approx \frac{e^{\beta U(x)/2}}{2} \int_{-\infty}^{+\infty} dx' \delta_\sigma(x + a - x') \rho_\nu(x') e^{-\beta U(x')} \\
 &\approx e^{-\beta \Delta U(x+a, x)} \frac{1}{2} e^{-\beta U(x)/2} \rho_\nu(x + a) + O(\sigma^2). \tag{A5}
 \end{aligned}$$

Note that equation (A5) is subdominant against equation (A4), due to the presence of the exponential factor (A3), so it is neglected in the following. The above analysis implies that the eigenvalue equation (15) can be approximated by

$$\frac{1}{2} e^{-\beta U(x)/2} \left[\rho_\nu(x - a) + \frac{\sigma^2}{2} \rho_\nu''(x - a) \right] + \frac{1}{2} e^{-\beta U(x)/2} \rho_\nu(x) = \lambda_\nu e^{-\beta U(x)/2} \rho_\nu(x), \tag{A6}$$

where $U'(x) > 0$, i.e. $x > 0$, provided that $x - a > \ell(x)$. A similar analysis for $U'(x) < 0$ leads to a completely analogous expression for $x < 0$, with the change $a \rightarrow -a$, provided that $x + a < h(x)$ —note that $\ell(x) = x$ for $x < 0$.

Summarising our findings, in the limit $\sigma \ll a$ we have the approximate equations

$$\frac{\sigma^2}{4} \rho_\nu''(x - a) + \frac{1}{2} [\rho_\nu(x - a) + \rho_\nu(x)] = \lambda_\nu \rho_\nu(x), \quad x > 0, x - a > \ell(x); \quad (\text{A7a})$$

$$\frac{\sigma^2}{4} \rho_\nu''(x + a) + \frac{1}{2} [\rho_\nu(x + a) + \rho_\nu(x)] = \lambda_\nu \rho_\nu(x), \quad x < 0, x + a < h(x). \quad (\text{A7b})$$

For even potentials, $\ell(x) = -x < 0$ for $x > 0$ ($h(x) = -x > 0$ for $x < 0$). Then, for even potentials, equation (A7a) (equation (A7b)) is limited to $x > a/2$ ($x < -a/2$). A more careful analysis would be needed in the region $-a/2 < x < a/2$ for even potentials—and a certain region of small x for potentials without symmetry. This could have been anticipated after the approximations made in equation (A4): whereas the starting expression clearly vanishes in the limit as $x \rightarrow 0$, the final approximate expression does not.

As a check of the consistency of equations (A7), note that a uniform solution corresponds to the equilibrium eigenvalue $\lambda_0 = 1$. Moreover, as mentioned in section 3 from the main text, for infinitely sharp jump distributions, i.e. in the case $\sigma = 0$, the Metropolis algorithm loses ergodicity and any function $\phi_\nu(x)$ of the form (26) with $\rho_\nu(x)$ being periodic with period a is expected to be a stationary solution of the master equation, i.e. an eigenfunction corresponding to the eigenvalue $\lambda_0 = 1$. Note that this is indeed predicted by equations (A7), as setting $\lambda_\nu = 1$ for $\sigma = 0$ yields $\rho_\nu(x \pm a) = \rho_\nu(x)$. Interestingly, equations (A7) still admit spatially periodic solutions, with period a , for $\sigma \neq 0$. Hence, by introducing the CDW ansatz (28), both equations (A7) simplify to

$$\rho_\nu''(x) + \frac{4(1 - \lambda_\nu)}{\sigma^2} \rho_\nu(x) = 0, \quad \forall x, \quad (\text{A8})$$

corresponding to equation (29) from the main text.

Appendix B. Robustness of the CDW eigenfunctions

We stress the main assumptions employed above to derive the novel CDW regime. First, we have assumed that the potential is convex, $U''(x) > 0$, so that it has only one minimum. As already stated, this assumption has been introduced to simplify the presentation but it is not essential: the result applies to a generic confining potential in one dimension, which may have multiple minima—the proof is lengthy and this is the reason why we do not present it here. The spirit of the proof for a general potential $U(x)$ lies in the fact that, beyond the lengths $l(x)$ and $h(x)$, additional x' -values satisfying $U(x') = U(x)$ may arise due to the presence of multiple minima. In such a scenario, the integrals over x' in equation (15) would be split into several intervals. Depending on the sign of $U'(x)$ and the position x , either the leftmost or rightmost intervals could be handled using the same reasoning as in appendix A. By carefully selecting assumptions analogous to those in equations (27) and (A3), the additional integral terms could be

neglected in favour of the dominant ones, ultimately leading to equation (29). However, these additional assumptions would further restrict the range of a and x values that support the CDW regime. The latter analysis is highly dependent on the explicit form of $U(x)$ and the number of minima, which justifies our choosing of a convex potential with a unique global minimum as a paradigmatic example to highlight the main feature behind the emergence of the CDW regime.

Second, we have assumed that, in the regions where $U' > 0$, the potential increases sufficiently fast, as expressed by equation (A3). Although this assumption might seem rather restrictive, we have shown in section 3.2, for the paradigmatic example of harmonic confinement, that it is not so. As is often the case with singular perturbation schemes, the small parameters do not need to be very small for their theoretical predictions be accurate.

First, it must be stressed that equation (A7a) is only valid for $x > a/2$ and equation (A7b) for $x < -a/2$, and we have extended the equation to all x . Then, we have that a must be ‘small’ enough for the CDW to emerge. On the other hand, equation (A3) entails that a must be ‘large’ enough. Since $U''(x) > 0$, for $x > 0$ one has

$$1 \gg e^{-\beta \Delta U(x+a,x)} = e^{-\beta a U'(\tilde{x})} > e^{-\beta a U'(x+a)}, \quad (\text{B1})$$

where $\tilde{x} \in [x, x+a]$. The convexity condition implies that it is enough to ensure the above relation for $x=0$, i.e. it is enough to have that $e^{-\beta a U'(a)} \ll 1$. For example, in the case of the harmonic potential, this is equivalent to having $e^{-\beta \chi a^2} \ll 1$.

Second, we may consider the higher-order terms in the width of the distribution that have been neglected in equation (A4). This would result in the inclusion of the $O(\sigma^4)$ contribution into the term in brackets in equation (A6), i.e.

$$\begin{aligned} & \frac{1}{2} e^{-\beta U(x)/2} \left[\rho_\nu(x-a) + \frac{\sigma^2}{2} \rho_\nu''(x-a) + \xi \frac{\sigma^4}{4!} \rho_\nu''''(x-a) \right] + \frac{1}{2} e^{-\beta U(x)/2} \rho_\nu(x) \\ & = \lambda_\nu e^{-\beta U(x)/2} \rho_\nu(x), \end{aligned} \quad (\text{B2})$$

where ξ is a order of unity parameter, which gives the ratio between the fourth central moment of δ_σ and its variance squared, i.e.

$$\int_{-\infty}^{+\infty} dz z^4 \delta_\sigma(z) \equiv \xi \sigma^4. \quad (\text{B3})$$

Therefore, equations (A7) would be generalised to

$$\xi \frac{\sigma^4}{48} \rho_\nu''''(x-a) + \frac{\sigma^2}{4} \rho_\nu''(x-a) + \frac{1}{2} [\rho_\nu(x-a) + \rho_\nu(x)] = \lambda_\nu \rho_\nu(x), \quad x > a/2; \quad (\text{B4a})$$

$$\xi \frac{\sigma^4}{48} \rho_\nu''''(x+a) + \frac{\sigma^2}{4} \rho_\nu''(x+a) + \frac{1}{2} [\rho_\nu(x+a) + \rho_\nu(x)] = \lambda_\nu \rho_\nu(x), \quad x < -a/2. \quad (\text{B4b})$$

Along the same line of reasoning followed before, it is shown that we have periodic solutions with period a . It is only the dispersion relation that changes, i.e. equation (31) holds but now $\lambda_\nu = 1 - \sigma^2 k_\nu^2/4 + \xi \sigma^4 k_\nu^4/48$. See also appendix C for a further check of the robustness of the emergence of periodic solutions.

Appendix C. Ergodicity breaking and periodic solutions

In this section, we provide more details on the emergence of periodic solutions in the limit $\sigma \rightarrow 0^+$. In this limit, the jump distribution reduces to the double Dirac-delta peak in equation (10). To be consistent with the discussion on the robustness of the CDW modes from section 3.2, we restrict ourselves to even potentials $U(x)$, although a similar analysis can be carried out for a non-symmetric potential.

For $\sigma = 0$, the eigenvalue/eigenfunction equation for the modes of the master equation (15) reduces exactly to

$$e^{-\beta|\Delta U(x,x-a)|/2}\phi_\nu(x-a) + e^{-\beta|\Delta U(x,x+a)|/2}\phi_\nu(x+a) + 2 - p(x-a,x) - p(x+a,x) = 2\lambda_\nu\phi_\nu(x), \quad \forall x. \quad (\text{C1})$$

It is clearer to split the above equation into three intervals, (i) $x > a/2$, (ii) $-a/2 < x < a/2$, and (iii) $x < -a/2$. In this way, $U(x) - U(x \pm a)$ has a definite sign on each interval, and we get

$$e^{\beta U(x-a)/2}\phi_\nu(x-a) + e^{-\beta\Delta U(x+a,x)} \left[e^{\beta U(x+a)/2}\phi_\nu(x+a) - e^{\beta U(x)/2}\phi_\nu(x) \right] = (2\lambda_\nu - 1)e^{\beta U(x)/2}\phi_\nu(x), \quad x > a/2, \quad (\text{C2a})$$

$$e^{-\beta\Delta U(x+a,x)} \left[e^{\beta U(x+a)/2}\phi_\nu(x+a) - e^{\beta U(x)/2}\phi_\nu(x) \right] \quad (\text{C2b})$$

$$+ e^{-\beta\Delta U(x-a,x)} \left[e^{\beta U(x-a)/2}\phi_\nu(x-a) - e^{\beta U(x)/2}\phi_\nu(x) \right] = 2(\lambda_\nu - 1)e^{\beta U(x)/2}\phi_\nu(x), \quad -a/2 < x < a/2, \quad (\text{C2c})$$

$$e^{\beta U(x+a)/2}\phi_\nu(x+a) + e^{-\beta\Delta U(x-a,x)} \left[e^{\beta U(x-a)/2}\phi_\nu(x-a) - e^{\beta U(x)/2}\phi_\nu(x) \right] = (2\lambda_\nu - 1)e^{\beta U(x)/2}\phi_\nu(x), \quad x < -a/2.$$

Indeed, the structure of this equation suggests the change of variables in equation (26). Introducing it into equation (C2),

$$\rho_\nu(x-a) + e^{-\beta\Delta U(x+a,x)} [\rho_\nu(x+a) - \rho_\nu(x)] = (2\lambda_\nu - 1)\rho_\nu(x), \quad x > a/2, \quad (\text{C3a})$$

$$e^{-\beta\Delta U(x+a,x)} [\rho_\nu(x+a) - \rho_\nu(x)] + e^{-\beta\Delta U(x-a,x)} [\rho_\nu(x-a) - \rho_\nu(x)] = 2(\lambda_\nu - 1)\rho_\nu(x), \quad -a/2 < x < a/2, \quad (\text{C3b})$$

$$\rho_\nu(x+a) + e^{-\beta\Delta U(x-a,x)} [\rho_\nu(x-a) - \rho_\nu(x)] = (2\lambda_\nu - 1)\rho_\nu(x), \quad x < -a/2. \quad (\text{C3c})$$

It is worth stressing that equation (C3) for $\rho_\nu(x)$ is exact for $\sigma = 0$. Interestingly, any periodic function $\rho_\nu(x)$ with period a , i.e. $\rho_\nu(x \pm a) = \rho_\nu(x) \forall x$, satisfies this equation with eigenvalue $\lambda_\nu = 1$. As already discussed in the main text, this is intimately related to the breaking of ergodicity of the one-dimensional Monte Carlo algorithm for $\sigma = 0$: any point x is only connected to $x_n = x \pm na$, $n \in \mathbb{N}$, and therefore the probability density of this lattice becomes distributed following the Boltzmann weight, i.e. it is

proportional to $\exp(-\beta U(x_n))$, but the lattices corresponding to different values of x are disconnected and the system does not reach equilibrium.

The consideration of a small width σ makes it possible to introduce an expansion in powers of σ^2 around equation (C3), as carried out in appendices A and B—see equations (A7) and (B4). Therein, we showed that these additional terms still allow for the emergence of periodic solutions with period a , but restrict their shape to the cosine and sine CDW modes in equation (30). Note that the shape of the CDW modes so obtained does not depend on the potential $U(x)$, due to our assumption (A3) that allowed us to neglect the terms proportional to $e^{-\beta \Delta U(x \pm a, x)}$ in equations (C3a) and (C3c), and analogous terms in the higher-order terms in our expansion in powers of σ^2 .

References

- [1] Frenkel D and Smit B 2023 *Understanding Molecular Simulation: From Algorithms to Applications* (Elsevier Science)
- [2] Newman M and Barkema G 1999 *Monte Carlo Methods In Statistical Physics* (Clarendon)
- [3] Mode C J 2011 Applications of Monte Carlo methods in biology *Med. Other Fields Sci.*
- [4] Glasserman P 2004 Monte Carlo methods in financial engineering *Applications of Mathematics: Stochastic Modelling and Applied Probability* (Springer)
- [5] Bishop C 2006 *Pattern Recognition and Machine Learning* (Springer)
- [6] Gilks W, Richardson S and Spiegelhalter D 1995 *Markov Chain Monte Carlo in Practice* (Chapman & Hall/CRC Interdisciplinary Statistics) (Taylor & Francis)
- [7] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 Equation of state calculations by fast computing machines *J. Chem. Phys.* **21** 1087
- [8] Gardiner C 2009 *Stochastic Methods: A Handbook for the Natural and Social Sciences* (Springer Series in Synergetics) (Springer)
- [9] Van Kampen N G 1992 *Stochastic Processes in Physics and Chemistry* ((North-Holland)
- [10] In the physical literature, ergodicity is often identified with irreducibility; the stochastic matrix of a Markov chain is said to be irreducible if any final state can be reached from an arbitrary initial state by means of a chain of transitions with non-zero probability.
- [11] Gelman A, Roberts G O and Gilks W R 1996 Efficient Metropolis Jumping Rules *Bayesian statistics 5: Proc. Fifth Valencia International Meeting* (Oxford University Press)
- [12] Gelman A, Gilks W R and Roberts G O 1997 Weak convergence and optimal scaling of random walk Metropolis algorithms *Ann. Appl. Probab.* **7** 110
- [13] Schmon S M and Gagnon P 2022 Optimal scaling of random walk Metropolis algorithms using Bayesian large-sample asymptotics *Stat. Comput.* **32** 28
- [14] Krauth W 2006 Statistical mechanics: algorithms and computations *Oxford Master Series in Physics* (Oxford University Press)
- [15] Talbot J, Tarjus G and Viot P 2003 Optimum Monte Carlo simulations: some exact results *J. Phys. A: Math. Gen.* **36** 9009
- [16] Guéry-Odelin D, Ruschhaupt A, Kiely A, Torrontegui E, Martínez-Garaot S and Muga J G 2019 Shortcuts to adiabaticity: concepts, methods and applications *Rev. Mod. Phys.* **91** 045001
- [17] Guéry-Odelin D, Jarzynski C, Plata C A, Prados A and Trizac E 2023 Driving rapidly while remaining in control: classical shortcuts from Hamiltonian to stochastic dynamics *Rep. Prog. Phys.* **86** 035902
- [18] Chepelianskii A D, Majumdar S N, Schawe H and Trizac E 2023 Metropolis Monte Carlo sampling: convergence, localization transition and optimality *J. Stat. Mech.* **123205**
- [19] Boscain U, Chittaro F C, Mason P, Pacqueau R and Sigalotti M 2010 *Motion Planning in Quantum Control via Intersection of Eigenvalues (49th IEEE Conf. on Decision and Control (CDC))* (IEEE) pp 3028–33
- [20] Wang Z H, Zheng Q, Wang X and Li Y 2016 The energy-level crossing behavior and quantum Fisher information in a quantum well with spin-orbit coupling *Sci. Rep.* **6** 22347
- [21] Teza G, Yaacoby R and Raz O 2023 Eigenvalue crossing as a phase transition in relaxation dynamics *Phys. Rev. Lett.* **130** 207103

- [22] Gruner G and Zettl A 1985 Charge density wave conduction: a novel collective transport phenomenon in solids *Phys. Rep.* **119** 117
- [23] Gruner G 1988 The dynamics of charge-density waves *Rev. Modern Phys.* <https://doi.org/10.1103/RevModPhys.60.1129> **60** <https://doi.org/10.1103/RevModPhys.60.1129>
- [24] Chen C-W, Choe J and Morosan E 2016 Charge density waves in strongly correlated electron systems *Rep. Prog. Phys.* **79** 084505
- [25] Bapat R B and Raghavan T E S 1997 Nonnegative matrices and applications *Encyclopedia of Mathematics and its Applications* (Cambridge University Press)
- [26] To simplify our notation, we do not show the explicit dependence on the temperature. In simple cases like the harmonic potential, the temperature can be absorbed in the definition of the unit of length
- [27] Lu Z and Raz O 2017 Nonequilibrium thermodynamics of the Markovian Mpemba effect and its inverse *Proc. Natl Acad. Sci.* **114** 5083
- [28] Lasanta A, Vega Reyes F, Prados A and Santos A 2017 When the hotter cools more quickly: Mpemba effect in granular fluids *Phys. Rev. Lett.* **119** 148001
- [29] Patrón A, Sánchez-Rey B, Plata C A and Prados A 2023 Non-equilibrium memory effects: granular fluids and beyond *Europhys. Lett.* **143** 61002
- [30] Peierls R E 1955 *Quantum Theory of Solids* (Oxford University Press) p 108
- [31] Bénichou O, Loverdo C, Moreau M and Voituriez R 2011 Intermittent search strategies *Rev. Mod. Phys.* **83** 81
- [32] The following numerical procedure has been employed for all the examples presented throughout the paper. To obtain the eigenvalues and eigenfunctions of the master equation, we have employed the FORTRAN 90 subroutine dsyev.f90 from LAPACK [51]. To avoid spurious boundary effects, the spatial variable x is restricted to a finite interval $[-L, L]$, with a sufficiently large L . Moreover, space is discretised with a mesh of N_d points. For the harmonic potential, we have fixed units by choosing $k_B T = 1$ and $\chi = 1$, and taken $L = 10$ and $N_d = 2000$.
- [33] A similar issue was found in [18] for the Schrödinger eigenvalues.
- [34] To numerically estimate Λ , as defined in equation (41), we have employed the open source nonlinear optimisation libraries from *NLOpt* [52] with $N_s = 100$ modes for each basis.
- [35] Since the numerics involve a spatial discretisation, σ must be larger than the lattice spacing in this discretisation.
- [36] Of course, more complex functionals like $\hat{\Lambda}_6[\omega]$, which also include the CDW basis, also fix the instability and improve the estimate for the next-to-largest eigenvalue.
- [37] Evans M R, Majumdar S N and Schehr G 2020 Stochastic resetting and applications *J. Phys. A: Math. Theor.* **53** 193001
- [38] García-Valladares G, Plata C A, Prados A and Manacorda A 2023 Optimal resetting strategies for search processes in heterogeneous environments *New J. Phys.* **25** 113031
- [39] Glauber R J 1963 Time-dependent statistics of the Ising model *J. Math. Phys.* **4** 294
- [40] Kirkpatrick S, Gelatt C D and Vecchi M P 1983 Optimization by simulated annealing *Science* **220** 671
- [41] Aarts E and Korst J 1989 *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing* (Wiley Series in Discrete Mathematics) (Optimization Wiley)
- [42] Van Laarhoven P J M and Aarts E H L 1987 *Simulated Annealing: Theory and Applications* (Springer)
- [43] Klich I, Raz O, Hirschberg O and Vucelja M 2019 Mpemba index and anomalous relaxation *Phys. Rev. X* **9** 021060
- [44] Kumar A and Bechhoefer J 2020 Exponentially faster cooling in a colloidal system *Nature* **584** 64
- [45] Busiello D M, Gupta D and Maritan A 2021 Inducing and optimizing Markovian Mpemba effect with stochastic reset *New J. Phys.* **23** 103012
- [46] Kumar A, Chétrite R and Bechhoefer J 2022 Anomalous heating in a colloidal system *Proc. Natl Acad. Sci.* **119** e2118484119
- [47] Zhang S and Hou J-X 2022 Theoretical model for the Mpemba effect through the canonical first-order phase transition *Phys. Rev. E* **106** 034131
- [48] Biswas A, Rajesh R and Pal A 2023 Mpemba effect in a Langevin system: population statistics, metastability and other exact results *J. Chem. Phys.* **159** 044120
- [49] Gal A and Raz O 2020 Precooling strategy allows exponentially faster heating *Phys. Rev. Lett.* **124** 124 060602
- [50] Chittari S S and Lu Z 2023 Geometric approach to nonequilibrium hasty shortcuts *J. Chem. Phys.* **159** 084106
- [51] Anderson E *et al* 1999 *Lapack Users' Guide* 3rd edn (Society for Industrial and Applied Mathematics)
- [52] Johnson S G 2007 The NLOpt nonlinear-optimization package (available at: github.com/stevengj/nlopt)