

<https://helda.helsinki.fi>

Cutoff Thermalization for Ornstein-Uhlenbeck Systems with
Small Levy Noise in the Wasserstein Distance : Cutoff
pý Thermalization for Ornstein Uhlenbeck Systems.

Barrera, Gerardo

2021-09

Barrera , G , Högele , M A & Pardo , J C 2021 , ' Cutoff Thermalization for
Ornstein-Uhlenbeck Systems with Small Levy Noise in the Wasserstein Distance : Cutoff
pý Thermalization for Ornstein Uhlenbeck Systems... ' , Journal of Statist
, no. 3 , 27 . <https://doi.org/10.1007/s10955-021-02815-0>

<http://hdl.handle.net/10138/333819>

<https://doi.org/10.1007/s10955-021-02815-0>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.



Cutoff Thermalization for Ornstein–Uhlenbeck Systems with Small Lévy Noise in the Wasserstein Distance

G. Barrera¹ · M. A. Högele² · J. C. Pardo³

Received: 2 January 2021 / Accepted: 16 August 2021
© The Author(s) 2021

Abstract

This article establishes *cutoff thermalization* (also known as the *cutoff phenomenon*) for a class of generalized Ornstein–Uhlenbeck systems $(X_t^\varepsilon(x))_{t \geq 0}$ with ε -small additive Lévy noise and initial value x . The driving noise processes include Brownian motion, α -stable Lévy flights, finite intensity compound Poisson processes, and red noises, and may be highly degenerate. *Window cutoff thermalization* is shown under mild generic assumptions; that is, we see an asymptotically sharp $\infty/0$ -collapse of the renormalized Wasserstein distance from the current state to the equilibrium measure μ^ε along a time window centered on a precise ε -dependent time scale t_ε . In many interesting situations such as reversible (Lévy) diffusions it is possible to prove the existence of an explicit, universal, deterministic *cutoff thermalization profile*. That is, for generic initial data x we obtain the stronger result $\mathcal{W}_p(X_{t_\varepsilon+r}^\varepsilon(x), \mu^\varepsilon) \cdot \varepsilon^{-1} \rightarrow K \cdot e^{-qr}$ for any $r \in \mathbb{R}$ as $\varepsilon \rightarrow 0$ for some spectral constants $K, q > 0$ and any $p \geq 1$ whenever the distance is finite. The existence of this limit is characterized by the absence of non-normal growth patterns in terms of an orthogonality condition on a computable family of generalized eigenvectors of \mathcal{Q} . Precise error bounds are given. Using these results, this article provides a complete discussion of the cutoff phenomenon for the classical linear oscillator with friction subject to ε -small Brownian motion or α -stable Lévy flights. Furthermore, we cover the highly degenerate case of a linear chain of oscillators in a generalized heat bath at low temperature.

Communicated by Alessandro Giuliani.

✉ G. Barrera
gerardo.barreravargas@helsinki.fi
M. A. Högele
ma.hoegele@uniandes.edu.co
J. C. Pardo
jcpardo@ciamat.mx

¹ Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, 00014 Helsingin yliopisto, Finland

² Departamento de Matemáticas, Facultad de Ciencias, Universidad de los Andes, Bogotá, Colombia

³ CIMAT, Jalisco S/N, Valenciana, CP 36240 Guanajuato, Guanajuato, Mexico

Keywords The cutoff phenomenon · Abrupt convergence · Ornstein–Uhlenbeck processes · Wasserstein distance · Nonnormal growth · Shift linearity of the Wasserstein distance

1 Introduction

The notion of cutoff thermalization (also known as the cutoff phenomenon or abrupt thermalization in the literature) has gained growing attention in recent years in the physics literature with applications to quantum Markov chains [72], chemical kinetics [8], quantum information processing [73], the Ising model [79], coagulation-fragmentation equations [83,84], dissipative quantum circuits [67] and open quadratic fermionic systems [104]. The term “cutoff” was originally coined in 1986 by Aldous and Diaconis in their celebrated paper [4] on card shuffling, where they observed and conceptualized the asymptotically abrupt collapse of the total variation distance between the current state of their Markov chain and the uniform limiting distribution at a precise deterministic time scale.

At this point we refrain from giving a full account on the mathematical literature on the cutoff phenomenon and refer to the overview article [41] and the introduction of [16]. Standard references in the mathematics literature on the cutoff phenomenon for discrete time and space include [1,3,4,9,15,18,32,33,40,42,74–78,80,103,109]. As introductory texts on the cutoff phenomenon in discrete time and space we recommend [68] and Chapter 18 in the monograph [78].

Although shown to be present in many important Markov chain models, cutoff thermalization is not universal. For instance, for reversible Markov chains Y. Peres formulated the widely used *product condition*, that is, the divergence of the product between the mixing time and the spectral gap for growing dimension, see introduction of [57]. The product condition is a necessary condition for pre-cutoff in total variation (see Proposition 18.4 in [78]), and a necessary and sufficient condition for cutoff in the L^2 sense (see [32]). This condition can be used to characterize cutoff for a large class of Markov chains, but it fails in general, see Chapter 18 in [78] for the details. The alternative condition that the product of the spectral gap and the maximal (expected) hitting time diverges is studied in [2] and [[56], Theorem 1]. In [17] p. 1454 the authors explain the limitations of the hitting time approach to characterize cutoff in general. To the best of our knowledge, there is no well-developed general theory as in the reversible case. However, Theorem 1.1 in [76] yields an abstract sufficient hitting time condition for the detection of the cutoff phenomenon, which is valid in reversible and non-reversible settings, see Section 3, Example 3.4 in [76].

This article establishes just such a criterion for the class of general (reversible and non-reversible) ergodic multidimensional Lévy-driven Ornstein–Uhlenbeck processes in continuous space and time for small noise amplitude with respect to the (Kantorovich-Rubinstein-) Wasserstein-distance. Recall that the classical d -dimensional Ornstein–Uhlenbeck process is given as the unique strong solution of

$$dX_t^\varepsilon = -\mathcal{Q}X_t^\varepsilon dt + \varepsilon dB_t, \quad X_0^\varepsilon = x, \quad \varepsilon > 0, \quad (1.1)$$

where \mathcal{Q} is a square matrix and $B = (B_t)_{t \geq 0}$ a given d -dimensional Brownian motion. For the definitions see for instance [89,93]. The marginal $X_t^\varepsilon(x)$ at a fixed time $t > 0$ has the Gaussian distribution $N(0, \varepsilon^2 \Sigma_t)$, where the covariance matrix Σ_t has an integral representation given in Theorem 3.1 of [98] or Sect. 6.1 of this article. Furthermore, if \mathcal{Q} has eigenvalues with positive real parts, the process $(X_t^\varepsilon(x))_{t \geq 0}$ has the unique limiting distribution $\mu^\varepsilon = N(0, \varepsilon^2 \Sigma_\infty)$, where $\Sigma_\infty = \lim_{t \rightarrow \infty} \Sigma_t$, see Theorem 4.1 and 4.2 in

[98]. Since \mathcal{Q} has full rank, Σ_∞ is known to be invertible. Moreover, the Gaussianity of the marginals and the limiting distribution leads to an explicit formula for the relative entropy

$$H(X_t^\varepsilon(x) \mid \mu^\varepsilon) = \frac{1}{2} \left(\frac{1}{\varepsilon^2} (e^{-\mathcal{Q}t} x)^* \Sigma_\infty^{-1} (e^{-\mathcal{Q}t} x) + \text{Tr}(\Sigma_\infty^{-1} \Sigma_t) - d + \log \frac{\|\Sigma_\infty\|}{\|\Sigma_t\|} \right), \tag{1.2}$$

where $\|\cdot\| = \det$. Note that $\text{Tr}(\Sigma_\infty^{-1} \Sigma_t) - d + \log(\|\Sigma_\infty\|/\|\Sigma_t\|) \rightarrow 0$ for any time scale $t \rightarrow \infty$, thus the first term in formula (1.2) turns out to be asymptotically decisive, when t is replaced by some $t_\varepsilon \rightarrow \infty$ as $\varepsilon \rightarrow 0$. In particular, for a positive multiple of the identity, $\mathcal{Q} = q \cdot I_d$, $q > 0$, and $t_\varepsilon := q^{-1} |\ln(\varepsilon)|$, the following dichotomy holds for any $x \neq 0$:

$$H(X_{\delta \cdot t_\varepsilon}^\varepsilon(x) \mid \mu^\varepsilon) \approx_\varepsilon \frac{1}{2} \left(\frac{|\Sigma_\infty^{-\frac{1}{2}} e^{-\mathcal{Q} \delta \cdot t_\varepsilon} x|}{\varepsilon} \right)^2 \propto \varepsilon^{2(\delta-1)} \xrightarrow{\varepsilon \rightarrow 0} \begin{cases} \infty & \text{for } \delta \in (0, 1) \\ 0 & \text{for } \delta > 1 \end{cases}. \tag{1.3}$$

The discussion of formula (1.3) for a general asymptotically exponentially stable matrix $-\mathcal{Q}$ is given in Sect. 6.1 of this article. The fine study of the dichotomy in (1.3) and its dependence on x for general \mathcal{Q} , is the core of *cutoff thermalization* for relative entropy in the context of continuous time and space. The main shortcoming of formula (1.2) is that it is not robust and hard to generalize to

- (I) general degenerate noise such as the linear oscillator with noise only in the position and
- (II) non-Gaussian white Lévy noise processes or red noise processes, such as α -stable Lévy flights, Poissonian jumps, Ornstein–Uhlenbeck processes, or even deterministic drifts.

Additionally, it is not obvious in general how formula (1.2) would imply an analogous dichotomy to the asymptotics in (1.3) for

- (III) statistically more tractable distances such as the total variation or the Wasserstein distance.

In [14] items (I) and (II) have been addressed for smooth density situations in the technically demanding total variation distance under natural but statistically hardly verifiable regularization conditions. In this article, we study the generalized Ornstein–Uhlenbeck process $X^\varepsilon(x) = (X_t^\varepsilon(x))_{t \geq 0}$ given as the unique strong solution of the linear ordinary stochastic differential equation with additive Lévy noise

$$dX_t^\varepsilon = -\mathcal{Q}X_t^\varepsilon dt + \varepsilon dL_t, \quad X_0^\varepsilon = x, \tag{1.4}$$

with the cutoff parameter $\varepsilon > 0$, where \mathcal{Q} is a general d -dimensional square matrix that has eigenvalues with positive real parts and $L = (L_t)_{t \geq 0}$ is a general (possibly degenerate) Lévy process with values in \mathbb{R}^d . The purpose of this article is twofold. First, it establishes *window cutoff thermalization* in the limit of small ε for the family of processes $(X^\varepsilon(x))_{\varepsilon \in (0, 1]}$ in terms of the renormalized Wasserstein distance whenever the latter is finite and $X^\varepsilon(x)$ has a unique limiting distribution μ^ε for each ε . The notion of window cutoff thermalization turns out to be a refined and robust analogue of the dichotomy (1.3) which addresses the issues (I)–(III) for the renormalized Wasserstein distance, that is, informally, with a limit of the following type

$$\lim_{\varepsilon \rightarrow 0} \mathcal{W}_p(X_{\delta \cdot t_\varepsilon}^\varepsilon(x), \mu^\varepsilon) \cdot \varepsilon^{-1} = \begin{cases} \infty & \text{for } \delta \in (0, 1) \\ 0 & \text{for } \delta > 1 \end{cases}. \tag{1.5}$$

Secondly, we study the stronger notion of a *cutoff thermalization profile*, that is, the existence of the limit for any fixed $r \in \mathbb{R}$

$$\lim_{\varepsilon \rightarrow 0} \mathcal{W}_p(X_{t_\varepsilon+r}^\varepsilon(x), \mu^\varepsilon) \cdot \varepsilon^{-1} = \mathcal{P}_x(r). \tag{1.6}$$

The presence of a cutoff thermalization profile for generic x turns out to be characterized by the absence of non-normal growth effects, that is, the orthogonality of asymptotic ($t \rightarrow \infty$) generalized eigenvectors of the exponential matrix $e^{-\mathcal{Q}t}$. In [12, 14] such limits have been studied and characterized for the total variation distance. The limit there, however, turns out to be hard to calculate or even to simulate numerically, while in our setting for $p \geq 1$ the limit (1.6) is shown to take the elementary explicit shape

$$\mathcal{P}_x(r) = K_x \cdot e^{-q_x r}, \quad r \in \mathbb{R},$$

where the positive constants K_x and q_x in general depend on the initial condition x . For generic values of x , that is, x having a non-trivial projection on one of the eigenspaces of the eigenvalues of \mathcal{Q} with smallest real part and highest multiplicity, it turns out to be the spectral gap of \mathcal{Q} . In addition, our normal growth characterization is applicable in concrete examples of interest such as the linear oscillator. The Markovian dynamics of (1.4) implies (whenever regularity assumptions, such as hypoellipticity, are satisfied) that the probability densities p_t^ε of the marginals $X_t^\varepsilon(x)$ are governed by the Fokker–Planck or master equation

$$\partial_t p_t^\varepsilon = (\mathcal{A}^\varepsilon)^* p_t^\varepsilon,$$

where the generator \mathcal{A}^ε in general amounts to a full-blown unbounded linear integro-differential operator. Therefore state-of-the-art analytic methods, at best, are capable of studying the spectrum of \mathcal{A}^ε (numerically), which yield an upper bound for exponential convergence to the equilibrium μ^ε for sufficiently large time in the case of the spectrum lying in the left open complex half-plane. See for instance [88] Section “Hypoelliptic Ornstein–Uhlenbeck semigroups” or Theorem 3.1 in [10]. However, these types of results can only establish (qualitative) upper bounds, which do not reflect the real convergence of p_t^ε to the equilibrium distribution μ^ε . It is with more flexible probabilistic techniques (coupling or replica) that it is possible to show cutoff thermalization in this level of generality.

The first work on cutoff thermalization covering certain equations of the type (1.1) is by Barrera and Jara [12] in 2015 for scalar nonlinear dissipative SDEs with a stable state and ε -small Brownian motion in the unnormalized total variation distance d_{TV} using coupling techniques. The authors show that for this natural ($d = 1$) gradient system, there always is a cutoff thermalization profile which can be given explicitly in terms of the Gauss error function. The follow-up work [13] covers cutoff thermalization with respect to the total variation distance for (1.1) in higher dimensions, where the picture is considerably richer, due to the presence of strong and complicated rotational patterns. Window cutoff thermalization is proved for the general case. In addition, the authors precisely characterize the existence of a cutoff thermalization profile in terms of the omega limit sets appearing in the long-term behavior of the matrix exponential function $e^{-\mathcal{Q}t}x$ in Lemma B.2 [13], which plays an analogous role in this article. We note that in (1.1) and [13] the Brownian perturbation is nondegenerate, and hence the examples of the linear oscillator or linear chains of oscillators subject to small Brownian motion are not covered there. The results of [14] mentioned above cover cutoff thermalization for (1.4) for nondegenerate noise dL in the total variation distance and yield many important applications such as the sample processes and the sample mean process. The proof methods are based on concise Fourier inversion techniques. Due to the mentioned regularity issue concerning the total variation distance the authors state their results under the hypothesis of continuous densities of the marginals, which to date is mathematically not characterized in simple terms. Their profile function is naturally given as a shift error of the Lévy–Ornstein–Uhlenbeck limiting measure for $\varepsilon = 1$ and measured in the total variation distance. These quantities are theoretically highly insightful, but almost impossible

to calculate and simulate in examples. Their abstract characterization of the existence of a cutoff-profile given in [13], which assesses the behavior of the mentioned profile function on a suitably defined omega limit set, is shown to be also valid in our setting (see Theorem 3.3).

While the total variation distance with which the cutoff phenomenon was originally stated is equivalent to the convergence in distribution in finite spaces, it is much more difficult to analyze in continuous space and is not robust to small non-smooth perturbations. There have also been attempts to describe the cutoff phenomenon for quantum systems in other types of metrics such as the trace norm, see for instance [72]. In this context the Wasserstein setting of the present article has the following four advantages in contrast to the original total variation distance.

(1) It does not require any regularity except some finite p th moment, $p > 0$. This allows us to treat degenerate noise and to cover second order equations. As an illustration we give a complete discussion of cutoff thermalization of the damped linear oscillator in the Wasserstein distance subject to Brownian motion, Poissonian jumps without any regularizing effect, α -stable processes including the Cauchy process and a deterministic perturbation. In the same sense we cover chains of linear oscillators in a generalized heat bath at low temperature.

(2) In contrast to the relative entropy and the total variation distance the Wasserstein distance has the particular property of *shift linearity* for $p \geq 1$, which reduces the rather complicated profile functions of [12–14] to a *simple exponential function* with no need for costly and complex simulation. In addition, the profile is *universal* and does not depend on which Wasserstein distance is applied nor on the statistical properties of the noise. For $p \in (0, 1)$ shift linearity seems not to be feasible, however we give upper and lower bounds which essentially account for the same. Therefore we may cover the case of the linear oscillator under ε -small α -stable perturbations including the Cauchy process for $\alpha = 1$.

(3) We also obtain cutoff thermalization for the physical observable finite p th moments, which cannot be directly deduced from any result in [12–14]. Our findings also naturally extend to small red noise and general ergodic perturbations as explained in Sect. 6.2.

(4) Due to the homogeneity structure of the Wasserstein distance we give meaningful asymptotic error estimates and estimates on the smallness of ε needed in order to observe cutoff thermalization on a finite interval $[0, T]$.

The Wasserstein distance also entails certain minor drawbacks. First, a price to pay is to pass from the unnormalized total variation distance (due to 0-homogeneity $d_{\text{TV}}(\varepsilon U_1, \varepsilon U_2) = d_{\text{TV}}(U_1, U_2)$) to the renormalized Wasserstein distance $\mathcal{W}_p/\varepsilon$. This is fairly natural to expect for any distance based on norms such as the L^p -norm, $p \geq 1$ due to the 1-homogeneity $\mathcal{W}_p(\varepsilon U_1, \varepsilon U_2) = \varepsilon \mathcal{W}_p(U_1, U_2)$. The second issue is that concrete evaluations of the Wasserstein distance are complicated in general. For $d = 1$ and $1 \leq p < \infty$ the Wasserstein distance has the explicit shape of an L^p -distance for the quantiles $F_{U_1}^{-1}$ and $F_{U_2}^{-1}$

$$\mathcal{W}_p(U_1, U_2) = \left(\int_0^1 |F_{U_1}^{-1}(\theta) - F_{U_2}^{-1}(\theta)|^p d\theta \right)^{\frac{1}{p}}.$$

However, there are no known higher dimensional counterparts of this formula. While by definition Wasserstein distances are minimizers of L^p -distances, they are always bounded above by the L^p -distance (by the natural coupling); however, lower bounds are typically hard to establish.

The dynamics of models (1.1) with small Brownian motion have been studied since the early days of Arrhenius [7], Ornstein and Uhlenbeck [85], Eyring [46] Kramers [70]. Since then, an enormous body of physics literature has emerged and we refer to the overview articles [61] on the exponential rates and [53] on the related phenomenon of stochastic

resonance. For an overview on the Ornstein–Uhlenbeck process see [66]. However, in many situations Brownian motion alone is too restrictive for modeling the driving noise, as laid out in the article by Penland and Ewald [91], where the authors identify the physical origins of stochastic forcing and discuss the trade off between Gaussian vs. non-Gaussian white and colored noises. In particular, heavy-tailed Lévy noise has been found to be present in physical systems such as for instance [22,31,43–45,52,101]. In the mathematics literature the dynamics of the exit times of ordinary, delay and partial differential equations with respect to such kind of Brownian perturbations is often referred to as Freidlin–Wentzell theory. It was studied in [19–21,23,26,34,35,47,49,100] and serves as the base on which metastability and stochastic resonance results are derived, for instance in [24,25,48,51,99]. More recent extensions of this literature for the non-Brownian Lévy case often including polynomial instead of exponential rates include [38,55,58–60,62–65] and references therein. A different, recent line of research starting with the works of [27–30,36] treats ε -small and simultaneously $1/\varepsilon$ -intensity accelerated Poisson random measures which yield large deviations for ε -parametrized Lévy processes, also in the context of Lévy processes, where this behavior typically fails to hold true.

The paper is organized as follows. After the setup and preliminary results the cutoff thermalization phenomenon is derived in Sect. 3.1. The main results on the stronger notion of profile cutoff thermalization are presented in Sect. 3.2 followed by the generic results on the weaker notion of window cutoff thermalization in Sect. 3.3. Section 4 is devoted to the applications in physics such as gradient systems and a complete discussion of the linear oscillator and numerical results of a linear Jacobi chains coupled to a heat bath. In Sect. 5 several conceptual examples illustrate certain mathematical features such as the fact that leading complex eigenvalues not necessarily destroy the profile thermalization. Moreover, we highlight the dependence of the thermalization time scale on the initial data x , and Jordan block multiplicities of \mathcal{Q} . In Sect. 6 we discuss the pure Brownian case for relative entropy, the validity of the results for general ergodic driven noises such as red noise and derive conditions on ε for observing the cutoff thermalization on a given finite time horizon. The proofs of the main results are given in the appendix.

2 The Setup

2.1 The Lévy Noise Perturbation dL

Let $L = (L_t)_{t \geq 0}$ be a Lévy process with values in \mathbb{R}^d , that is, a process with stationary and independent increments starting from 0 almost surely, and càdlàg paths (right-continuous with left limits). The most prominent examples are the Brownian motion and the compound Poisson process. For an introduction to the subject we refer to [6,97]. The characteristic function of the marginal L_t has the following (Lévy–Khintchine) representation for any $t \geq 0$

$$\begin{aligned} u &\mapsto \mathbb{E}[e^{i\langle u, L_t \rangle}] \\ &= \exp \left(t \left[i \langle u, b \rangle - \frac{1}{2} \langle \Sigma u, u \rangle + \int_{\mathbb{R}^d} (e^{i\langle u, y \rangle} - 1 - i \langle u, y \rangle \mathbf{1}\{|y| \leq 1\}) \nu(dy) \right] \right), \end{aligned}$$

for a drift vector $b \in \mathbb{R}^d$, Σ a $d \times d$ covariance matrix and ν a sigma-finite measure on \mathbb{R}^d with

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} (1 \wedge |z|^2) \nu(dz) < \infty.$$

Hypothesis 2.1 (Finite p th moment) *For $p > 0$ the Lévy process L has finite p th moments, which is equivalent to*

$$\int_{|z|>1} |z|^p \nu(dz) < \infty,$$

where ν is the Lévy jump measure.

Remark 2.1 (1) In case of $L = B$ being a Brownian motion Hypothesis 2.1 is true for any $p > 0$.

(2) For $p \in (0, 2)$ it also covers the case of α -stable noise for $\alpha \in (p, 2)$. Note that the latter only has moments of order $p < \alpha$ and hence no finite variance.

2.2 The Ornstein–Uhlenbeck Process $(X_t^\varepsilon(x))_{t \geq 0}$

We consider the following Ornstein–Uhlenbeck equation subject to ε -small Lévy noise

$$dX_t^\varepsilon = -\mathcal{Q}X_t^\varepsilon dt + \varepsilon dL_t, \quad X_0^\varepsilon = x, \tag{2.1}$$

where \mathcal{Q} is a deterministic $d \times d$ matrix. For $\varepsilon > 0$ and any $x \in \mathbb{R}^d$ the SDE (2.1) has a unique strong solution. By the variation of constant formula

$$X_t^\varepsilon(x) = e^{-\mathcal{Q}t}x + \varepsilon \int_0^t e^{-\mathcal{Q}(t-s)} dL_s =: e^{-\mathcal{Q}t}x + \varepsilon \mathcal{O}_t, \tag{2.2}$$

where \mathcal{O}_t is a stochastic integral which is defined in our setting by the integration by parts formula

$$\mathcal{O}_t = L_t - \int_0^t e^{-\mathcal{Q}(t-s)} \mathcal{Q}L_s ds.$$

In general, for $t > 0$ the marginals $X_t^\varepsilon(x)$ may not have densities and are only given in terms of its characteristics due to the irregular non-Gaussian jump component, see Proposition 2.1 in [82]. For the case of pure Brownian noise, the marginal $X_t^\varepsilon(x)$ exhibits a Gaussian density. Its mean and covariance matrix are given explicitly in Section 3.7 in [89].

2.3 Asymptotic Exponential Stability of $-\mathcal{Q}$

Hypothesis 2.2 (Asymptotic exponential stability of $-\mathcal{Q}$) *The real parts of all eigenvalues of \mathcal{Q} are positive.*

By formula (2.2) it is clearly seen, that the fine structure of $e^{-\mathcal{Q}t}x$ determines its dynamics. In general, calculating matrix exponentials is complicated. For basic properties and some explicit formulas we refer to [5], Chapters 7.10 and 7.14. Roughly speaking, for symmetric \mathcal{Q} and generic $x \in \mathbb{R}^d, x \neq 0$, the behavior of $e^{-\mathcal{Q}t}x$ is given by $e^{-\lambda t} \langle v, x \rangle v + o(e^{-\lambda t})$ where $\lambda > 0$ is the smallest eigenvalue of \mathcal{Q} and v is its corresponding eigenvector. For asymmetric \mathcal{Q} the picture is considerably blurred by the occurrence of multiple rotations. The complete analysis reads as follows and is carried out in detail in the examples.

Lemma 2.1 *Assume Hypothesis 2.2. Then for any initial value $x \in \mathbb{R}^d, x \neq 0$, there exist a rate $\eta := \eta(x) > 0$, multiplicities $\ell := \ell(x), m := m(x) \in \{1, \dots, d\}$, angles $\theta_1 :=$*

$\theta_1(x), \dots, \theta_m := \theta_m(x) \in [0, 2\pi)$ and a family of linearly independent vectors $v_1 := v_1(x), \dots, v_m := v_m(x)$ in \mathbb{C}^d such that

$$\lim_{t \rightarrow \infty} \left| \frac{e^{qt}}{t^{\ell-1}} e^{-Q_t} x - \sum_{k=1}^m e^{it\theta_k} v_k \right| = 0. \tag{2.3}$$

Moreover,

$$0 < \liminf_{t \rightarrow \infty} \left| \sum_{k=1}^m e^{it\theta_k} v_k \right| \leq \limsup_{t \rightarrow \infty} \left| \sum_{k=1}^m e^{it\theta_k} v_k \right| \leq \sum_{k=1}^m |v_k|. \tag{2.4}$$

The numbers $\{q \pm i\theta_k, k = 1, \dots, m\}$ are eigenvalues of the matrix Q and the vectors $\{v_k, k = 1, \dots, m\}$ are generalized eigenvectors of Q .

The lemma is established as Lemma B.1 in [13], p. 1195-1196, and proved there. It is stated there under the additional hypothesis of coercivity $\langle Qx, x \rangle \geq \delta|x|^2$ for some $\delta > 0$ and any $x \in \mathbb{R}^d$. However, inspecting the proof line by line it is seen that the authors only use Hypothesis 2.2 of the matrix Q . Hence the result is valid under the sole Hypothesis 2.2. For a detailed understanding of the computation of the exponential matrix we refer to the notes of [107], in particular, Theorem 22 and Section 3.

Remark 2.2 The precise properties (2.3) and (2.4) turn out to be crucial for the existence of a cutoff thermalization profile. Note that, in general, the limit

$$\lim_{t \rightarrow \infty} \left| \sum_{k=1}^m e^{it\theta_k} v_k \right|$$

does not exist. However, if in addition Q is symmetric we have $\theta_1 = \dots = \theta_m = 0$ and consequently,

$$\lim_{t \rightarrow \infty} \left| \sum_{k=1}^m e^{it\theta_k} v_k \right| = \left| \sum_{k=1}^m v_k \right| \neq 0.$$

2.4 The Wasserstein Distance \mathcal{W}_p

Given two probability distributions μ_1 and μ_2 on \mathbb{R}^d with finite p th moment for some $p > 0$, we define the Wasserstein distance of order p as follows

$$\mathcal{W}_p(\mu_1, \mu_2) = \inf_{\Pi} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |u - v|^p \Pi(du, dv) \right)^{\min\{1/p, 1\}}, \tag{2.5}$$

where the infimum is taken over all joint distributions (also called couplings) Π with marginals μ_1 and μ_2 . The Wasserstein distance quantifies the distance between probability measures, for an introduction we refer to [106]. For convenience of notation we do not distinguish a random variable U and its law \mathbb{P}_U as an argument of \mathcal{W}_p . That is, for random variables U_1, U_2 and probability measure μ we write $\mathcal{W}_p(U_1, U_2)$ instead of $\mathcal{W}_p(\mathbb{P}_{U_1}, \mathbb{P}_{U_2})$, $\mathcal{W}_p(U_1, \mu)$ instead of $\mathcal{W}_p(\mathbb{P}_{U_1}, \mu)$ etc.

Lemma 2.2 (Properties of the Wasserstein distance) *Let $p > 0, u_1, u_2 \in \mathbb{R}^d$ be deterministic vectors, $c \in \mathbb{R}$ and U_1, U_2 be random vectors in \mathbb{R}^d with finite p th moment. Then we have:*

- (a) *The Wasserstein distance is a metric, in the sense of being definite, symmetric and satisfying the triangle inequality in the sense of Definition 2.15 in [94].*

(b) *Translation invariance:* $\mathcal{W}_p(u_1 + U_1, u_2 + U_2) = \mathcal{W}_p(u_1 - u_2 + U_1, U_2)$.

(c) *Homogeneity:*

$$\mathcal{W}_p(cU_1, cU_2) = \begin{cases} |c| \mathcal{W}_p(U_1, U_2) & \text{for } p \in [1, \infty), \\ |c|^p \mathcal{W}_p(U_1, U_2) & \text{for } p \in (0, 1). \end{cases}$$

(d) *Shift linearity:* For $p \geq 1$ it follows

$$\mathcal{W}_p(u_1 + U_1, U_1) = |u_1|. \tag{2.6}$$

For $p \in (0, 1)$ equality (2.6) is false in general. However we have the following inequality

$$\max\{|u_1|^p - 2\mathbb{E}[|U_1|^p], 0\} \leq \mathcal{W}_p(u_1 + U_1, U_1) \leq |u_1|^p. \tag{2.7}$$

(e) *Domination:* For any given coupling $\tilde{\Pi}$ between U_1 and U_2 it follows

$$\mathcal{W}_p(U_1, U_2) \leq \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |v_1 - v_2|^p \tilde{\Pi}(dv_1, dv_2) \right)^{\min\{1/p, 1\}}.$$

(f) *Characterization:* Let $(U_n)_{n \in \mathbb{N}}$ be a sequence of random vectors with finite p th moments and U a random vector with finite p th moment the following are equivalent:

(1) $\mathcal{W}_p(U_n, U) \rightarrow 0$ as $n \rightarrow \infty$.

(2) $U_n \xrightarrow{d} U$ as $n \rightarrow \infty$ and $\mathbb{E}[|U_n|^p] \rightarrow \mathbb{E}[|U|^p]$ as $n \rightarrow \infty$.

(g) *Contraction:* Let $T : \mathbb{R}^d \rightarrow \mathbb{R}^k$, $k \in \mathbb{N}$, be Lipschitz continuous with Lipschitz constant 1. Then for any $p > 0$

$$\mathcal{W}_p(T(U_1), T(U_2)) \leq \mathcal{W}_p(U_1, U_2). \tag{2.8}$$

The proof of Lemma 2.2 is given in Appendix A.

Remark 2.3 (1) Property d) is less widely known and turns out to be crucial to simplify the thermalization profile for $p \geq 1$ from a complicated stochastic quantity to a deterministic exponential function, while still being useful for $p \in (0, 1)$.

(2) In general, the projection of a vector-valued Markov process to single coordinates is known to be non-Markovian. However, not surprisingly property g) allows to estimate the Wasserstein distance of its projections. This is used in Sect. 6.2 for degenerate systems and mimics the analogous property for the total variation distance given in Theorem 5.2 in [39].

Lemma 2.3 (Wasserstein approximation of the total variation distance) *Let U_1 and U_2 be two random variables taking values on \mathbb{R}^d . Assume that there exists $p \in (0, 1)$ small enough such that U_1 and U_2 possesses finite p th moments. Then*

$$d_{TV}(U_1, U_2) = \lim_{p' \rightarrow 0} \mathcal{W}_{p'}(U_1, U_2).$$

The content of this lemma is announced in Section 2.1 in [86]. The proof is given in Appendix A.

Remark 2.4 Assume that for any $x \neq 0$ and $p \in (0, 1)$ the formula $\mathcal{W}_p(x + \mathcal{O}_\infty, \mathcal{O}_\infty) = |x|^p$ is valid. By Lemma 2.3 we have

$$d_{TV}(x + \mathcal{O}_\infty, \mathcal{O}_\infty) = \lim_{p' \rightarrow 0} \mathcal{W}_{p'}(x + \mathcal{O}_\infty, \mathcal{O}_\infty) = \lim_{p' \rightarrow 0} |x|^{p'} = 1.$$

Hence for any $x \neq 0$, $d_{TV}(x + \mathcal{O}_\infty, \mathcal{O}_\infty) = 1$ which in general false whenever \mathcal{O}_∞ has a continuous positive density in \mathbb{R}^d , for instance, for \mathcal{O}_∞ being α -stable with index $\alpha \in (0, 2]$. In other words, $\mathcal{W}_p(x + \mathcal{O}_\infty, \mathcal{O}_\infty) = |x|^p$ breaks down for p sufficiently small in all smooth density situations.

2.5 Limiting Distribution μ^ε

We fix $\varepsilon > 0$. By Proposition 2.2 in [82], Hypotheses 2.1 and 2.2 yield the existence of a unique equilibrium distribution μ^ε and its characteristics are given there. Moreover, the limiting distribution μ^ε has finite p th moments. It is the distribution of $\varepsilon\mathcal{O}_\infty$, where \mathcal{O}_∞ is the limiting distribution of \mathcal{O}_t as $t \rightarrow \infty$ (with respect to the weak convergence). In fact, it follows the stronger property.

Lemma 2.4 *Let Hypotheses 2.1 and 2.2 be satisfied. Then for any $x \neq 0$, $\varepsilon > 0$ and $0 < p' \leq p$ we have $\mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon) \rightarrow 0$ as $t \rightarrow \infty$.*

Proof First note, there exist positive constants q_* and C_0 such that $|e^{-\mathcal{Q}t}| \leq C_0e^{-q_*t}$ for any $t \geq 0$ due to the usual Jordan decomposition and the estimate

$$|e^{-\mathcal{Q}t}| \leq \left(\sup_{s \geq 0} \max_{0 \leq j \leq d-1} \frac{s^j}{j!} e^{-(q-q_*)s} \right) e^{-q_*t} = C_0e^{-q_*t}, \tag{2.9}$$

where q is the minimum of the real parts of the eigenvalues of \mathcal{Q} . Then for any $t \geq 0$ and $x, y \in \mathbb{R}^d$

$$|X_t^\varepsilon(x) - X_t^\varepsilon(y)|^{p'} \leq |e^{-\mathcal{Q}t}|^{p'} |x - y|^{p'} \leq C_0^{p'} e^{-q_*p't} |x - y|^{p'}.$$

Hence

$$\mathcal{W}_{p'}(X_t^\varepsilon(x), X_t^\varepsilon(y)) \leq \left(C_0e^{-q_*t} |x - y| \right)^{\min\{1, p'\}}.$$

By disintegration of the invariant distribution μ^ε we have

$$\begin{aligned} \mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon) &= \mathcal{W}_{p'}(X_t^\varepsilon(x), X_t^\varepsilon(\mu^\varepsilon)) \leq \int_{\mathbb{R}^d} \mathcal{W}_{p'}(X_t^\varepsilon(x), X_t^\varepsilon(y)) \mu^\varepsilon(dy) \\ &\leq \left(C_0e^{-q_*t} \right)^{\min\{1, p'\}} \int_{\mathbb{R}^d} |x - y|^{\min\{1, p'\}} \mu^\varepsilon(dy) \\ &\leq \left(C_0e^{-q_*t} |x| \right)^{\min\{1, p'\}} + \left(C_0e^{-q_*t} \right)^{\min\{1, p'\}} \int_{\mathbb{R}^d} |y|^{\min\{1, p'\}} \mu^\varepsilon(dy). \end{aligned} \tag{2.10}$$

Since $X_\infty^\varepsilon = \varepsilon\mathcal{O}_\infty$, it follows

$$\int_{\mathbb{R}^d} |y|^{\min\{1, p'\}} \mu^\varepsilon(dy) = \varepsilon^{\min\{1, p'\}} \mathbb{E}[|\mathcal{O}_\infty|^{\min\{1, p'\}}] < \infty.$$

As a consequence, $\limsup_{t \rightarrow \infty} \sup_{|x| \leq R} \mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon) = 0$ for any $R > 0$ and $\varepsilon > 0$. □

Observe that

$$X_t^\varepsilon(x) = e^{-\mathcal{Q}t}x + \varepsilon\mathcal{O}_t, \quad \text{where } \mathcal{O}_t := \int_0^t e^{-\mathcal{Q}(t-s)} dL_s.$$

In particular,

$$\mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty) \leq \left(C_0 e^{-q_* t}\right)^{\min\{1, p'\}} \mathbb{E}[|\mathcal{O}_\infty|^{\min\{1, p'\}}]. \tag{2.11}$$

By the exponential stability hypothesis we have $e^{-\mathcal{Q}t}x \rightarrow 0$ as $t \rightarrow \infty$. Therefore, Slutsky’s theorem yields $X_t^\varepsilon(x) \xrightarrow{d} \varepsilon\mathcal{O}_\infty$ as $t \rightarrow \infty$, where \mathcal{O}_∞ has distribution μ^1 .

Remark 2.5 It is not difficult to see that Hypothesis 2.1 yields for $1 \leq p' \leq p$

$$\begin{aligned} \mathbb{E}[|\mathcal{O}_\infty|^{\min\{1, p'\}}] &\leq |b| \frac{C_0}{q_*} + |\Sigma|^{1/2} \frac{C_0}{\sqrt{2q_*}} \\ &+ \frac{C_0}{q_*} \sqrt{\int_{|z| \leq 1} |z|^2 \nu(dz)} + \exp\left(\frac{C_0}{q_*} \int_{|z| > C_0^{-1}} |z| \nu(dz)\right), \end{aligned} \tag{2.12}$$

where $q_* > 0$ is given at the beginning of the proof of the preceding Lemma 2.4. The proof of the jump part and drift is elementary and given in [108] p.1000-1001. The Brownian component can be easily estimated by the Itô isometry, see for instance [71] Section 5.6.

3 The Main Results

3.1 The Derivation of Cutoff Thermalization

3.1.1 The Key Estimates for $p \geq 1$

Recall that μ^ε has the distribution of $\varepsilon\mathcal{O}_\infty$. For transparency we start with $1 \leq p' \leq p$. On the one hand, by Lemma 2.2 properties a), b), c) and d) we have

$$\begin{aligned} \mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon) &= \mathcal{W}_{p'}(e^{-\mathcal{Q}t}x + \varepsilon\mathcal{O}_t, \varepsilon\mathcal{O}_\infty) \\ &\leq \mathcal{W}_{p'}(e^{-\mathcal{Q}t}x + \varepsilon\mathcal{O}_t, e^{-\mathcal{Q}t}x + \varepsilon\mathcal{O}_\infty) + \mathcal{W}_{p'}(e^{-\mathcal{Q}t}x + \varepsilon\mathcal{O}_\infty, \varepsilon\mathcal{O}_\infty) \\ &= \mathcal{W}_{p'}(\varepsilon\mathcal{O}_t, \varepsilon\mathcal{O}_\infty) + |e^{-\mathcal{Q}t}x| \\ &= \varepsilon\mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty) + |e^{-\mathcal{Q}t}x|. \end{aligned} \tag{3.1}$$

On the other hand, since $p' \geq 1$, property d) in Lemma 2.2 with the help of properties a), b) and c) yields

$$\begin{aligned} |e^{-\mathcal{Q}t}x| &= \mathcal{W}_{p'}(e^{-\mathcal{Q}t}x + \varepsilon\mathcal{O}_\infty, \varepsilon\mathcal{O}_\infty) \\ &\leq \mathcal{W}_{p'}(e^{-\mathcal{Q}t}x + \varepsilon\mathcal{O}_\infty, e^{-\mathcal{Q}t}x + \varepsilon\mathcal{O}_t) + \mathcal{W}_{p'}(e^{-\mathcal{Q}t}x + \varepsilon\mathcal{O}_t, \varepsilon\mathcal{O}_\infty) \\ &= \varepsilon\mathcal{W}_{p'}(\mathcal{O}_\infty, \mathcal{O}_t) + \mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon). \end{aligned} \tag{3.2}$$

Combining the preceding inequalities we obtain

$$\frac{|e^{-\mathcal{Q}t}x|}{\varepsilon} - \mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty) \leq \frac{\mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon)}{\varepsilon} \leq \frac{|e^{-\mathcal{Q}t}x|}{\varepsilon} + \mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty). \tag{3.3}$$

Since the $\mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty) \rightarrow 0$ as $t \rightarrow \infty$, for any $t_\varepsilon \rightarrow \infty$ as $\varepsilon \rightarrow 0$ we have $\mathcal{W}_{p'}(\mathcal{O}_{t_\varepsilon}, \mathcal{O}_\infty) \rightarrow 0$ as $\varepsilon \rightarrow 0$. It remains to show abrupt convergence of $\frac{|e^{-\mathcal{Q}t_\varepsilon}x|}{\varepsilon}$ for the correct choice of t_ε^x . Therefore, the refined analysis of the linear system $e^{-\mathcal{Q}t}x$ carried out in Lemma 2.1 is necessary.

Remark 3.1 Note that the preceding formula (3.3) is valid for any p of Hypothesis 2.1. If L has finite moments of all orders, that is, formally $p = \infty$, we may pass to the limit in (3.3) and obtain

$$\frac{|e^{-Q_t}x|}{\varepsilon} - \mathcal{W}_\infty(\mathcal{O}_t, \mathcal{O}_\infty) \leq \frac{\mathcal{W}_\infty(X_t^\varepsilon(x), \mu^\varepsilon)}{\varepsilon} \leq \frac{|e^{-Q_t}x|}{\varepsilon} + \mathcal{W}_\infty(\mathcal{O}_t, \mathcal{O}_\infty). \tag{3.4}$$

This is satisfied for instance in the case of pure Brownian motion or uniformly bounded jumps. Moreover

$$\begin{aligned} \mathcal{W}_\infty(\mathcal{O}_t, \mathcal{O}_\infty) &= \lim_{p \rightarrow \infty} \mathcal{W}_p(\mathcal{O}_t, \mathcal{O}_\infty) \leq \lim_{p \rightarrow \infty} \int_{\mathbb{R}^d} \mathcal{W}_p(\mathcal{O}_t, \mathcal{O}_t(z)) \mathbb{P}(\mathcal{O}_\infty \in dz) \\ &\leq |e^{-Q_t}| \cdot \mathbb{E}[|\mathcal{O}_\infty|] \rightarrow 0, \quad \text{as } t \rightarrow \infty. \end{aligned}$$

3.1.2 The Key Estimates for $p \in (0, 1)$

We point out that for $0 < p' \leq p$ the distance $\mathcal{W}_{p'}$ satisfies all properties of Lemma 2.2, however, with modified versions of c) and d). Therefore, the upper bound (3.1) has the shape

$$\begin{aligned} \mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon) &\leq \mathcal{W}_{p'}(e^{-Q_t}x + \varepsilon\mathcal{O}_t, e^{-Q_t}x + \varepsilon\mathcal{O}_\infty) + \mathcal{W}_{p'}(e^{-Q_t}x + \varepsilon\mathcal{O}_\infty, \varepsilon\mathcal{O}_\infty) \\ &= \varepsilon^{p'} \mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty) + \varepsilon^{p'} \mathcal{W}_{p'}\left(\frac{e^{-Q_t}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty\right) \end{aligned}$$

and the lower bound (3.2) reads

$$\begin{aligned} \varepsilon^{p'} \mathcal{W}_{p'}\left(\frac{e^{-Q_t}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty\right) &= \mathcal{W}_{p'}(e^{-Q_t}x + \varepsilon\mathcal{O}_\infty, \varepsilon\mathcal{O}_\infty) \\ &\leq \mathcal{W}_{p'}(e^{-Q_t}x + \varepsilon\mathcal{O}_\infty, e^{-Q_t}x + \varepsilon\mathcal{O}_t) \\ &\quad + \mathcal{W}_{p'}(e^{-Q_t}x + \varepsilon\mathcal{O}_t, \varepsilon\mathcal{O}_\infty) \\ &= \varepsilon^{p'} \mathcal{W}_{p'}(\mathcal{O}_\infty, \mathcal{O}_t) + \mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon). \end{aligned}$$

The combination of the preceding inequalities yields

$$\begin{aligned} \mathcal{W}_{p'}\left(\frac{e^{-Q_t}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty\right) - \mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty) &\leq \frac{\mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{p'}} \\ &\leq \mathcal{W}_{p'}\left(\frac{e^{-Q_t}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty\right) + \mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty). \end{aligned} \tag{3.5}$$

Remark 3.2 For $p' \in (0, 1)$, property d) in Lemma 2.2 yields

$$\frac{|e^{-Q_t}x|^{p'}}{\varepsilon^{p'}} - 2\mathbb{E}[|\mathcal{O}_\infty|^{p'}] \leq \mathcal{W}_{p'}\left(\frac{e^{-Q_t}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty\right) \leq \frac{|e^{-Q_t}x|^{p'}}{\varepsilon^{p'}}. \tag{3.6}$$

3.2 The First Main Result: Characterizations of Profile Cutoff Thermalization

This subsection presents the first cutoff thermalization results of in the sense of (1.6) for the system (1.4) with $x \neq 0$.

Remark 3.3 Note that for initial value $x = 0$ there is no cutoff thermalization. Indeed, by property c) in Lemma 2.2 we have

$$\mathcal{W}_{p'}(X_t^\varepsilon(0), \mu^\varepsilon) = \mathcal{W}_{p'}(\varepsilon \mathcal{O}_t, \varepsilon \mathcal{O}_\infty) = \varepsilon^{\min\{1, p'\}} \mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty).$$

Hence for any $t_\varepsilon \rightarrow \infty$ as $\varepsilon \rightarrow 0$ we have

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon}^\varepsilon(0), \mu^\varepsilon)}{\varepsilon^{\min\{1, p'\}}} = \lim_{\varepsilon \rightarrow 0} \mathcal{W}_{p'}(\mathcal{O}_{t_\varepsilon}, \mathcal{O}_\infty) = 0$$

excluding a cutoff time scale separation.

3.2.1 Explicit Cutoff Thermalization Profile in Case of First Moments $p \geq 1$

The first main result characterizes the convergence of $\mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon)/\varepsilon$ to a profile function for $x \neq 0$ and $1 \leq p' \leq p$.

Theorem 3.1 (Cutoff thermalization profile) *Let v satisfy Hypothesis 2.1 for some $1 \leq p \leq \infty$. Let \mathcal{Q} satisfy Hypothesis 2.2 and $x \in \mathbb{R}^d$, $x \neq 0$, with the spectral representation $q > 0$, $\ell, m \in \{1, \dots, d\}$, $\theta_1, \dots, \theta_m \in [0, 2\pi)$ and $v_1, \dots, v_m \in \mathbb{C}^d$ of Lemma 2.1.*

Then the following statements are equivalent.

(i) *The ω -limit set*

$$\omega(x) := \left\{ \text{accumulation points of } \sum_{k=1}^m e^{it\theta_k} v_k \text{ as } t \rightarrow \infty \right\} \tag{3.7}$$

is contained in a sphere, that is, the function

$$\omega(x) \ni u \mapsto |u| \text{ is constant.} \tag{3.8}$$

(ii) *For the time scale*

$$t_\varepsilon^x = \frac{1}{q} |\ln(\varepsilon)| + \frac{\ell - 1}{q} \ln(|\ln(\varepsilon)|) \tag{3.9}$$

the system $(X_t^\varepsilon(x))_{t \geq 0}$ exhibits for all asymptotically constant window sizes $w_\varepsilon \rightarrow w > 0$ the abrupt thermalization profile for any $1 \leq p' \leq p$ in the following sense

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon} = \mathcal{P}_x(r) \text{ for any } r \in \mathbb{R},$$

where

$$\mathcal{P}_x(r) := \frac{e^{-rqw}}{q^{\ell-1}} |v| \text{ for any representative } v \in \omega(x). \tag{3.10}$$

Under either of the conditions, for ε sufficiently small, we have the error estimate

$$\left| \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon} - \mathcal{P}_x(r) \right| \leq \mathcal{W}_{p'}(\mathcal{O}_{t_\varepsilon^x}, \mathcal{O}_\infty) + \left| \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} |x|}{\varepsilon} - \mathcal{P}_x(r) \right| \tag{3.11}$$

which for generic x yields a constant C_x such that

$$\left| \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon} - \mathcal{P}_x(r) \right| \leq C_x \varepsilon^{1 \wedge \frac{q}{q}}. \tag{3.12}$$

The proof of Theorem 3.1 is given at the end of Appendix B. In the sequel, we essentially characterize when the function

$$\omega(x) \ni u \mapsto |u|$$

is constant. We enumerate v_1, \dots, v_m as follows. Without loss of generality we assume that $\theta_1 = 0$, that is, $v_1 \in \mathbb{R}^d$. Otherwise we take $v_1 = 0$ and eliminate it from the sum $\sum_{k=1}^m e^{i\theta_k t} v_k$. Without loss of generality we assume that $m = 2n + 1$ for some $n \in \mathbb{N}$. We assume that v_k and $v_{k+1} = \bar{v}_k$ are complex conjugate for all even number $k \in \{2, \dots, m\}$. For $k \in \{2, \dots, m\}$ we write $v_k = \hat{v}_k + i\check{v}_k$ where $\hat{v}_k, \check{v}_k \in \mathbb{R}^d$. Since

$$e^{i\theta_k t} v_k = [\cos(\theta_k t)\hat{v}_k - \sin(\theta_k t)\check{v}_k] + i[\sin(\theta_k t)\hat{v}_k + \cos(\theta_k t)\check{v}_k],$$

the decomplexification given in Lemma E.1 yields the representation

$$\sum_{k=1}^m e^{i\theta_k t} v_k = v_1 + 2 \sum_{k=1}^n \left(\cos(\theta_{2k} t)\hat{v}_{2k} - \sin(\theta_{2k} t)\check{v}_{2k} \right), \tag{3.13}$$

where $v_1 \in \mathbb{R}^d$.

Remark 3.4 Note that the angles $\theta_2, \dots, \theta_{2n}$ in (3.13) coming from Lemma 2.1 are rationally 2π -independent for generic matrices Q and initial values x . In other words, they satisfy the non-resonance condition

$$h_1\theta_2 + \dots + h_n\theta_{2n} \notin 2\pi \cdot \mathbb{Z} \tag{3.14}$$

for all $(h_1, \dots, h_n) \in \mathbb{Z}^n \setminus \{0\}$.

Theorem 3.2 *Let the assumptions of Theorem 3.1 be satisfied. In addition, we assume that the angles $\theta_2, \dots, \theta_{2n}$ are rationally 2π -independent according to (3.14) in Remark 3.4. Then i) and ii) in Theorem 3.1 are equivalent to the following normal growth condition: the family of \mathbb{R}^d -valued vectors*

$$(v_1, \hat{v}_2, \check{v}_2, \dots, \hat{v}_{2n}, \check{v}_{2n}) \text{ is orthogonal and satisfies } |\hat{v}_{2k}| = |\check{v}_{2k}| \text{ for } k = 1, \dots, n. \tag{3.15}$$

In this case the profile function has the following shape

$$\mathcal{P}_x(r) = \frac{e^{-rqw}}{q^{\ell-1}} \left| \sum_{j=1}^m v_j \right|, \quad r \in \mathbb{R}. \tag{3.16}$$

The proof is given in Appendix E. It consists of a characterization of Theorem 3.1 item i), that is, the property of $\omega(x)$ being contained in a sphere. This characterization is carried out in two consecutive steps in Appendix E under the non-resonance condition (3.14) given in Remark 3.4. Lemma E.2 yields the necessary implication, while Lemma E.3 states the sufficiency.

Remark 3.5 (1) It is clear that under item i) of Theorem 3.1 the profile can be defined as

$$\mathcal{P}_x(r) = \frac{e^{-rqw}}{q^{\ell-1}} |u|$$

for any representative $u \in \omega(x)$. Under the assumption of non-resonance of Remark 3.4 we have that $u = \sum_{j=1}^m v_j \in \omega(x)$. Indeed, since the $\theta_2, \dots, \theta_{2n}$ are rationally independent $((e^{i\theta_2 t}, \dots, e^{i\theta_{2n} t}))_{t \geq 0}$ is dense in the torus S_1^n , see Corollary 4.2.3 in [105]. Hence we approximate the point $(1, \dots, 1)$ for a subsequence $t_k \rightarrow \infty$ as $k \rightarrow \infty$ and hence $u = \sum_{j=1}^m v_j \in \omega(x)$ and (3.16) is valid.

- (2) The existence of a thermalization profile boils down to the precise geometric structure of the complicated limit set $\omega(x)$. However, it is not difficult to cover several cases of interest. In particular, in the case \mathcal{Q} being symmetric $\omega(x) = \{\sum_{j=1}^m v_k\}$ since all the rotation angles vanish, the function (3.8) is trivially constant.
- (3) The shape of the thermalization profile given in (3.10) is surprisingly universal:

$$\mathcal{P}_x(r) = \frac{e^{-r\mathfrak{q}w}}{q^{\ell-1}}|u| = \mathcal{W}_{p'}\left(\frac{e^{-r\mathfrak{q}w}}{q^{\ell-1}}u + \mathcal{O}_\infty, \mathcal{O}_\infty\right).$$

It does not depend on the parameters $1 \leq p' \leq p \in [1, \infty]$ (beyond finite moments of order p) nor on the statistical properties of the driving noise v due to the shift linearity (2.6), item d) of Lemma 2.2. For $p = 2$ item d) of Lemma 2.2 is well-known and a direct consequence of Pythagoras’ theorem, see for instance [87], Section 2, p. 412. We give the proof for general $p \geq 1$ in Appendix A.

- (4) The statistical information of L enters in the rate of convergence on the right-hand side of (3.11). Indeed, by (2.10) we have generically

$$\mathcal{W}_{p'}(\mathcal{O}_{t_\varepsilon^x}, \mathcal{O}_\infty) \leq \mathbb{E}[|\mathcal{O}_\infty|] \cdot |e^{-\mathcal{Q}t_\varepsilon^x}| \leq C_0\mathbb{E}[|\mathcal{O}_\infty|]\varepsilon, \tag{3.17}$$

where C_0 is given in (2.9). Moreover, $\mathbb{E}[|\mathcal{O}_\infty|]$ is bounded explicitly in terms of the characteristic of the noise and the matrix \mathcal{Q} , see (2.12).

- (5) The order of the asymptotic error $|\mathcal{P}_x(r) - \frac{|e^{-\mathcal{Q}(t_\varepsilon^x+r\cdot w)x}|}{\varepsilon}|$ depends inherently on the spectral structure of \mathcal{Q} . In the worst case its rates of convergence are of logarithmic order $1/t_\varepsilon^x$, see formula (5.1) in the example of Sect. 5.3. In Sect. 5.1 we see the optimal rate of convergence where this error is zero. However, this is not the generic picture. Generically all eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_d \in \mathbb{C}$ have different real parts (up to pairs of complex conjugate eigenvalues) with multiplicity 1. Without loss of generality we label $\lambda_1, \lambda_2, \dots, \lambda_d \in \mathbb{C}$ by ascending (positive) real parts. Moreover, $\mathfrak{q} = \text{Re}(\lambda_1)$ in the generic case. Under the assumption of a thermalization profile we count with the speed of convergence of order $e^{-\mathfrak{g}t_\varepsilon^x} = K(x)\varepsilon^{\mathfrak{g}/\mathfrak{q}}$, where

$$\mathfrak{g} = \begin{cases} \text{Re}(\lambda_2) - \mathfrak{q}, & \text{if } \lambda_2 \neq \bar{\lambda}_1, \\ \text{Re}(\lambda_3) - \mathfrak{q}, & \text{if } \lambda_2 = \bar{\lambda}_1, \end{cases} \tag{3.18}$$

and since any initial datum x has the unique representation $x = \sum_{j=1}^d c_j(x)v_j$ and hence $K(x)$ can be taken as

$$K(x) = \max_{j=1, \dots, d} |c_j(x)v_j|, \tag{3.19}$$

where v_j are the eigenvectors associated to the eigenvalue λ_j .

- (6) By (3.11), item (3) and item (5) we obtain the generic order of magnitude of ε such that the asymptotic approximation holds for concrete systems in terms of the noises characteristics, \mathcal{Q} , the long-term dynamics of $|e^{-\mathcal{Q}t}|$, and the initial value x .

3.2.2 Abstract Cutoff Thermalization Profile in Case of $p \in (0, 1)$

This result is stated in order to cover perturbations of the Cauchy process, where $p < 1$ and other stable processes such as the Holtsmark process $p < \frac{1}{2}$. Here the profile function does exist but remains abstract.

Theorem 3.3 (Abstract cutoff thermalization profile for any $p > 0$) *Let the assumptions (and the notation) of Theorem 3.1 be valid for some $0 < p \leq \infty$. Then for any $0 < p' \leq p$ the following statements are equivalent.*

- (i) *For any $\lambda > 0$, the function $\omega(x) \ni u \mapsto \mathcal{W}_{p'}(\lambda u + \mathcal{O}_\infty, \mathcal{O}_\infty)$ is constant, where $\omega(x)$ is given in (3.7).*
- (ii) *For the time scale t_ε^x given in (3.9) the system $(X_t^\varepsilon(x))_{t \geq 0}$ exhibits for all asymptotically constant window sizes $w_\varepsilon \rightarrow w > 0$ the abrupt thermalization profile for any $0 < p' \leq p$ in the following sense*

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{1, p'\}}} = \mathcal{P}_{x, p'}(r) \quad \text{for any } r \in \mathbb{R},$$

where

$$\mathcal{P}_{x, p'}(r) := \mathcal{W}_{p'}\left(\frac{e^{-r \cdot q w}}{q^{\ell-1}} v + \mathcal{O}_\infty, \mathcal{O}_\infty\right) \quad \text{for any representative } v \in \omega(x). \quad (3.20)$$

The proof is given in Appendix B.

Remark 3.6 (1) For $1 \leq p' \leq p$ Theorem 3.3 with the help of property d) in Lemma 2.2 recovers an abstract version of Theorem 3.1 which also extends to $p < 1$.

- (2) The (asymptotic) error estimates of Theorem 3.1 are harder to obtain for $0 < p' < 1$.
- (3) For (nondegenerate) pure Brownian motion, the existence of a cutoff thermalization profile in total variation distance is equivalent to the set $\Sigma^{-1/2} \omega(x)$ being contained in a sphere, where Σ is the covariance matrix of the invariant distribution, see Corollary 2.11 in [13]. In Corollary 4.14 of [11] it is shown that a corresponding geometric condition is at least sufficient. For further unexpected properties in the pure α stable case, see [101].

3.3 The Second Main Result: Generic Window Cutoff Thermalization

Roughly speaking, condition (3.8) in item i) of Theorem 3.1 (as well as item i) in Theorem 3.3) fails to hold if the rotational part of \mathcal{Q} is too strong. However, for the general case we still have abrupt thermalization in the following weaker sense.

Theorem 3.4 *Let the assumptions (and the notation) of Theorem 3.1 be valid for some $0 < p \leq \infty$. Then the system $(X_t^\varepsilon(x))_{t \geq 0}$ exhibits window cutoff thermalization on the time scale*

$$t_\varepsilon^x = \frac{1}{q} |\ln(\varepsilon)| + \frac{\ell - 1}{q} \ln(|\ln(\varepsilon)|)$$

and in the sense that for all asymptotically constant window sizes $w_\varepsilon \rightarrow w > 0$ it follows

$$\lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{1, p'\}}} = \infty \quad \text{and} \quad \lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{1, p'\}}} = 0.$$

for all $0 < p' \leq p$.

The proof is given Appendix C. In contrast to other distances, the Wasserstein distance also implies the cutoff thermalization for the physical observables as follows.

Corollary 3.1 *Let the assumptions (and the notation) of Theorem 3.1 be valid for some $p > 0$. Then we have for any $0 < p' \leq p < \infty$ and $x \neq 0$*

$$\lim_{r \rightarrow \infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \mathbb{E}[|X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x)|^{p'}] = \lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \mathbb{E}[|X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x)|^{p'}] = \mathbb{E}[|\mathcal{O}_\infty|^{p'}]$$

and

$$\lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \mathbb{E}[|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)|^{p'}] = \lim_{r \rightarrow -\infty} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \mathbb{E}[|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)|^{p'}] = \infty.$$

For the proof we refer to Appendix D.

Corollary 3.2 *Let the assumptions (and the notation) of Theorem 3.1 be valid for some $p > 0$. Then we have*

$$\lim_{\varepsilon \rightarrow 0} \mathcal{W}_p(X_{\delta \cdot t_\varepsilon}^\varepsilon(x), \mu^\varepsilon) \cdot \varepsilon^{-1} = \begin{cases} \infty & \text{for } \delta \in (0, 1) \\ 0 & \text{for } \delta > 1 \end{cases}.$$

This corollary justifies formula (1.5) in the introduction. For the proof we refer to Appendix D.

Remark 3.7 (1) In general for $1 \leq p' \leq p$ there is no thermalization profile in the sense of Theorem 3.1 (and Theorem 3.3). However, it is easy to see that a cutoff thermalization profile implies window cutoff thermalization. The contrary not always holds. For instance, for different values of $u \in \omega(x)$ there is no unique candidate for the profile. To be more precise,

$$\limsup_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon} = \frac{e^{-r qw}}{q^{\ell-1}} |\hat{u}|$$

and

$$\liminf_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon} = \frac{e^{-r qw}}{q^{\ell-1}} |\check{u}|,$$

where $\check{u}, \hat{u} \in \omega(x)$. The discussion of the linear oscillator given in Sect. 4.2 yields an example where $\omega(z)$ is not contained in a sphere for any $z \neq 0$. The case of subcritical damping always exhibits complex eigenvalues which together with the precise structure of the dynamics excludes a thermalization profile and only window thermalization remains valid.

(2) The error estimate in Remark 3.5 item (4) remains untouched and item (5) is slightly adapted as follows. Here we consider the error term

$$\left| \frac{e^{-\mathcal{Q} t_\varepsilon^x} x}{\varepsilon} - \sum_{k=1}^m e^{i t_\varepsilon^x \theta_k} v_k \right|, \tag{3.21}$$

which analogously depends on the spectral structure of \mathcal{Q} . Generically all eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_d \in \mathbb{C}$ have different real parts (up to pairs of complex conjugate eigenvalues) with multiplicity 1. Without loss of generality we label $\lambda_1, \lambda_2, \dots, \lambda_d \in \mathbb{C}$ by ascending (positive) real parts. Moreover, $q = \text{Re}(\lambda_1)$ in the generic case. The speed of convergence of (3.21) is of order $e^{-g t_\varepsilon^x} = K(x) \varepsilon^{g/q}$, where g is given in (3.18) and $K(x)$ given in (3.19) is estimated identically.

- (3) In Sect. 5.1 we give an (ad hoc) linear (2×2) -system showing a thermalization profile in the presence of complex (conjugate) eigenvalues for all initial values.
- (4) The example in Sect. 5.2 represents a system where the presence of a thermalization profile depends on the initial value x .
- (5) We finally construct in Sect. 5.3 a system with arbitrary high logarithmic corrections terms.

4 Physics Examples

4.1 Gradient Systems

For a symmetric \mathcal{Q} , Theorem 3.1 applies to the gradient case

$$dX_t^\varepsilon = -\nabla\mathcal{U}(X_t^\varepsilon)dt + \varepsilon dL_t \quad \text{with } X_0^\varepsilon = x \neq 0$$

for the quadratic potential form $\mathcal{U}(z) = (1/2)z^*Qz$. Indeed, by the spectral decomposition we have an orthonormal basis $v_1, v_2, \dots, v_d \in \mathbb{R}^d$ with corresponding eigenvalues $0 < \lambda_1 \leq \dots \leq \lambda_d$ such that

$$e^{-\mathcal{Q}t}x = \sum_{j=1}^d e^{-\lambda_j t} \langle v_j, x \rangle v_j.$$

Let $\tau(x) = \min\{j \in \{1, \dots, d\} : \langle v_j, x \rangle \neq 0\}$ and $J = \{j \in \{\tau(x), \dots, d\} : \lambda_j = \lambda_{\tau(x)}\}$. Hence

$$\lim_{t \rightarrow \infty} e^{\lambda_{\tau(x)}t} e^{-\mathcal{Q}t}x = \sum_{j \in J} \langle v_j, x \rangle v_j \neq 0.$$

That is, for $p \geq 1$ the cutoff thermalization profile is $\mathcal{P}_x(r) = e^{-r\lambda_{\tau(x)}w} |\sum_{j \in J} \langle v_j, x \rangle v_j|$. More generally, [14], Proposition A.4.ii) yields a complete description of the spectral decomposition of non-symmetric \mathcal{Q} with real spectrum.

4.2 The Linear Oscillator

In this subsection we provide a complete discussion of the cutoff thermalization of the damped linear oscillator driven by different noises at small temperature. We consider

$$\begin{cases} dX_t^\varepsilon = Y_t^\varepsilon dt, \\ dY_t^\varepsilon = -\kappa X_t^\varepsilon dt - \gamma Y_t^\varepsilon dt + \varepsilon dL_t \end{cases} \tag{4.1}$$

with initial conditions $X_0^\varepsilon = x, Y_0^\varepsilon = y$ and Lévy noise $L = (L_t)_{\geq 0}$ satisfying Hypothesis 2.1 for some $p > 0$. Examples of interest are the following:

For $p \leq \infty$ we cover

- (1) standard Brownian motion,
- (2) deterministic (linear) drift,
- (3) discontinuous compound Poisson process with finitely many point increments.

For $p < \alpha$ for some $\alpha > 0$

- (4) α -stable Lévy flight with finite first moment for index $\alpha \in (1, 2)$,
- (5) α -stable Lévy flight with index $\alpha \in (0, 1]$ including the Cauchy flight when $\alpha = 1$. See [54] and [101] for a thorough discussion.

We rewrite the system (4.1) as a vector valued Ornstein–Uhlenbeck process

$$d \begin{pmatrix} X_t^\varepsilon \\ Y_t^\varepsilon \end{pmatrix} = -\mathcal{Q} \begin{pmatrix} X_t^\varepsilon \\ Y_t^\varepsilon \end{pmatrix} dt + \varepsilon d\mathcal{L}_t,$$

where

$$\mathcal{Q} := \begin{pmatrix} 0 & -1 \\ \kappa & \gamma \end{pmatrix} \quad \text{and} \quad \mathcal{L}_t := \begin{pmatrix} 0 \\ L_t \end{pmatrix}. \tag{4.2}$$

Let $z = (x, y)^* \neq (0, 0)^*$. In the sequel, we compute $e^{-Q_t}z$. The eigenvalues of $-Q$ are given by

$$\lambda_{\pm} = \frac{-\gamma \pm \sqrt{\gamma^2 - 4\kappa}}{2}.$$

Note that for any $\gamma, \kappa > 0$ the respective real parts are strictly negative.

4.2.1 Overdamped Linear Oscillator

$\Delta = \gamma^2 - 4\kappa > 0$. In this case, $-Q$ has two real different eigenvalues

$$\lambda_- := \frac{1}{2}(-\gamma - \sqrt{\Delta}) < \frac{1}{2}(-\gamma + \sqrt{\Delta}) =: \lambda_+ < 0$$

with the respective eigenvectors v_- and v_+ . The exponential matrix is given by

$$e^{-Q_t} = (v_- \ v_+) \operatorname{diag} (e^{-\lambda_- t}, e^{-\lambda_+ t}) (v_- \ v_+)^* \quad \text{for } t \geq 0.$$

Recall $z = (x, y)^* \neq (0, 0)^*$. We denote by $\tilde{z} := (v_- \ v_+)^* z = (\tilde{z}_1, \tilde{z}_2)^*$ the coordinate change of z . Note that

$$e^{-Q_t}z = (v_- \ v_+)(\tilde{z}_1 e^{-\lambda_- t}, \tilde{z}_2 e^{-\lambda_+ t})^*.$$

This formula yields that for any $z \neq 0$ there exist an explicit $q > 0$ and $u \neq 0$ such that

$$\lim_{t \rightarrow \infty} e^{qt} e^{-Q_t}z = u_z. \tag{4.3}$$

Indeed, if $\tilde{z}_1 = 0$ then $\tilde{z}_2 \neq 0$ and we have $\lim_{t \rightarrow \infty} e^{\lambda_+ t} e^{-Q_t}z = (v_- \ v_+)(0, \tilde{z}_2)^* =: u_z \neq 0$. If $\tilde{z}_2 = 0$ then $\tilde{z}_1 \neq 0$ and we have $\lim_{t \rightarrow \infty} e^{\lambda_- t} e^{-Q_t}z = (v_- \ v_+)(\tilde{z}_1, 0)^* =: u_z \neq 0$. Finally, if $\tilde{z}_1 \neq 0$ and $\tilde{z}_2 \neq 0$. Then $\lim_{t \rightarrow \infty} e^{\lambda_- t} e^{-Q_t}z = (v_- \ v_+)(\tilde{z}_1, 0)^* =: u_z \neq 0$. In particular, for any $z \neq 0$, the omega limit set $\omega(z)$ defined in (3.7) consists of a single point u_z .

Hence for the noises (1)–(4), Theorem 3.1 applies for $1 \leq p' \leq p$ and thermalization profile holds at the time scale

$$t_\varepsilon^x = \frac{1}{q} |\ln(\varepsilon)|,$$

with profile

$$\mathcal{P}_z(r) = e^{-r q w} |u_z|$$

for all window sizes $w > 0$. Roughly speaking, for any $1 \leq p' \leq p$ Theorem 3.1 yields for some positive $K_{p,p'}$

$$\mathcal{W}_{p'}(X_{t_\varepsilon^x+r \cdot w}^\varepsilon(x), \mu^\varepsilon) \approx_\varepsilon \varepsilon e^{-r q w} |u| + \varepsilon^2 K_{p,p'} \mathbb{E}[\mathcal{O}_\infty].$$

For the noise (5), Theorem 3.3 still implies profile thermalization, however, the profile is given by the abstract formula

$$\mathcal{P}_{x,p'}(r) = \mathcal{W}_{p'}(e^{-r q w} u_z + \mathcal{O}_\infty, \mathcal{O}_\infty)$$

for any $p' \leq p < \alpha$.

4.2.2 Critically Damped Linear Oscillator: $\Delta = \gamma^2 - 4\kappa = 0$

In this case, $-Q$ has a repeated real eigenvalue $\lambda := \lambda_- = \lambda_+ = -\gamma/2 < 0$ and the matrix exponential is given by

$$e^{-Qt} = e^{\lambda t} I_2 + e^{\lambda t} t(-Q - \lambda I_2) \quad \text{for } t \geq 0.$$

Let $z = (x, y)^* \neq (0, 0)^*$. On the one hand, if $z \in \text{Ker}(-Q - \lambda I_2)$, that is, in Lemma 2.1 we have $\ell = 1$, and $e^{-\lambda t} e^{-Qt} z = u_z$ for $u_z = z$. On the other hand, $z \notin \text{Ker}(-Q - \lambda I_2)$ which corresponds to $\ell = 2$ yields

$$\lim_{t \rightarrow \infty} \frac{e^{-\lambda t}}{t} e^{-Qt} z = (-Q - \lambda I_2)z =: u_z \neq 0.$$

In particular, for any $z \neq 0$, the omega limit set $\omega(z)$ defined in (3.7) consists of the a single point u_z . Hence for the noises (1)–(4) Theorem 3.1 still applies for $1 \leq p' \leq p$ and profile thermalization holds true at the modified time scale

$$t_\varepsilon^x = \frac{1}{q} |\ln(\varepsilon)| + \frac{\ell - 1}{q} \ln(|\ln(\varepsilon)|)$$

with the modified profile

$$\mathcal{P}_z(r) = \frac{e^{-rw(\gamma/2)}}{(\gamma/2)^{\ell-1}} |u|$$

for all window sizes $w > 0$. Roughly speaking, for any $1 \leq p' \leq p$ Theorem 3.1 yields for some positive constant $K_{p,p'}$

$$\mathcal{W}_{p'}(X_{t_\varepsilon^x+r \cdot w}^\varepsilon(x), \mu^\varepsilon) \approx_\varepsilon \varepsilon \cdot \frac{e^{-rw(\gamma/2)}}{(\gamma/2)^{\ell-1}} |u| + \varepsilon^2 |\ln(\varepsilon)|^{\ell-1} \cdot K_{p,p'} \mathbb{E}[|\mathcal{O}_\infty|].$$

For the noise (5), Theorem 3.3 still implies profile thermalization, however, the modified profile is also given by the abstract formula

$$\mathcal{P}_{x,p'}(r) = \mathcal{W}_{p'}\left(e^{-r qw} u_z + \mathcal{O}_\infty, \mathcal{O}_\infty\right)$$

for any $p' \leq p < \alpha$.

In the sequel, we discuss the general case of complex conjugate eigenvalues in order to treat the subcritical case.

4.2.3 Non-normal Growth of the Linear Oscillator for Complex Eigenvalues

Recall that the eigenvalues of Q are given by

$$\lambda_- = \hat{\lambda} - i\check{\lambda} \quad \text{and} \quad \lambda_+ = \hat{\lambda} + i\check{\lambda}, \quad \check{\lambda} \neq 0.$$

By the Jacobi formula, see for instance Theorem 1 in [81], Part Three, Sec 8.3, we have

$$\det(e^{\hat{\lambda}t} e^{-Qt}) = e^{2\hat{\lambda}t} e^{-\text{trace}(Qt)} = 1. \tag{4.4}$$

By the Lagrange interpolation theorem (see Theorem 7.11, p.209, in [5]) we have

$$e^{-Qt} = \frac{e^{-t\lambda_-}}{-2i\check{\lambda}} (-Q + \lambda_+ I_2) + \frac{e^{-t\lambda_+}}{2i\check{\lambda}} (-Q + \lambda_- I_2)$$

$$\begin{aligned}
 &= \frac{e^{-t\check{\lambda}} e^{it\check{\lambda}}}{2i\check{\lambda}} (\mathcal{Q} - \lambda_+ I_2) - \frac{e^{-t\check{\lambda}} e^{-it\check{\lambda}}}{2i\check{\lambda}} (\mathcal{Q} - \lambda_- I_2) \\
 &= -\frac{e^{-t\check{\lambda}}}{\check{\lambda}} \operatorname{Re} \left(e^{it\check{\lambda}} (\mathcal{Q} - \lambda_+ I_2) \right).
 \end{aligned} \tag{4.5}$$

Let $z = (x, y)^* \neq (0, 0)^*$. The preceding equality yields

$$|e^{t\check{\lambda}} e^{-\mathcal{Q}t} z| = \frac{1}{|\check{\lambda}|} |\operatorname{Re} (e^{it\check{\lambda}} (\mathcal{Q} - \lambda_+ I_2) z)|.$$

Moreover, by (4.4) we deduce

$$\liminf_{t \rightarrow \infty} |e^{\hat{\lambda}t} e^{-\mathcal{Q}t} z| > 0. \tag{4.6}$$

Additionally by the periodicity we have that

$$\limsup_{t \rightarrow \infty} |e^{\hat{\lambda}t} e^{-\mathcal{Q}t} z| < \infty. \tag{4.7}$$

Note that $|e^{t\check{\lambda}} e^{-\mathcal{Q}t} z|$ is a constant function if and only if $|\operatorname{Re} (e^{it\check{\lambda}} (-\mathcal{Q} + \lambda_+ I_2) z)|$ is so, too. In the sequel, we characterize when the function

$$t \mapsto |\operatorname{Re} (e^{it\check{\lambda}} (\mathcal{Q} - \lambda_+ I_2) z)| \text{ is constant.}$$

Let

$$a(z) := \operatorname{Re}((\mathcal{Q} - \lambda_+ I_2)z) \quad \text{and} \quad b(z) := \operatorname{Im}((\mathcal{Q} - \lambda_+ I_2)z). \tag{4.8}$$

Note that

$$\operatorname{Re}(e^{it\check{\lambda}} (\mathcal{Q} - \lambda_+ I_2)z) = \cos(\check{\lambda}t)a(z) - \sin(\check{\lambda}t)b(z). \tag{4.9}$$

Combining (4.5) with (4.9) yields

$$e^{\hat{\lambda}t} e^{-\mathcal{Q}t} z = -\frac{1}{\check{\lambda}} \left(\cos(\check{\lambda}t)a(z) - \sin(\check{\lambda}t)b(z) \right). \tag{4.10}$$

As a consequence, the Pythagoras theorem yields

$$\begin{aligned}
 |\operatorname{Re}(e^{it\check{\lambda}} (\mathcal{Q} - \lambda_+ I_2)z)|^2 &= |\cos(\check{\lambda}t)a(z) - \sin(\check{\lambda}t)b(z)|^2 \\
 &= \cos^2(\check{\lambda}t)|a(z)|^2 + \sin^2(\check{\lambda}t)|b(z)|^2 - 2 \cos(\check{\lambda}t) \sin(\check{\lambda}t) \langle a(z), b(z) \rangle.
 \end{aligned} \tag{4.11}$$

Remark 4.1 Note that equation (4.11) does not require any specific structure of \mathcal{Q} . It only uses that $d = 2$, $\mathcal{Q} \in \mathbb{R}^{2 \otimes 2}$ and the existence of conjugate complex, non-real eigenvalues. For this (more general) case we state the following lemma.

Lemma 4.1 (Profile cutoff characterization by the absence of non-normal growth) *For $d = 2$ the following statements are equivalent.*

- (i) *The function $t \mapsto |\operatorname{Re}(e^{it\check{\lambda}} (\mathcal{Q} - \lambda_+ I_2)z)|$ is constant.*
- (ii) *$|a(z)|^2 = |b(z)|^2$ and $\langle a(z), b(z) \rangle = 0$, where $a(z)$ and $b(z)$ are given in (4.8).*
- (iii) *For some $R > 0$*

$$\omega(z) \subset \{|u| = R\},$$

where

$$\omega(z) = \{u \in \mathbb{R}^2 \mid e^{\hat{\lambda}t_n} e^{-\mathcal{Q}t_n} z \rightarrow u \text{ for some } (t_n)_{n \in \mathbb{N}}, t_n \rightarrow \infty\}.$$

Proof The proofs of $\text{ii} \implies \text{i}$ and $\text{ii} \implies \text{iii}$ are immediately from (4.11).

We continue with $\text{i} \implies \text{ii}$ and assume that $t \mapsto |\operatorname{Re}(e^{it\check{\lambda}}(\mathcal{Q} - \lambda_+ I_2)z)|$ is constant. Evaluating in $t_n = \frac{\pi n}{\check{\lambda}}$, $n \in \mathbb{N}$ we obtain

$$|\operatorname{Re}(e^{it_n\check{\lambda}}(\mathcal{Q} - \lambda_+ I_2)z)|^2 = |a(z)|^2.$$

Now, we evaluate in $s_n = \frac{\pi + 2n\pi}{2\check{\lambda}}$, $n \in \mathbb{N}$ and deduce

$$|\operatorname{Re}(e^{is_n\check{\lambda}}(\mathcal{Q} - \lambda_+ I_2)z)|^2 = |b(z)|^2.$$

Hence $|a(z)|^2 = |b(z)|^2$ as in ii). Inserting the preceding equalities in (4.11), we have for any $t \geq 0$

$$|a(z)|^2 = |\operatorname{Re}(e^{it\check{\lambda}}(\mathcal{Q} - \lambda_+ I_2)z)|^2 = |a(z)|^2 - 2 \cos(\check{\lambda}t) \sin(\check{\lambda}t) \langle a(z), b(z) \rangle.$$

Since $\check{\lambda} \neq 0$ the latter implies $\langle a(z), b(z) \rangle = 0$.

We continue with $\text{iii} \implies \text{ii}$. For the sequence $t_n = \frac{2\pi n}{\check{\lambda}}$, $n \in \mathbb{N}$, applied to (4.10) yields $-\frac{a(z)}{\check{\lambda}} \in \omega(z)$. This implies $R = |\frac{a(z)}{\check{\lambda}}|$. For the sequence $t_n = \frac{1}{\check{\lambda}}(2\pi n + \frac{\pi}{2})$, $n \in \mathbb{N}$, we have $\frac{b(z)}{\check{\lambda}} \in \omega(z)$, and hence also $R = |\frac{b(z)}{\check{\lambda}}|$. This and $\check{\lambda} \neq 0$ gives $|a(z)| = |b(z)|$. For the inner product we use that (4.5) and (4.11) imply

$$\begin{aligned} |e^{t\hat{\lambda}} e^{-\mathcal{Q}t} z|^2 &= \frac{1}{\check{\lambda}^2} |\cos(\check{\lambda}t)a(z) - \sin(\check{\lambda}t)b(z)|^2 \\ &= \frac{1}{\check{\lambda}^2} \cos^2(\check{\lambda}t)|a(z)|^2 + \frac{1}{\check{\lambda}^2} \sin^2(\check{\lambda}t)|b(z)|^2 - \frac{2}{\check{\lambda}^2} \cos(\check{\lambda}t) \sin(\check{\lambda}t) \langle a(z), b(z) \rangle \\ &= \cos^2(\check{\lambda}t)R^2 + \sin^2(\check{\lambda}t)R^2 - \frac{2}{\check{\lambda}^2} \cos(\check{\lambda}t) \sin(\check{\lambda}t) \langle a(z), b(z) \rangle \\ &= R^2 - \frac{2}{\check{\lambda}^2} \cos(\check{\lambda}t) \sin(\check{\lambda}t) \langle a(z), b(z) \rangle. \end{aligned}$$

For $t_n = \frac{\pi/4 + 2\pi n}{\check{\lambda}}$ in (4.10) we have $\frac{\sqrt{2}}{2\check{\lambda}}(b(z) - a(z)) \in \omega(z)$. In addition, for this t_n we obtain

$$R^2 = \lim_{n \rightarrow \infty} |e^{t_n\hat{\lambda}} e^{-\mathcal{Q}t_n} z|^2 = R^2 - \frac{4}{\check{\lambda}} \langle a(z), b(z) \rangle,$$

which yields that $\langle a(z), b(z) \rangle = 0$. □

With this result at hand we complete the discussion of the linear oscillator in the sequel.

4.2.4 Subcritically Damped Linear Oscillator: $\Delta = \gamma^2 - 4\kappa < 0$

Recall that the eigenvalues of \mathcal{Q} in the case of (4.1) are given by

$$\lambda_- = \hat{\lambda} - i\check{\lambda} \quad \text{and} \quad \lambda_+ = \hat{\lambda} + i\check{\lambda},$$

where

$$\hat{\lambda} = \frac{\gamma}{2} \quad \text{and} \quad \check{\lambda} = \frac{\sqrt{4\kappa - \gamma^2}}{2} \neq 0.$$

By (4.6) and (4.7) for the noises (1)–(5) Theorem 3.4 implies window thermalization for any $0 < p' \leq p$ at time scale

$$t_\varepsilon^x = \frac{2}{\gamma} |\ln(\varepsilon)| \quad \text{for any initial condition } (x, y) \in \mathbb{R}^2, (x, y) \neq (0, 0).$$

In the sequel, by using the shift linearity property d) in Lemma 2.2 we exclude the existence of a cutoff thermalization profile for any $1 \leq p' \leq p$ and noises (1)–(4).

Lemma 4.2 *Let $1 \leq p' \leq p$. For any $\gamma > 0$ and $\kappa > 0$ such that $\gamma^2 - 4\kappa < 0$, there is no cutoff thermalization profile for any $(x, y) \neq (0, 0)$.*

Proof We apply Lemma 4.1 to (4.2). Recall $z = (x, y)^* \neq 0$. A straightforward calculation yields

$$a(z) = \left(-\frac{\gamma}{2}x - y, \kappa x + \frac{\gamma}{2}y \right)^* \quad \text{and} \quad b(z) = -\frac{\sqrt{4\kappa - \gamma^2}}{2}(x, y)^*.$$

The condition $\langle a(z), b(z) \rangle = 0$ reads as

$$\begin{aligned} 0 &= -\left(\frac{\gamma}{2}x + y\right)x + \left(\kappa x + \frac{\gamma}{2}y\right)y = -\frac{\gamma}{2}x^2 - xy + \kappa xy \\ &\quad + \frac{\gamma}{2}y^2 = -\frac{\gamma}{2}x^2 + (\kappa - 1)xy + \frac{\gamma}{2}y^2, \end{aligned}$$

that is,

$$\gamma x^2 - \gamma y^2 - 2(\kappa - 1)xy = 0. \tag{4.12}$$

Since $(x, y) \neq (0, 0)$, the preceding equality yields $x \neq 0$ and $y \neq 0$. The condition $|a(z)|^2 = |b(z)|^2$ is equivalent to

$$\frac{\gamma^2}{4}x^2 + y^2 + \gamma xy + \kappa^2 x^2 + \frac{\gamma^2}{4}y^2 + \kappa \gamma xy = \left(\kappa - \frac{\gamma^2}{4}\right)(x^2 + y^2).$$

Simplifying we obtain

$$(\gamma^2 + 2\kappa(\kappa - 1))x^2 + (\gamma^2 + 2(1 - \kappa))y^2 + 2\gamma(\kappa + 1)xy = 0. \tag{4.13}$$

For $\kappa = 1$ we have that (4.12) yields $x^2 = y^2$. Substituting in (4.13) we obtain

$$0 = \gamma^2 x^2 \pm 2\gamma x^2 = (\gamma^2 \pm 2\gamma)x^2.$$

Since $x^2 > 0$ and $\gamma > 0$, the unique solution is $\gamma = 2$, which implies $\gamma^2 - 4\kappa = 4 - 4 = 0$ and gives a contradiction to the subcritical damping $\gamma^2 - 4\kappa < 0$ of this case. As a consequence $\kappa = 1$ excludes profile thermalization for any parameters $\gamma > 0$ and $(x, y) \neq (0, 0)$. In the sequel, we assume $\kappa \neq 1$. Multiplying (4.13) by $(\kappa - 1)$ we have

$$((\kappa - 1)\gamma^2 + 2\kappa(\kappa - 1)^2)x^2 + ((\kappa - 1)\gamma^2 - 2(\kappa - 1)^2)y^2 + 2\gamma(\kappa - 1)(\kappa + 1)xy = 0. \tag{4.14}$$

Inserting the expression for $2(\kappa - 1)xy$ being given in (4.12) into (4.14) we obtain

$$\begin{aligned} 0 &= ((\kappa - 1)\gamma^2 + 2\kappa(\kappa - 1)^2)x^2 + ((\kappa - 1)\gamma^2 - 2(\kappa - 1)^2)y^2 + (\kappa + 1)\gamma^2(x^2 - y^2) \\ &= ((\kappa - 1)\gamma^2 + 2\kappa(\kappa - 1)^2 + (\kappa + 1)\gamma^2)x^2 + ((\kappa - 1)\gamma^2 - 2(\kappa - 1)^2 - (\kappa + 1)\gamma^2)y^2 \\ &= 2(\kappa(\gamma^2 + (\kappa - 1)^2))x^2 - 2(\gamma^2 + (\kappa - 1)^2)y^2. \end{aligned}$$

Since $\gamma^2 + (\kappa - 1)^2 > 0$, we obtain $\kappa x^2 - y^2 = 0$. In other words, $y = \pm\sqrt{\kappa}x$. Substituting in (4.12) we have

$$0 = \gamma x^2 - \gamma^2 y^2 - 2(\kappa - 1)xy = \left(\gamma - \gamma\kappa \mp 2(\kappa - 1)\sqrt{\kappa}\right)x^2$$

which implies $\gamma - \gamma\kappa \mp 2(k - 1)\sqrt{\kappa} = 0$. Hence $\gamma(1 - \kappa) = \pm 2(\kappa - 1)\sqrt{\kappa}$. Since $\kappa \neq 1$ we have $\gamma = \mp 2\sqrt{\kappa}$ which implies $\gamma^2 - 4\kappa = 0$ and gives a contradiction to the subcritical damping $\gamma^2 - 4\kappa < 0$. As a consequence for any $\kappa > 0$ and $\gamma > 0$ such that $\gamma^2 - 4\kappa < 0$, there is no profile thermalization for any initial condition $(x, y) \neq (0, 0)$. \square

This concludes the complete analysis of the linear oscillator (4.1).

4.3 Linear Chain of Oscillators in a Thermal Bath at Low Temperature

4.3.1 Window Cutoff Thermalization for the Linear Chain of Oscillators

Our results cover the setting of Jacobi chains of n oscillators with nearest neighbor interactions coupled to heat baths at its two ends, as discussed in Section 4.1 in [92] and Section 4.2 in [69]. For the sake of simplicity we show window cutoff thermalization for n oscillators with the Hamiltonian

$$\mathcal{H} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

$$(p, q) \mapsto \mathcal{H}(p, q) := \frac{1}{2} \sum_{i=1}^n p_i^2 + \frac{1}{2} \sum_{i=1}^n \gamma q_i^2 + \frac{1}{2} \sum_{i=1}^{n-1} \kappa (q_{i+1} - q_i)^2.$$

Coupling the first and the n th oscillator to a Langevin heat bath each with positive temperature ε^2 and positive coupling constants ς_1 and ς_n yields for $X^\varepsilon = (X^{1,\varepsilon}, \dots, X^{2n,\varepsilon}) = (p^\varepsilon, q^\varepsilon) = (p_1^\varepsilon, \dots, p_n^\varepsilon, q_1^\varepsilon, \dots, q_n^\varepsilon)$ the system

$$dX_t^\varepsilon = -QX_t^\varepsilon dt + \varepsilon dL_t,$$

where Q is a $2n \times 2n$ -dimensional real matrix of the following shape

$$Q = \left(\begin{array}{cccc|cccc} \varsigma_1 & 0 & \dots & 0 & \kappa + \gamma & -\kappa & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & -\kappa & 2\kappa + \gamma & -\kappa & 0 & 0 & \dots & 0 \\ & & & \ddots & \dots & \ddots & & \ddots & \ddots & & \\ & & & \ddots & \dots & \ddots & & \ddots & \ddots & & \\ 0 & \dots & 0 & 0 & \dots & & 0 & -\kappa & 2\kappa + \gamma & -\kappa & \\ 0 & 0 & \dots & 0 & \varsigma_n & 0 & \dots & 0 & -\kappa & \kappa + \gamma & \\ \hline -1 & 0 & \dots & 0 & 0 & 0 & \dots & \dots & \dots & 0 & \\ 0 & -1 & \dots & 0 & 0 & & & & & & \\ \vdots & & \ddots & \ddots & \vdots & & & \ddots & & \vdots & \\ & & & \ddots & \vdots & & & \ddots & & \vdots & \\ & & & \ddots & -1 & 0 & & & & & \\ 0 & \dots & 0 & -1 & 0 & 0 & \dots & \dots & \dots & 0 & \end{array} \right),$$

and $L_t = (L_t^1, 0, \dots, L_t^n, 0, \dots, 0)^*$. Here L^1, L^n are one dimensional independent Lévy processes satisfying Hypothesis 2.1 for some $p > 0$. By Section 4.1 in [92] Q satisfies Hypothesis 2.2. Consequently by Theorem 3.4 the system exhibits window cutoff thermalization for any initial condition $x \neq 0$. The presence of a thermalization profile depends highly on the choice of the parameters $\kappa, \varsigma_1, \varsigma_2, \gamma$.

4.3.2 Numerical Example of a Linear Chain of Oscillators

In the sequel, we set $\zeta_1 = \zeta_n = \kappa = 1$, $\gamma = 0.01$ and $n = 5$. The following computations are carried out in Wolfram Mathematica 12.1. The interaction matrix Q is given by

$$\begin{bmatrix}
 1 & 0 & 0 & 0 & 0 & 1.01 & -1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & -1 & 2.01 & -1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2.01 & -1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 2.01 & -1 \\
 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 1.01 \\
 -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0
 \end{bmatrix}$$

with the following eigenvalues

$$\begin{bmatrix}
 \lambda_1 \\
 \bar{\lambda}_1 \\
 \lambda_2 \\
 \bar{\lambda}_2 \\
 \lambda_3 \\
 \bar{\lambda}_3 \\
 \lambda_4 \\
 \bar{\lambda}_4 \\
 \lambda_5 \\
 \lambda_6
 \end{bmatrix}
 =
 \begin{bmatrix}
 0.0263377 + 1.88656 \cdot i \\
 0.0263377 - 1.88656 \cdot i \\
 0.104782 + 1.55549 \cdot i \\
 0.104782 - 1.55549 \cdot i \\
 0.234099 + 1.06262 \cdot i \\
 0.234099 - 1.06262 \cdot i \\
 0.395218 + 0.517319 \cdot i \\
 0.395218 - 0.517319 \cdot i \\
 0.452655 + 0 \cdot i \\
 0.0264706 + 0 \cdot i
 \end{bmatrix}
 .$$

Since we have 10 complex eigenvalues, we obtain a base of 10 eigenvectors $v_1, \bar{v}_1, v_2, \bar{v}_2, v_3, \bar{v}_3, v_4, \bar{v}_4, v_5, v_6$ where $v_1, v_2, v_3, v_4 \in \mathbb{C}^{10} \setminus \mathbb{R}^{10}$ and $v_5, v_6 \in \mathbb{R}^{10}$, maintaining the natural ordering. Hence for the initial condition x we have the unique representation

$$x = \sum_{j=1}^4 (c_j(x)v_j + \bar{c}_j(x)\bar{v}_j) + c_5(x)v_5 + c_6(x)v_6,$$

where $c_1(x), c_2(x), c_3(x), c_4(x) \in \mathbb{C}$ and $c_5(x), c_6(x) \in \mathbb{R}$. We note that the minimum of real parts of the eigenvalues is taken by the eigenvalues $\lambda_1, \bar{\lambda}_1$. Let $q = \text{Re}(\lambda_1) = 0.0263377$ and $\theta = \text{arg}(\lambda_1) = 1.55684$. Hence, for generic x (not properly contained in any eigenspace) we have

$$\begin{aligned}
 e^{qt} e^{-Qt} x &\approx e^{i\theta t} c_1(x)v_1 + \bar{e}^{i\theta t} \bar{c}_1(x)\bar{v}_1 + 2e^{(q-0.0264706)t} \text{Re}(c_6(x)v_6) \\
 &= 2\text{Re}(e^{i\theta t} c_1(x)v_1) + 2e^{(q-0.0264706)t} \text{Re}(c_6(x)v_6),
 \end{aligned}$$

where $c_6(x)$ is a constant depending on x and

$$v_1 = \begin{bmatrix} 0.112319 - 0.0891416 \cdot i \\ -0.448508 + 0.0287844 \cdot i \\ 0.579305 + 0 \cdot i \\ -0.448508 + 0.0287844 \cdot i \\ 0.112319 - 0.0891416 \cdot i \\ 0.0464105 + 0.060184 \cdot i \\ -0.0119363 - 0.237904 \cdot i \\ -0.00428606 + 0.307009 \cdot i \\ -0.0119363 - 0.237904 \cdot i \\ 0.0464105 + 0.060184 \cdot i \end{bmatrix}.$$

The vector of $c_j = c_j(x)$ is given by

$$c(x) = (c_1, \bar{c}_1, c_2, \bar{c}_2, c_3, \bar{c}_3(x), c_4, \bar{c}_4, c_5, c_6)^* = [v_1|\bar{v}_1|v_2|\bar{v}_2|v_3|\bar{v}_3|v_4|\bar{v}_4|v_5|v_6]^{-1}x.$$

For instance for $x = e_1 = (1, 0, \dots, 0)$ we obtain

$$c(e_1) = \begin{bmatrix} 0.0800993 - 0.0495081 \cdot i \\ 0.0800993 + 0.0495081 \cdot i \\ 0.186213 - 0.0967681 \cdot i \\ 0.186213 + 0.0967681 \cdot i \\ 0.371378 - 0.158507 \cdot i \\ 0.371378 + 0.158507 \cdot i \\ -0.0619613 + 0.787062 \cdot i \\ -0.0619613 - 0.787062 \cdot i \\ 1.62062 + 5.543 \cdot 10^{-16} \cdot i \\ 1.2715 + 2.33866 \cdot 10^{-16} \cdot i \end{bmatrix}.$$

Hence

$$c_1(e_1)v_1 = \hat{w} + i \cdot \check{w} = \begin{bmatrix} 0.00396485 \\ 0.00886695 \\ 0.0101247 \\ 0.0197063 \\ 0.0218997 \\ 0.0375945 \\ 0.0340029 \\ -0.043929 \\ 0.12981 \\ 0.101846 \end{bmatrix} + i \cdot \begin{bmatrix} -0.00793113 \\ -2.28066 \cdot 10^{-9} \\ -0.0169701 \\ -0.001468 \\ -0.0310825 \\ -0.00568992 \\ 0.0661107 \\ -0.0599756 \\ -0.0802337 \\ -0.0629493 \end{bmatrix}$$

and consequently

$$e^{qt} e^{-Qt} e_1 \approx 2 \cos(\theta t) \hat{w} - 2 \sin(\theta t) \check{w}.$$

In the sequel, we check the existence of a thermalization profile for $p \geq 1$. A straightforward computations show that $|\hat{w}| = 0.181073 \neq 0.140425 = |\check{w}|$ and $\langle \hat{w}, \check{w} \rangle = -0.0130705 \neq 0$. Hence Theorem 3.2 yields the absence of a thermalization profile.

Recall that $e^{-qt_\varepsilon} = \varepsilon$. Hence

$$\frac{e^{-Qt_\varepsilon} x}{\varepsilon} = e^{qt_\varepsilon} e^{-Qt_\varepsilon} \approx (e^{i\theta t_\varepsilon} c_1(x)v_1 + \bar{e}^{i\theta t_\varepsilon} \bar{c}_1(x)\bar{v}_1) + 2e^{(q-0.0264706)t_\varepsilon} \text{Re}(c_6(x)v_6)$$

$$\begin{aligned}
 &= 2\operatorname{Re}(e^{i\theta t_\varepsilon} c_1(x) v_1) + 2\varepsilon^{(0.0264706-q)/q} \operatorname{Re}(c_6(x) v_6) \\
 &= 2\operatorname{Re}(e^{i\theta t_\varepsilon} c_1(x) v_1) + 2\varepsilon^{0.0001329} \operatorname{Re}(c_6(x) v_6).
 \end{aligned}$$

The low order of the error is essentially due to the relative spectral gap $(0.0264706-q)/q$.

5 Conceptual Examples

In the sequel, we give mathematical examples illustrating typical features of linear systems. We start with a non-symmetric linear system with complex (conjugate) eigenvalues exhibiting always a thermalization profile. This is followed by an ad hoc example illustrating the sensitive dependence of a thermalization profile on the initial condition. Finally we provide an example of repeated eigenvalues, where a log-log correction in the thermalization time scale appears.

5.1 Example: Leading Complex Eigenvalues Do Not Exclude Profile

Let

$$\mathcal{Q} := \begin{pmatrix} \lambda & \theta \\ -\theta & \lambda \end{pmatrix} \quad \text{with } \lambda > 0 \text{ and } \theta \neq 0.$$

The eigenvalues of $-\mathcal{Q}$ are given by $-\lambda \pm i\theta$. A straightforward computation yields

$$e^{-\mathcal{Q}t} = e^{-\lambda t} \begin{pmatrix} \cos(\theta t) & -\sin(\theta t) \\ \sin(\theta t) & \cos(\theta t) \end{pmatrix}.$$

Hence for any $z = (x, y)^*$ we have

$$|e^{\lambda t} e^{-\mathcal{Q}t} z| = \left| \begin{pmatrix} \cos(\theta t) & -\sin(\theta t) \\ \sin(\theta t) & \cos(\theta t) \end{pmatrix} z \right| = |z|.$$

As a consequence, Theorem 3.1 implies a thermalization profile for any initial value $z \neq 0$.

5.2 Example: The Initial Value Strongly Determines the Cutoff

We consider an embedding of the linear oscillator (4.1) in \mathbb{R}^3 . Assume the case of subcritical damping $\gamma^2 < 4\kappa$ for positive parameters γ, κ, λ ,

$$\mathcal{Q} := \begin{pmatrix} \lambda & 0 \\ 0 & \mathcal{Q}_1 \end{pmatrix} \quad \text{and} \quad \mathcal{Q}_1 = \begin{pmatrix} 0 & -1 \\ \kappa & \gamma \end{pmatrix}.$$

The matrix \mathcal{Q}_1 is precisely the one for the linear oscillator (4.2) analyzed in Example 4.2. A straightforward computation shows

$$e^{-\mathcal{Q}t} = \begin{pmatrix} e^{-\lambda t} & 0 \\ 0 & e^{-\mathcal{Q}_1 t} \end{pmatrix}.$$

For any initial value $z = (z_1, 0, 0)$ with $z_1 \neq 0$ we have $|e^{\lambda t} e^{-\mathcal{Q}t} z| = |z_1|$ and therefore a thermalization profile is valid due to Theorem 3.1. However, for any $z = (0, z_2, z_3)$ with $(z_2, z_3) \neq (0, 0)$ we have $|e^{\frac{\gamma}{2}t} e^{-\mathcal{Q}t} z| = |e^{-\mathcal{Q}_1 t} (z_2, z_3)^*|$ which by the case of subcritical damping ($\gamma^2 < 4\kappa$) discussed in Sect. 4.2 does not have a cutoff thermalization profile. Instead, by Theorem 3.4 only window cutoff thermalization is valid. That is, the presence of a thermalization profile is sensitive with respect to the initial condition.

In the sequel, we emphasize the presence of a threshold effect for the existence of a thermalization profile with respect to the parameters due to competing real parts of the eigenvalues. Let $z = (z_1, z_2, z_3)$ with $z_1 \neq 0, (z_2, z_3) \neq (0, 0)$. If in addition $\frac{\gamma}{2} < \lambda$ we have

$$\lim_{t \rightarrow \infty} e^{\lambda t} |e^{-Q t} z| = |z_1|,$$

which implies that $\omega(z) \subset \{|u| = |z_1|\}$ and therefore by Theorem 3.1 a thermalization profile. However, if $\frac{\gamma}{2} \geq \lambda$ we have

$$\lim_{t \rightarrow \infty} e^{\frac{\gamma}{2} t} |e^{-Q t} z| = \lim_{t \rightarrow \infty} e^{\frac{\gamma}{2} t} e^{-\lambda t} |z_1| + \lim_{t \rightarrow \infty} e^{\frac{\gamma}{2} t} |e^{-Q_1 t} (z_2, z_3)^*|,$$

which is not constant, as discussed in Sect. 4.2, and has only window thermalization (by Theorem 3.4), but no profile due the negative result in Lemma 4.2.

5.3 Example: Multiplicities in the Jordan block Yield Logarithmic Corrections

Let Q be a d -squared matrix with all its eigenvalues equal to $\lambda > 0$. Theorem 7.10 p.209 in [5] yields

$$e^{-Q t} = e^{-\lambda t} \sum_{j=0}^{d-1} \frac{t^j}{j!} (-Q + \lambda I_d)^j.$$

For $z \in \mathbb{R}^d, z \neq 0$ let

$$l(z) = \max\{k \in \{0, \dots, d - 1\} : (-Q + \lambda I_d)^k z \neq 0\}.$$

Then

$$e^{\lambda t} e^{-Q t} z = \sum_{k=0}^{d-1} \frac{t^k}{k!} (-Q + \lambda I_d)^k z = \sum_{k=0}^{l(z)} \frac{t^k}{k!} (-Q + \lambda I_d)^k z.$$

On the one hand, if $l(z) = 0$ we have $e^{\lambda t} e^{-Q t} z = z$. On the other hand, if $l(z) \geq 1$ we obtain

$$\frac{e^{\lambda t}}{t^{l(z)}} e^{-Q t} z = \sum_{k=0}^{l(z)-1} \frac{t^{k-l(z)}}{k!} (-Q + \lambda I_d)^k z + \frac{1}{l(z)!} (-Q + \lambda I_d)^{l(z)} z. \tag{5.1}$$

Hence

$$\lim_{t \rightarrow \infty} \frac{e^{\lambda t}}{t^{l(z)}} e^{-Q t} z = \frac{1}{l(z)!} (-Q + \lambda I_d)^{l(z)} z \neq 0,$$

and in this case $\ell(z) = l(z) + 1 \geq 2$, where $\ell(z)$ is the constant given in Lemma 2.1. Due to Theorem 3.1 there is always a thermalization profile. However, if $\ell(z) \geq 2$, the log-log correction in (3.9) appears. Note that a log-log correction and the presence of a thermalization profile are independent properties. It is not complicated to construct an example with no thermalization profile and log-log correction.

6 Extensions and Applications

This section contains the cutoff phenomenon for the relative entropy in the Brownian case, for the Wasserstein distance with stationary red noises and comments about the computational observation of the cutoff phenomenon.

6.1 Cutoff Thermalization in the Relative Entropy

In this subsection we discuss the asymptotics in the explicit formula (1.2) of the relative entropy for general exponentially asymptotically stable $-Q$.

The strongest notion of thermalization of interest is given in terms of the Kullback-Leibler divergence also called relative entropy. For pure Brownian perturbations the marginals of

$$dX_t^\varepsilon(x) = -QX_t^\varepsilon(x)dt + \varepsilon\sigma dB_t, \quad X_0^\varepsilon = x,$$

for some $\sigma \in \mathbb{R}^{d \times d}$ are known to be

$$X_t^\varepsilon(x) \stackrel{d}{=} N(e^{-Qt}x, \varepsilon^2\Sigma_t),$$

where

$$\Sigma_t = e^{-Qt} \left(\int_0^t e^{Qs} \sigma \sigma^* e^{Q^*s} ds \right) e^{-Q^*t}$$

is a symmetric and non-negative definite square matrix, see Proposition 3.5 in [89]. Since Q satisfies Hypothesis 2.2, we have $e^{-Qt}x \rightarrow 0$ and $\varepsilon^2\Sigma_t \rightarrow \varepsilon^2\Sigma_\infty$ as $t \rightarrow \infty$ which implies the existence of a unique limiting distribution $\mu^\varepsilon = N(0, \varepsilon^2\Sigma_\infty)$. A priori, Σ_t and Σ_∞ may degenerate, however, Theorem 3.4 applies for $p = \infty$ and Theorem 3.3 is valid under condition (3.8). If additionally we assume that $-Q$ and σ are controllable, i.e. $\text{Rank}[-Q, \sigma] = d$, the matrices Σ_t and Σ_∞ turn out to be non-singular. Moreover, Σ_∞ is the unique symmetric positive definite solution of the Lyapunov matrix equation $Q\Sigma_\infty + \Sigma_\infty Q^* = \sigma\sigma^*$. The relative entropy is given explicitly by formula (1.2), which we rewrite as

$$H(X_t^\varepsilon(x)|\mu^\varepsilon) - \frac{1}{2\varepsilon^2} (e^{-Qt}x)^* \Sigma_\infty^{-1} e^{-Qt}x = \frac{1}{2} \left(\text{Tr}(\Sigma_\infty^{-1}\Sigma_t) - d + \ln \left(\frac{\det(\Sigma_\infty)}{\det(\Sigma_t)} \right) \right).$$

For any $t_\varepsilon \rightarrow \infty$ as $\varepsilon \rightarrow 0$ we have that the error term in the right-hand side of the preceding equality tends to zero as $\varepsilon \rightarrow 0$. In the sequel, we analyze the asymptotic quadratic form

$$\frac{1}{2\varepsilon^2} (e^{-Qt_\varepsilon}x)^* \Sigma_\infty^{-1} (e^{-Qt_\varepsilon}x) = \frac{1}{2} \left| \Sigma_\infty^{-1/2} \frac{e^{-Qt_\varepsilon}x}{\varepsilon} \right|^2.$$

By Lemma 2.1 it has the spectral decomposition

$$\lim_{t \rightarrow \infty} \left| \frac{e^{qt}}{t^{\ell-1}} e^{-Qt}x - \sum_{k=1}^m e^{it\theta_k} \nu_k \right| = 0. \tag{6.1}$$

For the scale t_ε^x and w_ε given in Theorem 3.1 we have

$$\lim_{\varepsilon \rightarrow 0} \frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1} e^{-q(t_\varepsilon^x + r \cdot w_\varepsilon)}}{\varepsilon} = q^{1-\ell} e^{-qwr}. \tag{6.2}$$

Lemma 1 implies

$$\begin{aligned} & \left| \left| \Sigma_\infty^{-1/2} \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} \right| - \left| \Sigma_\infty^{-1/2} \frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k \right| \right| \\ & \leq \left| \Sigma_\infty^{-1/2} \frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \left(e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)} \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}} - \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k \right) \right|. \end{aligned}$$

Combining the preceding inequality with (6.1) and (6.2), yields

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \left| \Sigma_\infty^{-1/2} \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} \right| &= q^{1-\ell} e^{-qwr} \limsup_{\varepsilon \rightarrow 0} \left| \Sigma_\infty^{-1/2} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k \right| \\ &\leq q^{1-\ell} e^{-qwr} \left| \Sigma_\infty^{-1/2} \right| \sum_{k=1}^m |v_k| \end{aligned} \tag{6.3}$$

and

$$\begin{aligned} \liminf_{\varepsilon \rightarrow 0} \left| \Sigma_\infty^{-1/2} \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} \right| &= q^{1-\ell} e^{-qwr} \liminf_{\varepsilon \rightarrow 0} \left| \Sigma_\infty^{-1/2} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k \right| \\ &\geq q^{1-\ell} e^{-qwr} |\Sigma_\infty^{1/2}|^{-1} \liminf_{\varepsilon \rightarrow 0} \left| \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k \right| \\ &\geq q^{1-\ell} e^{-qwr} |\Sigma_\infty^{1/2}|^{-1} \liminf_{t \rightarrow \infty} \left| \sum_{k=1}^m e^{it\theta_k} v_k \right|. \end{aligned} \tag{6.4}$$

Hence the analogue of Theorem 3.4 is valid for the relative entropy, that is,

$$\lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \left| \Sigma_\infty^{-1/2} \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} \right| = 0$$

and by (2.4) in Lemma 2.1

$$\lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \left| \Sigma_\infty^{-1/2} \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} \right| = \infty.$$

Moreover, we have the analogue of Theorem 3.1 with the following modification. By (6.3) and (6.4) the existence of a cutoff thermalization profile holds if and only if the geometric condition of $|\Sigma_\infty^{-1/2} \omega(x)|$ being contained in a sphere is satisfied, where $\omega(x)$ is given in (3.7). Recall that the normal growth condition (3.15) in Theorem 3.2 under the non-resonance hypothesis (3.14) in Remark 3.4 is given by

$$\begin{aligned} & (v_1, \hat{v}_2, \check{v}_2, \dots, \hat{v}_{2n}, \check{v}_{2n}) \text{ being orthogonal and satisfying} \\ & |\hat{v}_{2k}| = |\check{v}_{2k}| \text{ for } k = 1, \dots, n. \end{aligned}$$

This characterization of the thermalization profile changes for the relative entropy to the following *weighted normal growth condition*:

$$\begin{aligned} & (\Sigma_\infty^{-1/2} v_1, \Sigma_\infty^{-1/2} \hat{v}_2, \Sigma_\infty^{-1/2} \check{v}_2, \dots, \Sigma_\infty^{-1/2} \hat{v}_{2n}, \Sigma_\infty^{-1/2} \check{v}_{2n}) \text{ is orthogonal and satisfies} \\ & |\Sigma_\infty^{-1/2} \hat{v}_{2k}| = |\Sigma_\infty^{-1/2} \check{v}_{2k}| \text{ for } k = 1, \dots, n. \end{aligned} \tag{6.5}$$

In case of (6.5) the thermalization profile is given by

$$\tilde{\mathcal{P}}_x(r) = q^{1-\ell} e^{-qwr} |\Sigma_\infty^{-1/2} u|, \tag{6.6}$$

where u is any representative of $\omega(x)$. To our knowledge, this result is original and not known in the literature due to the lack of Lemma 2.1.

6.2 Cutoff Thermalization for Small Red and More General Ergodic Noises

In this subsection we show that our results remain intact if we replace the driving Lévy noise by red noise or more general ergodic noises.

In the sequel, we consider the generalized Ornstein–Uhlenbeck process $(X_t^\varepsilon(x))_{t \geq 0}$

$$dX_t^\varepsilon = -\mathcal{Q}X_t^\varepsilon dt + \varepsilon dU_t, \quad X_0^\varepsilon = x, \tag{6.7}$$

where the matrix \mathcal{Q} satisfies Hypothesis 2.2. Equation (6.7) is driven by (i) a stationary multidimensional Ornstein–Uhlenbeck process $(U_t)_{t \geq 0}$ given by

$$dU_t = -\Lambda U_t dt + dL_t, \quad U_0 \stackrel{d}{=} \tilde{\mu}, \tag{6.8}$$

where the matrix Λ satisfies Hypothesis 2.2, $L = (L_t)_{t \geq 0}$ fulfills Hypothesis 2.1 for some $p > 0$ and U_0 is independent of L . We point out that $U_t \stackrel{d}{=} \tilde{\mu}$ for all $t \geq 0$. We stress that with more technical effort the subordinated linear process $(U_t)_{t \geq 0}$ can be replaced by virtually any ergodic (Feller-) Markov process which is sufficiently integrable. For illustration of the ideas we focus on the stationary driving noise given by (6.8). By the variation of constant formula we have

$$X_t^\varepsilon = e^{-\mathcal{Q}t} x + \varepsilon e^{-\mathcal{Q}t} \int_0^t e^{\mathcal{Q}s} dU_s =: e^{-\mathcal{Q}t} x + \varepsilon \mathcal{U}_t.$$

Note that

$$d \begin{pmatrix} X_t^\varepsilon \\ U_t \end{pmatrix} = \Gamma_\varepsilon \begin{pmatrix} X_t^\varepsilon \\ U_t \end{pmatrix} dt + \begin{pmatrix} \varepsilon \\ 1 \end{pmatrix} dL_t.$$

where

$$\Gamma_\varepsilon := \begin{pmatrix} -\mathcal{Q} & -\varepsilon \Lambda \\ 0 & -\Lambda \end{pmatrix}. \tag{6.9}$$

Since \mathcal{Q} and Λ satisfy Hypothesis 2.2 and Γ_ε is an upper block matrix, we have that Γ_ε also satisfies Hypothesis 2.2. In particular, the vector process $((X_t^\varepsilon, U_t))_{t \geq 0}$ is an Ornstein–Uhlenbeck process and hence Markovian. As a consequence, Theorem 4.1 in [98] yields the existence and uniqueness of an invariant and limiting distribution (for the weak convergence) $(X_\infty^\varepsilon, U_0)$ of $((X_t^\varepsilon, U_t))_{t \geq 0}$. Hence $X_\infty^\varepsilon \stackrel{d}{=} \varepsilon \mathcal{U}_\infty$. We continue with the estimate of $\mathcal{W}_{p'}(X_t^\varepsilon, X_\infty^\varepsilon)$. For any $1 \leq p' \leq p$ the analogous computations used in (3.3) yield

$$\left| \frac{\mathcal{W}_{p'}(X_t^\varepsilon, X_\infty^\varepsilon)}{\varepsilon} - \frac{|e^{-\mathcal{Q}t} x|}{\varepsilon} \right| \leq \mathcal{W}_{p'}(\mathcal{U}_t, \mathcal{U}_\infty).$$

Hence cutoff thermalization occurs whenever $\mathcal{W}_{p'}(\mathcal{U}_t, \mathcal{U}_\infty) \rightarrow 0$ as $t \rightarrow \infty$. Properties a), b) and d) of Lemma 2.2 imply

$$\begin{aligned} \mathcal{W}_{p'}(\mathcal{U}_t, \mathcal{U}_\infty) &= \mathcal{W}_{p'}(X_t^1 - e^{-\mathcal{Q}t} x, X_\infty^1) \\ &\leq \mathcal{W}_{p'}(X_t^1 - e^{-\mathcal{Q}t} x, X_\infty^1 - e^{-\mathcal{Q}t} x) + \mathcal{W}_{p'}(X_\infty^1 - e^{-\mathcal{Q}t} x, X_\infty^1) \end{aligned}$$

$$= \mathcal{W}_{p'}(X_t^1, X_\infty^1) + |e^{-\mathcal{Q}t}x|. \tag{6.10}$$

We point out that the vector process $(X_t^1, U_t)_{t \geq 0}$ is a Markov process. Since in general projections of Markovian processes are not Markovian, we study the process X_t in more detail. Due to the triangle structure we have the dependences $X_t^1(x, U_0)$ and $U_t(U_0)$. In case of initial data (x, u) instead of (x, U_0) in (6.7) and (6.8) we write (for $\varepsilon = 1$) $X_t^1(x, u)$ and $U_t(u)$. Analogously to the total variation distance, Theorem 5.2 in [39] the Wasserstein exhibits a contraction property which for completeness is shown here. By the contraction property g) in Lemma 2.2 for the projection $T(x, u) = x$ we have

$$\mathcal{W}_{p'}(X_t^1(x, U_0), X_\infty^1) \leq \mathcal{W}_{p'}((X_t^1(x, U_0), U_t(U_0)), (X_\infty^1, U_t(U_0))) \tag{6.11}$$

for any $1 \leq p' \leq p$. We note that

$$\mathcal{W}_{p'}((X_t^1(x, U_0), U_t(U_0)), (X_\infty^1, U_t(U_0))) = \mathcal{W}_{p'}((X_t^1(x, U_0), U_t(U_0)), (X_\infty^1, U_0)). \tag{6.12}$$

Lemma 2.4 applied to the vector-valued Ornstein–Uhlenbeck process $((X_t^1(x, U_0), U_t(U_0)))_{t \geq 0}$ instead of $(X_t^\varepsilon(x))_{t \geq 0}$ where Γ_1 replaces \mathcal{Q} yields the limit

$$\lim_{t \rightarrow \infty} \mathcal{W}_{p'}((X_t^1(x, U_0), U_t(U_0)), (X_\infty^1, U_0)) = 0.$$

The preceding limit with the help of (6.11) implies

$$\lim_{t \rightarrow \infty} \mathcal{W}_{p'}(X_t^1(x, U_0), X_\infty^1) = 0.$$

Therefore the cutoff thermalization behavior of the Ornstein–Uhlenbeck driven system (6.7) is the same as the white noise driven system (2.1) given in Theorem 3.1 and Theorem 3.4. This is not surprising since the shift-linearity property of the Wasserstein distance for $p \geq 1$ cancels out the specific invariant distribution.

6.3 Conditions on ε for the Observation of the Cutoff on a Fixed Interval $[0, T]$

This subsection provides bounds on the size of ε in order to observe cutoff on a fixed (large) interval $[0, T]$. Similar observations have been made in Section IV of [8] in order illustrate the optimal tuning of the parameter ε .

Our main results contain the time scale $t_\varepsilon \rightarrow \infty$ as $\varepsilon \rightarrow 0$ at which thermalization occurs. However, the computational resources can only cover up to finite time horizon $T > 0$. In the sequel, we line out estimates on the smallest of ε in order to observe the cutoff thermalization before time $T/2 \in [0, T]$, that is, $T \geq 2t_\varepsilon^x$ for $\varepsilon \ll 1$. In other words, we have the lower bound

$$\varepsilon \geq e^{-(\mathcal{Q}T/2)}. \tag{6.13}$$

Since our results are asymptotic, it is required that $\varepsilon < \varepsilon_0$, where ε_0 typically depends of \mathcal{Q}, x and $\mathbb{E}[\mathcal{O}_\infty]$.

Given an error $\eta > 0$. In the light of estimates (3.3) ($p \geq 1$) and (3.5) ($p \in (0, 1)$) we carry out the following error analysis. In the sequel, we always consider a generic initial condition x . Formula (3.17) in item (4) of Remark 3.5 yields the following upper bound of the error

$$C_0 \mathbb{E}[\mathcal{O}_\infty] \varepsilon \leq \eta/2, \tag{6.14}$$

where the constant C_0 is given in terms of the spectral gap in (2.9) and an upper bound of $\mathbb{E}[\mathcal{O}_\infty]$ is expressed in terms of the noise parameters in the estimate (2.12). By Remark 3.7

item (2) we have an error of order

$$K(x)\varepsilon^{\mathfrak{g}/q} \leq \eta/2, \tag{6.15}$$

where the spectral gap \mathfrak{g} is given in (3.18) and the constant $K(x)$ is estimated in (3.19). Combining (6.13) (6.14) and (6.15) and solving for ε yields

$$e^{-(qT/2)} \leq \varepsilon \leq \min \left\{ \frac{\eta}{2C_0\mathbb{E}[|\mathcal{O}_\infty|]}, \left(\frac{\eta}{2K(x)} \right)^{q/\mathfrak{g}} \right\}. \tag{6.16}$$

Acknowledgements The research of GBV has been supported by the Academy of Finland, via the Matter and Materials Profi4 university profiling action. GBV also would like to express his gratitude to University of Helsinki for all the facilities used along the realization of this work. The research of MAH has been supported by the proyecto de la Convocatoria 2020-2021: “Stochastic dynamics of systems perturbed with small Markovian noise with applications in biophysics, climatology and statistics” of the School of Sciences (Facultad de Ciencias) at Universidad de los Andes. JCP acknowledges support from CONACyT-MEXICO CB-250590. The authors would like to thank prof. J. M. Pedraza (Physics Department, Universidad de los Andes) for useful comments which have led to the Sect. 6.3 and prof. J. Goodrick (Mathematics department, Universidad de los Andes) for helpful comments on the introduction. The authors are grateful to the reviewers for the thorough examination of the paper, which has led to a significant improvement.

Funding Open access funding provided by University of Helsinki including Helsinki University Central Hospital.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A. Proof of Lemma 2.2 (Properties of the Wasserstein Distance)

Proof of Lemma 2.2 Property a) is shown for $p \geq 1$ in [106] p. 94. The proof for $p \in (0, 1)$ follows by the same reasoning with the help of the subadditivity of the map $\mathbb{R}_+ \ni r \rightarrow r^p \in \mathbb{R}_+$. Item b) is straightforward for any $p > 0$ due to the translation invariance in formula (2.5). The homogeneity property of item c) follows directly from (2.5) for any $p > 0$.

In the sequel we show item d). Since we are not aware of a proof in the literature the statement is shown here. Synchronous replica (U_1, U_1) with joint law $\Pi(du, du)$ (natural coupling) yields the upper bound for any $p > 0$ as follows

$$\mathcal{W}_p(u_1 + U_1, U_1) \leq \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |u_1 + u - u|^p \Pi(du, du) \right)^{\min\{1, 1/p\}} = |u_1|^{\min\{1, p\}}. \tag{A.1}$$

We continue with the lower bound for $p \geq 1$. Let π any coupling (joint law) between $u_1 + U_1$ and U_1 . Note that

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathbb{R}^d} (u - v)\pi(du, dv) &= \int_{\mathbb{R}^d \times \mathbb{R}^d} u\pi(du, dv) \\ &- \int_{\mathbb{R}^d \times \mathbb{R}^d} v\pi(du, dv) = \mathbb{E}[u_1 + U_1] - \mathbb{E}[U_1] = u_1. \end{aligned}$$

Then the triangular inequality yields

$$|u_1| = \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} (u - v)\pi(du, dv) \right| \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |u - v|\pi(du, dv).$$

Minimizing over all possible coupling between $u_1 + U_1$ and U_1 we obtain

$$|u_1| \leq \mathcal{W}_1(u_1 + U_1, U_1). \tag{A.2}$$

For $p \geq 1$, Jensen’s inequality with the help of (A.1) and (A.2) yields

$$|u_1| \leq \mathcal{W}_1(u_1 + U_1, U_1) \leq \mathcal{W}_p(u_1 + U_1, U_1) \leq |u_1|,$$

such that $\mathcal{W}_p(u_1 + U_1, U_1) = |u_1|$.

For $p \in (0, 1)$, the triangle inequality and the translation invariance b) imply

$$\begin{aligned} |x|^p &= \mathcal{W}_p(x, 0) \leq \mathcal{W}_p(x, x + U) + \mathcal{W}_p(x + U, U) + \mathcal{W}_p(U, 0) \\ &= \mathcal{W}_p(x + U, U) + 2\mathbb{E}[|U|^p], \end{aligned}$$

and hence

$$\mathcal{W}_p(x + U, U) \geq |x|^p - 2\mathbb{E}[|U|^p]. \tag{A.3}$$

Combining (A.1) and (A.3) we obtain (2.7). This finishes the proof of item d).

Property e) is straightforward. The characterization in item f) is proven Theorem 6.9 in [106] for $p \geq 1$. For $p \in (0, 1)$ we refer to Remark 1.4 in [50].

For completeness we give a proof of item g). We apply the Kantorovich duality (Theorem 5.10 p. 57-58 in [106]) for the cost function $c(x, y) = |x - y|^p$ for any $x, y \in \mathbb{R}^k$ and some $p > 0$. Let $\tilde{X} = T(X)$ and $\tilde{Y} = T(Y)$. By item iii) of Theorem 5.10 in [106] we have

$$\mathcal{W}_p(\tilde{X}, \tilde{Y}) = \max_{(\psi, \varphi)} \left(\mathbb{E}[\varphi(\tilde{Y}) - \psi(\tilde{X})] \right)^{\min\{1, \frac{1}{p}\}},$$

where the maximum is running over all integrable functions ψ and φ such that

$$\varphi(x) - \psi(y) \leq |x - y|^p \tag{A.4}$$

for all $x, y \in \mathbb{R}^k$. In addition, item iii) of Theorem 5.10 in [106] states the existence of a respective maximizer (φ_*, ψ_*) . The preceding equality yields

$$\begin{aligned} \mathcal{W}_p(\tilde{X}, \tilde{Y}) &= \left(\mathbb{E}[\varphi_*(\tilde{Y}) - \psi_*(\tilde{X})] \right)^{\min\{1, \frac{1}{p}\}} = \left(\mathbb{E}[\varphi_*(T(Y)) - \psi_*(T(X))] \right)^{\min\{1, \frac{1}{p}\}} \\ &= \left(\mathbb{E}[\varphi_* \circ T(Y) - \psi_* \circ T(X)] \right)^{\min\{1, \frac{1}{p}\}}. \end{aligned} \tag{A.5}$$

Using (A.4) and the fact that T is Lipschitz continuous with Lipschitz constant 1, we have for any $u, v \in \mathbb{R}^d$

$$\varphi_* \circ T(u) - \psi_* \circ T(v) = \varphi_*(T(u)) - \psi_*(T(v)) \leq |T(u) - T(v)|^p \leq |u - v|^p$$

and hence

$$\mathcal{W}_p(X, Y) \geq \left(\mathbb{E}[\varphi_* \circ T(Y) - \psi_* \circ T(X)] \right)^{\min\{1, \frac{1}{p}\}}. \tag{A.6}$$

The statement of item g) is a direct consequence of (A.5) and (A.6). □

Proof of Lemma 2.3 Let $p' \in (0, p]$. Given U_1, U_2 let π be some joint distribution of (U_1, U_2) . By definition

$$\begin{aligned} \mathcal{W}_{p'}(U_1, U_2) &= \inf_{\pi} \int_{\mathbb{R}^d \times \mathbb{R}^d} |u_1 - u_2|^{p'} \pi(du_1, du_2) \\ &= \inf_{\pi} \int_{\mathbb{R}^d \times \mathbb{R}^d} |u_1 - u_2|^{p'} \mathbf{1}\{u_1 \neq u_2\} \pi(du_1, du_2) \\ &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |u_1 - u_2|^{p'} \mathbf{1}\{u_1 \neq u_2\} \pi(du_1, du_2). \end{aligned}$$

Note that for any $u_1, u_2 \in \mathbb{R}^d$

$$|u_1 - u_2|^{p'} \leq \mathbf{1}\{|u_1 - u_2| \leq 1\} + |u_1 - u_2|^p \mathbf{1}\{|u_1 - u_2| > 1\} \leq 1 + |u_1 - u_2|^p$$

and $|u_1 - u_2|^{p'} \mathbf{1}\{u_1 \neq u_2\} \rightarrow \mathbf{1}\{u_1 \neq u_2\}$ as $p' \rightarrow 0$ for any $u_1, u_2 \in \mathbb{R}^d$. By the dominated convergence theorem we obtain

$$\begin{aligned} \lim_{p' \rightarrow 0} \mathcal{W}_{p'}(U_1, U_2) &\leq \lim_{p' \rightarrow 0} \int_{\mathbb{R}^d \times \mathbb{R}^d} |u_1 - u_2|^{p'} \mathbf{1}\{u_1 \neq u_2\} \pi(du_1, du_2) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbf{1}\{u_1 \neq u_2\} \pi(du_1, du_2). \end{aligned}$$

Minimizing over all joint distributions π of U_1, U_2 we obtain

$$\lim_{p' \rightarrow 0} \mathcal{W}_{p'}(U_1, U_2) \leq d_{\text{TV}}(U_1, U_2).$$

Moreover, the dominated convergence theorem also yields the lower bound

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbf{1}\{u_1 \neq u_2\} \pi(du_1, du_2) &= \lim_{p' \rightarrow 0} \int_{\mathbb{R}^d \times \mathbb{R}^d} |u_1 - u_2|^{p'} \mathbf{1}\{u_1 \neq u_2\} \pi(du_1, du_2) \\ &\geq \lim_{p' \rightarrow 0} \mathcal{W}_{p'}(U_1, U_2). \end{aligned}$$

Minimizing π as above we deduce

$$d_{\text{TV}}(U_1, U_2) \geq \lim_{p' \rightarrow 0} \mathcal{W}_{p'}(U_1, U_2)$$

and consequently

$$d_{\text{TV}}(U_1, U_2) = \lim_{p' \rightarrow 0} \mathcal{W}_{p'}(U_1, U_2).$$

□

Appendix B. Proof of Theorem 3.1 and Theorem 3.3 (Cutoff Thermalization Profile)

The following proposition presents the core arguments of the subsequent proofs of Theorem 3.1 and Theorem 3.3.

Proposition B.1 For any $0 < p' \leq p$ it follows

$$\limsup_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_x^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}}$$

$$= \limsup_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \tag{B.1}$$

and

$$\begin{aligned} & \liminf_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} \\ &= \liminf_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right). \end{aligned} \tag{B.2}$$

In particular, the limit

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} \text{ exists} \tag{B.3}$$

if and only if the limit

$$\lim_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \text{ exists.} \tag{B.4}$$

Proof of Proposition B.1 Let $0 < p' \leq p$. We first treat the case $0 < p' \leq 1$. By (3.5), for any $0 < p' \leq 1$, we have

$$\left| \frac{\mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{p'}} - \mathcal{W}_{p'} \left(\frac{e^{-\mathcal{Q}t}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \right| \leq \mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty).$$

We continue with the case $p' \geq 1$. By (3.3) and property d) in Lemma 2.2 for any $p' \geq 1$, we obtain

$$\left| \frac{\mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon)}{\varepsilon} - \mathcal{W}_{p'} \left(\frac{e^{-\mathcal{Q}t}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \right| \leq \mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty).$$

Combining the preceding inequalities we obtain for any $0 < p' \leq p$

$$\left| \frac{\mathcal{W}_{p'}(X_t^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} - \mathcal{W}_{p'} \left(\frac{e^{-\mathcal{Q}t}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \right| \leq \mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty). \tag{B.5}$$

Let t_ε^x be the time scale given in Theorem 3.1 and $w_\varepsilon \rightarrow w > 0$, as $\varepsilon \rightarrow 0$. By (2.11) we have $\mathcal{W}_{p'}(\mathcal{O}_t, \mathcal{O}_\infty) \rightarrow 0$ whenever $t \rightarrow \infty$. Consequently,

$$\limsup_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), X_\infty^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} = \limsup_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty \right)$$

and

$$\liminf_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), X_\infty^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} = \liminf_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)}x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty \right).$$

In the sequel, we study the asymptotics of the drift term $\frac{e^{-\mathcal{Q}t}x}{\varepsilon}$. By Lemma 2.1 it has the spectral decomposition

$$\lim_{t \rightarrow \infty} \left| \frac{e^{\mathcal{Q}t}}{t^{\ell-1}} e^{-\mathcal{Q}t}x - \sum_{k=1}^m e^{it\theta_k} v_k \right| = 0. \tag{B.6}$$

A straightforward calculation shows

$$\lim_{\varepsilon \rightarrow 0} \frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1} e^{-q(t_\varepsilon^x + r \cdot w_\varepsilon)}}{\varepsilon} = q^{\ell-1} e^{-qwr}. \tag{B.7}$$

With the help of the spectral decomposition (B.6) and the triangle inequality we have

$$\begin{aligned} & \mathcal{W}_{p'} \left(\frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\ & \leq \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\ & \quad + \mathcal{W}_{p'} \left(\left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k - \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} \right) + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\ & = \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) + R_\varepsilon^x, \end{aligned} \tag{B.8}$$

where

$$R_\varepsilon^x := \mathcal{W}_{p'} \left(\left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k - \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} \right) + \mathcal{O}_\infty, \mathcal{O}_\infty \right).$$

On the other hand, analogous reasoning yields

$$\begin{aligned} & \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\ & \leq \mathcal{W}_{p'} \left(\frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty \right) + R_\varepsilon^x. \end{aligned} \tag{B.9}$$

Combining (B.8) and (B.9) we have

$$\begin{aligned} & \left| \mathcal{W}_{p'} \left(\frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \right. \\ & \quad \left. - \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \right| \leq R_\varepsilon^x. \end{aligned} \tag{B.10}$$

In the sequel, we show $R_\varepsilon^x \rightarrow 0$ as $\varepsilon \rightarrow 0$. By continuity of $\mathcal{W}_{p'}$, it is enough to prove

$$\left| \frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k - \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} \right| \rightarrow 0, \quad \varepsilon \rightarrow 0.$$

By limit (B.6) and limit (B.7) we obtain for $\varepsilon \rightarrow 0$

$$\begin{aligned} & \left| \frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k - \frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} \right| \\ & = \frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \left| \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k - \frac{e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)} e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}} \right| \rightarrow 0. \end{aligned}$$

Inequality (B.10) with the help of the preceding limit yields

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{e^{-\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)} x}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\ &= \limsup_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right). \end{aligned}$$

The analogous result holds for the lower limit. □

Proof of Theorem 3.3 We start with the proof of i) \implies ii). Statement i) implies that for any $\lambda \geq 0$ the map

$$\omega(x) \ni u \mapsto \mathcal{W}_{p'}(\lambda u + \mathcal{O}_\infty, \mathcal{O}_\infty)$$

defined on

$$\omega(x) = \left\{ \text{accumulation points of } \sum_{k=1}^m e^{it\theta_k} v_k \text{ as } t \rightarrow \infty \right\}.$$

is constant. Hence

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\ &= \mathcal{W}_{p'} \left(q^{1-\ell} e^{-qwr} \hat{u} + \mathcal{O}_\infty, \mathcal{O}_\infty \right). \end{aligned} \tag{B.11}$$

Indeed, by definition of upper limits, there is a sequence $(\varepsilon_n)_{n \in \mathbb{N}}, \varepsilon_n \rightarrow 0, n \rightarrow \infty$ for which the upper limit is the true limit, i.e.,

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\ &= \lim_{n \rightarrow \infty} \mathcal{W}_{p'} \left(\frac{(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})^{\ell-1}}{\varepsilon_n e^{\mathcal{Q}(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})}} \sum_{k=1}^m e^{i(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right). \end{aligned}$$

By (B.7) we have that the sequence

$$\left(\frac{(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})^{\ell-1}}{\varepsilon_n e^{\mathcal{Q}(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})}} \sum_{k=1}^m e^{i(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})\theta_k} v_k \right)_{n \in \mathbb{N}}$$

is uniformly bounded. Then the Bolzano–Weierstrass theorem yields a subsequence $(\varepsilon_{n_j})_{j \in \mathbb{N}}$ of $(\varepsilon_n)_{n \in \mathbb{N}}$ such that

$$\lim_{j \rightarrow \infty} \frac{(t_{\varepsilon_{n_j}}^x + r \cdot w_{\varepsilon_{n_j}})^{\ell-1}}{\varepsilon_{n_j} e^{\mathcal{Q}(t_{\varepsilon_{n_j}}^x + r \cdot w_{\varepsilon_{n_j}})}} \sum_{k=1}^m e^{i(t_{\varepsilon_{n_j}}^x + r \cdot w_{\varepsilon_{n_j}})\theta_k} v_k = q^{1-\ell} e^{-qwr} \hat{u}$$

for some $\hat{u} \in \omega(x)$. By continuity we deduce

$$\limsup_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{\mathcal{Q}(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right)$$

$$\begin{aligned}
 &= \lim_{n \rightarrow \infty} \mathcal{W}_{p'} \left(\frac{(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})^{\ell-1}}{\varepsilon_n e^{q(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})}} \sum_{k=1}^m e^{i(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\
 &= \lim_{j \rightarrow \infty} \mathcal{W}_{p'} \left(\frac{(t_{\varepsilon_{n_j}}^x + r \cdot w_{\varepsilon_{n_j}})^{\ell-1}}{\varepsilon_{n_j} e^{q(t_{\varepsilon_{n_j}}^x + r \cdot w_{\varepsilon_{n_j}})}} \sum_{k=1}^m e^{i(t_{\varepsilon_{n_j}}^x + r \cdot w_{\varepsilon_{n_j}})\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\
 &= \mathcal{W}_{p'}(q^{1-\ell} e^{-qwr} \hat{u} + \mathcal{O}_\infty, \mathcal{O}_\infty).
 \end{aligned}$$

An analogous reasoning also justifies

$$\begin{aligned}
 &\liminf_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\
 &= \mathcal{W}_{p'}(q^{1-\ell} e^{-qwr} \check{u} + \mathcal{O}_\infty, \mathcal{O}_\infty),
 \end{aligned} \tag{B.12}$$

where $\check{u} \in \omega(x)$. For $\lambda = q^{1-\ell} e^{-qwr}$ we obtain

$$\begin{aligned}
 &\lim_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(t_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\
 &= \mathcal{W}_{p'}(q^{1-\ell} e^{-qwr} u + \mathcal{O}_\infty, \mathcal{O}_\infty),
 \end{aligned} \tag{B.13}$$

where u is any element in $\omega(x)$. It follows the proof of ii) \implies i). Statement ii) yields the existence of the limit (B.3). By Proposition B.1 we have that the limit (B.4) also exists. Pick an arbitrary element $u \in \omega(x)$. Then there exists a sequence of $(\varepsilon_n)_{n \in \mathbb{N}}$, $\varepsilon_n \rightarrow 0$, $n \rightarrow \infty$ such that

$$\lim_{n \rightarrow \infty} \sum_{k=1}^m e^{i(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})\theta_k} v_k = u.$$

By (B.7) we have

$$\lim_{n \rightarrow \infty} \frac{(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})^{\ell-1}}{\varepsilon_n e^{q(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})}} \sum_{k=1}^m e^{i(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})\theta_k} v_k = q^{1-\ell} e^{-qwr} u. \tag{B.14}$$

On the other hand, the existence of the limit (B.4) implies for any $(\tilde{\varepsilon}_n)_{n \in \mathbb{N}}$ such that $\tilde{\varepsilon}_n \rightarrow 0$, as $n \rightarrow \infty$ the limit

$$\begin{aligned}
 &\lim_{n \rightarrow \infty} \mathcal{W}_{p'} \left(\frac{(t_{\tilde{\varepsilon}_n}^x + r \cdot w_{\tilde{\varepsilon}_n})^{\ell-1}}{\tilde{\varepsilon}_n e^{q(t_{\tilde{\varepsilon}_n}^x + r \cdot w_{\tilde{\varepsilon}_n})}} \sum_{k=1}^m e^{i(t_{\tilde{\varepsilon}_n}^x + r \cdot w_{\tilde{\varepsilon}_n})\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\
 &= \mathcal{W}_{p'}(q^{1-\ell} e^{-qwr} \tilde{u} + \mathcal{O}_\infty, \mathcal{O}_\infty)
 \end{aligned}$$

for some $\tilde{u} \in \omega(x)$. In particular for $\tilde{\varepsilon}_n = \varepsilon_n$ we have

$$\begin{aligned}
 \mathcal{W}_{p'}(q^{1-\ell} e^{-qwr} \tilde{u} + \mathcal{O}_\infty, \mathcal{O}_\infty) &= \lim_{n \rightarrow \infty} \mathcal{W}_{p'} \left(\frac{(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})^{\ell-1}}{\varepsilon_n e^{q(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})}} \sum_{k=1}^m e^{i(t_{\varepsilon_n}^x + r \cdot w_{\varepsilon_n})\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \\
 &= \mathcal{W}_{p'}(q^{1-\ell} e^{-qwr} u + \mathcal{O}_\infty, \mathcal{O}_\infty),
 \end{aligned}$$

where the last equality follows by (B.14) and continuity. As a consequence, the function

$$\omega(x) \ni u \mapsto \mathcal{W}_{p'}(q^{1-\ell} e^{-qwr} u + \mathcal{O}_\infty, \mathcal{O}_\infty) = \mathcal{W}_{p'}(q^{1-\ell} e^{-qwr} \tilde{u} + \mathcal{O}_\infty, \mathcal{O}_\infty)$$

is constant. Since $r \in \mathbb{R}$ is arbitrary, we obtain the statement i) for arbitrary $\lambda > 0$. \square

Proof of Theorem 3.1 By Theorem 3.3 we have that profile thermalization holds if and only if for any $\lambda \geq 0$ the shift map $\mathcal{S}_\lambda^\infty : \omega(x) \rightarrow [0, \infty)$ defined by $\mathcal{S}_\lambda^\infty(u) = \mathcal{W}_{p'}(\lambda u + \mathcal{O}_\infty, \mathcal{O}_\infty)$ is constant. Since $1 \leq p' \leq p$, property d) of Lemma 2.2 yields

$$\mathcal{W}_{p'}(v + \mathcal{O}_\infty, \mathcal{O}_\infty) = |v|$$

for any $v \in \mathbb{R}^d$. Therefore, $\mathcal{S}_\lambda^\infty(u) = |\lambda u|$ for any $u \in \omega(x)$. Consequently, the map $\mathcal{S}_\lambda^\infty$ is constant if and only if $\omega(x)$ is contained in a sphere. In this case, (B.13) yields the cutoff thermalization profile

$$\mathcal{P}_x(r) = q^{1-\ell} e^{-qwr} |u|,$$

where u is any element in $\omega(x)$. \square

Appendix C. Proof of Theorem 3.4 (Window Cutoff Thermalization)

Proof of Theorem 3.4 Let $0 < p' \leq p$. By Proposition B.1 we have

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} \\ &= \limsup_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right) \end{aligned}$$

and

$$\begin{aligned} & \liminf_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} \\ &= \liminf_{\varepsilon \rightarrow 0} \mathcal{W}_{p'} \left(\frac{(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k + \mathcal{O}_\infty, \mathcal{O}_\infty \right). \end{aligned}$$

In particular, Property d) in Lemma 2.2 yields

$$|u_\varepsilon|^{\min\{p', 1\}} - 2\mathbb{E}[\mathcal{O}_\infty | p'] \leq \mathcal{W}_{p'}(u_\varepsilon + \mathcal{O}_\infty, \mathcal{O}_\infty) \leq |u_\varepsilon|^{\min\{p', 1\}}, \tag{C.1}$$

where

$$u_\varepsilon = \frac{(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k.$$

We start with the upper limit. Therefore by (B.7) and (C.1) we have

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} &\leq \limsup_{\varepsilon \rightarrow 0} |u_\varepsilon|^{\min\{p', 1\}} \\ &\leq q^{1-\ell} e^{-qwr} \limsup_{\varepsilon \rightarrow 0} \left| \sum_{k=1}^m e^{i(\mathfrak{t}_\varepsilon^x + r \cdot w_\varepsilon)\theta_k} v_k \right| \\ &\leq e^{-qwr} q^{1-\ell} \sum_{k=1}^m |v_k|. \end{aligned}$$

Hence,

$$\lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} = 0.$$

We continue with the lower limit. By (B.7) and (C.1) we have

$$\begin{aligned} \liminf_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} &\geq \liminf_{\varepsilon \rightarrow 0} |u_\varepsilon|^{\min\{p', 1\}} - 2\mathbb{E}[|\mathcal{O}_\infty|^{p'}] \\ &= q^{1-\ell} e^{-qwr} \liminf_{\varepsilon \rightarrow 0} \left| \sum_{k=1}^m e^{i(t_\varepsilon^x+r \cdot w_\varepsilon)\theta_k} v_k \right| - 2\mathbb{E}[|\mathcal{O}_\infty|^{p'}] \\ &\geq q^{1-\ell} e^{-qwr} \liminf_{t \rightarrow \infty} \left| \sum_{k=1}^m e^{i\theta_k t} v_k \right| - 2\mathbb{E}[|\mathcal{O}_\infty|^{p'}] \end{aligned}$$

By (2.4) in Lemma 2.1 we have $\liminf_{t \rightarrow \infty} |\sum_{k=1}^m e^{i\theta_k t} v_k| > 0$. Hence,

$$\lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x), \mu^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} = \infty.$$

Appendix D. Proof of Corollary 3.1 and Corollary 3.2 (Moment Thermalization Cutoff)

The following lemma is used in the proof of Corollary 3.1 and in Sect.6.1.

Lemma 1 For any $d \times d$ matrix A it follows

$$\begin{aligned} &\left| \left| A \frac{e^{-Q(t_\varepsilon^x+r \cdot w_\varepsilon)} x}{\varepsilon} \right| - \left| A \frac{(t_\varepsilon^x+r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x+r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x+r \cdot w_\varepsilon)\theta_k} v_k \right| \right| \\ &\leq |A| \frac{(t_\varepsilon^x+r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x+r \cdot w_\varepsilon)}} \left| \left(e^{q(t_\varepsilon^x+r \cdot w_\varepsilon)} \frac{e^{-Q(t_\varepsilon^x+r \cdot w_\varepsilon)} x}{(t_\varepsilon^x+r \cdot w_\varepsilon)^{\ell-1}} - \sum_{k=1}^m e^{i(t_\varepsilon^x+r \cdot w_\varepsilon)\theta_k} v_k \right) \right|. \end{aligned}$$

Proof The inverse triangle inequality and the submultiplicativity of matrix norm imply

$$\begin{aligned} \left| \left| A e^{-Qt} x \right| - \left| A \frac{t^{\ell-1}}{e^{qt}} \sum_{k=1}^m e^{it\theta_k} v_k \right| \right| &\leq \left| A e^{-Qt} x - A \frac{t^{\ell-1}}{e^{qt}} \sum_{k=1}^m e^{it\theta_k} v_k \right| \\ &\leq |A| \left| e^{-Qt} x - \frac{t^{\ell-1}}{e^{qt}} \sum_{k=1}^m e^{it\theta_k} v_k \right|. \end{aligned}$$

In particular, we have

$$\begin{aligned} &\left| \left| A \frac{e^{-Q(t_\varepsilon^x+r \cdot w_\varepsilon)} x}{\varepsilon} \right| - \left| A \frac{(t_\varepsilon^x+r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x+r \cdot w_\varepsilon)}} \sum_{k=1}^m e^{i(t_\varepsilon^x+r \cdot w_\varepsilon)\theta_k} v_k \right| \right| \\ &\leq |A| \frac{(t_\varepsilon^x+r \cdot w_\varepsilon)^{\ell-1}}{\varepsilon e^{q(t_\varepsilon^x+r \cdot w_\varepsilon)}} \left| \left(e^{q(t_\varepsilon^x+r \cdot w_\varepsilon)} \frac{e^{-Q(t_\varepsilon^x+r \cdot w_\varepsilon)} x}{(t_\varepsilon^x+r \cdot w_\varepsilon)^{\ell-1}} - \sum_{k=1}^m e^{i(t_\varepsilon^x+r \cdot w_\varepsilon)\theta_k} v_k \right) \right|. \end{aligned}$$

□

We show the moment thermalization of Corollary 3.1.

Proof of Corollary 3.1 For convenience we start with the case $p' \geq 1$ and we write $\|X\|_{p'} = (\mathbb{E}[|X|^{p'}])^{1/p'}$. We start with the proof of

$$\lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \mathbb{E}[|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)|^{p'}] = \mathbb{E}[|\mathcal{O}_\infty|^{p'}].$$

Note that

$$\frac{1}{\varepsilon} \|X_t^\varepsilon(x)\|_{p'} = \frac{1}{\varepsilon} \|e^{-\mathcal{Q}t}x + \varepsilon \mathcal{O}_t\|_{p'} \leq \frac{1}{\varepsilon} |e^{-\mathcal{Q}t}x| + \|\mathcal{O}_t\|_{p'}.$$

Hence

$$\frac{1}{\varepsilon^{p'}} \|X_t^\varepsilon(x)\|_{p'}^{p'} \leq \left(\frac{1}{\varepsilon} |e^{-\mathcal{Q}t}x| + \|\mathcal{O}_t\|_{p'}\right)^{p'}.$$

Since $\mathcal{W}_p(\mathcal{O}_t, \mathcal{O}_\infty) \rightarrow 0$ as $t \rightarrow \infty$, we have $\|\mathcal{O}_t\|_{p'} \rightarrow \|\mathcal{O}_\infty\|_{p'} < \infty$ as $t \rightarrow \infty$. The preceding inequality yields

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}^{p'} &\leq \limsup_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^x+r \cdot w_\varepsilon)}x| + \|\mathcal{O}_{t_\varepsilon^x+r \cdot w_\varepsilon}\|_{p'}\right)^{p'} \\ &= \left(\limsup_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^x+r \cdot w_\varepsilon)}x| + \|\mathcal{O}_{t_\varepsilon^x+r \cdot w_\varepsilon}\|_{p'}\right)\right)^{p'} \\ &= \left(\limsup_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^x+r \cdot w_\varepsilon)}x| + \|\mathcal{O}_\infty\|_{p'}\right)\right)^{p'}. \end{aligned}$$

By Lemma 1 and the continuity of $y \mapsto |y|^{p'}$, sending $r \rightarrow \infty$ we obtain

$$\begin{aligned} &\lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}^{p'} \\ &\leq \left(\lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^x+r \cdot w_\varepsilon)}x| + \|\mathcal{O}_\infty\|_{p'}\right)\right)^{p'} = \|\mathcal{O}_\infty\|_{p'}^{p'}. \end{aligned} \tag{D.1}$$

We continue with the proof of

$$\lim_{r \rightarrow \infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}^{p'} \geq \|\mathcal{O}_\infty\|_{p'}^{p'}.$$

Note that

$$\|\mathcal{O}_t\|_{p'} = \frac{1}{\varepsilon} \|\varepsilon \mathcal{O}_t\|_{p'} \leq \frac{1}{\varepsilon} \|X_t^\varepsilon(x)\|_{p'} + \frac{1}{\varepsilon} |e^{-\mathcal{Q}t}x|.$$

Hence

$$\liminf_{\varepsilon \rightarrow 0} \|\mathcal{O}_{t_\varepsilon^x+r \cdot w_\varepsilon}\|_{p'} + \liminf_{\varepsilon \rightarrow 0} \left(-\frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^x+r \cdot w_\varepsilon)}x|\right) \leq \liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}\right).$$

Since for a general family $(a_\varepsilon)_{\varepsilon>0}$ we have $\liminf_{\varepsilon \rightarrow 0} (-a_\varepsilon) = -\limsup_{\varepsilon \rightarrow 0} a_\varepsilon$ it follows

$$\|\mathcal{O}_\infty\|_{p'} - \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^x+r \cdot w_\varepsilon)}x| \leq \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}.$$

By continuity we have

$$\left(\|\mathcal{O}_\infty\|_{p'} - \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^x+r \cdot w_\varepsilon)}x|\right)^{p'} \leq \liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}^{p'}\right).$$

By Lemma 1, and sending $r \rightarrow \infty$ we have

$$\|\mathcal{O}_\infty\|_{p'}^{p'} \leq \lim_{r \rightarrow \infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}^{p'}. \tag{D.2}$$

Combining (D.1) and (D.2) we obtain

$$\lim_{r \rightarrow \infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}^{p'} = \lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}^{p'} = \|\mathcal{O}_\infty\|_{p'}^{p'}.$$

In the sequel, we show

$$\lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}^{p'} = \infty.$$

Note that

$$\frac{1}{\varepsilon} |e^{-\mathcal{Q}t}x| = \frac{1}{\varepsilon} \|X_t^\varepsilon(x) - \varepsilon \mathcal{O}_t\|_{p'} \leq \frac{1}{\varepsilon} \|X_t^\varepsilon(x)\|_{p'} + \|\mathcal{O}_t\|_{p'}$$

and hence

$$\begin{aligned} \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}t}x|^{p'} &\leq \left(\frac{1}{\varepsilon} \|X_t^\varepsilon(x)\|_{p'} + \|\mathcal{O}_t\|_{p'} \right)^{p'}. \\ \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}(t_\varepsilon^x+r \cdot w_\varepsilon)}x|^{p'} &\leq \liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} + \|\mathcal{O}_{t_\varepsilon^x+r \cdot w_\varepsilon}\|_{p'} \right)^{p'} \\ &= \left(\liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} + \|\mathcal{O}_{t_\varepsilon^x+r \cdot w_\varepsilon}\|_{p'} \right) \right)^{p'} \\ &= \left(\liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} \right) + \|\mathcal{O}_\infty\|_{p'} \right)^{p'}. \end{aligned}$$

Sending $r \rightarrow -\infty$, Lemma 1 yields

$$\begin{aligned} \infty &= \lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}(t_\varepsilon^x+r \cdot w_\varepsilon)}x|^{p'} \\ &\leq \left(\lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} \right) + \|\mathcal{O}_\infty\|_{p'} \right)^{p'}. \end{aligned}$$

Hence

$$\lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}^{p'} = \infty.$$

We continue with the case $p' \in (0, 1)$ and we write $\|X\|_{p'} = \mathbb{E}[|X|^{p'}]$. We start with the proof of

$$\lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \mathbb{E}[|X_{t_\varepsilon^x+r \cdot w_\varepsilon}^\varepsilon(x)|^{p'}] = \mathbb{E}[|\mathcal{O}_\infty|^{p'}].$$

Note that

$$\frac{1}{\varepsilon^{p'}} \|X_t^\varepsilon(x)\|_{p'} = \frac{1}{\varepsilon^{p'}} \|e^{-\mathcal{Q}t}x + \varepsilon \mathcal{O}_t\|_{p'} \leq \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}t}x|^{p'} + \|\mathcal{O}_t\|_{p'}.$$

Hence

$$\frac{1}{\varepsilon^{p'}} \|X_t^\varepsilon(x)\|_{p'} \leq \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}t}x| + \|\mathcal{O}_t\|_{p'}.$$

Since $\mathcal{W}_p(\mathcal{O}_t, \mathcal{O}_\infty) \rightarrow 0$ as $t \rightarrow \infty$, we have $\|\mathcal{O}_t\|_{p'} \rightarrow \|\mathcal{O}_\infty\|_{p'} < \infty$ as $t \rightarrow \infty$. The preceding inequality yields

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} &\leq \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}(t_\varepsilon^\varepsilon + r \cdot w_\varepsilon)x}|^{p'} + \|\mathcal{O}_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}\|_{p'} \\ &= \limsup_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^\varepsilon + r \cdot w_\varepsilon)x}|\right)^{p'} + \|\mathcal{O}_\infty\|_{p'} \\ &= \left(\limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^\varepsilon + r \cdot w_\varepsilon)x}|\right)^{p'} + \|\mathcal{O}_\infty\|_{p'}. \end{aligned}$$

By Lemma 1 and the continuity of $y \mapsto |y|^{p'}$, sending $r \rightarrow \infty$ we obtain

$$\begin{aligned} &\lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} \\ &\leq \left(\lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} |e^{-\mathcal{Q}(t_\varepsilon^\varepsilon + r \cdot w_\varepsilon)x}|\right)^{p'} + \|\mathcal{O}_\infty\|_{p'} = \|\mathcal{O}_\infty\|_{p'}^{p'}. \end{aligned} \tag{D.3}$$

We continue with the proof of

$$\lim_{r \rightarrow \infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} \geq \|\mathcal{O}_\infty\|_{p'}.$$

Note that

$$\|\mathcal{O}_t\|_{p'} = \frac{1}{\varepsilon^{p'}} \|\varepsilon \mathcal{O}_t\|_{p'} \leq \frac{1}{\varepsilon^{p'}} \|X_t^\varepsilon(x)\|_{p'} + \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}_t x}|^{p'}.$$

Hence

$$\liminf_{\varepsilon \rightarrow 0} \|\mathcal{O}_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}\|_{p'} + \liminf_{\varepsilon \rightarrow 0} \left(-\frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}(t_\varepsilon^\varepsilon + r \cdot w_\varepsilon)x}|^{p'}\right) \leq \liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}\right).$$

Since for a general sequence $(a_\varepsilon)_{\varepsilon > 0}$ we have $\liminf_{\varepsilon \rightarrow 0} (-a_\varepsilon) = -\limsup_{\varepsilon \rightarrow 0} a_\varepsilon$ it follows

$$\|\mathcal{O}_\infty\|_{p'} - \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}(t_\varepsilon^\varepsilon + r \cdot w_\varepsilon)x}|^{p'} \leq \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}.$$

By Lemma 1, and sending $r \rightarrow \infty$ we obtain

$$\|\mathcal{O}_\infty\|_{p'} \leq \lim_{r \rightarrow \infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'}. \tag{D.4}$$

Combining (D.3) and (D.4) we obtain

$$\lim_{r \rightarrow \infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} = \lim_{r \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} = \|\mathcal{O}_\infty\|_{p'}.$$

In the sequel, we show

$$\lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} = \infty.$$

Note that

$$\frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}_t x}|^{p'} = \frac{1}{\varepsilon^{p'}} \|X_t^\varepsilon(x) - \varepsilon \mathcal{O}_t\|_{p'} \leq \frac{1}{\varepsilon^{p'}} \|X_t^\varepsilon(x)\|_{p'} + \|\mathcal{O}_t\|_{p'}$$

and hence

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}(t_\varepsilon^\varepsilon + r \cdot w_\varepsilon)x}|^{p'} \leq \liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} + \|\mathcal{O}_{t_\varepsilon^\varepsilon + r \cdot w_\varepsilon}\|_{p'}\right)$$

$$= \liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\delta + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} \right) + \|\mathcal{O}_\infty\|_{p'}.$$

Sending $r \rightarrow -\infty$, Lemma 1 yields

$$\infty = \lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} |e^{-\mathcal{Q}(t_\varepsilon^\delta + r \cdot w_\varepsilon)x}|_{p'} \leq \lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \left(\frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\delta + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} \right) + \|\mathcal{O}_\infty\|_{p'}.$$

Hence

$$\lim_{r \rightarrow -\infty} \liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^{p'}} \|X_{t_\varepsilon^\delta + r \cdot w_\varepsilon}^\varepsilon(x)\|_{p'} = \infty.$$

□

In the sequel we prove Corollary 3.2.

Proof of Corollary 3.2 Let t_ε^δ be the time scale given in Theorem 3.1. By (B.5) we have for any $\delta > 0$

$$\limsup_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{\delta t_\varepsilon^\delta}^\varepsilon(x), X_\infty^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} = \limsup_{\varepsilon \rightarrow 0} \mathcal{W}_{p'}\left(\frac{e^{-\mathcal{Q}(\delta t_\varepsilon^\delta)x}}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty\right) \tag{D.5}$$

and

$$\liminf_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{\delta t_\varepsilon^\delta}^\varepsilon(x), X_\infty^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} = \liminf_{\varepsilon \rightarrow 0} \mathcal{W}_{p'}\left(\frac{e^{-\mathcal{Q}(\delta t_\varepsilon^\delta)x}}{\varepsilon} + \mathcal{O}_\infty, \mathcal{O}_\infty\right). \tag{D.6}$$

By (2.4) given in Lemma 2.1 we have

$$0 < \liminf_{\varepsilon \rightarrow 0} \left| \frac{e^{q(\delta t_\varepsilon^\delta)}}{(\delta t_\varepsilon^\delta)^{\ell-1}} e^{-\mathcal{Q}(\delta t_\varepsilon^\delta)x} \right| \leq \limsup_{\varepsilon \rightarrow 0} \left| \frac{e^{q(\delta t_\varepsilon^\delta)}}{(\delta t_\varepsilon^\delta)^{\ell-1}} e^{-\mathcal{Q}(\delta t_\varepsilon^\delta)x} \right| < \infty \tag{D.7}$$

for any $\delta > 0$. A straightforward calculation shows

$$\lim_{\varepsilon \rightarrow 0} \frac{(\delta t_\varepsilon^\delta)^{\ell-1} e^{-q(\delta t_\varepsilon^\delta)}}{\varepsilon} = \begin{cases} \infty & \text{for } \delta \in (0, 1), \\ 0 & \text{for } \delta > 1. \end{cases} \tag{D.8}$$

Since

$$\frac{e^{-\mathcal{Q}(\delta t_\varepsilon^\delta)x}}{\varepsilon} = \frac{(\delta t_\varepsilon^\delta)^{\ell-1} e^{-q(\delta t_\varepsilon^\delta)}}{\varepsilon} \frac{e^{q(\delta t_\varepsilon^\delta)}}{(\delta t_\varepsilon^\delta)^{\ell-1}} e^{-\mathcal{Q}(\delta t_\varepsilon^\delta)x},$$

the relations (D.5), (D.6), (D.7) and the limit (D.8) imply with the help of the continuity of $\mathcal{W}_{p'}$ the desired result:

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{W}_{p'}(X_{\delta t_\varepsilon^\delta}^\varepsilon(x), X_\infty^\varepsilon)}{\varepsilon^{\min\{p', 1\}}} = \begin{cases} \infty & \text{for } \delta \in (0, 1), \\ 0 & \text{for } \delta > 1. \end{cases}$$

□

Appendix E. Proof of Theorem 3.2 (Normal Growth Characterization)

We start with the following lemma, which shows that the \mathbb{C} -linear independence of a family of pairs complex conjugate vectors implies the \mathbb{R} -linear independence of the family of real and imaginary parts, in which the characterization in Theorem 3.2 is stated. The lemma is used in the representation (3.13). The proof is given for completeness.

Lemma E.1 (Complexification) *For any $d \in \mathbb{N}$ let $d \geq 2n + r$ for some $n, r \in \mathbb{N}_0$. Consider an arbitrary family of linearly independent vectors $(w_1, \dots, w_r, v_1, \bar{v}_1, \dots, v_n, \bar{v}_n)$ in \mathbb{C}^d , where \bar{v}_j denotes the complex conjugate of v_j . Assume that $w_1, \dots, w_r \in \mathbb{R}^d$. For short we write $v_j = \hat{v}_j + i\check{v}_j$, with $\hat{v}_j, \check{v}_j \in \mathbb{R}^d$. Then the family of vectors*

$$(w_1, \dots, w_r, \hat{v}_1, \check{v}_1, \dots, \hat{v}_n, \check{v}_n)$$

is linear independent in \mathbb{R}^d .

Proof Let $\alpha_1, \beta_1, \dots, \alpha_n, \beta_n, \gamma_1, \dots, \gamma_r \in \mathbb{R}$ such that

$$\gamma_1 w_1 + \dots + \gamma_r w_r + \alpha_1 \hat{v}_1 + \beta_1 \check{v}_1 + \dots + \alpha_n \hat{v}_n + \beta_n \check{v}_n = 0.$$

Let $\hat{g}_j = \gamma_j$ and $\check{g}_j = 0$, and for $j \in \{1, \dots, n\}$ $\hat{c}_j = \hat{d}_j = \alpha_j/2$ and $\check{c}_j = -\check{d}_j = -\beta_j/2$. Then for $c_j = \hat{c}_j + i\check{c}_j$ and $d_j = \hat{d}_j + i\check{d}_j$ we have

$$g_1 w_1 + \dots + g_r w_r + c_1 v_1 + d_1 \bar{v}_1 + \dots + c_n v_n + d_n \bar{v}_n = 0.$$

By Hypothesis $w_1, \dots, w_r, v_1, \bar{v}_1, \dots, v_n, \bar{v}_n$ is a family of linearly independent vectors in \mathbb{C}^{2d} and hence $g_1 = \dots = g_r = c_1 = d_1 = \dots = c_n = d_n = 0$ which is equivalent to $\gamma_1 = \dots = \gamma_r = \alpha_1 = \beta_1 = \dots = \alpha_n = \beta_n = 0$ as desired. \square

In the sequel, we characterize in Lemmas E.2 and E.3 under the non-resonance hypothesis (3.14) given in Remark 3.4 when the function $\omega(x) \ni u \mapsto |u|$ is constant, which is the statement of Theorem 3.1 item i), and compute the norm of (3.13). Using representation (3.13) of $\omega(x)$ defined by (3.7) we prepare the statement of Lemmas E.2 and E.3. Lemma E.2 yields the necessity, while Lemma E.3 states the sufficiency of the normal growth condition (3.15) of Theorem 3.3. The Pythagoras theorem yields

$$\begin{aligned} \left| \sum_{k=1}^m e^{i\theta_k t} v_k \right|^2 &= |v_1|^2 + 4 \left\langle v_1, \sum_{k=1}^n \left(\cos(\theta_{2k} t) \hat{v}_{2k} - \sin(\theta_{2k} t) \check{v}_{2k} \right) \right\rangle \\ &+ 4 \left| \sum_{k=1}^n \left(\cos(\theta_{2k} t) \hat{v}_{2k} - \sin(\theta_{2k} t) \check{v}_{2k} \right) \right|^2 \\ &= |v_1|^2 + 4 \sum_{k=1}^n \left(\cos(\theta_{2k} t) \langle v_1, \hat{v}_{2k} \rangle - \sin(\theta_{2k} t) \langle v_1, \check{v}_{2k} \rangle \right) \\ &+ 4 \left| \sum_{k=1}^n \left(\cos(\theta_{2k} t) \hat{v}_{2k} - \sin(\theta_{2k} t) \check{v}_{2k} \right) \right|^2. \end{aligned}$$

We continue with the last term on the right-hand side of the preceding expression omitting the prefactor

$$\begin{aligned} &\left| \sum_{k=1}^n \left(\cos(\theta_{2k} t) \hat{v}_{2k} - \sin(\theta_{2k} t) \check{v}_{2k} \right) \right|^2 \\ &= \sum_{k=1}^n |\cos(\theta_{2k} t) \hat{v}_{2k} - \sin(\theta_{2k} t) \check{v}_{2k}|^2 \\ &+ \sum_{k \neq k'} \langle \cos(\theta_{2k} t) \hat{v}_{2k} - \sin(\theta_{2k} t) \check{v}_{2k}, \cos(\theta_{2k'} t) \hat{v}_{2k'} - \sin(\theta_{2k'} t) \check{v}_{2k'} \rangle \end{aligned}$$

$$\begin{aligned}
 &= \sum_{k=1}^n \left(\cos^2(\theta_{2k}t) |\hat{v}_{2k}|^2 + \sin^2(\theta_{2k}t) |\check{v}_{2k}|^2 - 2 \cos(\theta_{2k}t) \sin(\theta_{2k}t) \langle \hat{v}_{2k}, \check{v}_{2k} \rangle \right) \\
 &+ \sum_{k \neq k'} \cos(\theta_{2k}t) \cos(\theta_{2k'}t) \langle \hat{v}_{2k}, \hat{v}_{2k'} \rangle + \sum_{k \neq k'} \sin(\theta_{2k}t) \sin(\theta_{2k'}t) \langle \check{v}_{2k}, \check{v}_{2k'} \rangle \\
 &- \sum_{k \neq k'} \sin(\theta_{2k}t) \cos(\theta_{2k'}t) \langle \check{v}_{2k}, \hat{v}_{2k'} \rangle - \sum_{k \neq k'} \cos(\theta_{2k}t) \sin(\theta_{2k'}t) \langle \hat{v}_{2k}, \check{v}_{2k'} \rangle.
 \end{aligned}$$

Combining the preceding equalities we deduce

$$\begin{aligned}
 &\left| \sum_{k=1}^m e^{i\theta_k t} v_k \right|^2 = |v_1|^2 \\
 &+ 4 \sum_{k=1}^n \cos(\theta_{2k}t) \langle v_1, \hat{v}_{2k} \rangle - 4 \sum_{k=1}^n \sin(\theta_{2k}t) \langle v_1, \check{v}_{2k} \rangle \\
 &+ 4 \left| \sum_{k=1}^n \left(\cos(\theta_{2k}t) \hat{v}_{2k} - \sin(\theta_{2k}t) \check{v}_{2k} \right) \right|^2 \\
 &= |v_1|^2 + 4 \sum_{k=1}^n \cos(\theta_{2k}t) \langle v_1, \hat{v}_{2k} \rangle - 4 \sum_{k=1}^n \sin(\theta_{2k}t) \langle v_1, \check{v}_{2k} \rangle \\
 &+ 4 \sum_{k=1}^n \left(\cos^2(\theta_{2k}t) |\hat{v}_{2k}|^2 + \sin^2(\theta_{2k}t) |\check{v}_{2k}|^2 - 2 \cos(\theta_{2k}t) \sin(\theta_{2k}t) \langle \hat{v}_{2k}, \check{v}_{2k} \rangle \right) \\
 &+ 4 \sum_{k \neq k'} \cos(\theta_{2k}t) \cos(\theta_{2k'}t) \langle \hat{v}_{2k}, \hat{v}_{2k'} \rangle + 4 \sum_{k \neq k'} \sin(\theta_{2k}t) \sin(\theta_{2k'}t) \langle \check{v}_{2k}, \check{v}_{2k'} \rangle \\
 &- 4 \sum_{k \neq k'} \sin(\theta_{2k}t) \cos(\theta_{2k'}t) \langle \check{v}_{2k}, \hat{v}_{2k'} \rangle - 4 \sum_{k \neq k'} \cos(\theta_{2k}t) \sin(\theta_{2k'}t) \langle \hat{v}_{2k}, \check{v}_{2k'} \rangle. \tag{E.1}
 \end{aligned}$$

After rearrangement of the sums, we obtain

$$\begin{aligned}
 \left| \sum_{k=1}^m e^{i\theta_k t} v_k \right|^2 &= |v_1|^2 + 4 \sum_{k=1}^n \cos(\theta_{2k}t) \langle v_1, \hat{v}_{2k} \rangle - 4 \sum_{k=1}^n \sin(\theta_{2k}t) \langle v_1, \check{v}_{2k} \rangle \\
 &+ 4 \left| \sum_{k=1}^n \cos(\theta_{2k}t) \hat{v}_{2k} \right|^2 + 4 \left| \sum_{k=1}^n \sin(\theta_{2k}t) \check{v}_{2k} \right|^2 \\
 &- 8 \sum_{k,k'} \sin(\theta_{2k}t) \cos(\theta_{2k'}t) \langle \check{v}_{2k}, \hat{v}_{2k'} \rangle. \tag{E.2}
 \end{aligned}$$

Lemma E.2 *Assume that the family $(v_1, \hat{v}_2, \check{v}_2, \dots, \hat{v}_{2n}, \check{v}_{2n})$ is orthogonal and $|\hat{v}_{2k}| = |\check{v}_{2k}|$ for all k . Then the function*

$$t \mapsto \left| \sum_{k=1}^m e^{i\theta_k t} v_k \right|$$

is constant and has the value $|v_1|^2 + 4 \sum_{k=1}^n |\hat{v}_{2k}|^2$ and consequently $\omega(x) \ni u \mapsto |u|$ is constant.

Proof The orthogonality hypothesis in relation (E.1) yields

$$\left| \sum_{k=1}^m e^{i\theta_k t} v_k \right|^2 = |v_1|^2 + 4 \sum_{k=1}^n \cos^2(\theta_{2k} t) |\hat{v}_{2k}|^2 + 4 \sum_{k=1}^n \sin^2(\theta_{2k} t) |\check{v}_{2k}|^2.$$

Since $|\hat{v}_{2k}| = |\check{v}_{2k}|$ for all k , the Pythagoras identity yields the desired result. □

Lemma E.3 *If the function*

$$\omega(x) \ni u \mapsto |u| \tag{E.3}$$

is constant and the angles $\theta_2, \dots, \theta_{2n}$ are rationally independent according to Remark 3.4 then the family of \mathbb{R}^d -valued vectors $(v_1, \hat{v}_2, \check{v}_2, \dots, \hat{v}_{2n}, \check{v}_{2n})$ is orthogonal and satisfies $|\hat{v}_{2k}| = |\check{v}_{2k}|$ for all $k = 1, \dots, n$.

Proof By (E.2) we have

$$\begin{aligned} \left| \sum_{k=1}^m e^{i\theta_k t} v_k \right|^2 &= |v_1|^2 + 4 \sum_{k=1}^n \cos(\theta_{2k} t) \langle v_1, \hat{v}_{2k} \rangle - 4 \sum_{k=1}^n \sin(\theta_{2k} t) \langle v_1, \check{v}_{2k} \rangle \\ &\quad + 4 \left| \sum_{k=1}^n \cos(\theta_{2k} t) \hat{v}_{2k} \right|^2 + 4 \left| \sum_{k=1}^n \sin(\theta_{2k} t) \check{v}_{2k} \right|^2 \\ &\quad - 8 \sum_{k,k'} \sin(\theta_{2k} t) \cos(\theta_{2k'} t) \langle \check{v}_{2k}, \hat{v}_{2k'} \rangle. \end{aligned}$$

Since the angles $\theta_2, \dots, \theta_{2n}$ are rationally independent, Corollary 4.2.3 in [105] and the assumption that the function (E.3) is constant implies that the following function F is constant:

$$\begin{aligned} F(x) &= |v_1|^2 + 4 \sum_{k=1}^n x_k \langle v_1, \hat{v}_{2k} \rangle - 4 \sum_{k=1}^n y_k \langle v_1, \check{v}_{2k} \rangle + 4 \left| \sum_{k=1}^n x_k \hat{v}_{2k} \right|^2 + 4 \left| \sum_{k=1}^n y_k \check{v}_{2k} \right|^2 \\ &\quad - 8 \sum_{k,k'} y_k x_{k'} \langle \check{v}_{2k}, \hat{v}_{2k'} \rangle, \tag{E.4} \end{aligned}$$

where $x = (x_1, \dots, x_n) \in [-1, 1]^n$ and $x_k^2 + y_k^2 = 1$ for all k . We point out that for each x_k there are two solutions $y_k = \pm\sqrt{1 - x_k^2}$ for the equation $x_k^2 + y_k^2 = 1$ and all combinations of signs of y_k are admitted.

Step 1: We start with the proof of $\langle v_1, \hat{v}_{2k} \rangle = 0$ for all k . Since the function F is constant, comparing the choices $x_k = 1$ and $y_k = 0$ for all k with $x_k = -1$ and $y_k = 0$ for all k yields

$$|v_1|^2 + 4 \sum_{k=1}^n \langle v_1, \hat{v}_{2k} \rangle + \sum_{k=1}^n |\hat{v}_{2k}|^2 = |v_1|^2 - 4 \sum_{k=1}^n \langle v_1, \hat{v}_{2k} \rangle + \sum_{k=1}^n |\hat{v}_{2k}|^2,$$

which implies

$$\sum_{k=1}^n \langle v_1, \hat{v}_{2k} \rangle = 0. \tag{E.5}$$

Comparing the choice $x_1 = -1, x_k = 1$ for $k \geq 2$, and consequently $y_k = 0$ for all k with $x_1 = 1, x_k = -1$ for $k \geq 2$, and consequently $y_k = 0$ for all k we have

$$|v_1|^2 - 4 \langle v_1, \hat{v}_2 \rangle + 4 \sum_{k=2}^n \langle v_1, \hat{v}_{2k} \rangle + 4| - \hat{v}_2 + \sum_{k=2}^n \hat{v}_{2k} |^2$$

$$\begin{aligned}
 &= |v_1|^2 + 4\langle v_1, \hat{v}_2 \rangle - 4 \sum_{k=2}^n \langle v_1, \hat{v}_{2k} \rangle + 4|\hat{v}_2 - \sum_{k=2}^n \hat{v}_{2k}|^2 \\
 &= |v_1|^2 + 4\langle v_1, \hat{v}_2 \rangle - 4 \sum_{k=2}^n \langle v_1, \hat{v}_{2k} \rangle + 4|-\hat{v}_2 + \sum_{k=2}^n \hat{v}_{2k}|^2,
 \end{aligned}$$

which implies

$$\langle v_1, \hat{v}_2 \rangle - \sum_{k=2}^n \langle v_1, \hat{v}_{2k} \rangle = 0. \tag{E.6}$$

Combining (E.5) and (E.6) we obtain $\langle v_1, \hat{v}_2 \rangle = 0$. Analogously it is shown $\langle v_1, \hat{v}_{2k} \rangle = 0$ for all $k \geq 2$. Switching the role of x_k and y_k in the preceding reasoning also shows $\langle v_1, \check{v}_{2k} \rangle = 0$ for all k .

Step 2: By Step 1 formula (E.4) boils down to

$$F(x) = |v_1|^2 + 4 \left| \sum_{k=1}^n x_k \hat{v}_{2k} \right|^2 + 4 \left| \sum_{k=1}^n y_k \check{v}_{2k} \right|^2 - 8 \sum_{k,k'} x_k y_{k'} C_{k,k'},$$

where $C_{k,k'} = \langle \check{v}_{2k}, \hat{v}_{2k'} \rangle$. The choice $x_1 = \frac{\sqrt{2}}{2}$, $x_k = \pm 1$ for all $k \geq 2$ implies $y_1 = \pm \frac{\sqrt{2}}{2}$ and $y_k = 0$ for $k \geq 2$. Hence

$$\begin{aligned}
 &|v_1|^2 + 4 \left| \frac{\sqrt{2}}{2} \hat{v}_2 + \sum_{k=2}^n (\pm) \hat{v}_{2k} \right|^2 + 2|\check{v}_2|^2 - 4C_{1,1} - 4\sqrt{2} \sum_{k \geq 2} (\pm) C_{k,1} \\
 &= |v_1|^2 + 4 \left| \frac{\sqrt{2}}{2} \hat{v}_2 + \sum_{k=2}^n (\pm) \hat{v}_{2k} \right|^2 + 2|\check{v}_2|^2 + 4C_{1,1} + 4\sqrt{2} \sum_{k \geq 2} (\pm) C_{k,1}
 \end{aligned}$$

for any sequence of signs \pm . Consequently,

$$C_{1,1} + \sqrt{2} \sum_{k \geq 2} (\pm) C_{k,1} = 0$$

for any sequence of signs \pm . Therefore $C_{k,1} = 0$ for all $k \geq 1$. Analogously, we infer that $C_{k,k'} = 0$ for all $k, k' \geq 1$. This proves that $\langle \hat{v}_{2k}, \check{v}_{2k'} \rangle = 0$ for all k, k' .

Step 3: By Step 2 we obtain

$$F(x) = |v_1|^2 + 4 \left| \sum_{k=1}^n x_k \hat{v}_{2k} \right|^2 + 4 \left| \sum_{k=1}^n y_k \check{v}_{2k} \right|^2.$$

For any choice $x_1 \in \{-1, 1\}$, $x_k = 1$ (which implies $y_k = 0$) for all $k \geq 2$, the Pythagoras theorem yields

$$F(x) = |v_1|^2 + 4|x_1 \hat{v}_2 + \sum_{k=2}^n \hat{v}_{2k}|^2 = |v_1|^2 + 4|\hat{v}_2|^2 + 4 \left| \sum_{k=2}^n \hat{v}_{2k} \right|^2 + 8 \left\langle x_1 \hat{v}_2, \sum_{k=2}^n \hat{v}_{2k} \right\rangle,$$

which implies

$$\left\langle \hat{v}_2, \sum_{k=2}^n \hat{v}_{2k} \right\rangle = 0,$$

due to x_1 can be chosen ± 1 . An analogous reasoning yields

$$\sum_{\substack{k=1 \\ k \neq k'}}^n \langle \hat{v}_{2k'}, \hat{v}_{2k} \rangle = 0. \tag{E.7}$$

for any $k' \geq 1$.

For any choice $x_k \in \{-1, 1\}$ (which implies $y_k = 0$) for all k , the Pythagoras theorem yields

$$\begin{aligned} F(x) &= |v_1|^2 + 4 \left| \sum_{k=1}^n x_k \hat{v}_{2k} \right|^2 = |v_1|^2 + 4|x_1 \hat{v}_2 + x_2 \hat{v}_4 + \sum_{k \geq 3} x_k \hat{v}_{2k}|^2 \\ &= |v_1|^2 + 4|x_1 \hat{v}_2 + x_2 \hat{v}_4|^2 + 4 \left| \sum_{k \geq 3} x_k \hat{v}_{2k} \right|^2 + 8 \left\langle x_1 \hat{v}_2 + x_2 \hat{v}_4, \sum_{k \geq 3} x_k \hat{v}_{2k} \right\rangle. \end{aligned}$$

Specifying $x_k = 1$ for all $k \geq 3$ and comparing it with $x_k = -1$ for all $k \geq 3$ implies

$$\left\langle x_1 \hat{v}_2 + x_2 \hat{v}_4, \sum_{k \geq 3} \hat{v}_{2k} \right\rangle = 0.$$

Then

$$\langle x_1 \hat{v}_2, \sum_{k \geq 3} \hat{v}_{2k} \rangle + \langle x_2 \hat{v}_4, \sum_{k \geq 3} \hat{v}_{2k} \rangle = 0.$$

Bearing in mind (E.7), summing and subtracting \hat{v}_4 and \hat{v}_2 in the corresponding sums above yields

$$\langle x_1 \hat{v}_2, -\hat{v}_4 \rangle + \langle x_2 \hat{v}_4, -\hat{v}_2 \rangle = 0.$$

Since $x_1, x_2 \in \{-1, 1\}$ is arbitrary, we infer $\langle \hat{v}_2, \hat{v}_4 \rangle = 0$. By analogous reasoning it is shown that $\langle \hat{v}_{2k}, \hat{v}_{2k'} \rangle = 0$ for all $k \neq k'$. Switching the role of x_k and y_k in the preceding reasoning also shows $\langle \check{v}_{2k}, \check{v}_{2k'} \rangle = 0$ for all $k \neq k'$.

Combining Step 1-3 shows that the family $(v_1, \hat{v}_2, \check{v}_2, \dots, \hat{v}_{2n}, \check{v}_{2n})$ is orthogonal.

Step 4: In the sequel, we prove $|\hat{v}_{2k}| = |\check{v}_{2k}|$ for all k . By Step 3 we obtain

$$F(x) = |v_1|^2 + 4 \sum_{k=1}^n x_k^2 |\hat{v}_{2k}|^2 + 4 \sum_{k=1}^n y_k^2 |\check{v}_{2k}|^2.$$

The choice $x_1 = \zeta, x_k = 1$ for all $k \geq 2$ implies $y_1 = \pm\sqrt{1 - \zeta^2}, y_k = 0$ for all $k \geq 2$. Therefore the quadratic polynomial

$$\begin{aligned} \zeta \mapsto F((\zeta, 1, \dots, 1)) &= |v_1|^2 + 4 \sum_{k=2}^n |\hat{v}_{2k}|^2 + 4\zeta^2 |\hat{v}_2|^2 + 4(1 - \zeta^2) |\check{v}_2|^2 \\ &= |v_1|^2 + 4 \sum_{k=2}^n |\hat{v}_{2k}|^2 + 4|\check{v}_2|^2 + 4\zeta^2 (|\hat{v}_2|^2 - |\check{v}_2|^2) \end{aligned}$$

is constant. Consequently, $|\hat{v}_2|^2 = |\check{v}_2|^2$. Analogously it is shown that $|\hat{v}_{2k}|^2 = |\check{v}_{2k}|^2$ for all $k \geq 2$. □

References

1. Aldous, D.: Random walks on finite groups and rapidly mixing Markov chains. In: Seminar on Probability, XVII. Lecture Notes in Math, vol. 986, pp. 243–297. Springer, Berlin (1983)
2. Aldous, D.: Hitting times for random walks on vertex-transitive graphs. *Math. Proc. Camb. Philos. Soc.* **106**(1), 179–191 (1989)
3. Aldous, D., Diaconis, P.: Strong uniform times and finite random walks. *Adv. Appl. Math.* **8**(1), 69–97 (1987)
4. Aldous, D., Diaconis, P.: Shuffling cards and stopping times. *Am. Math. Mon.* **93**(5), 333–348 (1986)
5. Apostol, T.M.: *Calculus*, vol. II, 2nd edn. Wiley, New York (1967)
6. Applebaum, D.: *Lévy Processes and Stochastic Calculus*. Cambridge University Press, Cambridge (2004)
7. Arrhenius, S.A.: Über die Dissociationswärme und den Einfluß der Temperatur auf den Dissociationsgrad der Elektrolyte. *Z. Phys. Chem.* **4**, 96–116 (1889)
8. Bayati, B., Owahi, H., Koumoutsakos, P.: A cutoff phenomenon in accelerated stochastic simulations of chemical kinetics via flow averaging (FLAVOR-SSA). *J. Chem. Phys.* **133**, 244–117 (2010)
9. Bayer, D., Diaconis, P.: Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.* **2**(2), 294–313 (1992)
10. Biler, P., Karch, G.: Generalized Fokker-Planck equations and convergence to their equilibria. *Banach Cent. Publ.* **60**(1), 307–318 (2003)
11. Barrera, G., Högele, M., Pardo, J.C.: The cutoff phenomenon in total variation for nonlinear Langevin systems with small layered stable noise. *Electron. J. Probab.* 2021. (to appear). [arXiv:2011.10806](https://arxiv.org/abs/2011.10806)
12. Barrera, G., Jara, M.: Abrupt convergence of stochastic small perturbations of one dimensional dynamical systems. *J. Stat. Phys.* **163**(1), 113–138 (2016)
13. Barrera, G., Jara, M.: Thermalisation for small random perturbations of dynamical systems. *Ann. Appl. Probab.* **30**(3), 1164–1208 (2020)
14. Barrera, G., Pardo, J.C.: Cut-off phenomenon for the Ornstein-Uhlenbeck processes driven by Lévy processes. *Electron. J. Probab.* **25**(15), 1–133 (2020)
15. Barrera, J., Lachaud, B., Ycart, B.: Cut-off for n-tuples of exponentially converging processes. *Stoch. Process. Appl.* **116**(10), 1433–1446 (2006)
16. Barrera, J., Bertoncini, O., Fernández, R.: Abrupt convergence and escape behavior for birth and death chains. *J. Stat. Phys.* **137**(4), 595–623 (2009)
17. Basu, R., Hermon, J., Peres, Y.: Characterization of cutoff for reversible Markov chains. *Ann. Probab.* **45**(3), 1448–1487 (2017)
18. Bertoncini, O., Barrera, J., Fernández, R.: Cut-off and exit from metastability: two sides of the same coin. *C. R. Acad. Sci. Paris Ser. I*(346), 691–696 (2008)
19. Berglund, N., Gentz, B.: On the noise-induced passage through an unstable periodic orbit I: Two-level model. *J. Stat. Phys.* **114**(5–6), 1577–1618 (2004)
20. Berglund, N., Gentz, B.: The Eyring-Kramers law for potentials with nonquadratic saddles. *Markov Process. Relat. Fields* **16**, 549–598 (2010)
21. Berglund, N., Gentz, B.: Sharp estimates for metastable lifetimes in parabolic SPDEs: Kramers’ law and beyond. *Electron. J. Probab.* **18**(24), 1–58 (2013)
22. Bódai, T., Franzke, C.: Predictability of fat-tailed extremes. *Phys. Rev. E* **3**(96), 032120 (2017)
23. Bovier, A., Eckhoff, M., Gaynard, V., Klein, M.: Metastability in reversible diffusion processes I: Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc.* **6**, 399–424 (2004)
24. Bovier, A., Gaynard, V., Klein, M.: Metastability in reversible diffusion processes II: Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc.* **7**, 69–99 (2005)
25. Barret, F., Bovier, A., Méléard, S.: Uniform estimates for metastable transitions in a coupled bistable system. *Electron. J. Probab.* **15**(12), 323–345 (2010)
26. Brassescio, S.: Unpredictability of an exit time. *Stoch. Process. Appl.* **63**, 55–65 (1996)
27. Budhiraja, A., Dupuis, P., Maroulas, V.: Variational representations for continuous time processes. *Ann. de l’Inst. Henri Poincaré (B) Probabilités et Statistiques* **47**(3), 725–747 (2011)
28. Budhiraja, A., Chen, J., Dupuis, P.: Large deviations for stochastic partial differential equations driven by a Poisson random measure. *Stoch. Process. Appl.* **123**(2), 523–560 (2013)
29. Budhiraja, A., Nyquist, P.: Large deviations for multidimensional state-dependent shot-noise processes. *J. Appl. Probab.* **52**(4), 1097–1114 (2015)
30. Budhiraja, A., Dupuis, P.: Analysis and Approximation of rare events. *Representations and Weak Convergence Methods. Probability Theory and Stochastic Modelling*, vol. 94. Springer (2019)
31. Chen, X., Wu, F., Duan, J., Kurths, J., Lic, X.: Most probable dynamics of a genetic regulatory network under stable Lévy noise. *Appl. Math. Comput.* **348**(1), 425–436 (2019)
32. Chen, G., Saloff-Coste, L.: The cutoff phenomenon for ergodic Markov processes. *Electron. J. Probab.* **13**(3), 26–78 (2008)

33. Diaconis, P., Graham, R.L., Morrison, J.A.: Asymptotic analysis of a random walk on a hypercube with many dimensions. *Random Struct. Algorithms* **1**(1), 51–72 (1990)
34. Day, M.V.: On the exponential exit law in the small parameter exit problem. *Stochastics* **8**, 297–323 (1983)
35. Day, M.V.: Exit cycling for the Van der Pol oscillator and quasipotential calculations. *J. Dyn. Differ. Equ.* **8**(4), 573–601 (1996)
36. de Oliveira Gomes, A., Högele, M.: The Kramers problem for SDEs driven by small, accelerated Lévy noise with exponentially light jumps. *Stoch. Dyn.* **2150019**, (2021)
37. Debussche, A., Högele, M., Imkeller, P.: Metastability for the Chafee-Infante equation with small heavy-tailed Lévy noise. *Electron. Commun. Probab.* **16**, 213–225 (2011)
38. Debussche, A., Högele, M., Imkeller, P.: The Dynamics of Nonlinear Reaction-Diffusion Equations with Small Lévy Noise. *Springer Lecture Notes in Mathematics* **2085**, 163 (2013)
39. Devroye, L., Lugosi, G.: *Combinatorial Methods in Density Estimation*. Springer, New York (2001)
40. Diaconis, P., Shahshahani, M.: Time to reach stationarity in the Bernoulli-Laplace diffusion model. *SIAM J. Math. Anal.* **18**, 208–218 (1987)
41. Diaconis, P.: The cut-off phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. USA* **93**(4), 1659–1664 (1996)
42. Diaconis, P.: *Group Representations in Probability and Statistics*. Lecture Notes-Monograph Series, vol. 11. Institute of Mathematical Statistics, Hayward (1988)
43. Ditlevsen, P.D.: Observation of α -stable noise induced millennial climate changes from an ice record. *Geophys. Res. Lett.* **10**(26), 1441–1444 (1999)
44. Ditlevsen, P.D.: Anomalous jumping in a double-well potential. *Phys. Rev. E* **1**(60), 172–179 (1999)
45. Ditlevsen, P.D., Ditlevsen, O.D.: On the stochastic nature of the rapid climate shifts during the last Ice Age. *J. Clim.* **2**(22) (2009)
46. Eyring, H.: The activated complex in chemical reactions. *J. Chem. Phys.* **3**, 107–115 (1935)
47. Freidlin, M.I., Wentzell, A.D.: *Random Perturbation of Dynamics Systems*, 3rd edn. Springer-Verlag, Berlin (2012)
48. Faris, G.W., Jona-Lasinio, G.: Large fluctuations for a nonlinear heat equation with noise. *J. Phys. A* **15**, 3025–3055 (1982)
49. Freidlin, M.I.: Random perturbations of reaction-diffusion equations: the quasideterministic approximation. *Trans. Am. Math. Soc.* **305**(2), 665–697 (1988)
50. Friesen, M., Jin, P., Rüdiger, B.: Stochastic equation and exponential ergodicity in Wasserstein distance for affine processes. *Ann. Appl. Probab.* **30**(5), 2165–2195 (2020)
51. Galves, A., Olivieri, E., Vares, M.E.: Metastability for a class of dynamical systems subject to small random perturbations. *Ann. Probab.* **15**, 1288–1305 (1987)
52. Gairing, J., Högele, M., Kosenkova, T., Monahan, A.: How close are time series to power tail Lévy diffusions? *Chaos AIP* **27**(11), 073112-1–073112-20 (2017)
53. Gammaitoni, L., Hänggi, P., Jung, P., Marchesoni, F.: Stochastic Resonance. *Rev. Mod. Phys.* **70**, 223–288 (1998)
54. Garbaczewski, P., Olkiewicz, R.: Ornstein-Uhlenbeck-Cauchy process. *J. Math. Phys.* **41**, 6843–6860 (2000)
55. Godovanchuk, V.V.: Asymptotic probabilities of large deviations due to large jumps of a Markov process. *Theory Probab. Appl.* **26**, 314–327 (1982)
56. Hermon, J.: Some inequalities for reversible Markov chains and branching random walks via spectral optimization (2021). [arXiv:1908.08525v3](https://arxiv.org/abs/1908.08525v3)
57. Hermon, J.: A spectral characterization for concentration of the cover time. *J. Theor. Probab.* **33**, 2167–2184 (2020)
58. Högele, M., Pavlyukevich, I.: The exit problem from the neighborhood of a global attractor for heavy-tailed Lévy diffusions. *Stoch. Anal. Appl.* **32**(1), 163–190 (2013)
59. Högele, M., Pavlyukevich, I.: Metastability in a class of hyperbolic dynamical systems perturbed by heavy-tailed Lévy type noise. *Stoch. Dyn.* **15**, no. 3, 1550019-1-1550019–26 (2015)
60. Högele, M., Pavlyukevich, I.: The first passage problem for stable linear delay equations perturbed by power law Lévy noise. *Chaos AIP* **29**, no. 6, 063104-1-063104–17 (2019)
61. Hänggi, P., Talkner, P., Borkovec, M.: Reaction rate theory: fifty years after kramers. *Rev. Mod. Phys.* **62**, 251–342 (1990)
62. Högele, M.: The first exit problem of reaction-diffusion equations for small multiplicative Lévy noise. *ALEA Lat. Am. J. Probab. Math. Stat.* **16**, 665–709 (2019)
63. Imkeller, P., Pavlyukevich, I.: Metastable behaviour of small noise Lévy-driven diffusions. *ESAIM: Prob. Stat.* **12**, 412–437 (2008)

64. Imkeller, P., Pavlyukevich, I.: First exit times of SDEs driven by stable Lévy processes. *Stoch. Proc. Appl.* **116**(4), 611–642 (2006)
65. Imkeller, P., Pavlyukevich, I., Stauch, M.: First exit times of non-linear dynamical systems in \mathbb{R}^d perturbed by multifractal Lévy noise. *J. Stat. Phys.* **141**(1), 94–119 (2010)
66. Jacobsen, M.: Laplace and the origin of the Ornstein-Uhlenbeck process. *Bernoulli* **2**(3), 271–286 (1996)
67. Johnson, P.D., Ticozzi, F., Viola, L.: Exact stabilization of entangled states in finite time by dissipative quantum circuits. *Phys. Rev. A* **96**, (2017)
68. Jonsson, G.F., Trefethen, L.N.: A numerical analysis looks at the ‘cut-off phenomenon’ in card shuffling and other Markov chains. In: *Numerical Analysis 1997 (Dundee 1997)*, pp. 150–178. Addison Wesley Longman, Harlow (1998)
69. Jakšić, V., Pillet, C., Shirikyan, A.: Entropic fluctuations in thermally driven harmonic networks. *J. Stat. Phys.* **166**, 926–1015 (2017)
70. Kramers, H.A.: Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284–304 (1940)
71. Karatzas, I., Shreve, S.E.: *Brownian Motion and Stochastic Calculus*, 2nd edn. Springer, New York (1998)
72. Kastoryano, M.J., Reeb, D., Wolf, M.M.: A cutoff phenomenon for quantum Markov chains. *J. Phys. A* **45**, 075307 (2012)
73. Kastoryano, M.J., Wolf, M.M., Eisert, J.: Precisely timing dissipative quantum information processing. *Phys. Rev. Lett.* **110**, 110501 (2013)
74. Labbé, C., Lacoïn, H.: Cutoff phenomenon for the asymmetric simple exclusion process and the biased card shuffling. *Ann. Probab.* **47**, 1541–1586 (2019)
75. Lacoïn, H.: The cutoff profile for the simple exclusion process on the circle. *Ann. Probab.* **44**, (2016)
76. Lancia, C., Nardi, F.R., Scoppola, B.: Entropy-driven cutoff phenomena. *J. Stat. Phys.* **149**(1), 108–141 (2012)
77. Levin, D., Luczak, M.J., Peres, Y.: Glauber dynamics for mean-field Ising model: cut-off, critical power law, and metastability. *Probab. Theory Relat. Fields* **146**(1), 223–265 (2010)
78. Levin, D., Peres, Y., Wilmer, E.: *Markov Chains and Mixing Times*. Amer. Math. Soc, Providence (2009)
79. Lubetzky, E., Sly, A.: Cutoff for the Ising model on the lattice. *Invent. Math.* **191**, 719–755 (2013)
80. Mélio, P.L.: The Cut-off phenomenon for Brownian motions on compact symmetric spaces. *Potential Anal.* **40**, 427–509 (2014)
81. Magnus, J.R., Neudecker, H.: *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, New York (1999)
82. Masuda, H.: On Multidimensional Ornstein-Uhlenbeck process driven by a general Lévy process. *Bernoulli* **10**(1), 97–120 (2004)
83. Murray, R.W., Pego, R.L.: Cutoff estimates for the Becker-Döring equations. *Commun. Math. Sci.* **15**, 1685–1702 (2017)
84. Murray, R.W., Pego, R.L.: Algebraic decay to equilibrium for the Becker-Döring equations. *SIAM J. Math. Anal.* **48**(4), 2819–2842 (2016)
85. Ornstein, L., Uhlenbeck, G.: On the theory of Brownian motion. *Phys. Rev.* **36**(5), 823–841 (1930)
86. Panaretos, V., Zemel, Y.: *An invitation to statistics in Wasserstein space*. SpringerBriefs in Probability and Mathematical Statistics (2020)
87. Panaretos, V., Zemel, Y.: Statistical aspects of Wasserstein distances. *Annu. Rev. Stat. Appl.* **6**, 405–31 (2019)
88. Patie, P., Vaidyanathan, A.: A spectral theoretical approach for hypocoercivity applied to some degenerate hypoelliptic, and non-local operators. *Kinet. Relat. Models* **13**(3), 479–506 (2020)
89. Pavliotis, G.A.: *Stochastic Processes and Applications Diffusion Processes, the Fokker-Planck and Langevin Equations*. Springer, New York (2014)
90. Pavlyukevich, I.: First exit times of solutions of stochastic differential equations driven by multiplicative Lévy noise with heavy tails. *Stoch. Dyn.* **11**(2–3), 495–519 (2011)
91. Penland, C., Ewald, B.E.: On modelling physical systems with stochastic models: diffusion versus Lévy processes. *Philos. Trans. R. Soc. Lond. A* **366**, 2457–2476 (2008)
92. Raquépas, R.: A note on Harris ergodic theorem, controllability and perturbations of harmonic networks. *Ann. Henri Poincaré* **20**, 605–629 (2019)
93. Risken, H.: *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer, New York (1989)
94. Rudin, W.: *Principles of Mathematical Analysis*, 3rd edn. McGraw-Hill Inc, New York (1976)
95. Saloff-Coste, L.: Random walks on finite groups. In: Kesten H. (eds) *Probability on Discrete Structures*, vol. 110, pp. 263–346. Springer (2004)

96. Saloff-Coste, L.: Precise estimates on the rate at which certain diffusions tend to equilibrium. *Math. Z.* **217**, 641–677 (1994)
97. Sato, K.: *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge (1999)
98. Sato, K., Yamazato, M.: Operator-self-decomposable distributions as limit distributions of processes of Ornstein-Uhlenbeck type. *Stoch. Process. Appl.* **17**(1), 73–100 (1984)
99. Salins, M., Spiliopoulos, K.: Metastability and exit problems for systems of stochastic reaction-diffusion equations. *Ann. Probab.* [arXiv:1903.06038](https://arxiv.org/abs/1903.06038), (2021)
100. Siegert, W.: *Local Lyapunov Exponents*. Springer Lecture Notes in Mathematics, vol. 1963 (2009)
101. Sokolov, I.M., Ebeling, W., Dybiec, B.: Harmonic oscillator under Lévy noise: unexpected properties in the phase space. *Phys. Rev. E* **83**, 041118-1–041118-7 (2011)
102. Trefethen, L.N., Embree, M.: *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press, Princeton (2005)
103. Trefethen, L.N., Trefethen, L.M.: How many shuffles to randomize a deck of cards? *Proc. R. Soc. Lond. A* **456**(8), 2561–2568 (2000)
104. Vernier, É.: Mixing times and cutoffs in open quadratic fermionic systems. *Sci. Post. Phys.* **9**(049), 1–30 (2020)
105. Viana, M., Oliveira, K.: *Foundations of Ergodic Theory*. Cambridge University Press, Cambridge (2016)
106. Villani, C.: *Optimal Transport. Old and New*. Springer, New York (2009)
107. Wahlén, E.: The matrix exponential. www.maths.lth.se/media/MATM14/2012vt12/exp.pdf. Accessed 25 Sept 2020
108. Wang, J.: On the exponential ergodicity of Lévy-driven Ornstein-Uhlenbeck processes. *J. Appl. Prob.* **49**, 990–1004 (2012)
109. Ycart, B.: Cutoff for samples of Markov chains. *ESAIM. Probab. Stat.* **3**, 89–106 (1999)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.