

PAPER • OPEN ACCESS

# On thermalization of a system with discrete phase space

To cite this article: Alberto Imparato 2025 *J. Phys. A: Math. Theor.* **58** 495001

View the [article online](#) for updates and enhancements.

You may also like

- [SED Fitting for Algol Binaries with Some Observations Taken during the Eclipses](#)  
Mikhail Yu. Kovalev and
- [An Independent Assessment of the International Sunspot Number since 1996](#)  
Peter Meadows
- [Revising the theory of cross phenomena](#)  
Zi-Kui Liu

# On thermalization of a system with discrete phase space

Alberto Imparato\* 

Department of Physics, University of Trieste, Strada Costiera 11, 34151 Trieste, Italy  
Istituto Nazionale di Fisica Nucleare, Trieste section, Via Valerio 2, 34127 Trieste, Italy

E-mail: [alberto.imparato@units.it](mailto:alberto.imparato@units.it)

Received 8 August 2025; revised 3 November 2025

Accepted for publication 26 November 2025

Published 5 December 2025



CrossMark

## Abstract

We investigate the thermalization of a stochastic system with a discrete phase space, initially at equilibrium at temperature  $T_i$  and then thermalizing in an environment at temperature  $T_f$ , considering both cases  $T_i > T_f$  and  $T_i < T_f$ . For the simple case of a system with constant energy gaps, we show that the relation between the time scales of the cooling and heating processes is not univocal and depends on the magnitude of the energy gap itself. Specifically, the eigenvalues of the corresponding stochastic matrix determine the relaxation time scales: for large energy gaps, the cooling process exhibits the shortest relaxation times to equilibrium, whereas for small gaps, the heating process is faster at all scales. To quantify the degree of thermalization, we consider both the Kullback–Leibler divergence and the Fisher information (along with related quantities). In the intermediate-to-long-time regime, both quantities convey the same type of information regarding the rate of thermalization and follow the ordering predicted by the dynamical eigenvalues. We then examine a more complex system with a more intricate stochastic matrix—namely, a 1D Ising model—and confirm the existence of two regimes, one in which cooling becomes faster than heating. Finally, we relate our findings to a previous work in which a harmonic oscillator was used as the working fluid, where the heating process was always found to be faster than the cooling one.

Keywords: thermalization, nonequilibrium relaxation, stochastic processes

\* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Thermalization refers to the process by which a physical system evolves towards thermal equilibrium when in contact with a surrounding environment at constant temperature. It arises across diverse contexts, from quantum systems to biological processes and remains a subject of active research due to its fundamental and often unexpected features. A typical example is the Mpemba effect [1], that refers to the counterintuitive phenomenon where, under certain conditions, a system initially prepared at a higher temperature equilibrates faster than an otherwise identical system starting at a lower temperature. This effect challenges the intuitive expectation that cooler systems should reach equilibrium more quickly and has been observed in both classical and quantum systems [2–6].

In stochastic systems initially prepared in a non equilibrium state and then coupled to a bath at *final* temperature  $T_f$  the thermalization process occurs along different pathways that are selected by the initial conditions, and it involves energy exchange with the bath and transient currents across the system at different length and time scales. Thus, while the approach to equilibrium may appear straightforward, and governed by few intuitive mechanisms, it often reveals rich dynamical behavior and abundance of transient phenomena.

An intriguing problem to investigate is the timescale over which thermalization occurs and the key features that drive this process. To explore this, one requires a set of operational tools capable of quantifying both the extent of equilibration and the associated timescales.

In this spirit, in [7], the authors introduced the concept of statistical velocity and statistical length for a stochastic process to study the constraints on the rates of observables of interest. As a result the authors found a speed limit on the evolution of stochastic observables involving the statistical velocity defined as the square root of the Fisher information.

Introduced in the context of information geometry, Fisher information on turn quantifies the rate of change of the Kullback–Leibler (KL) divergence along a stochastic trajectory. Recently the information geometry approach has been extended to the to quantum stochastic thermodynamics, in particular in [8] the concept of Fisher information has been extended to the quantum realm.

In [9] the authors used the definition of statistical length (or distance) introduced in [7] for stochastic dynamics to argue that in two opposite thermalization processes between two temperatures  $T_+$  and  $T_-$ , with  $T_+ > T_-$ , the heating process is always faster than the reverse cooling process. In this framework, the statistical length of a process is simply the integral of the instantaneous statistical velocity, thus suggesting the wording ‘thermal kinematics’, for its resemblance with the kinematics in classical mechanics. Specifically the authors of [9] use a classical harmonic oscillator as working fluid for their theoretical and experimental study of the thermalization processes between two temperatures  $T_+$  and  $T_-$ . In the heating process the system is initially at equilibrium at temperature  $T_-$  for  $t \leq 0$ , and at  $t > 0$  is put in contact with a bath at temperature  $T_+$ . In the reverse (cooling) process the system is initially at equilibrium at temperature  $T_+$  and at  $t > 0$  is put in contact with a bath at temperature  $T_-$ . The statistical lengths introduced in [7] turns out to be symmetric, i.e. equal for the two processes, with the heating process covering this *distance* in a shorter time thus for the harmonic oscillator heating turns out to be faster than cooling, as one would intuitively expect. It is worth noting that an analogous result has been obtained in the quantum case in [10] where a two-level system, the quantum harmonic oscillator, and a trapped quantum Brownian particle showed faster heating than cooling under appropriate conditions.

In the present paper, we take a different approach, and consider classical systems with discrete phase space that evolve through a master equation and undergo thermalization processes between two temperatures  $T_+ > T_-$ . This allows us to distinguish between the different time scales corresponding to the eigenvalues of the stochastic matrix setting the transition rates

between the different states. Such time scales must enter into any possible kinematic description of an equilibration process.

The main result of the present paper is that for a given system with discrete phase space, one can always find a region in parameter space where the heating process become slower than the cooling. This occurs in a regime where the temperatures are low or energy gaps between the states are large compared to the thermal energies  $T_+$  and  $T_-$  set by the two baths (here and in the following we set  $k_B = 1$ ). Furthermore, the statistical length introduced in [7, 9] turns out to be asymmetric between the two processes, and in general the heating is characterized by a longer statistical length, thus signaling different thermalization pathways between the heating and the cooling processes. The symmetry of the statistical length is restored for two paradigmatic systems: a two state system (e.g. a single spin) and the classical harmonic oscillator, the latter in accordance with [9].

In terms of the different stochastic matrix eigenvalues describing the convergence towards equilibrium it is known that the first eigenvalue is vanishing, corresponding to the equilibrium state, while the second slowest eigenvalue describes the global decay towards equilibrium, all the other eigenvalues describing faster processes occurring in the system during the relaxation towards equilibrium [11]. The main result of the present paper can be thus rephrased as follows: in the limit where the dynamics becomes slow, i.e. for low temperatures or large energy gaps, the relaxation rates towards equilibrium, given by the second and higher order slowest eigenvalues, can become smaller for a heating process than for the corresponding cooling process, resulting in a somehow counterintuitive effect, as one would expect the dynamics driven by the hot reservoir being faster than the one driven by the cold reservoir. While in reference [7] the formalism of the Fokker–Planck equation was used to introduce and discuss the properties of the statistical velocity and statistical length, two examples of Markov jump models were also considered. The formalism of Markov jump models is the one considered in the present paper, as discussed below.

## 1. Stochastic system

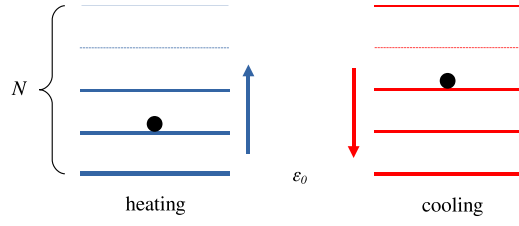
In the following we use the standard formalism of the master equation, and focus on relaxation processes between an initial ( $t \leq 0$ ) and a final ( $t > 0$ ) temperature,  $T_i$  and  $T_f$ , respectively. The theory of the master equation is briefly reviewed, and we refer the reader to, e.g. [11], for a more detailed discussion.

We consider a system with  $N$  discrete energy levels  $\epsilon_k$ , and we order the energy levels in ascending order  $\epsilon_0 = 0 \leq \epsilon_1 \leq \dots \leq \epsilon_{N-1}$ .

In this paper we want to study the thermalization dynamics between two temperatures  $T_-$  and  $T_+$ . More precisely, we compare the relaxation from  $T_-$  to  $T_+$ , where the system, initially at equilibrium at  $T_i = T_-$  ( $t \leq 0$ ) is placed in contact with a bath at  $T_f = T_+$  ( $t > 0$ ), this corresponds to the heating protocol. In the reverse protocol the system is initially at equilibrium at temperature  $T_i = T_+$  ( $t \leq 0$ ) and then cooled by placing it in contact with a bath at temperature  $T_f = T_-$  ( $t > 0$ ). The initial states of the two processes are depicted in figure 1.

We ask the following question: in which process is the new equilibrium thermal state reached first?

In the following we will use the notation  $\mathbf{p}_-^+(t)$  and  $\mathbf{p}_+^-(t)$  that indicate the probability vector for the heating and cooling process, respectively (from initial to final temperature  $\mathbf{p}_i^f(t)$ , where the subscript  $i = -, +$  indicates the initial temperature, and the superscript  $f = +, -$  indicates the final temperature).



**Figure 1.** Initial state for the heating (left) and the cooling (right) process for a system with  $N$  discrete energy levels. The different thickness of the single levels represents the different size of the initial populations in each level (not in scale). For  $T_i = T_-$  the lower energy levels are more populated with respect to the case  $T_i = T_+$ . The arrows represent the average directions of the initial probability currents in the two cases.

The dynamics is described by a master equation of the type

$$\dot{\mathbf{p}}(t) = \mathbf{W} \cdot \mathbf{p}(t), \tag{1}$$

where  $\mathbf{W}$  is a stochastic matrix with elements  $w_{ij}$ , whose dependence on the (inverse) temperature is encoded through the detailed balance condition for its elements

$$w_{kj} = w_{jk} e^{\beta(\epsilon_j - \epsilon_k)}. \tag{2}$$

Thus we indicate with  $\mathbf{W}^+$  and  $\mathbf{W}^-$  the stochastic matrices ruling the dynamics during the heating and the cooling process, respectively.

Let  $\lambda_n$ , with  $n = 0, \dots, N - 1$  be the eigenvalues of  $\mathbf{W}$  for a given temperature, and let us assume they have been ordered in ascending order  $\lambda_0 > \lambda_1 \geq \dots \lambda_{N-1}$ . Given that  $\mathbf{W}$  is a stochastic matrix, the largest eigenvalue is  $\lambda_0 = 0$ , with all the other eigenvalues which are negative, see Ch.V in [11]. Thus if  $\Phi_0$  is the eigenvector corresponding to the eigenvalue  $\lambda_0$  then it corresponds to the equilibrium thermal probability at the final inverse temperature,  $\Phi_0 = \mathbf{p}^{eq}(\beta_f)$ , while the other eigenvalues  $\lambda_n$ ,  $n > 0$  determine the rate of relaxation towards the equilibrium state  $\mathbf{p}^{eq}(\beta_f)$ .

Thus we have

$$\mathbf{p}_\pm^\pm(t) = e^{t\mathbf{W}^\pm} \cdot \mathbf{p}_\pm^{eq}, \quad \mathbf{p}_\pm^\mp(t) = e^{t\mathbf{W}^\mp} \cdot \mathbf{p}_\pm^{eq}, \tag{3}$$

with  $\mathbf{p}_\pm^{eq} = \mathbf{p}^{eq}(\beta_\pm)$ .

To analyze the difference in the dynamics of the two processes, the heating and the cooling one, one has at least two options.

- i) One is tempted to use the KL divergence (or relative entropy)

$$D(\mathbf{p}_a || \mathbf{p}_b) = \sum_k p_{a,k} \log \left( \frac{p_{a,k}}{p_{b,k}} \right), \tag{4}$$

to evaluate the convergence toward equilibrium, and evaluate  $D(\mathbf{p}_i^f(t) || \mathbf{p}_f^{eq})$  at any time  $t \geq 0$ . However, as detailed in [7, 9], the KL divergence is not a metric in the space of the probability vectors: it does not, for example, satisfy the triangle inequality, and thus cannot be used as a distance between probabilities, in particular it does not provide a proper *distance* from equilibrium.

Nevertheless, the KL relative entropy can provide useful information on the convergence rate of the equilibration process, as we will see in the following. Let us define the KL distances for the two processes

$$D_{-}^{+}(t) = D(\mathbf{p}_{-}^{+}(t) || \mathbf{p}_{+}^{eq}), \quad D_{+}^{-}(t) = D(\mathbf{p}_{+}^{-}(t) || \mathbf{p}_{-}^{eq}). \quad (5)$$

While the KL relative entropy is asymmetric, so at initial time one has  $D_{+}^{-}(0) \neq D_{-}^{+}(0)$ , we find that the inequality  $D_{+}^{-}(0) \geq D_{-}^{+}(0)$  holds (which becomes an equality in the trivial case  $\beta_{+} = \beta_{-}$ ), and a formal proof is given in appendix A. In appendix B we also prove that, as one would expect, the convergence to zero of the KL divergence is dominated by the second slowest eigenvalue  $\lambda_1$ :

$$D(\mathbf{p}_i^f(t) || \mathbf{p}_f^{eq}) \underset{t \rightarrow \infty}{\sim} e^{2\lambda_1 t}. \quad (6)$$

Thus the order relation between  $D_{+}^{+}(t)$  and  $D_{+}^{-}$  at long time is dictated by the relation between  $\lambda_1^{+}$  and  $\lambda_1^{-}$  which as we will see is not univocally fixed.

ii) One can use the approach introduced in [9] and consider the Fisher information, the instantaneous statistical velocity and the statistical length that read

$$I(t) = \langle (\partial_t \ln \mathbf{p}(t))^2 \rangle, \quad (7)$$

$$v(t) = \sqrt{I(t)}, \quad (8)$$

$$\mathcal{L}(t) = \int_0^t d\tau v(\tau), \quad (9)$$

respectively.

The Fisher information is related to the KL divergence through the relation

$$\begin{aligned} D(\mathbf{p}(t + \delta t) || \mathbf{p}(t)) &= \delta t^2 \sum_k \frac{(\partial_t p_k(t))^2}{p_k(t)} + O(\delta t^3) \\ &= \delta t^2 I(t) + O(\delta t^3) \end{aligned} \quad (10)$$

Therefore inspection of equations (8) suggests that the instantaneous statistical velocity is related to the rate of change of the KL divergence. We notice that in [9] the long time limit of the statistical length was found to be independent of the thermodynamic *direction*, i.e.  $\mathcal{L}_{+}^{+}(\infty) = \mathcal{L}_{+}^{-}(\infty)$ . For the systems with discrete energy spectrum considered in the present manuscript, we anticipate that in general this is the case only for  $N = 2$  states, or in the limit of small energy gaps.

In appendix B we also prove that the long time behavior of the Fisher entropy ( the statistical velocity ) is dominated by the largest eigenvalue of the stochastic matrix  $I(t) \sim e^{2\lambda_1 t}$  ( $v(t) \sim e^{\lambda_1 t}$ ). Thus, as for the case of the KL divergence, we will see that the order relation between  $I_{+}^{+}(t)$  and  $I_{+}^{-}(t)$ , and thus between  $v_{+}^{+}(t)$  and  $v_{+}^{-}(t)$  is not univocally fixed. Furthermore, given that the long time behavior of the KL divergence and the Fisher entropy are both governed by  $\lambda_1$ , there is in principle no reason to prefer one quantity over another to describe the convergence to equilibrium.

We also find numerical evidence that the following inequality holds  $I_{-}^{+}(0) \geq I_{+}^{-}(0)$ . Such an equality has been numerically verified for all the models considered in this paper (ladder model in section 2 and Ising model in section 3), and for a wide range of parameters. While the author of this paper has not succeeded in finding a general analytic proof of this inequality besides the case  $N = 2$ , a fair level of intuition can be gained by considering that the following equalities hold

$$\begin{aligned} \left. \frac{d}{dt} D(\mathbf{p}_+^-(t) || \mathbf{p}_+^{eq}) \right|_{t=0} &= \left. \frac{d}{dt} D(\mathbf{p}_-^+(t) || \mathbf{p}_-^{eq}) \right|_{t=0} = 0, \\ I_+^-(0) &= \left. \frac{d^2}{dt^2} D(\mathbf{p}_+^-(t) || \mathbf{p}_+^-(0)) \right|_{t=0}, \\ I_-^+(0) &= \left. \frac{d^2}{dt^2} D(\mathbf{p}_-^+(t) || \mathbf{p}_-^+(0)) \right|_{t=0}. \end{aligned}$$

where we notice that  $\mathbf{p}_+^-(0) = \mathbf{p}_+^{eq}$  and  $\mathbf{p}_-^+(0) = \mathbf{p}_-^{eq}$ . Thus the inequality  $I_-^+(0) \geq I_+^-(0)$  expresses the intuitive expectation that the system departs from its initial condition faster when in contact with the hot bath at temperature  $T_+$ .

A few comments on the transition rates in the stochastic matrices are now in order. As far as the symmetric part of the matrix elements  $w_{kj}$  is concerned, in stochastic thermodynamics it is customary to consider those terms as temperature independent  $w_{kj} = \omega_0 f(k, j)$ , with the asymmetric term  $f(k, j)$  satisfying the detailed balance condition (2). We will term the bath corresponding to this choice standard bath (SB). However, when one derives the dynamical equations starting from a detailed bath-system model, the *pre-factor* turns out to depend explicitly on the temperature. If one starts from a microscopic model of bath interacting with the system of interest, as one standardly does when studying open quantum systems, the transition rates exhibit prefactors that depend explicitly on the temperature, and for a bosonic bath the transition rate between a state  $j$  and a state  $k$  reads [12]

$$\omega_{kj} = \frac{\omega_0 |\Delta \epsilon_{kj}|}{1 - e^{-\beta |\Delta \epsilon_{kj}|}} \begin{cases} e^{-\beta \Delta \epsilon_{kj}}, & \text{if } \Delta \epsilon_{kj} > 0 \\ 1, & \text{otherwise} \end{cases} \quad (11)$$

with the symmetric prefactor that goes to  $\omega_0/\beta$  in the classical limit  $\beta \rightarrow 0$  [13], and where  $\epsilon_{kj} = \epsilon_k - \epsilon_j$ , so that the rates in equation (11) satisfy the detailed balance relation  $\omega_{kj}/\omega_{jk} = \exp[-\beta(\epsilon_k - \epsilon_j)]$ . We will term the bath corresponding to this choice bosonic bath (BB), and in the following we will set the time scale by taking  $\omega_0 = 1$  for both types of baths.

## 2. Equispaced energy levels

In the following we consider the case of equispaced energy levels  $\epsilon_k = k\epsilon$ , with  $\epsilon > 0$  and  $k = 0, \dots, N-1$ . We also use the notation  $\lambda_{N,k}$  to enumerate the eigenvalues of the corresponding  $N \times N$  stochastic matrix  $\mathbf{W}$ .

We will consider stochastic block tridiagonal matrices, i.e. matrices whose non-zero elements are located on the lower diagonal, the main diagonal and the upper diagonal, with  $w_{kj} = 0$  if  $|k-j| > 1$ . This is the case, for example, of the master equations for a harmonic oscillator where the particle interacts with the bath through the  $x$  coordinate ( $a$  and  $a^\dagger$  operators) or of a single spin system where the interaction with the bath occurs through the  $x$  (or  $y$ ) component of the spin ( $\sigma^+$  or  $\sigma^-$  operators).

The stochastic matrix thus reads

$$\mathbf{W} = \begin{pmatrix} -w^\uparrow & w^\downarrow & 0 & 0 & \dots & 0 \\ w^\uparrow & -(w^\downarrow + w^\uparrow) & w^\downarrow & 0 & \dots & 0 \\ 0 & w^\uparrow & -(w^\downarrow + w^\uparrow) & w^\downarrow & \dots & 0 \\ 0 & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & 0 & w^\uparrow & -w^\downarrow \end{pmatrix} \quad (12)$$

The case of system with degenerate energy levels and transition rates  $w_{kj}$  connecting two arbitrary microscopic states  $k$  and  $j$  will be considered in section 3.

For the SB we will set  $w^\downarrow = \omega_0$  and  $w^\uparrow = \omega_0 \exp(-\beta\epsilon)$ , while for the BB equation (11) gives

$$w^\uparrow = \frac{\omega_0 \epsilon}{1 - e^{-\beta\epsilon}} e^{-\beta\epsilon} \quad (13)$$

$$w^\downarrow = \frac{\omega_0 \epsilon}{1 - e^{-\beta\epsilon}} \quad (14)$$

### 2.1. $N = 2$ case

We study first the case of a particle with 2 energy levels (a 1/2 spin particle or a qubit), with stochastic matrix

$$\mathbf{W} = \begin{pmatrix} -w^\uparrow & w^\downarrow \\ w^\uparrow & -w^\downarrow \end{pmatrix}$$

and with  $w^\uparrow = w^\downarrow \exp(-\beta\epsilon)$ : it has two eigenvalues  $\lambda_{2,0} = 0$  and  $\lambda_{2,1} = -(w^\uparrow + w^\downarrow) = -w^\uparrow(1 + \exp(-\beta\epsilon))$ . Thus we see  $|\lambda_{2,1}^+| = |\lambda_{2,1}(\beta_+)| > |\lambda_{2,1}^-| = |\lambda_{2,1}(\beta_-)|$  for any pair of temperatures such that  $\beta_+ < \beta_-$  ( $T_+ > T_-$ ). The decay rate of the heating process is larger than the decay rate of the cooling process:  $1/|\lambda_{2,1}|$  is the only time scale that enters in the convergence towards equilibrium and as such heating is expected to be the faster process. This is confirmed by the analysis of the quantities introduced above to characterize the convergence towards equilibrium, see figure 2. One finds that  $D_+^-(t) > D_-^+(t)$  for any  $t \geq 0$ , indicating that, for a given  $t$ , the heating process brings the system *closer* to its final equilibrium state than the corresponding cooling process. Furthermore one finds that for the thermodynamic distance, as given by equation (9), the following inequality holds  $\mathcal{L}_-^+(t) \geq \mathcal{L}_+^-(t)$  for any  $t \geq 0$ , and is equal for the two processes in the long time regime  $\mathcal{L}_-^+(\infty) = \mathcal{L}_+^-(\infty)$ . Therefore such a distance is independent of the thermodynamic direction, as one would expect for a metric measuring distances between equilibrium states. This result for  $N = 2$  is derived in appendix C.1 and is in accordance with the findings of [9] for the harmonic oscillator. Thus, one finds that for the degree of completion, defined as  $\mathcal{L}_i^f(t)/\mathcal{L}_i^f(\infty)$  the following equality holds

$$\mathcal{L}_-^+(t)/\mathcal{L}_-^+(\infty) \geq \mathcal{L}_+^-(t)/\mathcal{L}_+^-(\infty) \quad (15)$$

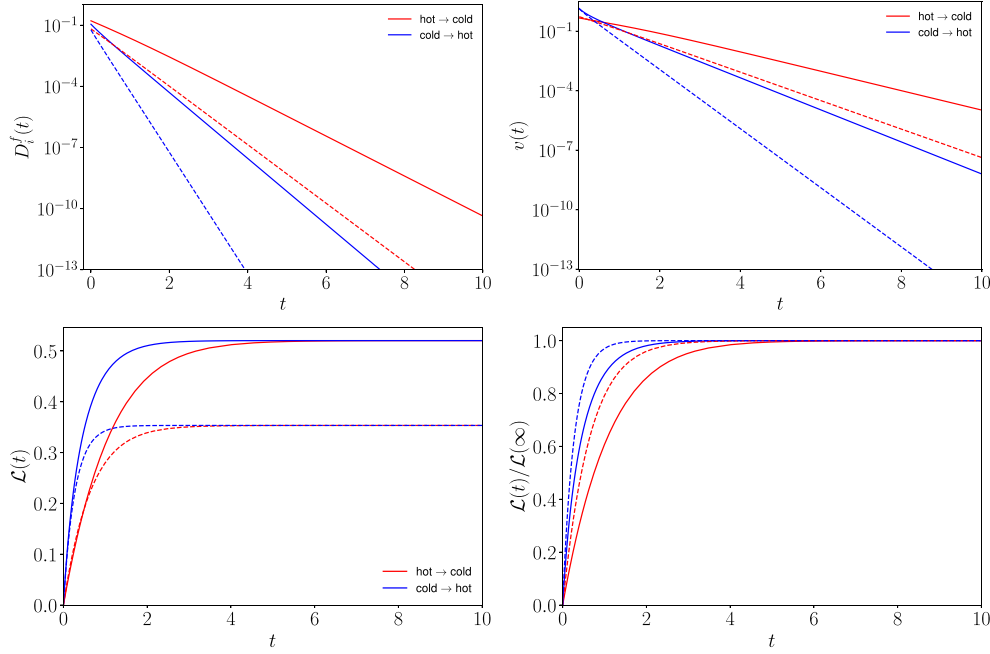
confirming the prediction based on the bare analysis of  $\lambda_{2,1}^\pm$  that heating is faster than cooling.

As far as the statistical velocity is concerned, equation (8), the order relation  $v_+^-(t)$  and  $v_+^-(t)$  requires a more detailed analysis.

As anticipated above at  $t = 0$  one finds  $v_+^-(0) > v_+^-(0)$ , with the same inequality holding at short time, i.e. the heating process is initially the faster one, however at intermediate and large  $t$  the inequality is inverted  $v_+^-(t) > v_+^-(t)$ , see figure 2. This can be understood by comparing the behavior of  $v_i^f(t)$  with  $D_i^f(t)$  and  $\mathcal{L}_i^f(t)$  in figure 2 and by noticing that at intermediate and large time the heating process is *closer* to the final equilibrium state. Given that  $v(t)$  is related to the rate of change of the instantaneous KL divergence, see equations (8) and (10), we conclude that statistical velocity decreases upon completion of the equilibration process. This is compatible with the inequality  $|\lambda_{2,1}^+| > |\lambda_{2,1}^-|$  and with the fact that in the long time regime  $v(t) \sim e^{\lambda_{2,1} t}$ . As discussed above the inequality at initial time  $v_+^-(0) > v_+^-(0)$  cannot be proved in general, however a proof for the  $N = 2$  case is provided in appendix C.1.

### 2.2. $N > 2$ case

The situation becomes more interesting if one considers 3 or more states. For the general case with  $N$  states the characteristic polynomial of the stochastic matrix (12) can be obtained



**Figure 2.** KL divergence  $D_+^+(t)$  and  $D_-^-(t)$  (equation (5), top left), statistical velocity  $v(t)$  (equation (8), top right), statistical velocity  $\mathcal{L}(t)$  (equation (9), bottom left), and degree of completion  $\mathcal{L}_i^f(t)/\mathcal{L}_i^f(\infty)$  (bottom right) for a two state system ( $N = 2$ ) with  $\beta_- = 1.4$ ,  $\beta_+ = 0.6$ ,  $\epsilon = 2$  (full lines) and  $\epsilon = 1$  (dashed lines). Notice that  $D_+^+(t) > D_-^-(t)$  for any  $t \geq 0$  and  $\mathcal{L}_-^-(\infty) = \mathcal{L}_+^+(\infty)$ . The BB is used to obtain the plots in this figure.

exactly, as well as the expression of the single eigenvalues  $\lambda_{N,k}$  that reads

$$\lambda_{N,k} = - (w^\uparrow + w^\downarrow) + (-1)^{m(k,N)} g_{N,k} \sqrt{w^\uparrow w^\downarrow} \tag{16}$$

with

$$m(k,N) = \begin{cases} 2, & \text{if } k = 1, \dots, \lfloor \frac{N}{2} \rfloor \\ 1, & \text{if } k = \lfloor \frac{N}{2} \rfloor + 1, \dots, N-1 \end{cases}$$

and where  $g_{N,k}$  is a decreasing function of  $k$  for  $k \leq \lfloor N/2 \rfloor$ ,

$$g_{N,k} = \frac{2}{\sqrt{1 + \tan^2\left(\frac{k\pi}{N}\right)}} \tag{17}$$

see appendix D for the derivation. In particular we want to study the second slowest decay rate in the equilibration processes, namely the second largest eigenvalues that reads  $\lambda_{N,1} = -(w^\uparrow + w^\downarrow) + g_{N,1} \sqrt{w^\uparrow w^\downarrow}$ , with  $1 \leq g_{N,1} < 2$  for  $N > 2$  and  $g_{N,1} \rightarrow 2$  in the limit  $N \rightarrow \infty$ . We now investigate whether and under which conditions the second slowest decay rate for the cooling process can become larger than the corresponding rate for the heating process. In the following we treat the inverse temperatures  $\beta_-$  and  $\beta_+$  as parameters, and the energy gap  $\epsilon$  as

a free variable and consider first the SB, the results for the BB turning out to be qualitatively identical. One finds that the condition  $|\lambda_{N,1}^+| \leq |\lambda_{N,1}^-|$  is equivalent to

$$g_{N,1} e^{(\beta_+ + \beta_-)\epsilon/2} \geq e^{\beta_+ \epsilon/2} + e^{\beta_- \epsilon/2} \quad (18)$$

For  $N = 2$  this condition is never satisfied for a non-vanishing  $\epsilon$  as  $g_{2,1} = 0$ , while above  $N = 2$  it is always possible to find a combination of parameters for which it holds. The functions on the two sides of equation (18) take both positive values, and are both increasing, with the left hand side smaller at small  $\epsilon$  but increasing faster than the right hand side, thus for some value of the gap  $\epsilon_{N,1}^*$  the two sides of equation (18) will become equal, and for  $\epsilon > \epsilon_{N,1}^*$  the strict inequality holds, resulting in  $|\lambda_{N,1}^+| < |\lambda_{N,1}^-|$  holds, i.e. the second slowest decay rate for cooling is faster than the corresponding rate for heating.

We see that for fixed  $\beta_+$  and  $\beta_-$ , since  $g_{N,1}$  increases with  $N$ , the value of  $\epsilon_{N,1}^*$  decreases as  $N$  increases. In particular as  $N \rightarrow \infty$ ,  $g_{\infty,1} = 2$ , and one immediately concludes from equation (18) that  $\epsilon_{\infty}^* = 0$ , thus  $|\lambda_{\infty,1}^+| < |\lambda_{\infty,1}^-|$  for any  $\epsilon > 0$ . Conversely for any arbitrary large but finite  $N$ , one has  $\epsilon_{N,1}^* > 0$ , and one can always choose an  $\epsilon$  small enough such that  $\epsilon_{N,1}^* > \epsilon > 0$ .

The difference between  $\lambda_{N,1}^+$  and  $\lambda_{N,1}^-$  vanishes as  $\epsilon \rightarrow 0$ , and one can prove that up to first order in  $\epsilon$  the statistical length for  $t \rightarrow \infty$  is symmetric  $\mathcal{L}^+(\infty) \simeq \mathcal{L}^-(\infty)$  with the difference being of order  $O(\epsilon^2)$ , see appendix C.2 and equation (C10) in particular. While the choice  $\epsilon = 0$  is meaningless, one can regard the classical harmonic oscillator considered in [9], as the limiting case of the present system with discrete energy levels, in the limit of large  $N$  and vanishing  $\epsilon$ . As such, the result of the present paper in this limit is in accordance with that of [9], namely that the statistical length is symmetric for a classical harmonic oscillator when  $t \rightarrow \infty$ .

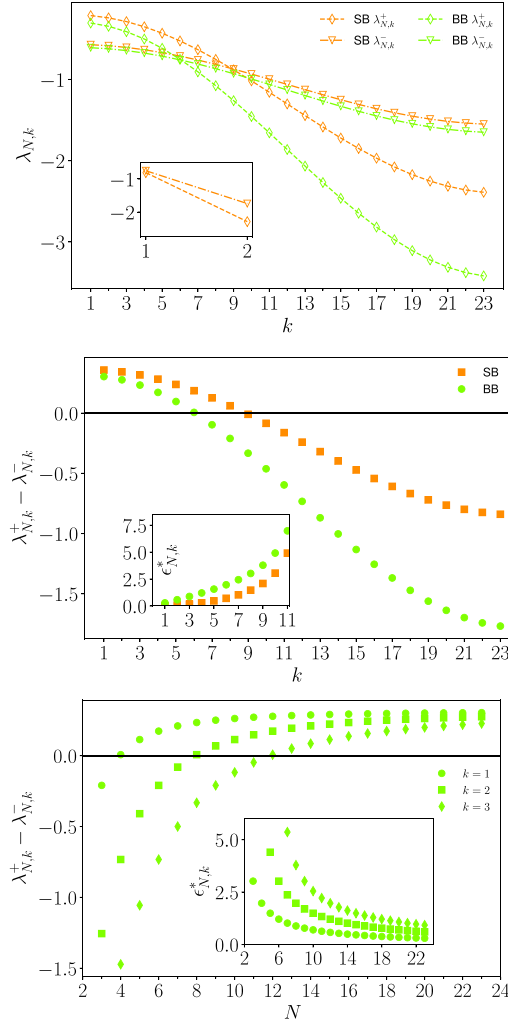
The parallelism finishes here as an harmonic oscillator with mass  $m$ , proper frequency  $\omega_0$  and friction  $\gamma$  only has one or two characteristic times describing the decay to equilibrium, which are the inverse of the frequencies  $\gamma/m$  or  $\gamma/m \pm \sqrt{\gamma^2/m^2 - 4\omega_0^2}$  depending on whether the oscillator is underdamped or not. In any case these characteristic times do not depend on the temperature, and are the same for the heating and cooling process. In other words for a Brownian particle in a harmonic potential the temperature only sets the size of the fluctuations, but not the decay rates towards equilibrium.

On the contrary the discrete model retains  $N$  (with  $N$  possibly large) time scales, that do depend on the bath temperature.

For higher order eigenvalues  $\lambda_{N,k}$  (equation (16)) with  $1 < k \leq \lfloor \frac{N}{2} \rfloor$  an inequality similar to equation (18) can be found by replacing  $g_{N,1}$  with  $g_{N,k}$ , and thus for each set of parameters in the stochastic matrix, one find a threshold value for the gap  $\epsilon_{N,k}^*$  for which  $|\lambda_{N,k}^+| \leq |\lambda_{N,k}^-|$  if  $\epsilon > \epsilon_{N,k}^*$ . Given that  $g_{N,k}$  is a decreasing function of  $k$  (as long as  $k \leq \lfloor \frac{N}{2} \rfloor$ ), one also has  $\epsilon_{N,k+1}^* > \epsilon_{N,k}^*$ : thus once one fixes all the parameters and the value of  $\epsilon$ , with  $\epsilon_{N,l+1}^* > \epsilon > \epsilon_{N,l}^*$ , the inequality  $|\lambda_{N,k}^+| \leq |\lambda_{N,k}^-|$  is satisfied only for the first  $l$  eigenvalues.

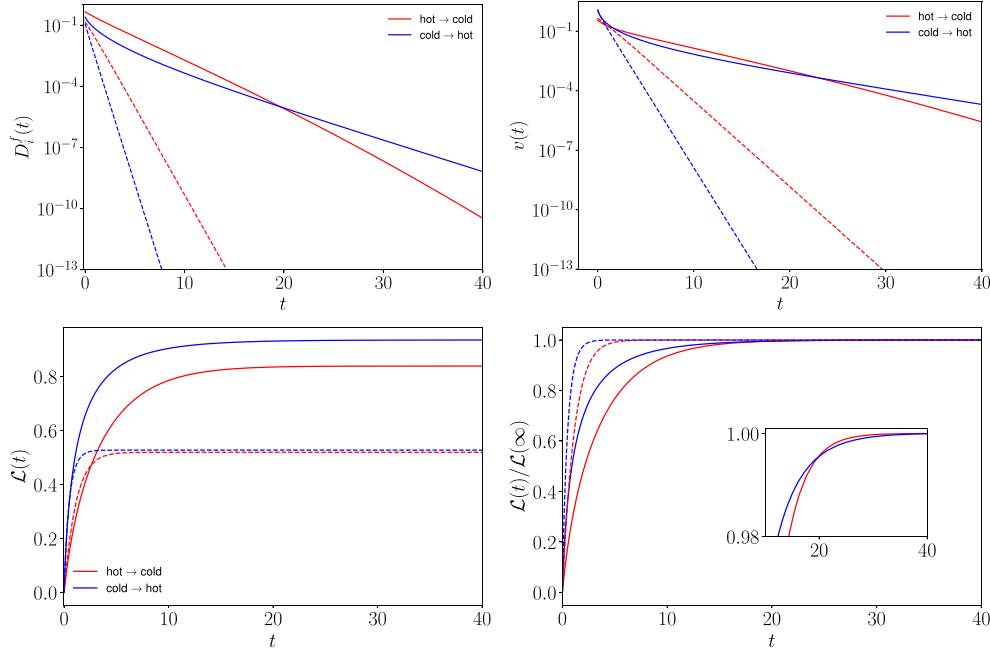
The above findings are exemplified by the panels of figure 3, where the eigenvalues  $\lambda_{N,k}^\pm$  are plotted for some choices of the system parameters.

Inspection of this figure confirms that, for fixed  $\beta_+$ ,  $\beta_-$ ,  $\epsilon$ , and  $N$ , the difference  $\lambda_{N,k}^+ - \lambda_{N,k}^-$  as a function of  $k$  is first positive for small  $k$ , and then becomes negative for some value of  $k$  that depends on the type of bath. Conversely, for fixed  $k$ , the difference  $\lambda_{N,k}^+ - \lambda_{N,k}^-$  as function of  $N$  is first negative for small  $N$  and then becomes positive as  $N$  increases. Such a difference



**Figure 3.** First panel: stochastic matrix eigenvalues as a function of the index  $k$  for the system with equispaced energy levels, with  $\epsilon = 2$ ,  $\beta_- = 1.4$ ,  $\beta_+ = 0.6$  and  $N = 24$  states (main panel),  $\epsilon = 1$  and  $N = 3$  (inset), for the two different types of baths SB and BB. Second panel: difference between the eigenvalues of the stochastic matrices  $\mathbf{W}^+$  and  $\mathbf{W}^-$  for the two types of baths and  $N = 24$ . One finds that the inequality  $\lambda_{N,k}^+ > \lambda_{N,k}^-$  holds for  $1 \leq k \leq 8$  (SB) and for  $1 \leq k \leq 6$  (BB). Inset: threshold energy spacing  $\epsilon_{N,k}^*$  as a function of  $k$  and for fixed  $N = 24$ , see equations (18) and (19) for the two different types of baths, SB and BB respectively. Third panel: Difference of the first 3 eigenvalues of the stochastic matrices  $\mathbf{W}^+$  and  $\mathbf{W}^-$  as functions of the system number of states  $N$ . Inset: threshold gap as a function of  $N$  and for fixed  $k = 1, 2, 3$  as given by the solution of the equality in equation (19) for the BB. One sees that the values of  $\epsilon_{N,k}^*$  as a function of  $N$  is a decreasing function of  $N$ .

becomes positive first for eigenvalues with lower index  $k$ , which is a consequence of the fact that the threshold value  $\epsilon_{N,k}^*$  is an increasing function of  $k$ , being the solution of the equality in equation (18) with  $g_{N,1}$  replaced by  $g_{N,k}$ .



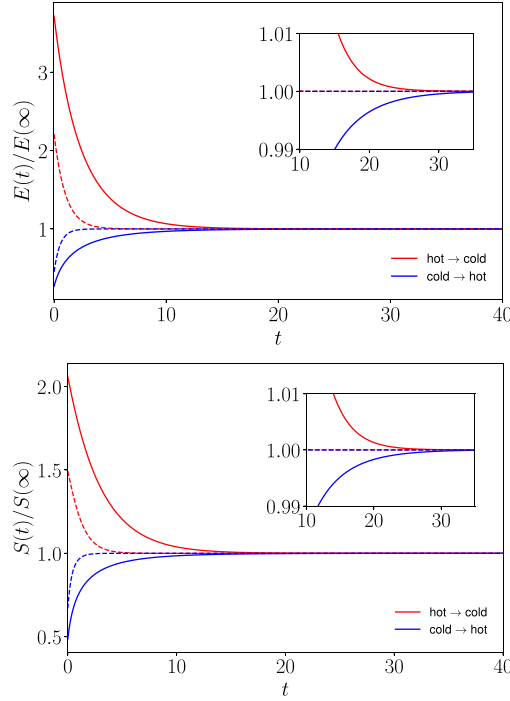
**Figure 4.** KL divergence, statistical velocity, statistical length and completion rate as functions of  $t$ , with  $\epsilon = 1$ ,  $\beta_- = 1.4$ ,  $\beta_+ = 0.6$  and  $N = 24$  states (full line),  $N = 3$  dashed line (same parameter as figure 3) and BB. For  $N = 24$  and in the intermediate and long time regime one clearly sees the inversion of the convergence rate, with the cooling process becoming the one converging faster to equilibrium. In the log-linear plots of  $D(t)$  and  $v(t)$  the slopes in the linear regions are  $2\lambda_1$  and  $\lambda_1$ , respectively (lines not shown), as predicted from long-time analysis discussed in appendix B.

In the case of BB, equation (18) becomes

$$g_{N,k} \left( e^{(\beta_+ + \beta_-)\epsilon/2} + 1 \right) \geq 2 \left( e^{\beta_+ \epsilon/2} + e^{\beta_- \epsilon/2} \right) \quad (19)$$

This gives a different value of the threshold gap  $\epsilon_{N,k}^*$ , but the qualitative behavior of the eigenvalues being identical to that obtained for the SB.

Having analyzed the behavior of the eigenvalues  $\lambda_{N,k}^\pm$  we can now study the convergence to equilibrium of the heating and the cooling processes, and in particular the KL divergence, the statistical velocity and the statistical length. A plot of such quantities is shown in figure 4. In particular, for  $N = 3$  the chosen value of  $\epsilon$  is too small to obtain the inversion of the order of the eigenvalues, thus one has  $|\lambda_{3,k}^+| > |\lambda_{3,k}^-|$ ,  $k = 1, 2$ , the KL relative entropy converges faster to its equilibrium value for the heating process, the same is found for statistical velocity, with  $v(t)$  going to zero faster for the heating process, and the statistical length satisfying at all times  $\mathcal{L}_+^\pm(t) \geq \mathcal{L}_-^\pm(t)$ . When  $N$  is large ( $N = 24$  in figure 4) a completely different behavior appears in the intermediate and long time regime: the KL relative entropy converges faster to its equilibrium value for the *cooling* process, thus the statistical velocity  $v(t)$  goes to zero faster than the one for the *heating* process. We remind the reader that the statistical velocity is indeed related to the rate of change of the instantaneous DL divergence, equations (8) and (10). In other words, since the *cooling* process is completed earlier, its pace, as measured by  $v(t)$  slows down. Conversely in the long time limit the *heating* process still proceeds at higher pace, as the



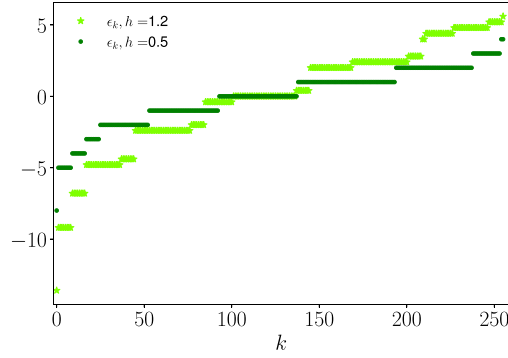
**Figure 5.** Rescaled average energy and stochastic entropy as functions of the time. The values of the parameters are the same as in figure 4. One notices that in the intermediate and long time regime the cooling process converges faster.

KL divergence is further away from the equilibrium value  $D_{-}^{+}(t) > D_{+}^{-}(t)$  and thus the system retains a higher velocity resulting in  $v_{-}^{+}(t) > v_{+}^{-}(t)$ . Mathematically this is a consequence of the inversion of the order of the eigenvalues, i.e.  $|\lambda_{N,k}^{+}| < |\lambda_{N,k}^{-}|$ , with the cooling process becoming the one with the shortest relaxation time and with the heating process taking more time to fill up the high energy levels. We remind the reader that in the long time limit we have proved  $D(\mathbf{p}_i^f(t) || \mathbf{p}_f^{eq}) \sim \exp(2\lambda_1^f t)$  (equation (6)) and a similar result has been obtained for  $I(t) = v^2(t)$ . The physical intuition behind these results is that, given the initial states as depicted in figure 1, for  $\epsilon$  large enough the *cooling* process becomes the faster one with the higher energy levels being emptied at a greater speed, while in the *heating* process it takes longer time for the dynamic to explore the phase space and in particular the states with higher energy.

While for the statistical length the equality  $\mathcal{L}_{-}^{+}(t) \geq \mathcal{L}_{+}^{-}(t)$  still holds at any time, the change at intermediate time scale is reflected in a different behavior of the completion rate, where the inequality now holding in the opposite direction  $\mathcal{L}_{-}^{+}(t)/\mathcal{L}_{-}^{+}(\infty) \leq \mathcal{L}_{+}^{-}(t)/\mathcal{L}_{+}^{-}(\infty)$ , signaling that the cooling process is completed first, see inset in the fourth panel of figure 4.

The above results are confirmed by the analysis of the average energy  $E(t) = \sum_k p_k(t) \epsilon_k$  and of the average stochastic entropy  $S(t) = -\sum_k p_k(t) \log p_k(t)$ : provided the energy gaps are large enough, in the intermediate and long time regime both  $E(t)$  and  $S(t)$  converge to their equilibrium values faster for the *cooling* process than for the heating one, see figure 5.

We conclude this section by noticing that in the long time limit we always find the strict inequality  $\mathcal{L}_{-}^{+}(\infty) > \mathcal{L}_{+}^{-}(\infty)$  for finite  $\epsilon$  and  $N > 2$  at variance with the findings of [9] for a classical harmonic oscillator, where the two lengths were found to be equal.



**Figure 6.** Energy eigenvalues of the model (20) with  $M = 8$  spins,  $J = 0.5$  and for two different values of  $h$ .

### 3. Thermalization in an Ising model

The stochastic matrix in equation (12) describes a simple topology of a linear network of states, where each state is only connected to two nearest neighbor states. In order to consider a more complex topology, where each state is connected to several states, we study the stochastic dynamics of a  $M$ -spin system, with Hamiltonian

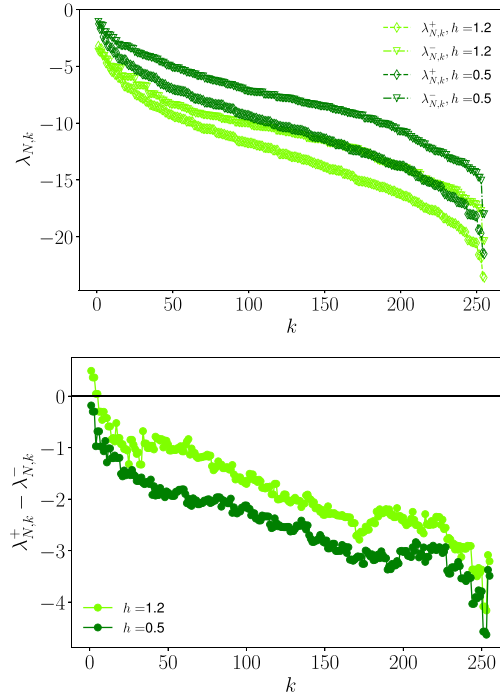
$$H = - \sum_{j=0}^{M-1} (J\sigma_j\sigma_{j+1} + h\sigma_j). \quad (20)$$

In order to investigate the thermalization of such a model, one can build on the findings of the previous section: upon increasing of the energy gap  $\epsilon$  in a model with equispaced energy levels there is a crossover from a regime where the heating process is faster than the cooling one on all the time scales to a regime where the cooling process becomes faster, at least for the first  $l$  slowest rates. In the model (20), increasing  $h$  would increase the gaps between the different energy shells of the model in equation (20) and reduce the degeneracy of the energy eigenvalues. This is exemplified in figure 6, where the eigenenergies  $\epsilon_k$  of (20) are plotted as a function of  $k$  for two different values of  $h$ .

We thus consider the dynamics generated by the stochastic matrix with elements (11), connecting two states differing by only one spin orientation, i.e. given a microscopic state  $\{\sigma_j\}$ , only one spin is allowed to flip in a single stochastic jump. With this choice, each microscopic state is connected to  $M$  other states.

The eigenvalues  $\lambda_{N,k}^{\pm}$  and their difference are plotted in figure 7 as a function of  $k$ , with  $N = 2^M$ . We see that by increasing  $h$  the eigenvalues show inversion of the order, i.e.  $|\lambda_{N,k}^+| < |\lambda_{N,k}^-|$ , at least for a few initial values of  $k$ , corresponding to the longest time scales.

This is then reflected in the KL divergence and in the statistical velocity, where for relatively small  $h$  one finds again that the heating process is the faster to converge to equilibrium, while for larger  $h$  the cooling process becomes the faster to converge to equilibrium in the intermediate and long time scale, see figure 8. The statistical length is found again to be non-symmetric between the two processes, signaling different pathways to equilibrium for the two opposite processes.



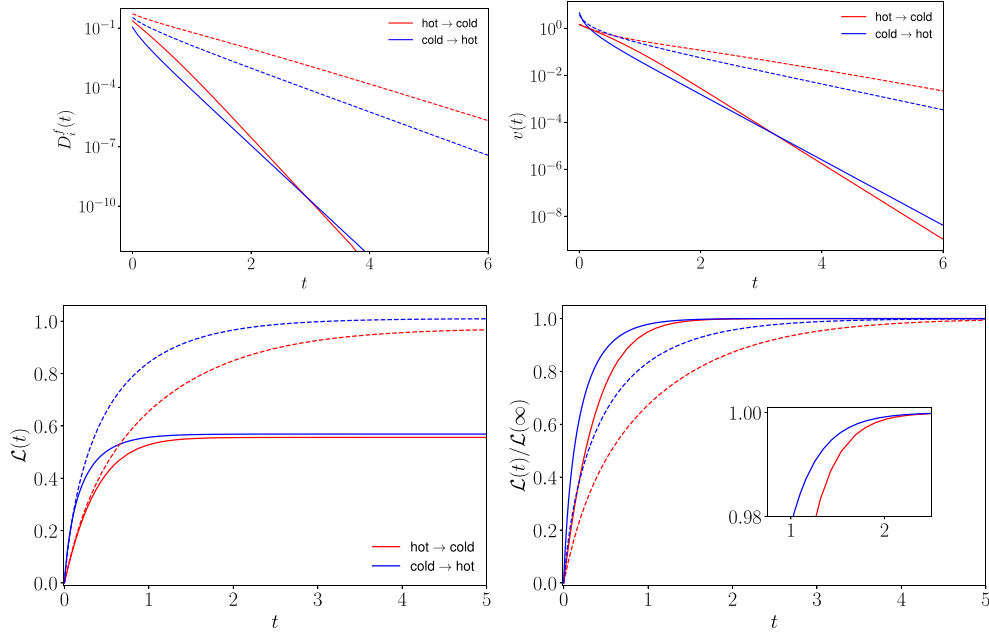
**Figure 7.** Top panel: stochastic matrix eigenvalues as a function of the index  $k$  for the spin system with Hamiltonian as given in equation (20), with  $M = 8$  spins,  $J = 0.5$ ,  $\beta_- = 1.4$ ,  $\beta_+ = 0.9$ , for two values of the external field  $h$  and BB rates, equation (11). Bottom panel: difference between the eigenvalues of the stochastic matrix with inverse temperature  $\beta_+$  and  $\beta_-$ .

#### 4. Conclusions

Our findings suggest that for physical systems with a detailed model of system-bath interaction, the heating or the cooling process can become faster, depending on the choice of system parameters. In particular choosing large energy gaps leads to slower relaxation when in contact with the hot bath, compared to the case when the equilibration occurs in the cold bath.

In this sense the single spin and the harmonic oscillator seem to represent exceptions and do not exhibit a crossover between these different behaviors. For these two systems the heating is always the faster process, as they are characterized by one or two characteristic times, which do not exhibit inversion of the order of eigenvalues describing convergence to equilibrium, and in the case of the harmonic oscillator are independent of the temperature. As such, the heating and the cooling process for these systems turn out to be intrinsically symmetric.

The use of the statistical length appears problematic in this context: while it is tempting to base a general formulation for a thermal kinematics theory on such a quantity, it turns out to be intrinsically asymmetric for two opposite simple, yet general, processes. This, in turn, signals that two opposite stochastic processes can be dominated by inherently different pathways. Furthermore, if one is interested in a measure of the completion of equilibration of a stochastic process, the KL divergence and the statistical velocity (or the Fisher information) seem to provide the same information starting from the intermediate time range, as they are



**Figure 8.** KL divergence  $D_{-}^{+}(t)$  and  $D_{+}^{-}(t)$  (equation (5), top left), statistical velocity  $v(t)$  (equation (8), top right), statistical velocity  $\mathcal{L}(t)$  (equation (9), bottom left), and degree of completion  $\mathcal{L}_{i}^{f}(t)/\mathcal{L}_{i}^{f}(\infty)$  (bottom right) for the spin system in equation (20) with  $M = 8$  spins,  $J = 0.5$ ,  $\beta_{-} = 1.4$ ,  $\beta_{+} = 0.9$ ,  $h = 1.2$  (full lines), and  $h = 0.5$  (dashed lines), BB. In the log-linear plots of  $D(t)$  and  $v(t)$  the slopes in the linear regions are  $2\lambda_{1}$  and  $\lambda_{1}$ , respectively (lines not shown), as predicted from long-time analysis discussed in appendix B.

both dominated by the second largest eigenvalue of the stochastic matrix. Thus in this range, where few slow processes dominates the convergence to equilibrium, there is no specific reason to prefer one indicator over another.

We emphasize that, while throughout this paper we have used the energy gaps as adjustable parameters to tune the inversion of the time scales, decreasing the temperatures  $T_{+}$  and  $T_{-}$  by the same factor is equivalent to increase the size of the gaps appearing in the stochastic matrices.

In the opinion of the author of the present paper the analysis of the entire spectrum of the stochastic matrix (or of the Fokker–Planck operator when relevant) represents the safest and more detailed method to assess whether one process is *faster* than the other, as it gives a way of comparing all the time scales in a transparent manner. This approach iii) can then be added up to the two introduced in section 1. Both the KL divergence and the Fisher information (and its related quantities) while useful and widely employed quantities in stochastic thermodynamics, represent global rather than detailed evaluators of the convergence to equilibrium, depending non-trivially on both the initial conditions and the stochastic matrix eigenvalues.

Thermalization of stochastic systems with discrete phase space was also considered in a recent paper [6], where the concept of thermomajorization was applied to the study of Mpemba effect. Interestingly, the thermomajorization criterion to evaluate the convergence to equilibrium is independent of the chosen monotone measure, such as the KL divergence considered in the present paper. Assessing the convergence to equilibrium of the cooling and

heating processes with the thermomajorization criterion would be an interesting extension of the present work.

As an extension of the present study it would also be interesting to examine systems with continuous phase space described by a Fokker–Planck equation as considered in [9], but with non-linear potentials in order to check whether the symmetry in the statistical length also disappears.

While the formalism of the present paper is the one of classical stochastic systems, we believe that the results presented are relevant also in quantum regime, where the dynamics is described by a quantum master equation and the energy levels are quantized. Indeed for a *global* quantum master equation with jump operators connecting eigenstates of the system Hamiltonian  $H$ , if the initial state is one of such eigenstates, the dynamics reduces to the one described by the classical master equation considered here, see, e.g. [14, 15].

Finally, the present theoretical study can be subject to experimental verification by using, e.g. a colloidal particle in a ladder-like potential, as realized with the setup discussed in [16] or [17].

### Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://github.com/albimparato/Thermalization>.

### Acknowledgments

I am grateful to Sascha Wald for interesting discussions at the beginning of this project. I am also grateful to Aljaz Godec for insightful comments, in particular on the convergence of the average energy and stochastic entropy.

### Appendix A. Initial values of the KL divergence

We want to prove the inequality

$$D_{+}^{-}(0) - D_{-}^{+}(0) = \sum_k p_{+,k}^{eq} \log \left( \frac{p_{+,k}^{eq}}{p_{-,k}^{eq}} \right) - p_{-,k}^{eq} \log \left( \frac{p_{-,k}^{eq}}{p_{+,k}^{eq}} \right) \geq 0$$

iff  $T_{+} \geq T_{-}$ . (A1)

One finds

$$\begin{aligned} D_{-}^{+}(0) &= \sum_k p_{-,k}^{eq} \log \left( \frac{p_{-,k}^{eq}}{p_{+,k}^{eq}} \right) \\ &= \sum_k \left[ p_{-,k}^{eq} (\beta_{+} - \beta_{-}) \epsilon_k \right] + \log Z_{+} - \log Z_{-} \\ &= (\beta_{+} - \beta_{-}) \langle \epsilon_k \rangle_{-} + \log Z_{+} - \log Z_{-}, \end{aligned} \quad (A2)$$

and similarly

$$\begin{aligned} D_{+}^{-}(0) &= \sum_k p_{+,k}^{eq} \log \left( \frac{p_{+,k}^{eq}}{p_{-,k}^{eq}} \right) \\ &= (\beta_{-} - \beta_{+}) \langle \epsilon_k \rangle_{+} + \log Z_{-} - \log Z_{+}, \end{aligned} \quad (A3)$$

and thus

$$D_+^-(0) - D_-^+(0) = (\beta_- - \beta_+) (\langle \epsilon_k \rangle_+ - \langle \epsilon_k \rangle_-) + 2 \log(Z_-/Z_+) \quad (\text{A4})$$

We treat the above difference as a function of  $\beta_+ = 1/T_+$  with  $\beta_- = 1/T_-$  a constant parameter, and introduce

$$\begin{aligned} f(\beta_+) &= D_+^-(0) - D_-^+(0) \\ &= \sum_k p_{+,i}^{eq} \log \left( \frac{p_{+,k}^{eq}}{p_{-,k}^{eq}} \right) - p_{-,k}^{eq} \log \left( \frac{p_{-,k}^{eq}}{p_{+,k}^{eq}} \right); \end{aligned} \quad (\text{A5})$$

and notice that

$$f(\beta_+ = \beta_-) = 0. \quad (\text{A6})$$

We also notice that

$$\begin{aligned} \partial_{\beta_+} f(\beta_+) &= \sum_k \left( \frac{p_{-,k}^{eq}}{p_{+,k}^{eq}} - \log \frac{p_{-,k}^{eq}}{p_{+,k}^{eq}} \right) \partial_{\beta_+} p_{+,k}^{eq} \\ &= \sum_k \left( \frac{p_{-,k}^{eq}}{p_{+,k}^{eq}} - \log \frac{p_{-,k}^{eq}}{p_{+,k}^{eq}} \right) (-\epsilon_k - \partial_{\beta_+} \ln Z_+) p_{+,k}^{eq} = \\ &= \sum_k \left( \frac{p_{-,k}^{eq}}{p_{+,k}^{eq}} + (\beta_- - \beta_+) \epsilon_k - \log \frac{Z_+}{Z_-} \right) (-\epsilon_k + \langle E \rangle_+^{eq}) p_{+,k}^{eq} \\ &= \langle E \rangle_+^{eq} - \langle E \rangle_-^{eq} - (\beta_- - \beta_+) (\langle E^2 \rangle_+^{eq} - (\langle E \rangle_+^{eq})^2), \end{aligned} \quad (\text{A7})$$

and finally

$$\begin{aligned} \frac{\partial^2}{\partial \beta_+^2} f(\beta_+) &= \partial_{\beta_+} \langle E \rangle_+^{eq} + \langle E^2 \rangle_+^{eq} - (\langle E \rangle_+^{eq})^2 \\ &\quad + (\beta_+ - \beta_-) \partial_{\beta_+} (\langle E^2 \rangle_+^{eq} - (\langle E \rangle_+^{eq})^2) \\ &= (\beta_+ - \beta_-) \partial_{\beta_+} (\langle E^2 \rangle_+^{eq} - (\langle E \rangle_+^{eq})^2). \end{aligned} \quad (\text{A8})$$

We now show that the second derivative of  $f(\beta_+)$  is always non-negative. By noticing that  $(\langle E^2 \rangle_+^{eq} - (\langle E \rangle_+^{eq})^2) = T_+^2 C_+$ , where  $C_+$  is the specific heat, we obtain

$$\partial_{\beta_+} (\langle E^2 \rangle_+^{eq} - (\langle E \rangle_+^{eq})^2) = -\frac{2}{\beta_+^3} C_+ + \frac{1}{\beta_+^2} \partial_{\beta_+} C_+, \quad (\text{A9})$$

which is negative given that  $\partial_{\beta_+} C_+ < 0$  for thermodynamic stability. Thus from equations (A8) and (A9) we conclude that

$$\frac{\partial^2}{\partial \beta_+^2} f(\beta_+) \geq 0 \quad \text{iff } \beta_+ \leq \beta_-, \quad (\text{A10})$$

the equality holding for equal temperatures. The last result tells us that  $f'(\beta_+)$ , as given by equation (A7), is an increasing function of  $\beta_+$  for  $\beta_+ < \beta_-$ , and by noticing that  $f'(\beta_+ = \beta_-) = 0$ , we conclude that  $f'(\beta_+) \leq 0$  for  $\beta_+ \leq \beta_-$ .

We now consider the function  $f(\beta_+)$  as given by equation (A5), and noticing that  $f(\beta_+ = \beta_-) = 0$  and that  $f(\beta_+)$  is decreasing for  $\beta_+ < \beta_-$ , we conclude that  $f(\beta_+) \geq 0$  for  $\beta_+ \leq \beta_-$ , which proves equation (A1).

It is worth noticing that throughout this paper we are considering finite temperatures, and non-critical systems.

## Appendix B. Expansion in eigenfunctions and long time behavior of the KL divergence and Fisher Information

Here we review the expansion of the solution of a master equation in the eigenvectors of the corresponding stochastic matrix, as discussed in, e.g. [11].

Given a stochastic matrix  $\mathbf{W}$  with eigenvalues  $\lambda_n$  and eigenvectors  $\Phi_n$ , the components of the vector  $\mathbf{p}$  solution to the master equation (1) read

$$p_k(t) = \Phi_{0,k} + \sum_{n=1}^{N-1} c_n e^{\lambda_n t} \Phi_{n,k}, \quad (\text{B1})$$

with the scalar product defined as

$$(\Phi_n, \Phi_m) = \sum_k \frac{\Phi_{n,k} \Phi_{m,k}}{\Phi_{0,k}} = \delta_{n,m}, \quad (\text{B2})$$

and

$$c_m = (\mathbf{p}(0), \Phi_m). \quad (\text{B3})$$

Here the letter  $n$  ( $m$ ) enumerates the eigenvector corresponding to a given eigenvalue  $\lambda_n$  ( $\lambda_m$ ), while the letter  $k$  enumerates the component of the vector  $\mathbf{p}$ . We remind the reader that the eigenvalues are enumerated in decreasing order  $\lambda_0 = 0 > \lambda_1 \geq \lambda_2 \dots \lambda_{N-1}$ .

We notice that the eigenvectors  $\Phi_n$  and eigenvalues  $\lambda_n$  of  $\mathbf{W}$  only depend on the temperature of the bath governing the dynamics, and for the case considered in this paper it is the final temperature  $\beta_f$ . This implies that in the above expansion (B1) the first eigenvector is

$$\Phi_0 = \mathbf{p}^{eq}(\beta_f). \quad (\text{B4})$$

We also notice that throughout this paper for the initial state we have chosen  $\mathbf{p}(0) = \mathbf{p}^{eq}(\beta_i)$ . Thus the coefficients  $c_m$  introduced in (B3) depend on both  $\beta_i$  and  $\beta_f$ .

We now consider the KL divergence for the two possible processes introduced in equation (5), and obtain

$$D(\mathbf{p}_i^f(t) || \mathbf{p}_f^{eq}) = \sum_k p_k(t) [\ln p_k(t) - \ln \Phi_{0,k}] \quad (\text{B5})$$

In order to study the long time behavior of this quantity we first consider

$$\begin{aligned} \log p_k(t) &= \log \Phi_{0,k} + \log \left( 1 + \sum_{n=1}^{N-1} c_n e^{\lambda_n t} \frac{\Phi_{n,k}}{\Phi_{0,k}} \right) \\ &= \log \Phi_{0,k} + \sum_{l=1}^{\infty} \frac{(-1)^{l+1}}{l} \left( \sum_{n=1}^{N-1} c_n e^{\lambda_n t} \frac{\Phi_{n,k}}{\Phi_{0,k}} \right)^l, \end{aligned} \quad (\text{B6})$$

Now in the long time limit we can approximate the above expression by only taking the term  $l = 1$  in the Taylor series, giving

$$\log p_k(t) \simeq \log \Phi_{0,k} + \sum_{n=1}^{N-1} c_n e^{\lambda_n t} \frac{\Phi_{n,k}}{\Phi_{0,k}}, \quad (\text{B7})$$

thus we can write

$$\begin{aligned} D(\mathbf{p}_i^f(t) \parallel \mathbf{p}_f^{eq}) &\simeq \sum_k \left( \sum_{n=1}^{N-1} c_n e^{\lambda_n t} \Phi_{n,k} \sum_{m=1}^{N-1} c_m e^{\lambda_m t} \frac{\Phi_{m,k}}{\Phi_{0,k}} \right) \\ &= \sum_{n=1}^{N-1} c_n^2 e^{2\lambda_n t} \end{aligned} \quad (\text{B8})$$

which is clearly dominated by the second eigenvalue  $\lambda_1$ . In obtaining the last result we have used the orthonormality of the eigenvectors equation (B2).

We are also interested in the Fisher information which in terms of the expansion (B1) can be written as

$$I(t) = \sum_k \frac{(\partial_t p_k(t))^2}{p_k(t)} = \sum_k \frac{\left( \sum_{n=1}^{N-1} \lambda_n c_n e^{\lambda_n t} \Phi_{n,k} \right)^2}{\Phi_{0,k} + \sum_{n=1}^{N-1} c_n e^{\lambda_n t} \Phi_{n,k}} \quad (\text{B9})$$

and taking the leading terms in the long time limit one obtains

$$\begin{aligned} I(t) &\simeq \sum_k \left( \sum_{n=1}^{N-1} c_n \lambda_n e^{\lambda_n t} \frac{\Phi_{n,k}}{\Phi_{0,k}} \sum_{m=1}^{N-1} c_m \lambda_m e^{\lambda_m t} \Phi_{m,k} \right) \\ &= \sum_{n=1}^{N-1} c_n^2 \lambda_n^2 e^{2\lambda_n t}. \end{aligned} \quad (\text{B10})$$

This expression is also dominated by the second eigenvalue  $\lambda_1$  of the stochastic matrix, and one finds that in the long time limit the following equality holds  $\partial_t^2 D(\mathbf{p}_i^f(t) \parallel \mathbf{p}_f^{eq}) = \dot{I}_i^f(t)/4$ .

### Appendix C. Symmetry of the statistical length

#### C.1. $N = 2$

For the case of  $N = 2$  states, the expression for the statistical length can be obtained analytically. We have

$$p_k(t) = \Phi_{0,k} + c_1 e^{\lambda_1 t} \Phi_{1,k}, \quad i = 0, 1 \quad (\text{C1})$$

with  $\lambda_1 = -w^\dagger(1 + \exp(-\beta\epsilon))$ ,  $\Phi_1 = (-1, 1)/\mathcal{N}_1$ ,  $\mathcal{N}_1 = 2 \cosh(\beta\epsilon/2)$ . The Fisher information reads

$$\begin{aligned} I(t) &= \sum_k \frac{(\partial_t p_k(t))^2}{p_k(t)} = \sum_k \frac{(c_1 \lambda_1 e^{-\lambda_1 t} \Phi_{1,k})^2}{\Phi_{0,k} + c_1 e^{-\lambda_1 t} \Phi_{1,k}} \\ &= \frac{\lambda_1^2 x_1^2(t)}{(\Phi_{0,0} - x_1(t))(\Phi_{0,1} + x_1(t))}, \end{aligned} \quad (\text{C2})$$

with

$$x_1(t) = \frac{c_1 e^{-\lambda_1 t}}{\mathcal{N}_1}. \quad (\text{C3})$$

Thus we have for the statistic length

$$\begin{aligned} \mathcal{L}(t) &= \int_0^t \frac{\lambda_1 |x_1(\tau)| d\tau}{\sqrt{(\Phi_{0,0} - x_1(t))(\Phi_{0,1} + x_1(t))}} \\ &= \pm \int_0^t \frac{\lambda_1 x_1(\tau) d\tau}{\sqrt{(\Phi_{0,0} - x_1(t))(\Phi_{0,1} + x_1(t))}} \\ &= \pm \int_{x_1(t)}^{x_1(0)} \frac{dx}{\sqrt{(\Phi_{0,0} - x)(\Phi_{0,1} + x)}} \end{aligned} \quad (\text{C4})$$

where the sign in the last expression depends on the sign of  $c_1$  in equation (C3).

From equation (B3) we obtain

$$\begin{aligned} c_1(- \rightarrow +) &= (\mathbf{p}^{eq}(\beta_-), \Phi_1(\beta_+)) = e^{-\epsilon\beta_-/2} \frac{e^{\epsilon(\beta_- - \beta_+)} - 1}{1 + e^{-\epsilon\beta_+}} > 0 \\ c_1(+ \rightarrow -) &= (\mathbf{p}^{eq}(\beta_+), \Phi_1(\beta_-)) = e^{-\epsilon\beta_+/2} \frac{e^{-\epsilon(\beta_- - \beta_+)} - 1}{1 + e^{-\epsilon\beta_-}} < 0 \end{aligned}$$

and thus

$$\begin{aligned} \mathcal{L}_+^-(t) &= 2 \left[ \arcsin \sqrt{\frac{e^{\epsilon\beta_-} + e^{\epsilon(\beta_- + \beta_+)} + e^{-\lambda_1^- t} (e^{\epsilon\beta_+} - e^{\epsilon\beta_-})}{(1 + e^{\epsilon\beta_+})(1 + e^{\epsilon\beta_-})}} - \arcsin \left( (1 + e^{-\epsilon\beta_+})^{-1/2} \right) \right], \\ \mathcal{L}_-^+(t) &= -2 \left[ \arcsin \sqrt{\frac{e^{\epsilon\beta_+} + e^{\epsilon(\beta_+ + \beta_-)} + e^{-\lambda_1^+ t} (e^{\epsilon\beta_-} - e^{\epsilon\beta_+})}{(1 + e^{\epsilon\beta_+})(1 + e^{\epsilon\beta_-})}} - \arcsin \left( (1 + e^{-\epsilon\beta_-})^{-1/2} \right) \right], \end{aligned}$$

which simplifies as

$$\begin{aligned} \mathcal{L}_+^-(t) &= 2 \left[ \arctan \left( \frac{(1 + e^{\epsilon\beta_+})(1 + e^{\epsilon\beta_-})}{1 + e^{\epsilon\beta_+} + e^{-\lambda_1^- t} (e^{\epsilon\beta_-} - e^{\epsilon\beta_+})} - 1 \right)^{1/2} - \arctan \left( e^{\epsilon\beta_+/2} \right) \right], \\ \mathcal{L}_-^+(t) &= -2 \left[ \arctan \left( \frac{(1 + e^{\epsilon\beta_+})(1 + e^{\epsilon\beta_-})}{1 + e^{\epsilon\beta_-} + e^{-\lambda_1^+ t} (e^{\epsilon\beta_+} - e^{\epsilon\beta_-})} - 1 \right)^{1/2} - \arctan \left( e^{\epsilon\beta_-/2} \right) \right], \end{aligned}$$

thus exhibiting the symmetry  $\mathcal{L}_+^-(\infty) = \mathcal{L}_-^+(\infty)$ .

We also notice that from equation (C2), one obtains the following inequality for the initial values of the Fisher information

$$I_-^+(0) - I_+^-(0) = e^{-2(\beta_+ + \beta_-)\epsilon} (e^{\beta_- \epsilon} - e^{\beta_+ \epsilon})^3 \geq 0 \quad (\text{C5})$$

for  $\beta_- \geq \beta_+$ .

## C.2. $N > 2$

For a system with an arbitrary number of states in general there is not the symmetry  $\mathcal{L}_+^-(\infty) = \mathcal{L}_-^+(\infty)$ . This is corroborated by the numerical findings of the main text. In the following we

show that the symmetry of the statistical length is restored in the limit of small  $\epsilon$ . In this limit we notice that we can write

$$\frac{p_k(0)}{\Phi_{0,i}} = \left[ 1 + \left( \frac{N-1}{2} - i \right) \epsilon (\beta_i - \beta_f) \right] + \mathcal{O}(\epsilon^2). \quad (\text{C6})$$

Introducing the expansions in  $\epsilon$ :  $\lambda_n = \lambda_n^{(0)} + \epsilon \lambda_n^{(1)} + \dots$  and  $\Phi_{n,i} = \Phi_{n,i}^{(0)} + \epsilon \Phi_{n,i}^{(1)} + \dots$ , from equations (B2) and (B3) we obtain for  $c_n$

$$\begin{aligned} c_n &= \sum_k \Phi_{n,k} + \epsilon (\beta_i - \beta_f) \sum_k \left( \frac{N-1}{2} - k \right) \Phi_{n,k} + \dots \\ &= \epsilon (\beta_i - \beta_f) \sum_k \left( \frac{N-1}{2} - k \right) \Phi_{n,k}^{(0)} + \mathcal{O}(\epsilon^2), \end{aligned} \quad (\text{C7})$$

where we have used the orthonormality condition equation (B2). Using this last result to analyze equation (B9), one realizes that the leading term in  $\epsilon$  in the Fisher information arises from the expansion of  $c_n$  (C7), and obtains the result

$$I = \epsilon^2 (\beta_i - \beta_f)^2 \left( \sum_n \tilde{c}_n \lambda_n^{(0)} e^{-\lambda_n^{(0)} t} \right)^2 + \mathcal{O}(\epsilon^3), \quad (\text{C8})$$

where

$$\tilde{c}_n = \sum_k k \cdot \Phi_{n,k}^{(0)}. \quad (\text{C9})$$

where neither  $\tilde{c}_n$  nor  $\lambda_n^{(0)}$  depend on the temperature.

Thus the statistical length for  $t \rightarrow \infty$  reads

$$\begin{aligned} \mathcal{L}(\infty) &= \int_0^\infty d\tau \sqrt{I(\tau)} \\ &= \epsilon |\beta_i - \beta_f| \int_0^\infty d\tau \left| \sum_n \tilde{c}_n \lambda_n^{(0)} e^{-\lambda_n^{(0)} \tau} \right| + \mathcal{O}(\epsilon^2). \end{aligned} \quad (\text{C10})$$

We conclude that to the lowest order in  $\epsilon$  the statistical length is symmetric. This is in agreement with the finding of [9] that for the classical harmonic oscillator with Brownian dynamics  $\mathcal{L}_+^-(\infty) = \mathcal{L}_-^+(\infty)$ . In the main text we provide numerical evidence that for finite  $\epsilon$  this symmetry disappears.

#### Appendix D. Characteristic polynomial and eigenvalues of the stochastic matrix (12)

The characteristic polynomial of the  $N \times N$  stochastic matrix (12) can be obtained through the Chebyshev recursion relations for tridiagonal matrices, resulting in the expression

$$\begin{aligned} P_N(\lambda) &= \frac{(-1)^N \lambda}{2^N \sqrt{(a+\lambda)^2 - 4bc}} \\ &\times \left[ \left( a + \lambda + \sqrt{(a+\lambda)^2 - 4bc} \right)^N - \left( a + \lambda - \sqrt{(a+\lambda)^2 - 4bc} \right)^N \right], \end{aligned}$$

with  $b = w^\uparrow$ ,  $c = w^\downarrow$ , and  $a = b + c$ . One immediately sees that  $\lambda = 0$  is an eigenvalue. A tedious but straightforward calculation leads to the remaining  $N - 1$  eigenvalues as given by equation (16).

## ORCID iD

Alberto Imparato  [0000-0002-7053-4732](https://orcid.org/0000-0002-7053-4732)

## References

- [1] Mpemba E B and Osborne D G 1969 Cool? *Phys. Educ.* **4** 172
- [2] Klich A and Schulman L S 2019 Mpemba index and anomalous relaxation *Phys. Rev. X* **9** 031003
- [3] Luçon E and Reygnier J 2022 The Mpemba effect in a kinetic model of granular media *J. Stat. Phys.* **186** 27
- [4] Holtzman R, Shapira D and Raz O 2022 The quantum Mpemba effect *Phys. Rev. Res.* **4** L022028
- [5] Das S, Dhar A, Saito K and Agarwalla B K 2022 A quantum thermodynamic route to the Mpemba effect *Phys. Rev. Lett.* **128** 080603
- [6] Van Vu T and Hayakawa H 2025 Thermomajorization Mpemba effect *Phys. Rev. Lett.* **134** 107101
- [7] Ito S and Dechant A 2020 Stochastic time evolution, information geometry and the cramer-rao bound *Phys. Rev. X* **10** 021056
- [8] Bettmann L P and Goold J 2025 Information geometry approach to quantum stochastic thermodynamics *Phys. Rev. E* **111** 014133
- [9] Ibáñez M, Dieball C, Lasanta A, Godec A and Rica R A 2024 Heating and cooling are fundamentally asymmetric and evolve along distinct pathways *Nat. Phys.* **20** 135
- [10] Tejero A, Sánchez R, Kaoutit L E, Manzano D and Lasanta A 2025 Asymmetries of thermal processes in open quantum systems *Phys. Rev. Res.* **7** 023020
- [11] van Kampen N G 1981 *Stochastic Processes in Physics and Chemistry* (North-Holland)
- [12] Breuer H P and Petruccione F 2002 *The Theory of Open Quantum Systems* (Oxford University Press)
- [13] Zwanzig R 2001 *Nonequilibrium Statistical Mechanics* (Oxford University Press, USA)
- [14] Hewgill A, De Chiara G and Imparato A 2021 Quantum thermodynamically consistent local master equations *Phys. Rev. Res.* **3** 013165
- [15] De Chiara G and Imparato A 2022 Quantum fluctuation theorem for dissipative processes *Phys. Rev. Res.* **4** 023230
- [16] Toyabe S, Sagawa T, Ueda M, Muneyuki E and Sano M 2010 Experimental demonstration of information-to-energy conversion and validation of the generalized jarzynski equality *Nat. Phys.* **6** 988
- [17] Lutz E and Ciliberto S 2015 Information: from Maxwell's demon to Landauer's eraser *Phys. Today* **68** 30