



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Numerics with coordinate transforms for efficient Brownian dynamics simulations

Citation for published version:

Phillips, D, Leimkuhler, B & Matthews, C 2024, 'Numerics with coordinate transforms for efficient Brownian dynamics simulations', *Molecular Physics*. <https://doi.org/10.1080/00268976.2024.2347546>

Digital Object Identifier (DOI):

[10.1080/00268976.2024.2347546](https://doi.org/10.1080/00268976.2024.2347546)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Molecular Physics

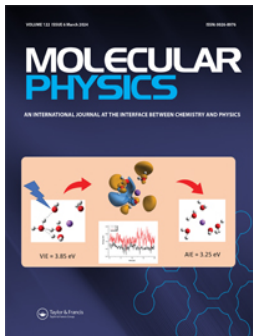
General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





Molecular Physics

An International Journal at the Interface Between Chemistry and Physics

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tmph20

Numerics with coordinate transforms for efficient Brownian dynamics simulations

Dominic Phillips, Benedict Leimkuhler & Charles Matthews

To cite this article: Dominic Phillips, Benedict Leimkuhler & Charles Matthews (02 May 2024): Numerics with coordinate transforms for efficient Brownian dynamics simulations, Molecular Physics, DOI: [10.1080/00268976.2024.2347546](https://doi.org/10.1080/00268976.2024.2347546)

To link to this article: <https://doi.org/10.1080/00268976.2024.2347546>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 02 May 2024.



Submit your article to this journal [↗](#)



Article views: 28



View related articles [↗](#)



View Crossmark data [↗](#)

Numerics with coordinate transforms for efficient Brownian dynamics simulations

Dominic Phillips^a, Benedict Leimkuhler^b and Charles Matthews^c

^aSchool of Informatics, University of Edinburgh, Edinburgh, UK; ^bSchool of Mathematics, University of Edinburgh, Edinburgh, UK; ^cAdobe Machine Learning, Adobe Systems, Edinburgh, UK

ABSTRACT

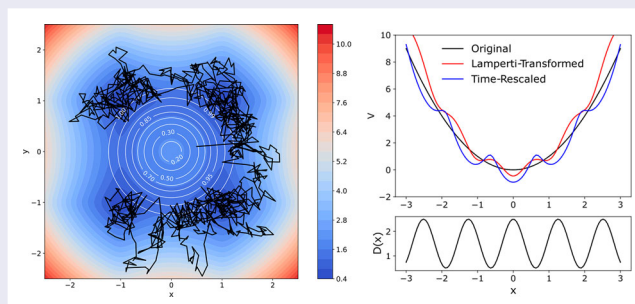
Many stochastic processes in the physical and biological sciences can be modelled as Brownian dynamics with multiplicative noise. However, numerical integrators for these processes can lose accuracy or even fail to converge when the diffusion term is configuration-dependent. One remedy is to construct a transform to a constant-diffusion process and sample the transformed process instead. In this work, we explain how coordinate-based and time-rescaling-based transforms can be used either individually or in combination to map a general class of variable-diffusion Brownian motion processes into constant-diffusion ones. The transforms are invertible, thus allowing recovery of the original dynamics. We motivate our methodology using examples in one dimension before then considering multivariate diffusion processes. We illustrate the benefits of the transforms through numerical simulations, demonstrating how the right combination of integrator and transform can improve computational efficiency and the order of convergence to the invariant distribution. Notably, the transforms that we derive are applicable to a class of multibody, anisotropic Stokes-Einstein diffusion that has applications in biophysical modelling.

ARTICLE HISTORY

Received 12 October 2023
Accepted 19 April 2024

KEYWORDS

Brownian dynamics;
Lamperti transform; time
rescaling; Stochastic
differential equations;
sampling



1. Introduction

Many problems in finance and the physical and biological sciences can be modelled as instances of Brownian dynamics. Examples include portfolio optimisation [1], options pricing [2], diffusion in biological membranes and nanocomposites [3, 4], cell migration [5], protein folding [6], neuronal dynamics [7], population genetics [8], MRI imaging [9], ecological modelling [10] and score-based diffusion for generative AI [11]. In these contexts, configuration-dependent diffusion is often critical to the modelling assumption but it can introduce problems for numerical modelling. It can make the problem stiffer by introducing unbounded noise or bounds on the state variables. Additionally, it can reduce the weak order of convergence of an integrator. This is a problem for simulation because sampling becomes more expensive.

It's also a problem for estimation, such as when fitting a Brownian dynamics 'grey-box' model, since high accuracy is required for the Extended Kalman Filter approximations to be meaningful [12].

One remedy for these problems is to design sophisticated, derivative-free numerical integrators that maintain high-accuracy convergence for certain classes of configuration-dependent diffusion. In recent years, many authors have contributed to a series of improvements and various integrators have been proposed [13–17]. However, a common drawback of these integrators is the requirement of multiple evaluations of the force and diffusion tensor per time step. This can be prohibitively expensive for multi-body simulations, where the evaluation of these terms is the computational bottleneck [18]. Furthermore, many of these integrators place restrictions

CONTACT Dominic Phillips  dominic.phillips@ed.ac.uk

on the functional form of the diffusion, often requiring commutative noise, which is not suitable for all applications.

An alternative approach, preferred whenever possible, is to transform the original process into a process with constant diffusion, thereby mitigating the sampling challenges introduced by multiplicative noise [19]. For certain classes of stochastic differential equations (SDEs), this is achieved through a Lamperti transform, a type of non-linear change of state variables [20, 21]. The resulting constant-diffusion process might exhibit enhanced numerical stability and can be sampled with computationally cheap, high weak-order integrators. Take for example the Black-Scholes model from financial mathematics, which describes geometric Brownian motion on the positive real axis. When simulated with sufficiently large step sizes, positivity can be violated which results in numerical instability. Here the Lamperti transform approach is especially valuable since it is possible to simultaneously construct a transform to unit diffusion whilst also removing the positivity constraint [22].

An alternative to a spatial coordinate transform is to apply a smooth, configuration-dependent time-rescaling [23, 24]. Recently, this has been explored as a method for adaptive step size control in Langevin dynamics sampling [25]. In this work, we take a different perspective and consider time-rescaling alongside the Lamperti transform as a strategy to remove multiplicative noise.

In this article, we derive conditions for applying the Lamperti and time-rescaling transforms, either separately or in combination, to achieve constant diffusion in multivariate Brownian dynamics with multiplicative noise. Through numerical experiments, we show how if the right choice of numerical integrator is used for the transformed process, then this leads to an efficient, second-order weak sampling method that involves just one force and one diffusion evaluation per time step. Furthermore, we show how the original autocorrelation function and evolving distribution can be accurately recovered by applying an inverse transform to the samples. We also draw connections with rare event sampling, illustrating by means of a one-dimensional example how coordinate transforms can be used in that context.

The article is structured as follows. Section 2 introduces Brownian dynamics and the Lamperti and time-rescaling transformations. Section 3 explores in detail how these transforms apply to one-dimensional Brownian dynamics, including rare event sampling. Section 4 extends the theory of transforms to multivariate Brownian dynamics. Numerical experiments in one dimension are presented in Section 5 and multivariate experiments are presented in Section 6. Conclusions are presented in Section 7.

2. Preliminaries

2.1. Brownian dynamics

Brownian dynamics is defined through an Itô stochastic differential equation (SDE), which in one dimension reads [26]

$$dx_t = -D(x_t) \frac{dV(x_t)}{dx} dt + kT \frac{dD(x_t)}{dx} dt + \sqrt{2kTD(x_t)} dW_t, \quad (1)$$

where $t \in \mathbb{R}_{>0}$ is time, $x_t \in \mathbb{R}$ is the state variable, W_t is a one-dimensional Wiener process, $V : \mathbb{R} \rightarrow \mathbb{R}$ is a potential energy function, $D : \mathbb{R} \rightarrow \mathbb{R}_{>0}$ is the diffusion coefficient, k is the Boltzmann constant and T is the temperature in degrees Kelvin. Note that the diffusion coefficient $D(x)$ is a function of x which means that we have configuration-dependent noise, also known as multiplicative noise.

In higher dimensions, (1) generalises to

$$d\mathbf{X}_t = -(\mathbf{D}(\mathbf{X}_t)\mathbf{D}(\mathbf{X}_t)^T) \nabla V(\mathbf{X}_t) dt + kT \operatorname{div}(\mathbf{D}\mathbf{D}^T)(\mathbf{X}_t) dt + \sqrt{2kT\mathbf{D}(\mathbf{X}_t)} d\mathbf{W}_t, \quad (2)$$

where $\mathbf{X}_t \in \mathbb{R}^n$ is the state variable, \mathbf{W}_t is an n -dimensional Wiener process, $V : \mathbb{R}^n \rightarrow \mathbb{R}$ is a potential function, and $\mathbf{D}\mathbf{D}^T : \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ is a configuration-dependent diffusion tensor that is everywhere positive definite. The matrix divergence in (2) is defined as the column vector resulting from applying the vector divergence to each matrix row. We identify $\mathbf{D}\mathbf{D}^T$ (not \mathbf{D}) as the diffusion tensor to avoid taking a matrix square root in the noise term.

We assume that V is confining in a way that ensures ergodicity of the dynamics. This is true for any V that grows sufficiently quickly as $|\mathbf{X}| \rightarrow \infty$, for details see Pavliotis (26) [26]. One consequence of ergodicity is that there exists a unique invariant distribution $\rho(\mathbf{X})$ – a probability distribution that does not change under the process dynamics. For Brownian dynamics, the invariant distribution is the canonical ensemble; $\rho(\mathbf{X}) \propto \exp(-V(\mathbf{X})/kT)$. Another consequence of ergodicity is that the long-time time average of any L1 integrable function f converges to its phase-space average as the simulation time goes to infinity, i.e.

$$\int_{\mathbb{R}^n} f(\mathbf{X}) \rho(\mathbf{X}) d\mathbf{X} = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(\mathbf{X}_t) dt. \quad (3)$$

In the remainder of this paper, we shall refer to (3) as ‘the ergodic theorem’.

2.2. The Lamperti transformation

Consider a time-homogeneous Itô SDE of the form

$$d\mathbf{X}_t = f(\mathbf{X}_t) dt + \sigma(\mathbf{X}_t)\mathbf{R} d\mathbf{W}_t, \quad (4)$$

where $\mathbf{X}_t \in \mathbb{R}^n$ is the state vector, $\mathbf{W}_t \in \mathbb{R}^m$ is an m -dimensional Wiener process, $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a drift function, $\sigma: \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^m$ is a configuration-dependent diffusion matrix and $\mathbf{R} \in \mathbb{R}^m \times \mathbb{R}^m$ is a constant matrix.

The *Lamperti transform* [20], $\zeta: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an invertible coordinate transformation that when applied to (4), results in a process $\mathbf{Y}_t := \zeta(\mathbf{X}_t)$ having unit diffusion. The required transform can be derived by applying the multivariate version of Itô's lemma and setting the coefficients of the noise terms to one. This gives a set of ODEs that ζ must satisfy. A consistent solution exists if: (i) the dimensions of the state variable and the noise are the same, (ii) \mathbf{R} is invertible, (iii) $\sigma(\mathbf{X}_t)$ has the diagonal form:

$$\sigma(\mathbf{X}_t) = \text{diag}(\sigma_1(X_{1,t}), \sigma_2(X_{2,t}), \dots, \sigma_n(X_{n,t})), \quad (5)$$

where $\sigma_i: \mathbb{R} \rightarrow \mathbb{R}_{>0}$ for all $i \in \{1, 2, \dots, n\}$. The solution is given by

$$\mathbf{Y}_t = \zeta(\mathbf{X}_t) = \mathbf{R}^{-1}\phi(\mathbf{X}_t), \quad (6)$$

where $\phi(\mathbf{X}_t) = [\phi_1(X_{1,t}), \phi_2(X_{2,t}), \dots, \phi_n(X_{n,t})]^T$ and $\phi_j: \mathbb{R} \rightarrow \mathbb{R}$ is the invertible function:

$$\phi_j(x) := \int_{x_{j,0}}^x \frac{1}{\sigma_j(z)} dz, \quad (7)$$

with $x_{j,0}$ being an arbitrary constant chosen from the state space of X_j . The governing equation of \mathbf{Y}_t then reads:

$$dY_{i,t} = \sum_{j=1}^n R_{ij}^{-1} \left(\frac{f_j(\phi^{-1}(\mathbf{R}\mathbf{Y}_t))}{\sigma_j(\phi_j^{-1}((\mathbf{R}\mathbf{Y}_t)_j))} - \frac{1}{2} \frac{\partial}{\partial x} \sigma_j(x) \Big|_{x=\phi_j^{-1}((\mathbf{R}\mathbf{Y}_t)_j)} \right) dt + dW_{i,t}.$$

The Lamperti transform can be used as a tool to find exact solutions for specific classes of SDEs [22] or to perform statistical inference for SDEs [27], but the extent to which the Lamperti transform is useful in practice is limited by the requirement in (5). We consider only time-homogeneous SDEs in this paper, although the Lamperti transform can also be extended to certain time-inhomogeneous problems [20].

2.3. The time-rescaling transform

An alternative approach for transforming an SDE to constant diffusion is the time-rescaling transformation (see,

for instance, [23] Chapter 8 and [24] Chapter 8). As before, consider an SDE of the form in Equation (4). We introduce a configuration-dependent time rescaling, denoted as $t \rightarrow \tau(t)$, with Jacobian $\frac{dt}{d\tau}(\mathbf{X}_t) = g(\mathbf{X}_t)$. The governing equation for the time-rescaled process becomes

$$d\mathbf{X}_\tau = f(\mathbf{X}_\tau)g(\mathbf{X}_\tau) d\tau + \sigma(\mathbf{X}_\tau)\mathbf{R}\sqrt{g(\mathbf{X}_\tau)} d\mathbf{W}_\tau, \quad (8)$$

where we have replaced dt with $\frac{dt}{d\tau}d\tau = g(\mathbf{X}_\tau)d\tau$ using a change of variables. The factor $\sqrt{g(\mathbf{X}_\tau)}$ in the noise arises from the scaling property of Brownian motion.

A transformation to unit diffusion is possible if and only if: (i) the dimensions of the state variable and the noise are the same, (ii) \mathbf{R} is invertible, (iii) the diffusion matrix $\sigma(\mathbf{X}_t)$ has the diagonal form:

$$\sigma(\mathbf{X}_t) = \text{diag}(D(\mathbf{X}_t), D(\mathbf{X}_t), \dots, D(\mathbf{X}_t)), \quad (9)$$

an isotropic matrix with arbitrary configuration dependence.

To remove the configuration dependence from the diffusion term, we choose $g(\mathbf{X}) = 1/D^2(\mathbf{X})$. Substituting this and the isotropic ansatz (9) into (8) simplifies the governing equations to

$$d\mathbf{X}_\tau = \frac{f(\mathbf{X}_\tau)}{D^2(\mathbf{X}_\tau)} d\tau + \mathbf{R} d\mathbf{W}_\tau.$$

We may then transform to unit diffusion through a linear transform

$$\mathbf{Y}_\tau = \mathbf{R}^{-1}\mathbf{X}_\tau.$$

Note that time-rescaling method can also be used to transform to an arbitrary isotropic diffusion $\tilde{D}(\mathbf{X})$ by making the choice $g(\mathbf{X}) = (\tilde{D}(\mathbf{X})/D(\mathbf{X}))^2$.

3. Transforms for one-dimensional Brownian dynamics

In this section, we consider the Lamperti and time-rescaling transforms applied to one-dimensional Brownian dynamics, comparing the two approaches. For detailed proofs of all results, see Appendix 4.

3.1. The Lamperti transform

In one dimension, the Lamperti transform emerges as an instance of a transformational symmetry inherent in Brownian dynamics. This symmetry states that, under an invertible coordinate transformation $x \rightarrow y(x)$, the one-dimensional Brownian dynamics (1) with potential $V(x)$ and diffusion function $D(x)$ is transformed into

another Brownian dynamics process with potential $\hat{V}(y)$ and diffusion function $\hat{D}(y)$ given by

$$\begin{aligned}\hat{V}(y) &= V(x(y)) + kT \ln \left| \frac{dy}{dx}(x(y)) \right|, \\ \hat{D}(y) &= D(x(y)) \left(\frac{dy}{dx}(x(y)) \right)^2,\end{aligned}\quad (10)$$

where $y \rightarrow x(y)$ is the inverse transformation (Theorem A.1).

Setting $\hat{D}(y) = 1$ and solving for $y(x)$ yields the one-dimensional Lamperti transform

$$y(x) = \int_{x_0}^x \left(\frac{1}{D(z)} \right)^{\frac{1}{2}} dz, \quad (11)$$

which is Equation (7) with $\sigma(z) = \sqrt{D(z)}$. From (11) we have $\frac{dy}{dx} = \left(\frac{1}{D(x)} \right)^{\frac{1}{2}}$. Substituting this result into (10), we arrive at the transformed, constant-diffusion dynamics:

$$dy_t = -\frac{d\hat{V}(y_t)}{dy} dt + \sqrt{2kT} dW_t, \quad (12)$$

with an effective potential given by

$$\hat{V}(y) = V(x(y)) - \frac{kT}{2} \ln D(x(y)).$$

Note that $\hat{V}(y)$ implicitly depends on x_0 in (11) through the inverse transform $x(y)$. Since x_0 changes the vertical offset of $y(x)$, it therefore changes the horizontal offset of $x(y)$. Changing x_0 thus corresponds to horizontally translating $\hat{V}(y)$, which shifts the mean position but otherwise has no physical consequence for the dynamics.

By writing down the ergodic theorem for the process (12) and transforming back to x -space, it can be shown that (Theorem A.2)

$$\int_{-\infty}^{\infty} f(x) \rho(x) dx = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(x(y_t)) dt, \quad (13)$$

so trajectories of the transformed process can be used directly to approximate ensemble averages with respect to $\rho(x)$, the invariant distribution of the original process. Furthermore, by choosing $f(x) = I(x \in [a, b])$ (the indicator function on the interval $[a, b]$) it can be shown that invariant distribution $\rho(x)$ of the original process and the invariant distribution $\hat{\rho}(y)$ of the Lamperti-transformed process are related by the equation $\rho(x) = \hat{\rho}(x(y)) \frac{dy}{dx}$ (Theorem A.3). Similarly, if we have a trajectory of discrete samples y_n with constant step size h , then choosing $f(x) = I(x \in [a, b])$ in (13) leads to a simple counting

formula to estimate finite-width integrals of the original invariant distribution:

$$\int_a^b \rho(x) dx \approx \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^N I(x(y_n) \in [a, b]). \quad (14)$$

This approximation becomes exact in the limit $h \rightarrow 0$.

3.2. The time-rescaling transform

Consider a configuration-dependent time rescaling $t \rightarrow \tau(t)$ with $\frac{dt}{d\tau}(x) = g(x)$. It can be shown that applying this to the original Brownian dynamics (1) results in another Brownian dynamics process but with a modified potential $\hat{V}(x)$ and diffusion coefficient $\hat{D}(x)$, given by (Theorem A.4):

$$\begin{aligned}\hat{V}(x) &= V(x) + kT \ln g(x), \\ \hat{D}(x) &= g(x)D(x).\end{aligned}\quad (15)$$

Setting $\hat{D}(x) = 1$ implies $g(x) = \frac{1}{D(x)}$. Substituting this result into (15), we arrive at the time-rescaled, constant-diffusion dynamics:

$$dx_\tau = -\frac{d\hat{V}(x_\tau)}{dx} d\tau + \sqrt{2kT} dW_\tau,$$

where

$$\hat{V}(x) = V(x) - kT \ln D(x)$$

is the effective potential. Notably, these dynamics differ from those obtained through the Lamperti transform.

By applying a time rescaling to the ergodic theorem of the original process x_t , it can be shown that (Theorem A.5):

$$\int_{-\infty}^{\infty} f(x) \rho(x) dx = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{\int_{\tau=0}^{T_{\text{sim}}} f(x_\tau) g(x_\tau) d\tau}{\int_{\tau=0}^{T_{\text{sim}}} g(x_\tau) d\tau}, \quad (16)$$

so trajectories of the transformed process can be used directly to approximate ensemble averages with respect to $\rho(x)$, the invariant distribution of the original process.

Discretising with a constant step size h in τ -time, and setting $f(x) = I(x \in [a, b])$, leads to a counting formula to estimate finite-width integrals of the original invariant distribution:

$$\int_a^b \rho(x) dx \approx \lim_{N \rightarrow \infty} \frac{\sum_{n=0}^N g(x_{\tau_n}) I(x_{\tau_n} \in [a, b])}{\sum_{n=0}^N g(x_{\tau_n})}. \quad (17)$$

This approximation becomes exact in the limit $h \rightarrow 0$.

Remark: The proof of (16) and (17) does not require Brownian dynamics, hence these results hold more generally for one-dimensional, time-homogeneous SDEs.

3.3. Comparing the two transform approaches

In one dimension, both the Lamperti and time-rescaling transforms are applicable for any $D(x) > 0$. However, whilst the time-rescaling transform can be computed exactly, the Lamperti transform often requires a numerical approximation due to the intractability of the integral (11). The two transforms also yield different effective potentials. The Lamperti transform tends to increase confinement of the potential in y -space wherever $D(x) > 1$ and decrease it wherever $0 < D(x) < 1$, whereas the time-rescaled effective potential is more confining where $\frac{dD}{dx} < 0$ and less confining where $\frac{dD}{dx} > 0$. Therefore, deciding which transform is more useful can depend on the functional form of $D(x)$. For rare-event sampling, one should choose whichever transform results in the least-confining effective potential as this improves numerical

stability for larger step sizes. Figure 1 compares the effective potentials resulting from the two transforms for various initial diffusion coefficients. In the next section, we provide a deeper insight into coordinate transforms for rare event sampling. We then derive multivariate versions of the Lamperti and time-rescaling transforms in Section 4.

3.4. Connection to variable-diffusion enhanced sampling

Recall that in this work, we generally consider that $D(x)$ is fixed by the application problem. Due to it being non-constant, it leads to less efficient numerical computations. Hence, the motivation to ‘transform it away’ into constant noise. Nonetheless, for problems where the primary

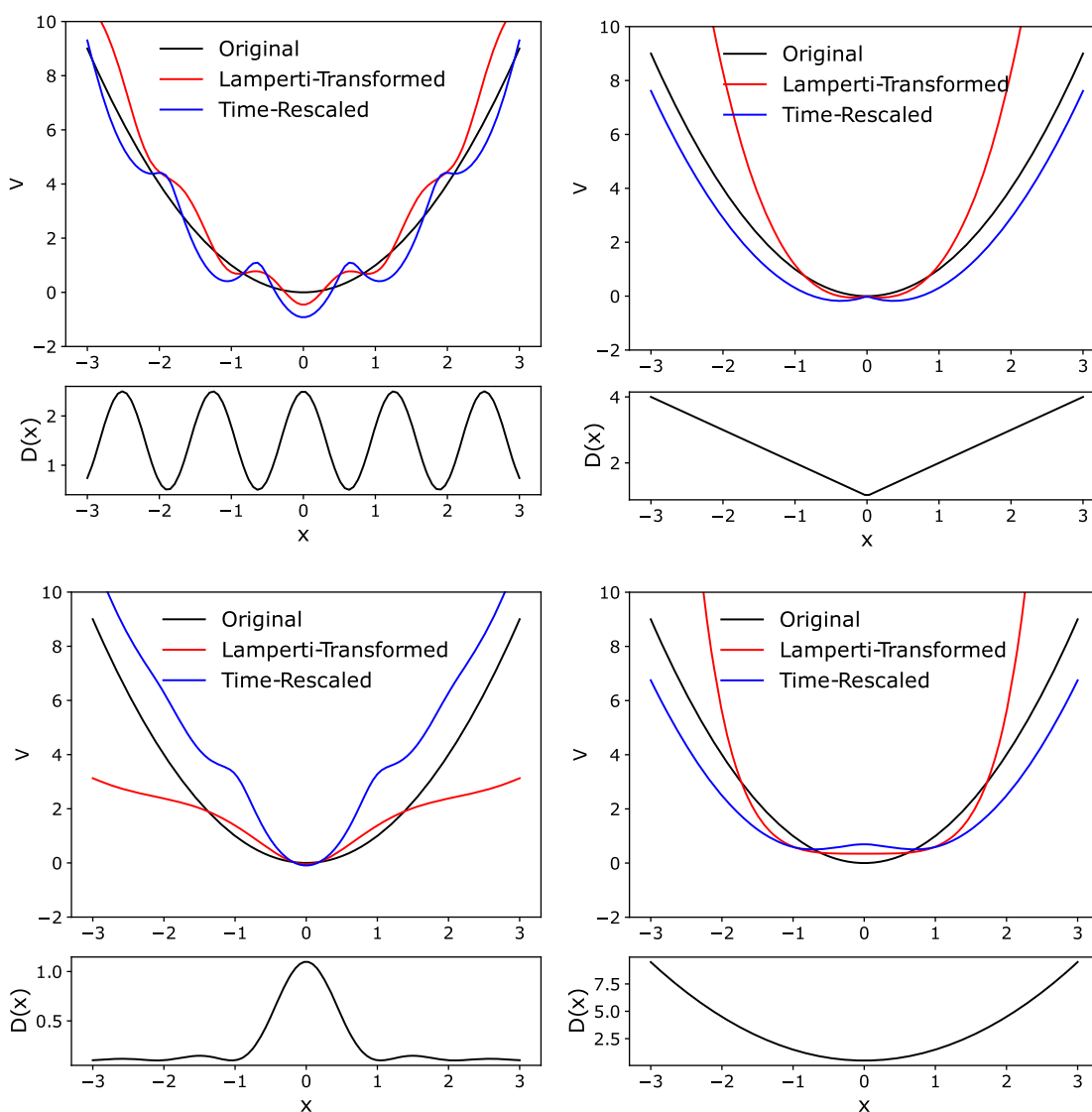


Figure 1. Comparison of the Lamperti and time-rescaling transforms when applied to the same quadratic potential $V(x) = x^2$ for a variety of different diffusion coefficients. The abscissa axis represents the original x coordinate for the time-rescaled potential and is the transformed $y(x)$ coordinate for the Lamperti-transformed potential.

goal is to explore a potential landscape, as is common in biomolecular modelling [28], it can be beneficial to take a different perspective and consider $D(x)$ as a free parameter that can be optimised. For this to be successful, the drawback of corrupting the dynamics must be offset by kinetic benefits, in the form of reduced metastability, gained by modifying the diffusion. A prototype of this idea dates back to the work of Roberts and Stramer (29), what the authors call *Langevin Tempered Diffusion* [29]. More recently, the optimisation problem for $D(x)$ was formulated precisely by Lelièvre et al. for dynamics without coordinate transforms [30]. Krivov and Karplus [31] explored how the Lamperti transform can be used to find natural reaction coordinates in small proteins [31]. The theoretical results in the previous section also offer new intuition for this class of sampling problem. We illustrate this through a numerical example below. For concreteness, we consider only the time-rescaling transform. Similar results could be obtained using the Lamperti transform.

Example 3.1 (Enhanced Sampling in a Double Well): Consider the double-well potential

$$V(x) = -\frac{1}{4}h^4x^2 + \frac{1}{2}c^2x^4 \quad (18)$$

with $h = c = 2$. This has a local maximum at $(x_{\max}, V_{\max}) = (0, 0)$, minima at $(x_{\pm}, V_{\min}) = (\pm 1, -2)$ and inflection points at $(\pm \frac{\sqrt{3}}{3}, -\frac{11}{9})$. Using this potential, we define a family of configuration-dependent diffusion functions:

$$D(x; \alpha) = \begin{cases} \exp\left(\alpha \cdot \frac{V(x)}{kT}\right) & \text{if } x_- < x < x_+ \\ \exp\left(\alpha \cdot \frac{V_{\min}}{kT}\right) & \text{otherwise,} \end{cases} \quad (19)$$

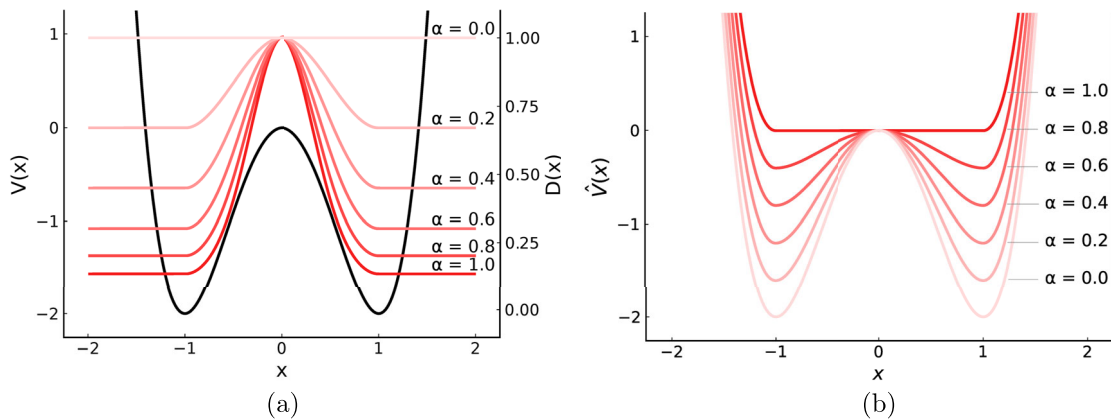


Figure 2. (a) Black: the double-well potential $V(x)$. Red: the configuration-dependent diffusion $D(x; \alpha)$ for various α with $kT = 1$. (b) Time-rescaled potentials after transforming to constant diffusion $D = 1$ with initial diffusion $D(x; \alpha)$. By design, the time-rescaled potentials are independent of kT . The transform has the effect of reducing the metastability of the well, removing it entirely when $\alpha = 1$.

where $0 \leq \alpha \leq 1$ controls the degree of configuration dependence, $\alpha = 0$ corresponds to constant noise. In Figure 2(a) we plot the potential and diffusion functions for $x \in [-2, 2]$ and $\alpha \in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ with $kT = 1$. Transforming (19) to constant diffusion using a time-rescaling leads to an effective potential

$$\hat{V}(x; \alpha) = \begin{cases} (1 - \alpha) V(x) & \text{if } x_- < x < x_+ \\ V(x) + \alpha |V_{\min}| & \text{otherwise.} \end{cases} \quad (20)$$

We plot this effective potential in Figure 2(b). As α increases, the rescaled potential flattens, gradually removing the metastability. This smooth removal of the metastability motivates the choice of diffusion considered in (19).

In Figure 3 we plot the α and kT dependence of the mean transition counts, N_T , between the wells for various numerical integrators (for details of the integrators, see Section 5.1). Unsurprisingly, simulations of the time-rescaled dynamics with $\alpha > 0$ result in enhanced sampling between the two wells, due to the lower energetic barrier. Furthermore, this enhanced-sampling effect is strongest at lower temperatures, since diffusion across the barrier is a kinetically-inhibited rare event.

For fixed step size, numerical integrator, and temperature, there exists an α value for which the sampling error of the transformed dynamics is minimum, bounded above by the sampling error with untransformed dynamics ($\alpha = 0$). In Figure 4 we plot this minimum sampling error as a function of kT with step size $h = 0.01$ for short simulations ($T_{\text{sim}}, \mathcal{T}_{\text{sim}} = 10$, Monte-Carlo error dominates the sampling error) and long simulations ($T_{\text{sim}}, \mathcal{T}_{\text{sim}} = 1000$, discretisation error dominates). For all integrators, sampling efficiency can be improved by considering an $\alpha > 0$ and applying a time rescaling to constant diffusion. Once more, we observe that the effect

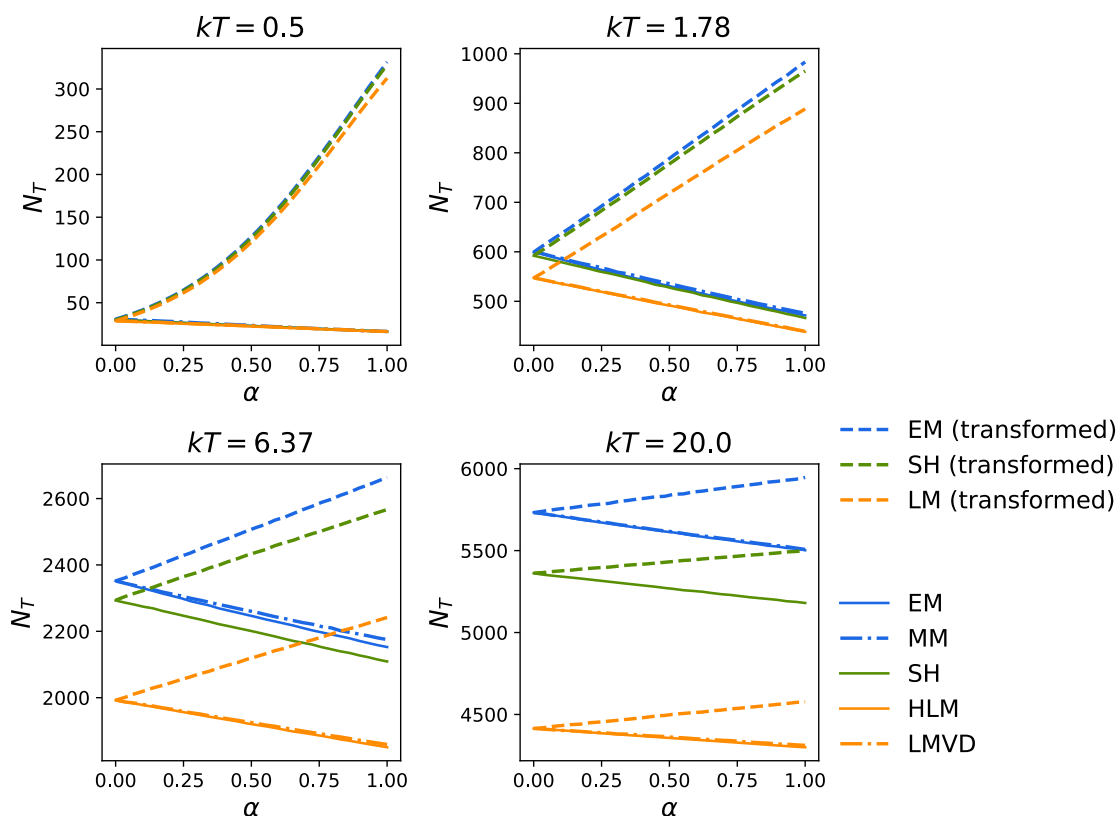


Figure 3. The mean transition counts (N_T) vs α for Brownian dynamics trajectories of length $T = 1000$ in a double-well potential with a constant step size $h = 0.01$. Each line represents an average over 5×10^3 independent repeats. Results are shown for four different values of kT , equally spaced in logarithmic scale. Solid lines represent integrators for the untransformed potential with diffusion function $D(x; \alpha)$, while dotted lines represent integrators for the time-transformed potential with $D(x) = 1$.

is most dramatic at low temperatures. See Appendix 1 for the full experimental methodology.

Brownian dynamics. Proofs of all results can be found in Appendix 4.

4. Transforms for multivariate Brownian dynamics

This section examines generalisations of the Lamperti and time-rescaling transforms to multivariate

4.1. The multivariate Lamperti transform

Consider multivariate Brownian dynamics with \mathbf{D} matrix

$$\mathbf{D}(\mathbf{X})_{ij} = D_i(X_i)R_{ij},$$

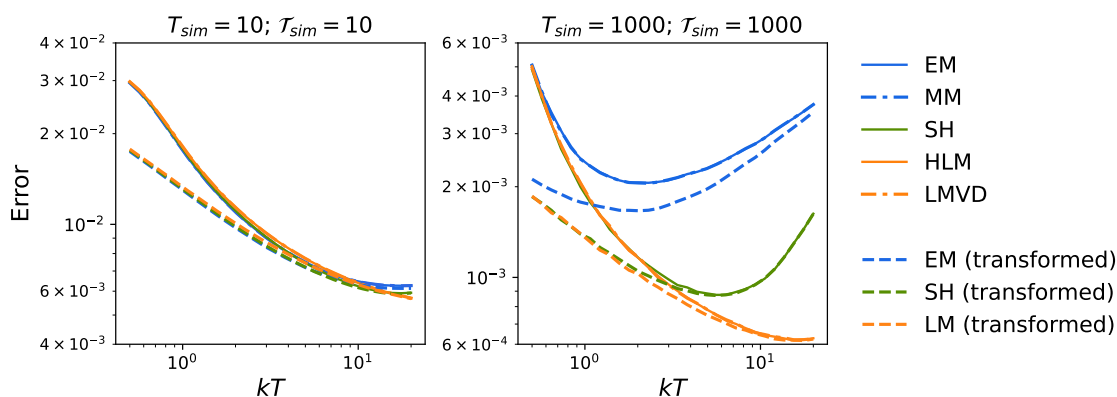


Figure 4. L_1 error in the invariant measure as a function of kT for the different integrators. Solid lines correspond to integrators of the original dynamics with $\alpha = 0$ (constant, unit diffusion). Dotted lines correspond to integrators of the transformed dynamics using the α value that gives the minimum L_1 error at that temperature. The left panel shows the error for trajectories to final time 10, where Monte Carlo error generally dominates. The right panel shows the error for trajectories to final time 1000, where discretisation error generally dominates.

where R_{ij} are the elements of an invertible, constant matrix. For this class of diffusion, although a Lamperti transform to unit diffusion can be constructed (Section 2.2), the transformed dynamics are only Brownian dynamics if \mathbf{R} is proportional to the identity. Specifically, when $\mathbf{D}(\mathbf{X})_{ij} = D_i(X_i)\delta_{ij}$, the Lamperti-transformed process is $Y_{i,t} = \sqrt{2kT} \int_{x_0}^{X_{i,t}} \frac{1}{D_i(x)} dx := \sqrt{2kT}\phi_i(X_{i,t})$, and obeys

$$dY_{i,t} = -\nabla_{Y_i} \hat{V}(\mathbf{Y}) dt + \sqrt{2kT} dW_i,$$

with an effective potential (Theorem A.6)

$$\hat{V}(\mathbf{Y}) = V(\phi^{-1}(\mathbf{Y})) - kT \sum_{k=1}^n \ln D_k(\phi_k^{-1}(Y_{k,t})). \quad (21)$$

In this case, the ergodic theorem generalises to (Theorem A.7)

$$\int_{\mathbb{R}^n} f(\mathbf{X})\rho(\mathbf{X}) d\mathbf{X} = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(\phi^{-1}(\mathbf{Y}_t)) dt, \quad (22)$$

where the map $\phi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}$ is constructed by individually applying ϕ_i^{-1} to each component of its argument, $1 \leq i \leq n$. We observe that there is an independent contribution to the effective potential for every diagonal component of \mathbf{D} .

4.2. The multivariate time-rescaling transform

Consider multivariate Brownian dynamics with \mathbf{D} matrix

$$\mathbf{D}(\mathbf{X}) = D(\mathbf{X})\mathbf{R},$$

where \mathbf{R} is invertible. For this class of \mathbf{D} matrix, a time-rescaling to Brownian dynamics unit diffusion can be constructed (Section 2.3). The time-rescaled process is given by $\mathbf{Y}_\tau = \mathbf{R}^{-1}\mathbf{X}_\tau$ where $\frac{dt}{d\tau} = g(\mathbf{X}) := 1/D^2(\mathbf{X})$ and it obeys

$$d\mathbf{Y}_\tau = -\nabla_{\mathbf{Y}} \hat{V}(\mathbf{Y}) dt + \sqrt{2kT} d\mathbf{W},$$

with an effective potential (Theorem A.8)

$$\hat{V}(\mathbf{Y}) = V(\mathbf{R}\mathbf{Y}) - 2kT \ln D(\mathbf{R}\mathbf{Y}). \quad (23)$$

The ergodic theorem generalises to (Theorem A.9)

$$\int_{\mathbb{R}^n} f(\mathbf{X})\rho(\mathbf{X}) d\mathbf{X} = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{\int_{\tau=0}^{T_{\text{sim}}} f(\mathbf{R}\mathbf{Y}_\tau)g(\mathbf{R}\mathbf{Y}_\tau)d\tau}{\int_{\tau=0}^{T_{\text{sim}}} g(\mathbf{R}\mathbf{Y}_\tau)d\tau}. \quad (24)$$

Remark: The proof of (24) does not require the assumption of Brownian dynamics and therefore it holds more generally for SDEs of the form considered in Section 2.3.

4.3. Combining multivariate transforms

The Lamperti and time-rescaling transforms can be combined. This allows transforming a wider class of diffusion processes to constant diffusion than is possible when applying either transform separately. Specifically, from the results of Sections 4.1 and 4.2, one expects that diffusion matrices of the form $\mathbf{D}(\mathbf{X}) = \mathbf{D}^{(1)}(\mathbf{X})\mathbf{R}\mathbf{D}^{(2)}(\mathbf{X})$, where

$$\mathbf{D}^{(1)}(\mathbf{X}) = \begin{bmatrix} D(\mathbf{X}) & & \\ & \ddots & \\ & & D(\mathbf{X}) \end{bmatrix},$$

$$\mathbf{D}^{(2)}(\mathbf{X}) = \begin{bmatrix} D_1(X_1) & & \\ & \ddots & \\ & & D_n(X_n) \end{bmatrix}, \quad (25)$$

can be transformed to constant diffusion. However, in this case, the transformed process has a non-conservative drift force unless \mathbf{R} is proportional to the identity (Theorem A.10). In particular, if $R_{ij} = \delta_{ij}$, then the process can be transformed to a constant-diffusion Brownian dynamics process \mathbf{Y}_τ through a time rescaling followed by a Lamperti transform, i.e.

$$\mathbf{X}_t \xrightarrow{\frac{dt}{d\tau} = g(\mathbf{X}) := D(\mathbf{X})^{-2}} \mathbf{X}_\tau \xrightarrow{\mathbf{Y}_\tau = \int_{i,t}^{\tau} D_i(x)^{-1} dx} \mathbf{Y}_\tau.$$

The effective potential of the transformed process is (Theorem A.11)

$$\hat{V}(\mathbf{Y}) = V(\mathbf{Y}) - 2kT \ln D(\phi^{-1}(\mathbf{Y})) - kT \sum_{i=1}^n \ln D_i(\phi_i^{-1}(\mathbf{Y})),$$

and the ergodic theorem is

$$\int_{\mathbb{R}^n} f(\mathbf{X})\rho(\mathbf{X}) d\mathbf{X} = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{\int_0^{T_{\text{sim}}} f(\phi^{-1}(\mathbf{Y}_\tau))g(\phi^{-1}(\mathbf{Y}_\tau))d\tau}{\int_0^{T_{\text{sim}}} g(\phi^{-1}(\mathbf{Y}_\tau))d\tau}.$$

5. Numerical experiments in one dimension

In this section, we simulate Brownian dynamics trajectories for a particular one-dimensional potential with variable noise. Specifically, we consider the potential $V(x) = \frac{x^2}{2} + \sin(1 + 3x)$ with diffusion profile $D(x) = 1 + |x|$ and $kT = 1$. Applying the results of Section 3, the

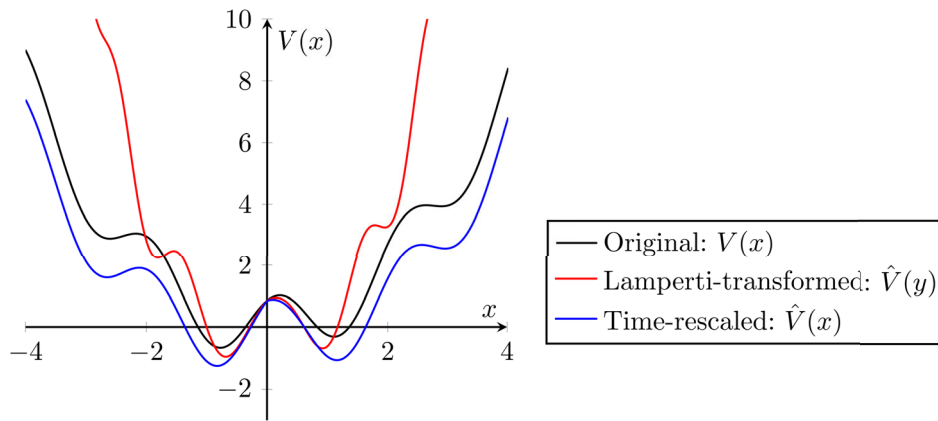


Figure 5. Comparison of effective potentials with original diffusion $D(x) = 1 + |x|$, $V(x) = \frac{x^2}{2} + \sin(1 + 3x)$ and $kT = 1$. The original potential is in black, the Lamperti-transformed potential is in red and the time-rescaled potential is in blue. Metastability in the potential makes this a challenging sampling problem. We observe that the Lamperti transform stiffens the potential, while the time-rescaling softens it.

Lamperti-transformed effective potential is

$$\hat{V}(y) = V\left(\frac{y}{4}(|y| + 4)\right) - \frac{kT}{2} \ln \left| 1 + |y| + \frac{y^2}{4} \right|$$

and the time-rescaled effective potential is

$$\hat{V}(x) = V(x) - kT \ln(1 + |x|).$$

In Figure 5, we sketch $V(x)$, $\hat{V}(y)$ and $\hat{V}(x)$ for $kT = 1$.

For this setup, we consider various numerical integrators (introduced below) both with and without the application of transforms. We compare the weak convergence to the invariant distribution, the sampling efficiency, and the effect of transforms on estimates of the autocorrelation function and the evolving distribution. All experiments are run on a Thinkpad P17 with a 12-core, 2.60GHz Intel i7-10750H CPU, using code implemented in Julia 1.8.5.¹

5.1. Numerical integrators

We consider the following numerical integrators: Euler-Maruyama (EM), Milstein Method (MM), Leimkuhler-Matthews (LM), Hummer-Leimkuhler-Matthews (HLM), Stochastic Heun (SH), and Limit Method with Variable Diffusion (LMVD). The defining equations of these integrators in the context of one-dimensional Brownian dynamics can be found in Appendix 2. The integrators can be summarised as follows:

The Euler-Maruyama (EM) integrator extends the Euler method to SDEs. It has a strong convergence order of 1/2 and a weak convergence order of 1 [32]. The Milstein Method (MM) modifies EM by incorporating a second-order correction derived from a stochastic Taylor series expansion. It is strong order 1 and weak order

1 and reduces to EM for constant diffusion [33]. The Leimkuhler-Matthews (LM) integrator is derived from the high-friction limit of the BAOAB-splitting method in the constant diffusion regime [34]. It has weak convergence order 2 for constant diffusion but is invalid (does not converge) for multiplicative noise. The Hummer-Leimkuhler-Matthews (HLM) integrator is an extension of LM that ensures that the expectation of position is exact in the case of locally linear diffusion, and is conjectured to improve convergence in the variable diffusion regime.² It reduces to LM for constant diffusion. The Stochastic Heun (SH) integrator is a two-stage Runge-Kutta method. It has weak convergence order of 2 for constant diffusion, otherwise order 1 for variable diffusion [35]. The accuracy gains of SH come at the cost of higher computational requirements, as it involves two force evaluations, two diffusion coefficient evaluations, and two diffusion gradient evaluations per iteration. The Limit Method with Variable Diffusion (LMVD) is a novel integrator that we introduce for the first time in this work. It has a weak convergence order of 2 for both constant and variable diffusion. It stems from the high-friction limit of the BAOAB-splitting method in the variable diffusion regime. Unlike SH, it requires one force evaluation per iteration, however, it requires two ODE solves per timestep, making it comparatively expensive compared to the related LM method. The derivation of LMVD can be found in Appendix 3. Note that LMVD reduces to LM for constant diffusion regime.

5.2. Error in infinite time

We compare weak convergence to the invariant distribution $\rho(x) \propto \exp(-V(x)/kT)$ with varying step size h , using trajectories generated by the different integrators

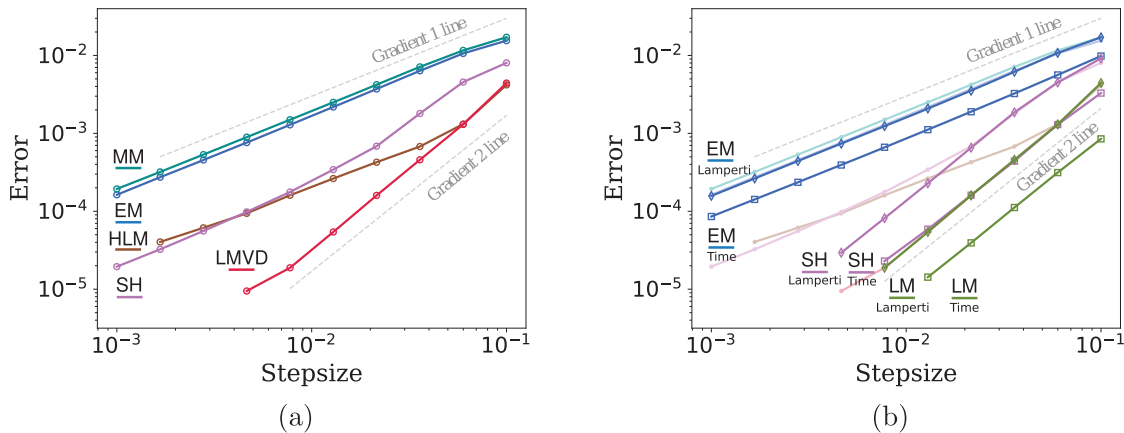


Figure 6. Rates of convergence to the invariant measure. The simulation time was fixed at $T_{\text{sim}}, \mathcal{T}_{\text{sim}} = 7.5 \times 10^7$ and 12 independent runs were averaged to further reduce sampling errors. (a) The untransformed methods. (b) When applying a transform to constant diffusion, either a Lamperti transform or a time rescaling. The untransformed methods are shown in faint in panel (b) to facilitate comparison.

both with and without transforms to constant diffusion. For untransformed dynamics, we compare EM, MM, HLM, SH, and LMVD. For the Lamperti-transformed dynamics and the time-rescaled dynamics, we compare the EM, LM, and SH integrators. We omit MM since it reduces to EM for constant diffusion, while LMVD and HLM both reduce to LM for constant diffusion. For each method, we run trajectories of length $T_{\text{sim}}, \mathcal{T}_{\text{sim}} = 7.5 \times 10^7$, and 12 independent runs are averaged to reduce sampling errors.

To assess the convergence of the invariant distribution, we divide a subset of the x domain into M equal-length intervals and compute the mean error between the empirical probabilities and the exact probabilities given by the invariant distribution. For Lamperti-transformed experiments, we derive empirical probabilities using Equation (14), for time-rescaled experiments we use Equation (17). We compute the L1 error:

$$\text{Error} := \frac{1}{M} \sum_{i=1}^M |\omega_i - \hat{\omega}_i|, \quad (26)$$

where ω_i is the exact occupancy probability of the i th interval and $\hat{\omega}_i$ is the empirical estimate. We use 30 equal-width intervals in the range -5 to 5 and run each integrator using 10 different step sizes, equally spaced in log space between 10^{-3} and 10^{-1} . Steps are in τ -time for time-rescaled methods, but t -time for all other methods. The error is plotted against the step size on a log-log scale, so first-order weak methods have a gradient of one, and second-order weak methods have a gradient of two. The results are shown in Figure 6.

Figure 6(a) confirms the expected orders of weak convergence for the untransformed methods. Notably,

MM has a larger error constant than EM, illustrating that improved strong convergence doesn't guarantee improved weak convergence. Examining Figure 6(b), we see that the effect of applying a transform is also integrator-dependent. Applying a transform to EM hardly changes the convergence properties, while applying a transform to constant diffusion for SH or LM restores their second-order convergence behaviour, with the transformed LM method closely following the convergence properties of LMVD. Note that the time-transformed methods are shifted relative to the Lamperti-transformed methods. This is expected because of the discrepancy between a constant step size in τ -time and t -time. It does not automatically imply that time-rescaled integrators result in more efficient samplers than Lamperti-transformed integrators. We investigate the question of efficiency in Section 5.3 below.

5.3. Computational efficiency and numerical stability

We assess the computational efficiency of each method by comparing the wall-clock time required to achieve a fixed L1 error in the invariant measure, as defined by Equation (26). We estimate the wall-clock cost per iteration of each method by timing 10^8 iterations with a fixed step size of $h = 0.01$, averaging over 12 runs. For transformed methods, any additional cost of applying the counting formulas (14) or (17) is included in these timings. We then fix a target error and run trajectories with various step sizes for each method, stopping when first reaching the target error. For each step size, we average 6000 repeats and find the minimum number of iterations to reach the specified error. The total wall-clock time is then estimated as the minimum number of iterations over

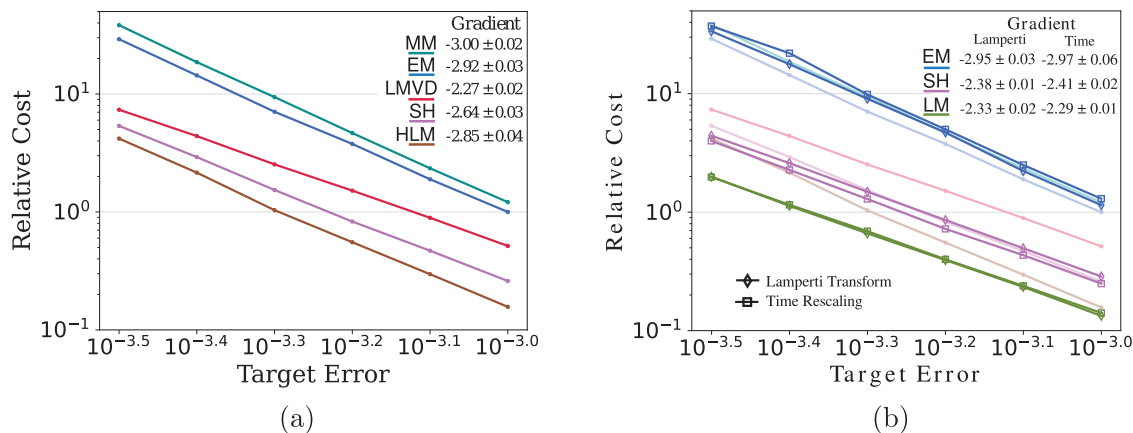


Figure 7. Cost-error diagram to compare numerical efficiency. Error is defined as per Equation (26). The cost is the wall-clock time to reach the target error (number of iterations times cost per iteration), relative to the wall-clock time for untransformed EM to reach a target error of 10^{-3} . (a) Untransformed methods. (b) When applying constant-diffusion transform: Lamperti or time rescaling. The untransformed methods are shown in faint for comparison.

Table 1. Stability thresholds (h^*) and the compute time required for 10^8 iterations (t) compared for untransformed and transformed methods.

Integrator	Untransformed		Lamperti		Time-rescaling	
	$t(s)$	h^*	$t(s)$	h^*	$t(s)$	h^*
EM	10.83(8)	0.20	12.77(8)	0.25	13.18(9)	0.25
SH	14.47(13)	0.25	17.22(8)	0.32	15.56(14)	2.5
LM	–	–	12.54(8)	0.25	13.09(4)	2.5
MM	11.97(18)	0.16	–	–	–	–
HLM	10.72(10)	0.20	–	–	–	–
LMVD	48.20(12)	0.25	–	–	–	–

Note: Standard errors in t were computed by averaging 12 runs with a constant step size of 0.01. Stability thresholds were determined as the first step size (in geometric increments of $10^{0.1}$) that resulted in numerical blow-up. Errors are in bracket notation, e.g. $10.83(8) = 10.83 \pm 0.08$.

the various step sizes times the cost per iteration. This calculation is repeated for 5 target errors logarithmically spaced between $10^{-3.5}$ and 10^{-3} . The resulting cost-error diagram is illustrated in Figure 7.

Numerical stability is estimated by determining the smallest step size, in logarithmic increments of $10^{0.1}$, where numerical blow-up occurs before $T_{\text{sim}}, T_{\text{sim}} = 10^6$. These stability threshold estimates as well as timing results for 10^8 iterations are shown in Table 1. We see that time-rescaling significantly improves the stability threshold of SH and LM in this case whereas the Lamperti transform does not. This is explained by the fact that, for this diffusion coefficient, the time-rescaled potential is the softer of the two transformed potentials (Figure 5). Thus, for this problem, the time-rescaling approach might be preferable if simulations with large step sizes are required.

As for the efficiency of the untransformed methods (Figure 7(a)), we note that the method with the best weak convergence (LMVD) isn't necessarily the most computationally efficient for a given range of target errors (HLM). Furthermore, Figure 7(b) shows how applying coordinate transforms can significantly

improve the computational efficiency of certain integrators but have a more modest impact on others. For example, the Lamperti-transformed/time-rescaled LM is the most computationally efficient method overall – approximately 5 times more efficient than LMVD. However, applying transforms only slightly improves the efficiency of SH, and even slightly reduces the efficiency of EM. In general, both types of transform have very similar effects on efficiency, and any differences can be attributed to small differences in the iteration cost (Table 1). Overall, we see that transformations only improve the efficiency of numerical integrators that have better convergence properties for constant noise.

5.4. Error in finite time

In the continuous limit, the Lamperti and time-rescaling transforms are invertible, allowing recovery of the original dynamics. However, the transforms of discredited dynamics could introduce additional bias for numerics. This section explores the effect of transforms on estimates of dynamic quantities, specifically the autocorrelation function and the evolving distribution.

5.4.1. Autocorrelation function

To obtain a reference estimate of the autocorrelation function, we run 200 randomised trajectories of length $T = 5000$ using the Stochastic Heun integrator with step size of $h = 0.01$, and with initial positions drawn from a standard normal distribution. For each trajectory, we use the Fast Fourier Transform (FFT) algorithm to estimate the normalised autocorrelation function and compute the mean and standard deviation of the best estimate. Additionally, we run trajectories under the same parameters but separately apply a Lamperti transform and time-rescaling transform. We then transform these trajectories back to x -space and t -time respectively and compute the normalised autocorrelation function. The three autocorrelation functions so obtained (reference, Lamperti and time-rescaling) are shown in Figure 8(a). We also calculate the mean and standard deviation in the mean of the difference between the Lamperti/time-rescaling autocorrelation functions and the reference estimate. These results are displayed in Figure 8(b).

Overall, at short times ($t \lesssim 8$) the differences in autocorrelation estimates are minimal and often not statistically significant. As time increases (between 10 and 20), the fractional error in the mean becomes more significant, but still much smaller than the standard deviation across runs (the standard deviation is $\sqrt{200} \approx 14$ times larger than the standard error in this case).

Note that while inverting the Lamperti transform is straightforward (apply the inverse coordinate transformation), undoing the time-rescaling process can be more challenging. This is because direct conversion from τ -time to t -time results in an irregular time series,

making it unsuitable for direct application of FFT-based methods. To address this, we perform linear interpolation on a t grid with the same regular spacing of $h = 0.01$ before applying the FFT. This interpolation step introduces bias. However, as we have seen in Figure 8, the overall bias remains small and unlikely to be practically significant. Alternatively, methods designed for unevenly spaced time series, such as least-squares spectral analysis, could be used but these add significant computational cost, negating any sampling efficiency benefits of the time transform.

5.4.2. Evolving distribution

With initial positions drawn from a standard normal distribution, we estimate the evolving distribution using 2.5×10^7 independent trajectories made with the SH integrator, step size $h = 10^{-4}$ and compare this to the evolving distribution estimates computed with step size $h = 0.02$ for each method. We compute the L1 errors with respect to the reference distribution at time snapshot intervals of $\delta t = 0.04$, using the same histogram binning as introduced in Section 5.2.

For methods involving the Lamperti transform, the initial condition is transformed to y -space, trajectories are evaluated and then transformed back to x -space. For methods involving time rescaling, trajectories are evaluated in τ -time and then transformed back to t -time. However, the conversion from τ -time to t -time is problematic since this transform depends on the unique history of each trajectory. To overcome this, we choose to simulate each trajectory in τ -time (in steps of $h = 0.02$) until slightly overshooting the $\delta t = 0.04$ interval. The

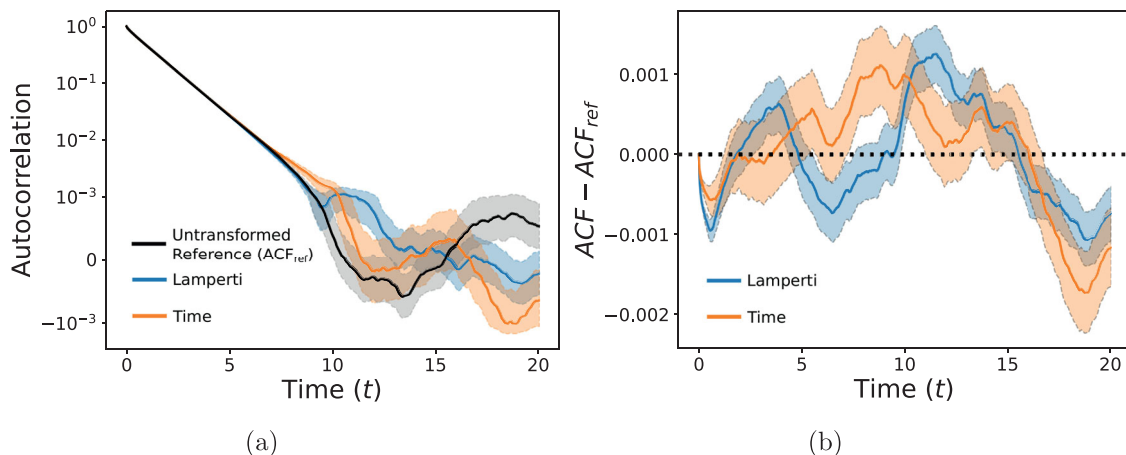


Figure 8. Comparing normalised autocorrelation function estimates with and without transforms. Panel (a) shows the mean and standard deviation in the mean of autocorrelation function estimates obtained using the Stochastic Heun integrator with 200 trajectories of length T_{sim} , $T_{sim} = 5000$ and step size $h = 0.01$. In black is the reference (no transform applied), in blue is when using a Lamperti transform and in orange is when using time rescaling. Panel (b) displays mean and standard error of the differences in autocorrelation function estimates using the Lamperti transform (blue) and time-rescaling transform (orange) compared to the reference (ACF_{ref}). Overall, biases introduced by the transforms are small and only practically significant at large times.

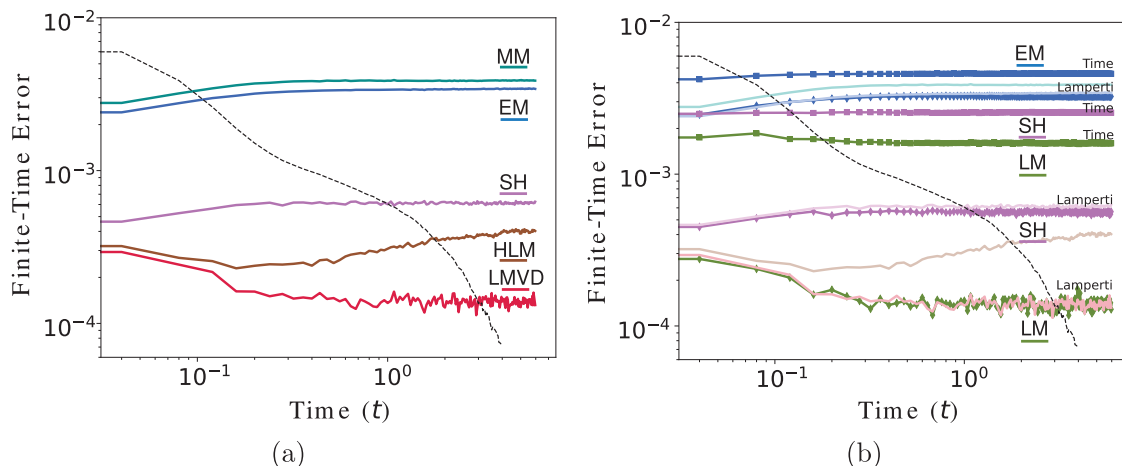


Figure 9. Finite-time errors of the evolving distribution in the interval $t \in [0, 6]$ for fixed step size $h = 0.02$. The reference distribution at time t is computed by averaging over 2.5×10^7 independent trajectories using the SH method with small step size $h = 10^{-4}$. In each plot, the dotted black line represents the L1 difference between the reference evolving distribution and the invariant distribution. (a) Errors of the untransformed methods. (b) Errors when applying a transform to constant diffusion, either a Lamperti transform or a time rescaling. The untransformed methods are shown in faint to facilitate comparison. Integrators are Leimkuhler-Matthews (LM), Milstein Method (MM), Euler-Maruyama (EM), Hummer-Leimkuhler-Matthews (HLM), Stochastic Heun (SH), Limit Method with Variable Diffusion (LMVD).

position at the required t -time is then estimated by linear interpolation with the previous sample. This approach is inexpensive but can introduce bias.

The resulting errors for the untransformed and transformed methods are shown in Figure 9(a,b), respectively. Additionally, the figures include a dotted black line representing the L1 difference between the reference evolving distribution and the invariant distribution. For points below this line, the evolving distribution is distinguishable from the invariant distribution.

Examining the untransformed methods in Figure 9(a), we observe that by $t = 6$, the errors for each method have already converged to their corresponding infinite-time errors depicted in Figure 6(a), which is consistent with the correlation timescale implied in Figure 8. However, we see that MM, EM and LM are unsuitable at this step size if high accuracy is required, as their errors soon exceed the difference between the evolving and invariant distributions.

Examining the transformed methods, we see that the Lamperti transform has no detrimental impact on finite-time errors. In particular, the Lamperti-transformed LM method has the same finite-time error as the more expensive LMVD method. By contrast, the time-rescaled methods show a noticeable bias, likely originating from the need for linear interpolation when computing the evolving distribution at fixed t . This bias can make these methods ill-suited for high-accuracy simulations of the evolving distribution in practice.

6. Multivariate numerical experiments

As an example of multivariate Brownian dynamics, we consider Stokes-Einstein diffusion, which models the diffusion of a low concentration of non-interacting, spherical particles suspended in a fluid. It has widespread applications, particularly in materials science [36, 37], and is also used for modelling water diffusion in MRI imaging applications [9]. In n dimensions, the Stokes-Einstein diffusion tensor is given by

$$\mathbf{D}_{SE} = \frac{k_B T}{6\pi\eta r} \mathbf{1}_n,$$

where T is Kelvin temperature, η is viscosity and $\mathbf{1}_n$ is the n -dimensional identity matrix.

If the temperature field or the fluid's material properties are non-homogeneous, then \mathbf{D}_{SE} is an isotropic, position-dependent matrix, and can be transformed to constant diffusion by time rescaling. Furthermore, if the medium is anisotropic, and certain diffusion directions are preferred over others, then the diffusion model can be generalised to $\mathbf{D}(\mathbf{X}) = \mathbf{D}_{SE}(\mathbf{X})\mathbf{D}^{(2)}(\mathbf{X})$, where $\mathbf{D}^{(2)}(\mathbf{X})$ is the diagonal matrix given by Equation (25). This kind of diffusion is common in biological tissues, for example in the brain, where positional alignment of white matter tracts results in preferred diffusion directions for water molecules [38]. Importantly, since this diffusion is of the form $\mathbf{D}(\mathbf{X}) = \mathbf{D}^{(1)}(\mathbf{X})\mathbf{D}^{(2)}(\mathbf{X})$ as per Equation (25), the process can be transformed to constant-diffusion through the combination of a time-rescaling and a Lamperti transform.

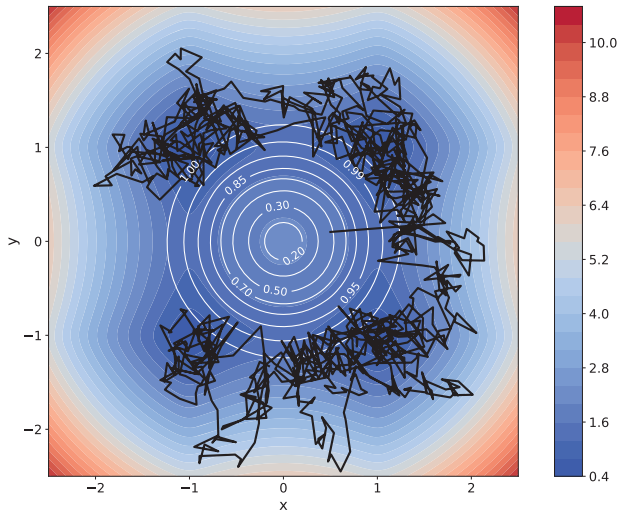
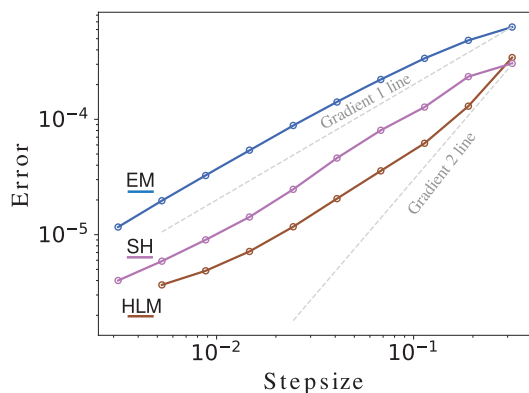


Figure 10. Heatmap of the quadruple-well potential function (27). White circles depict contours of the Frobenius norm of the Moro-Cadin diffusion tensor (28). The black path shows a Euler-Maruyama trajectory of Brownian dynamics of 1000 steps with time step of $h = 0.01$ and $kT = 1$. Note the small norm of the diffusion tensor in the vicinity of the origin. This inhibits hopping between the wells, making this a more challenging sampling problem.

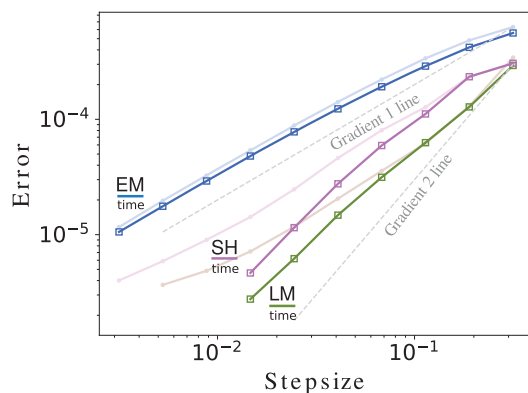
The example of Stokes-Einstein diffusion we consider is multivariate Brownian dynamics in a 2D quadruple-well potential given by

$$V(x, y) = \sqrt{\frac{17}{16} - 2x^2 + x^4} + \sqrt{\frac{17}{16} - 2y^2 + y^4}, \quad (27)$$

with a diffusion tensor given by the Moro-Cardin tensor [39]



(a)



(b)

Figure 11. Rates of convergence to the invariant measure for Brownian dynamics in a 2D quadruple well potential with Moro-Cadin diffusion tensor. The simulation time was fixed at $T_{\text{sim}}, \mathcal{T}_{\text{sim}} = 5 \times 10^6$ and 12 independent runs were averaged to further reduce sampling errors. (a) The untransformed methods. (b) When applying a transform to constant diffusion through a time rescaling. The untransformed methods are shown in faint to facilitate comparison. Integrators are Leimkuhler-Matthews (LM), Euler-Maruyama (EM), Hummer-Leimkuhler-Matthews (HLM) and Stochastic Heun (SH).

$$D(x, y) = \left(1 + A \exp \left(-\frac{x^2 + y^2}{2\sigma^2} \right) \right)^{-1} \mathbf{1}, \quad (28)$$

where $A = 5$ and $\sigma = 0.3$, see Figure 10. Since this diffusion tensor is isotropic, it can be mapped to constant diffusion through time rescaling.

Figure 11 illustrates the comparison of weak convergence to the invariant measure for the LM, EM, SH, and HLM integrators. Figure 11(a) is without any transforms, while Figure 11(b) is after a time-rescaling transform to constant diffusion has been applied. We follow the same general approach as first outlined in Section 5.2. We run trajectories of length $T_{\text{sim}}, \mathcal{T}_{\text{sim}} = 5 \times 10^6$ and average over 12 independent runs and we run each integrator using 10 different step sizes, equally spaced in log-space between $10^{-2.5}$ and $10^{-0.5}$. For histogram computation, we use a 30×30 grid of equal-width square bins covering the domain $[-3, 3] \times [-3, 3]$ in the x - y plane.

We observe similar behaviour to the one-dimensional numerical experiments discussed in Section 5.2. It is noteworthy that applying a time-rescaling transform enhances the convergence rate for both the SH and LM integrators. Similar to the one-dimensional case, the transformed LM integrator exhibits a lower error constant compared to SH, indicating its superior efficiency for this particular problem.

7. Conclusions

We examined two types of transform to constant diffusion for Brownian dynamics with multiplicative noise: the Lamperti transform and the time-rescaling transform. We derived conditions on the noise term for these

transforms to be applied and combined. Furthermore, through numerical experiments in one and two dimensions, we have shown how using these transforms, combined with an appropriate SDE integrator, can lead to a highly efficient sampling method for certain classes of multivariate noise.

For one-dimensional Brownian dynamics, we showed that both transforms are always applicable, regardless of the form of the diffusion coefficient. However, the two transforms affect the dynamics differently, so the choice of transform may depend on the specific problem at hand. We showed numerically that applying either transform with the Leimkuhler-Matthews (LM) integrator significantly improves the convergence to the invariant measure, resulting in a method that has approximately five times higher sampling efficiency than the Limit Method with Variable Diffusion (LMVD) – a highly-performant integrator for multiplicative noise that does not utilise transformations. This transformed method also significantly outperformed the popular Euler-Maruyama integrator, with a 10 to 25 times higher computational efficiency for the problem investigated. Crucially, this method only requires one force and one diffusion tensor evaluation per iteration, thus scaling better to high-dimensional problems than competing methods that require multiple force and/or diffusion evaluations per step.

In addition to investigating convergence to the invariant measure, we also verified whether dynamics information, in the form of the autocorrelation function and the evolving distribution, can be recovered after simulating a transformed process and then applying the inverse transform. We found that the Lamperti transform introduced no appreciable bias for estimates of either quantity, but that the time-rescaling transform is less suitable for recovering finite-time distributions.

For multivariate Brownian dynamics, the Lamperti and time-rescaling transforms have somewhat limited application. However, the two transformations can be combined to transform non-homogeneous, anisotropic Stokes-Einstein diffusion into a constant diffusion process. This is a broad class of diffusion tensors with applications in biological diffusion processes. Furthermore, in the context of rare event sampling, the flexibility of the coordinate transforms offers an additional advantage. In Section 3.4 we showed that they can be used to design diffusion profiles that, when transformed back to constant diffusion, lead to enhanced sampling. In ongoing research, we are actively exploring these techniques in diverse applications and are investigating further refinements to optimise their effectiveness for enhanced sampling. We anticipate that coordinate transforms will

improve the efficiency of Brownian dynamics simulations in various contexts.

Notes

1. GitHub repository: <https://github.com/dominicp6/Transforms-For-Brownian-Dynamics>
2. We would like to thank Gerhard Hummer for suggesting this method in personal correspondence.

Acknowledgments

We extend our gratitude to the reviewers for their insightful comments and constructive suggestions, which greatly improved the quality and clarity of this manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the United Kingdom Research and Innovation (grant EP/S02431X/1), UKRI Centre for Doctoral Training in Biomedical AI at the University of Edinburgh, School of Informatics.

Data availability statement

The software and data that support the findings of this study are openly available in Zenodo at <https://zenodo.org/badge/latest/doi/616585156>.

ORCID

Dominic Phillips  <http://orcid.org/0000-0002-5707-4778>

References

- [1] R.C. Merton, *Rev. Econ. Stat.* **51** (3), 247–257 (1969).
- [2] F. Black and M. Scholes, *J. Polit. Econ.* **81** (3), 637–654 (1973).
- [3] P.G. Saffman and M. Delbrück, *Proc. Natl. Acad. Sci. U.S.A.* **72** (8), 3111–3113 (1975).
- [4] N. Fakhri, F.C. MacKintosh, B. Lounis, L. Cagnet and M. Pasquali, *Science (New York, N.Y.)* **330** (6012), 1804–1807 (2010).
- [5] D. Selmeczi, L. Li, L.I. Pedersen, S.F. Nrrlykke, P.H. Hagedorn, S. Mosler, N.B. Larsen, E.C. Cox and H. Flyvbjerg, *Eur. Phys. J. Spec. Top.* **157** (1), 1–15 (2008).
- [6] R.B. Best and G. Hummer, *Proc. Natl. Acad. Sci.* **107** (3), 1088–1093 (2010).
- [7] D. Johnston and S.M.s. Wu, *Foundations of Cellular Neurophysiology* (MIT Press, Cambridge, Mass, 1995).
- [8] T. Maruyama, *Bull. Math. Biol.* **45** (4), 521–554 (1983).
- [9] A.L. Alexander, J.E. Lee, M. Lazar and A.S. Field, *Neurotherapeutics* **4** (3), 316–329 (2007).
- [10] M.M. Varughese and E.A.D. Pienaar, *Ecosphere* **4** (8), 1–14 (2013).

- [11] Y. Song, J. Sohl-Dickstein, D.P. Kingma, A. Kumar, S. Ermon and B. Poole, Internation Conference on Learning Representations, 2020.
- [12] N.R. Kristensen, H. Madsen and S.B. Jørgensen, *Automatica* **40** (2), 225–237 (2004).
- [13] G.N. Milstein and M.V. Tretyakov, *Stochastic Numerics for Mathematical Physics Scientific Computation* (Springer, Berlin, Heidelberg, 2004).
- [14] A. Rößler, *J. Comput. Appl. Math.* **164–165**, 613–627 (2004).
- [15] A. Rößler, *SIAM. J. Numer. Anal.* **47** (3), 1713–1738 (2009).
- [16] H. Lamba, J.C. Mattingly and A.M. Stuart, *IMA J. Numer. Anal.* **27** (3), 479–506 (2007).
- [17] D.F. Anderson and J.C. Mattingly, *Commun. Math. Sci.* **9** (1), 301–318 (2010).
- [18] B. Sprinkle, A. Donev, A.P.S. Bhalla and N. Patankar, *J. Chem. Phys.* **150** (16), 164116 (2019).
- [19] Y. Aït-Sahalia, *Ann. Stat.* **36** (2), 906–937 (2008).
- [20] J.K. Møller and H. Madsen, Technical University of Denmark, DTU Informatics, Building, 2010, 321.
- [21] H. Luschgy and G. Pagès, *Stoch. Process. Their. Appl.* **116** (2), 310–336 (2006).
- [22] S.G. de Boer, MSc thesis, Delft University of Technology, 2020.
- [23] B. Øksendal, *Stochastic Differential Equations*, Universitext (Springer, Berlin, Heidelberg, 2003).
- [24] A. Agazzi and J.C. Mattingly, Duke University Lecture Notes, 2023.
- [25] A. Leroy, B. Leimkuhler, J. Latz and D.J. Higham, arXiv:2403.11993, 2024.
- [26] G.A. Pavliotis, *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*, Texts in Applied Mathematics, 60 vols (Springer, New York, NY, 2014).
- [27] P. Craigmile, R. Herbei, G. Liu and G. Schneider, *WIREs. Comp. Stats.* **15** (2), e1585 (2023).
- [28] C. Abrams and G. Bussi, *Entropy* **16** (1), 163–199 (2014).
- [29] G.O. Roberts and O. Stramer, *Methodol. Comput. Appl. Probab.* **4** (4), 337–357 (2002).
- [30] T. Lelièvre, G.A. Pavliotis, G. Robin, R. Santet and G. Stoltz, arXiv:2404.12087 [cs, math], 2024.
- [31] S.V. Krivov and M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* **105** (37), 13841–13846 (2008).
- [32] M. Hutzenthaler, A. Jentzen and P.E. Kloeden, *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **467** (2130), 1563–1576 (2011).
- [33] G.N. Mil'shtejn, *Theory Probab. Appl.* **19** (3), 557–562 (1975).
- [34] B. Leimkuhler and C. Matthews, *Appl. Math. Res. Express.* **2013** (1), 34–56 (2013).
- [35] E.K. Blum, *Numerical Analysis and Computational Theory and Practice* (Addison Wesley, Boston, 1972).
- [36] R. Catlow, R. Bell, F. Cora and B. Slater, in *Introduction to Zeolite Science and Practice*, edited by J. Čejka, H. van Bekkum, A. Corma, and F. Schüth, Studies in Surface Science and Catalysis, 168 vols. (Elsevier, Amsterdam, 2007), pp. 659–700.
- [37] K.F. Kelton and A.L. Greer, in *Nucleation in Condensed Matter*, edited by K. F. Kelton, and A. L. Greer, Pergamon Materials Series, 15 vols (Pergamon, Oxford, 2010), pp. 279–329.
- [38] M. Moseley, J. Kucharczyk, H. Asgari and D. Norman, *Magn. Reson. Med.* **19** (2), 321–326 (1991).
- [39] G.J. Moro and F. Cardin, *Chem. Phys.* **235** (1), 189–200 (1998).
- [40] B. Leimkuhler, C. Matthews and J. Weare, *Stat. Comput.* **28** (2), 277–290 (2018).
- [41] T. Lelièvre, G. Stoltz and M. Rousset, *Free Energy Computations: A Mathematical Perspective*. (Imperial College Press, London, Hackensack, NJ, 2010). OCLC: ocn244765923.

Appendices

Appendix 1. Double-well enhanced sampling

Transition counts in a double well. Given a Brownian dynamics trajectory $X = (x_0, x_1, \dots, x_n)$, we define the transition counts N_T to be the total number of times that the trajectory's state passes directly from the left inflection point of the double well to the right inflection point (or vice versa) through a sequence of intermediate states.

L_1 error in the invariant measure. We use an identical method to the method of computation of L_1 discussed in Section 5.2. We use 30 equal-width intervals in the range -5 to 5.

Methodology details for Figure 4. For each kT value and integrator, experiments were run with 30 values for α equally spaced between 0 and 1 inclusive. For each α value, 5×10^5 independent repeats were run for the simulations of length $T_{\text{sim}}, T_{\text{sim}} = 10$ and 5×10^3 repeats for the simulations of length $T_{\text{sim}}, T_{\text{sim}} = 1000$ and the mean L_1 error was calculated. The minimum L_1 error across all considered α values is plotted in Figure 4.

Appendix 2. Numerical Integrators

We use the shorthand notation

$$a(x) := -D(x) \frac{dV}{dx} + kT \frac{dD}{dx},$$

$$\sigma(x) := \sqrt{2kTD(x)},$$

$$\tilde{a}(x) := a(x) - \frac{1}{2}\sigma(x) \frac{d\sigma}{dx} = -D(x) \frac{dV}{dx} + \frac{1}{2}kT \frac{dD}{dx},$$

where $a(x)$ is the drift term, $\sigma(x)$ the diffusion term, and $\tilde{a}(x)$ is the Stratonovich-corrected drift [26]. We consider the following integrators, where $w_n, w_{n+1} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, 1)$ and h is the step size:

- (i) Euler-Maruyama (EM)

$$x_{n+1} = x_n + a(x_n)h + \sigma(x_n)\sqrt{h}w_n;$$

- (ii) Milstein Method (MM)

$$x_{n+1} = x_n + a(x_n)h + \sigma(x_n)\sqrt{h}w_n + \frac{1}{2}kT \frac{dD}{dx}(x_n)(w_n^2 - 1)h;$$

- (iii) Leimkuhler-Matthews (LM)

$$x_{n+1} = x_n + a(x_n)h + \sigma(x_n)\sqrt{h} \frac{w_n + w_{n+1}}{2};$$

(iv) Hummer-Leimkuhler-Matthews (HLM)

$$x_{n+1} = x_n + \left(a(x_n) + \frac{1}{4}kT \frac{dD}{dx}(x_n) \right) h + \sigma(x_n) \sqrt{h} \frac{w_n + w_{n+1}}{2};$$

(v) Stochastic Heun (SH)

$$\begin{aligned} x_{n+1}^* &= x_n + \tilde{a}(x_n)h + \sigma(x_n)\sqrt{h}w_n \\ x_{n+1} &= x_n + \frac{1}{2}(\tilde{a}(x_n) + \tilde{a}(x_{n+1}^*))h \\ &\quad + \frac{1}{2}(\sigma(x_n) + \sigma(x_{n+1}^*))\sqrt{h}w_n; \end{aligned}$$

(vi) Limit Method with Variable Diffusion (LMVD)

$$\begin{aligned} \hat{x}_{n+1} &= \sqrt{kT}w_n - \sqrt{2hD(x_n)} \frac{dV}{dx}(x_n) \\ &\quad + kT \sqrt{\frac{h}{2D(x_n)}} \frac{dD}{dx}(x_n) \\ \tilde{x}_{n+1} &= \left\{ x \left(\sqrt{\frac{h}{2}} \right) \middle| x(0) = x_n, dx = \sqrt{D(x)} \hat{x}_{n+1} dt \right\} \\ x_{n+1} &= \left\{ x \left(\sqrt{\frac{h}{2}} \right) \middle| x(0) = \tilde{x}_{n+1}, dx \right. \\ &\quad \left. = \sqrt{kTD(x)}w_{n+1} dt \right\}. \end{aligned} \tag{A1}$$

Appendix 3. Derivation of the Limit Method with Variable Diffusion

Consider the dynamics originally proposed in [40, Equation 6]

$$\begin{aligned} d\mathbf{X}_t &= \mathbf{B}(\mathbf{X}_t)\mathbf{P}_t dt, \\ d\mathbf{P}_t &= -\mathbf{B}(\mathbf{X}_t)^T \nabla V(\mathbf{X}_t) dt + kT \operatorname{div}(\mathbf{B}^T)(\mathbf{X}_t) dt - \gamma \mathbf{P}_t dt \\ &\quad + \sqrt{2\gamma kT} d\mathbf{W}_t, \end{aligned} \tag{A2}$$

where $\mathbf{B}(\mathbf{X})$ is a positive definite matrix, $\gamma > 0$ is a friction parameter and $\mathbf{P} \in \mathbb{R}^n$ denotes the instantaneous system momentum. It is straightforward to check that these dynamics preserve the canonical distribution

$$\rho(\mathbf{X}, \mathbf{P}) \propto \exp(-V(\mathbf{X})/kT - \|\mathbf{P}\|^2/2kT),$$

for any positive definite matrix $\mathbf{B}(\mathbf{X})$, where the marginal distribution of position satisfies

$$\int \rho(\mathbf{X}, \mathbf{P}) d\mathbf{P} \propto \rho(\mathbf{X}).$$

We now consider discretisations of (A2) built via splitting the SDE into three pieces denoted A, B, and O:

$$\begin{aligned} d \begin{bmatrix} \mathbf{X}_t \\ \mathbf{P}_t \end{bmatrix} &= \underbrace{\begin{bmatrix} \mathbf{B}(\mathbf{X}_t)\mathbf{P}_t \\ \mathbf{0} \end{bmatrix}}_A dt \\ &\quad + \underbrace{\begin{bmatrix} \mathbf{0} \\ -\mathbf{B}(\mathbf{X}_t)^T \nabla V(\mathbf{X}_t) + kT \operatorname{div}(\mathbf{B}^T)(\mathbf{X}_t) \end{bmatrix}}_B dt \end{aligned}$$

$$+ \underbrace{\begin{bmatrix} \mathbf{0} \\ -\gamma \mathbf{P}_t dt + \sqrt{2\gamma kT} d\mathbf{W}_t \end{bmatrix}}_O.$$

Note that when \mathbf{B} is a constant matrix (A2) reduces to conventional Langevin dynamics, and the above splitting matches the pieces given in [34].

Taking any of the A, B or O pieces in isolation, we may solve the implied SDE exactly (in distribution) for time $t > 0$. Denoting the solution to each piece as $\phi_t(\mathbf{X}, \mathbf{P})$, given the initial conditions at $t = 0$ are (\mathbf{X}, \mathbf{P}) , we can write

$$\begin{aligned} \phi_t^A(\mathbf{X}, \mathbf{P}) &= (\{Y(t) | Y(0) = \mathbf{X}, dY = \mathbf{B}(Y)\mathbf{P} dt\}, \mathbf{P}), \\ \phi_t^B(\mathbf{X}, \mathbf{P}) &= (\mathbf{X}, \mathbf{P} - t\mathbf{B}(\mathbf{X})^T \nabla V(\mathbf{X}) + tkT \operatorname{div}(\mathbf{B}^T)(\mathbf{X})), \\ \phi_t^O(\mathbf{X}, \mathbf{P}) &= (\mathbf{X}, e^{-\gamma t} \mathbf{P} + \sqrt{kT} \sqrt{1 - e^{-2\gamma t}} \mathbf{R}), \end{aligned}$$

where $\mathbf{R} \sim N(\mathbf{0}, \mathbf{I})$ is a normal random vector. As ϕ^A has no explicit closed form, we write the update purely as the solution to the underlying ODE.

We now consider the overdamped limit $\gamma \rightarrow \infty$ with a time step $s > 0$, using the discretisation scheme

$$(\mathbf{X}_{n+1}, \mathbf{P}_{n+1}) := \phi_{s/2}^A \circ \phi_s^O \circ \phi_{s/2}^A \circ \phi_s^B(\mathbf{X}_n, \mathbf{P}_n).$$

Writing out the resulting steps, we obtain

$$\begin{aligned} \hat{\mathbf{P}}_n &= \mathbf{P}_n - s\mathbf{B}(\mathbf{X}_n)^T \nabla V(\mathbf{X}_n) + skT \operatorname{div}(\mathbf{B}^T)(\mathbf{X}_n) \\ \hat{\mathbf{X}}_n &= \left\{ Y(s/2) \middle| Y(0) = \mathbf{X}_n, dY = \mathbf{B}(Y)\hat{\mathbf{P}}_n \right\} \\ \mathbf{P}_{n+1} &= \sqrt{kT}\mathbf{R}_{n+1} \\ \mathbf{X}_{n+1} &= \left\{ Y(s/2) \middle| Y(0) = \hat{\mathbf{X}}_n, dY = \mathbf{B}(Y)\mathbf{P}_{n+1} \right\} \end{aligned}$$

which we may simplify by recognising that $\mathbf{P}_n \equiv \sqrt{kT}\mathbf{R}_n$.

We recover the LMVD method given in (A1) by considering the one-dimensional case where $B(x) = \sqrt{D(x)}$ and choosing $s = \sqrt{2h}$ for a time step of $h > 0$ to ensure consistency between schemes [41].

Appendix 4. Proofs

This section contains proofs of all results stated in Sections 3 and 4.

Theorem A.1: *Applying a continuous coordinate transform to a one-dimensional Brownian dynamics process results in another Brownian dynamics process with potential and diffusion function given by (10).*

Proof: Applying Itô's Lemma with $y = y(x)$, where x obeys (1), we have,

$$\begin{aligned} dy &= \left(-D(x(y)) \frac{dV}{dx}(x(y)) + kT \frac{dD}{dx}(x(y)) \right) \frac{dy}{dx}(x(y)) dt \\ &\quad + kTD(x(y)) \frac{d^2y}{dx^2}(x(y)) dt \\ &\quad + \sqrt{2kTD(x(y))} \frac{dy}{dx}(x(y)) dW_t. \end{aligned}$$

Now substituting in the transformations (10),

$$= -\hat{D} \frac{d}{dx} \left(\hat{V} dt - kT \ln \left| \frac{dy}{dx} \right| \right) \left(\frac{dy}{dx} \right)^{-1} dt$$

$$\begin{aligned}
& + kT \frac{d}{dx} \left(\hat{D} \frac{dy^{-2}}{dx} \right) \frac{dy}{dx} dt \\
& + kT \hat{D} \frac{dy^{-2}}{dx} \frac{d^2y}{dx^2} dt + \sqrt{2kT\hat{D}} dW_t \\
& = -\hat{D} \frac{d\hat{V}}{dy} dt + kT \hat{D} \frac{d^2y}{dx^2} \frac{dy}{dx} dt \\
& + kT \frac{d\hat{D}}{dy} dt - 2kT \hat{D} \frac{d^2y}{dx^2} \frac{dy}{dx} dt + kT \hat{D} \frac{dy^{-2}}{dx} \frac{d^2y}{dx^2} dt \\
& + \sqrt{2kT\hat{D}} dW_t \\
& = -\hat{D}(y) \frac{d\hat{V}}{dy} dt + kT \frac{d\hat{D}}{dy} dt + \sqrt{2kT\hat{D}(y)} dW_t,
\end{aligned}$$

which is Brownian motion with potential $\hat{V}(y)$ and diffusion function $\hat{D}(y)$. \blacksquare

Theorem A.2: *In one dimension, the Lamperti-transformed process can be used to compute phase space averages through (13).*

Proof: We assume that the transformed process y_t is ergodic and therefore satisfies

$$\lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(y_t) dt = \int_{-\infty}^{\infty} f(y) \hat{\rho}(y) dy, \quad (\text{A3})$$

where $\hat{\rho}(y) = \frac{1}{Z} \exp(-\frac{\hat{V}(y)}{kT})$ is the invariant distribution of the transformed process. Substituting in the effective potential from (10), the right-hand side of (A3) becomes

$$\begin{aligned}
& \int_{-\infty}^{\infty} f(y) \frac{\exp(-\frac{V(x(y))}{kT})}{\hat{Z}} \sqrt{D(x(y))} dy \\
& = \int_{-\infty}^{\infty} f(y(x)) \frac{\exp(-\frac{V(x)}{kT})}{\hat{Z}} \sqrt{D(x)} \frac{dy}{dx} dx.
\end{aligned}$$

Using the fact $\frac{dy}{dx} = \frac{1}{\sqrt{D(x)}}$, this equation simplifies to

$$\int_{-\infty}^{\infty} f(y(x)) \frac{\exp(-\frac{V(x)}{kT})}{\hat{Z}} dx = \frac{Z}{\hat{Z}} \int_{-\infty}^{\infty} f(y(x)) \rho(x) dx \quad (\text{A4})$$

where $Z = \int_{-\infty}^{\infty} \exp(-\frac{V(x)}{kT}) dx$ and $\hat{Z} = \int_{-\infty}^{\infty} \exp(-\frac{\hat{V}(y)}{kT}) dy$ are the partition functions of the original and transformed processes respectively.

But $Z = \hat{Z}$ since, by change of variables:

$$\begin{aligned}
Z & = \int_{-\infty}^{\infty} \exp\left(-\frac{V(x)}{kT}\right) dx \\
& = \int_{-\infty}^{\infty} \exp\left(-\frac{V(x(y))}{kT}\right) \left(\frac{dy}{dx}(x(y))\right)^{-1} dy \\
& = \int_{-\infty}^{\infty} \exp\left(-\frac{V(x(y)) + kT \ln \left|\frac{dy}{dx}(x(y))\right|}{kT}\right) dy \\
& = \int_{-\infty}^{\infty} \exp\left(-\frac{\hat{V}(y)}{kT}\right) dy \\
& = \hat{Z}.
\end{aligned}$$

Hence, from (A3) and (A4) we have

$$\int_{-\infty}^{\infty} f(y(x)) \rho(x) dx = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(y_t) dt.$$

Finally, if we redefine f as $f \circ x$, then we obtain:

$$\int_{-\infty}^{\infty} f(x) \rho(x) dx = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(x(y_t)) dt, \quad (\text{A5})$$

as required. \blacksquare

Theorem A.3: *In one dimension, the Lamperti-transformed invariant measure $\hat{\rho}(y)$ and the original invariant measure $\rho(x)$ are related by $\rho(x) = \hat{\rho}(x(y)) \frac{dy}{dx}$.*

Proof: Set $f(x) = I(x \in [a, b])$, the indicator function on the interval $[a, b]$, in (A5). This gives

$$\begin{aligned}
& \int_{-\infty}^{\infty} f(x) \rho(x) dx \\
& = \int_a^b \rho(x) dx = \lim_{T \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} I(x(y_t) \in [a, b]) dt \\
& = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} I(y_t \in [y(a), y(b)]) dt = \int_{y(a)}^{y(b)} \hat{\rho}(y) dy \\
& = \int_a^b \hat{\rho}(x(y)) \frac{dy}{dx} dx, \quad (\text{A6})
\end{aligned}$$

where in the second line we re-expressed the indicator function in terms of y_t and then applied the ergodic theorem for the y_t process. (A6) implies

$$\int_a^b \left(\rho(x) - \hat{\rho}(x(y)) \frac{dy}{dx} \right) dx = 0$$

which, from the arbitrariness of the constants a and b , proves the result. \blacksquare

Theorem A.4: *Applying a time-rescaling to a one-dimensional Brownian dynamics process results in another Brownian dynamics process with potential and diffusion given by (15).*

Proof: Applying a version of the time rescaling appearing in Equation (8) to one-dimensional Brownian dynamics we arrive at,

$$\begin{aligned}
dx_\tau & = -g(x)D(x) \frac{dV}{dx} d\tau + kTg(x) \frac{dD(x)}{dx} d\tau \\
& + \sqrt{2kTg(x)D(x)} dW_\tau. \quad (\text{A7})
\end{aligned}$$

Inserting the identities from Equation (15) into (A7) we obtain

$$\begin{aligned}
dx_\tau & = -\hat{D}(x) \frac{d}{dx} \left(\hat{V}(x) - kT \ln g(x) \right) d\tau \\
& + kTg(x) \frac{d}{dx} \left(\frac{\hat{D}(x)}{g(x)} \right) d\tau + \sqrt{2kT\hat{D}(x)} dW_\tau \\
& = -\hat{D}(x) \frac{d\hat{V}}{dx} d\tau + kT\hat{D}(x) \frac{g'(x)}{g(x)} d\tau \\
& + kT \frac{d\hat{D}(x)}{dx} d\tau - kT\hat{D}(x) \frac{g'(x)}{g(x)} d\tau + \sqrt{2kT\hat{D}(x)} dW_\tau
\end{aligned}$$

$$= -\hat{D}(x) \frac{d\hat{V}}{dx} d\tau + kT \frac{d\hat{D}(x)}{dx} d\tau + \sqrt{2kT\hat{D}(x)} dW_\tau,$$

which is a transformed version of the original one-dimensional Brownian dynamics but in an effective potential $\hat{V}(x)$ and a rescaled diffusion coefficient $\hat{D}(x)$, as required. ■

Theorem A.5: *In one dimension, the time-rescaled process can be used to compute phase space averages through (16).*

Proof: From the ergodic theorem applied to the original process,

$$\int_{-\infty}^{\infty} f(x) \rho(x) dx = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(x_t) dt.$$

Changing variables $t \rightarrow \tau$ in the integration, the right-hand side becomes,

$$\begin{aligned} & \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{\tau=0}^{\tau(T_{\text{sim}})} f(x_\tau) \frac{d\tau}{dt} d\tau \\ &= \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{\tau=0}^{\tau(T_{\text{sim}})} f(x_\tau) g(x_\tau) d\tau. \end{aligned}$$

Writing $\mathcal{T}_{\text{sim}} := \tau(T_{\text{sim}})$, this can be alternatively written as

$$\lim_{\mathcal{T}_{\text{sim}} \rightarrow \infty} \frac{1}{t(\mathcal{T}_{\text{sim}})} \int_{\tau=0}^{\mathcal{T}_{\text{sim}}} f(x_\tau) g(x_\tau) d\tau.$$

Finally, we note that by integrating $\frac{dt}{d\tau} = g(x)$ between 0 and \mathcal{T}_{sim} we can obtain an expression for $t(\mathcal{T}_{\text{sim}})$,

$$t(\mathcal{T}_{\text{sim}}) = \int_{\tau=0}^{\mathcal{T}_{\text{sim}}} g(x_\tau) d\tau.$$

Substituting this into the above equation completes the proof. ■

Theorem A.6: *A Lamperti-transformed process with original \mathbf{D} matrix of the form $\mathbf{D}(\mathbf{X})_{ij} = D_i(X_i)R_{ij}$ is an instance of Brownian dynamics if and only if \mathbf{R} is diagonal. Further, when $R_{ij} = \delta_{ij}$, the effective potential is given by (21).*

Proof: The stated transformation is a multivariate Lamperti transform (6) with

$$\begin{aligned} f(\mathbf{X}) &= -\mathbf{D}(\mathbf{X})\mathbf{D}(\mathbf{X})^T \nabla V(\mathbf{X}) + kT \text{div}(\mathbf{D}\mathbf{D}^T)(\mathbf{X}), \\ \sigma(\mathbf{X}) &= \sqrt{2kT\mathbf{D}(\mathbf{X})}. \end{aligned}$$

The transformed process therefore satisfies

$$\begin{aligned} dY_{i,t} &= \sum_{j=1}^n R_{ij}^{-1} \sqrt{2kT} \left(\frac{-\sum_{k=1}^n (\mathbf{D}\mathbf{D}^T)_{jk} \partial_k V}{\sqrt{2kT} D_j} \right. \\ &\quad \left. + \frac{kT \sum_{k=1}^n \partial_k (\mathbf{D}\mathbf{D}^T)_{jk}}{\sqrt{2kT} D_j} - \frac{1}{2} \sqrt{2kT} \partial_j D_j \right) dt \\ &\quad + \sqrt{2kT} dW_i, \end{aligned}$$

where $\partial_j := \frac{\partial}{\partial X_j}$ and V , \mathbf{D} and D_j are functions of \mathbf{Y}_t through the relations

$$\begin{aligned} V(\mathbf{X}_t) &= V(\phi^{-1}(\mathbf{R}\mathbf{Y}_t)), \quad \mathbf{D}(\mathbf{X}_t) = \mathbf{D}(\phi^{-1}(\mathbf{R}\mathbf{Y}_t)), \\ D(X_{j,t}) &= D(\phi_j^{-1}((\mathbf{R}\mathbf{Y})_{j,t})). \end{aligned}$$

Substituting $\mathbf{D}(\mathbf{X})_{ij} = D_i(X_i)R_{ij}$, this becomes

$$\begin{aligned} dY_{i,t} &= \sum_{j,k,l=1}^n R_{ij}^{-1} \left(\frac{-R_{jk} R_{kl} D_j D_k \partial_k V}{D_j} \right. \\ &\quad \left. + kT \frac{R_{jl} R_{kl} \partial_k (D_j D_k)}{D_j} \right) dt \\ &\quad - kT \sum_{j=1}^n R_{ij}^{-1} \partial_j D_j dt \\ &\quad + \sqrt{2kT} dW_i. \end{aligned}$$

Expanding,

$$\begin{aligned} dY_{i,t} &= \sum_{k=1}^n -R_{ki} D_k \partial_k V dt \\ &\quad + kT \left(\sum_{j,k,l=1}^n R_{ij}^{-1} R_{jl} R_{kl} \left(\partial_k D_j \frac{D_k}{D_j} + \partial_k D_k \right) \right. \\ &\quad \left. - \sum_{j=1}^n R_{ij}^{-1} \partial_j D_j \right) dt \\ &\quad + \sqrt{2kT} dW_i. \end{aligned}$$

Noting that $\partial_k D_j = \delta_{kj} \partial_j D_j$, this becomes

$$\begin{aligned} dY_{i,t} &= \sum_{k=1}^n -R_{ki} D_k \partial_k V dt \\ &\quad + kT \left(\sum_{j,l} R_{ij}^{-1} R_{jl} R_{jl} \partial_k D_k + \sum_k R_{ki} \partial_k D_k \right. \\ &\quad \left. - \sum_{j=1}^n R_{ij}^{-1} \partial_j D_j \right) dt \\ &\quad + \sqrt{2kT} dW_i. \end{aligned}$$

Changing variables so that the derivatives are with respect to \mathbf{Y} we get

$$\frac{\partial}{\partial X_k} = \sum_{l=1}^n \frac{\partial Y_l}{\partial X_k} \frac{\partial}{\partial Y_l} = \sum_{l=1}^n \frac{R_{lk}^{-1}}{D_k(X_k)} \frac{\partial}{\partial Y_l},$$

and the transformed equation becomes

$$\begin{aligned} dY_{i,t} &= -\partial_i V dt + kT \sum_{k=1}^n \left(\left(\sum_{l=1}^n R_{ik}^{-1} R_{kl} R_{kl} R_{kk}^{-1} \right) \right. \\ &\quad \left. + R_{ki} R_{kk}^{-1} + R_{ik}^{-1} R_{kk}^{-1} \right) \nabla_{Y_k} \ln D_k dt \\ &\quad + \sqrt{2kT} dW_i, \end{aligned}$$

or equivalently:

$$dY_{i,t} = -\partial_i V dt + kT \sum_{k=1}^n M_{ik} \nabla_{Y_k} \ln D_k dt + \sqrt{2kT} dW_i,$$

where \mathbf{M} is the matrix defined in the theorem statement. Note that only if the matrix \mathbf{R} is diagonal (and therefore, $M_{ij} = \delta_{ij}$) is

it possible to express the drift term as a gradient of an effective potential, given by

$$\hat{V}(\mathbf{Y}) = V(\phi^{-1}(\mathbf{R}\mathbf{Y})) - kT \sum_{i=1}^n \ln D_i(\phi_i^{-1}(\mathbf{R}\mathbf{Y}_i)). \quad (\text{A8})$$

In particular, if we set $R_{ij} = \delta_{ij}$, then we identify the transformed process as constant-diffusion Brownian dynamics with an effective potential

$$\hat{V}(\mathbf{Y}) = V(\phi^{-1}(\mathbf{Y})) - kT \sum_{k=1}^n \ln D_k(\phi_k^{-1}(Y_k)). \quad (\text{A9})$$

Remark: The functions D_i can be arbitrarily scaled in such a manner that for all i , $R_{ii} = 1$ in Equation (A8). The transformed process then becomes equivalent to the case $\mathbf{R} = \mathbf{I}$, as discussed in Section 4.1. ■

Theorem A.7: A Lamperti-transformed process with original \mathbf{D} matrix of the form $\mathbf{D}(\mathbf{X}) = D_i(X_i)\delta_{ij}$ can be used to compute phase-space averages through (22).

Proof: We assume that the effective potential (A9) is such that geometric ergodicity holds for \mathbf{Y}_t . Then, by applying the ergodic theorem to the transformed process, we obtain:

$$\lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(\mathbf{Y}_t) dt = \int_{\mathbb{R}^n} f(\mathbf{Y}) \hat{\rho}(\mathbf{Y}) d\mathbf{Y}.$$

Substituting in the effective potential, we have:

$$= \int_{\mathbb{R}^n} f(\mathbf{Y}) \frac{1}{\hat{Z}} \exp\left(-\frac{V(\phi^{-1}(\mathbf{Y}))}{kT}\right) \prod_{i=1}^n (D_i(\phi_i^{-1}(Y_{i,t}))) d\mathbf{Y}$$

Next, we change variables from \mathbf{Y} to \mathbf{X} . The Jacobian factor is given by:

$$J = \left| \frac{d\mathbf{Y}}{d\mathbf{X}} \right| = \left| \frac{1}{D_i(X_i)} \delta_{ij} \right| = \prod_{i=1}^n \frac{1}{D_i(X_i)},$$

which exactly cancels with the diffusion coefficients in the integral, and we have

$$\begin{aligned} &= \int_{\mathbb{R}^n} f(\phi(\mathbf{X})) \frac{1}{\hat{Z}} \exp\left(-\frac{V(\mathbf{X})}{kT}\right) d\mathbf{X} \\ &= \frac{Z}{\hat{Z}} \int_{\mathbb{R}^n} f(\phi(\mathbf{X})) \rho(\mathbf{X}) d\mathbf{X}. \end{aligned}$$

Choosing $f(\mathbf{Y}) = 1$ leads to $\hat{Z} = Z$, hence

$$\lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(\mathbf{Y}_t) dt = \int_{\mathbb{R}^n} f(\phi(\mathbf{X})) \rho(\mathbf{X}) d\mathbf{X}.$$

Finally, if we redefine f as $f \circ \phi^{-1}$, then we obtain:

$$\lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(\phi^{-1}(\mathbf{Y}_t)) dt = \int_{\mathbb{R}^n} f(\mathbf{X}) \rho(\mathbf{X}) d\mathbf{X},$$

as required. ■

Theorem A.8: The effective potential of a time-rescaled Brownian dynamics process with original \mathbf{D} matrix $\mathbf{D}(\mathbf{X}) = D(\mathbf{X})\mathbf{R}$ is given by (23).

Proof: The time-rescaling transform follows (8), (9) with $f(\mathbf{X}) = -\mathbf{D}(\mathbf{X})\mathbf{D}(\mathbf{X})^T \nabla V(\mathbf{X}) + kT \text{div}(\mathbf{D}(\mathbf{X})\mathbf{D}(\mathbf{X})^T)$, which gives (we transform to constant diffusion $\sqrt{2kT}$):

$$d\mathbf{Y}_\tau = -\mathbf{R}^{-1} \frac{\mathbf{D}\mathbf{D}^T \nabla_{\mathbf{X}} V - kT \text{div}(\mathbf{D}\mathbf{D}^T)}{D^2} dt + \sqrt{2kT} d\mathbf{W}_\tau.$$

Here, V , \mathbf{D} and D are functions of \mathbf{Y}_τ through the relations:

$$V(\mathbf{X}_\tau) = V(\mathbf{R}\mathbf{Y}_\tau), \quad \mathbf{D}(\mathbf{X}_\tau) = \mathbf{D}(\mathbf{R}\mathbf{Y}_\tau), \quad D(\mathbf{X}_\tau) = D(\mathbf{R}\mathbf{Y}_\tau).$$

Substituting $\mathbf{D}(\mathbf{X}) = D(\mathbf{X})\mathbf{R}$, in components this becomes

$$dY_{i,\tau} = -\frac{\sum_j D^2 R_{ji} \partial_j V - kT \sum_j R_{ji} \partial_j (D^2)}{D^2} dt + \sqrt{2kT} dW_{i,\tau},$$

which simplifies to

$$dY_{i,\tau} = \sum_j R_{ji} (-\partial_j V + 2kT \partial_j \ln D) dt + \sqrt{2kT} dW_{i,\tau}.$$

Changing variables so that the derivatives are with respect to \mathbf{Y} we get

$$\frac{\partial}{\partial X_i} = \sum_{j=1}^n \frac{\partial Y_j}{\partial X_i} \frac{\partial}{\partial Y_j} = \sum_{j=1}^n R_{ji}^{-1} \frac{\partial}{\partial Y_j}.$$

The \mathbf{R} matrix then cancels with its inverse, and the dynamics now reads

$$d\mathbf{Y}_\tau = (-\nabla_{\mathbf{Y}} V(\mathbf{R}\mathbf{Y}) + 2kT \nabla_{\mathbf{Y}} \ln D(\mathbf{R}\mathbf{Y})) dt + \sqrt{2kT} d\mathbf{W}_\tau,$$

which is constant-diffusion Brownian dynamics in an effective potential $\hat{V}(\mathbf{Y})$ given by

$$\hat{V}(\mathbf{Y}) = V(\mathbf{R}\mathbf{Y}) - 2kT \ln D(\mathbf{R}\mathbf{Y}).$$

This completes the proof. ■

Theorem A.9: A time-rescaled process with original \mathbf{D} matrix of the form $\mathbf{D}(\mathbf{X}) = D(\mathbf{X})\mathbf{R}$ can be used to compute phase-space averages through (24).

Proof: We begin with the ergodic theorem of the original process, which states

$$\int_{\mathbb{R}^n} f(\mathbf{X}) \rho(\mathbf{X}) d\mathbf{X} = \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{t=0}^{T_{\text{sim}}} f(\mathbf{X}_t) dt.$$

To express this in terms of the time-rescaled process, we change the variable $t \rightarrow \tau$ in the integration, resulting in:

$$\begin{aligned} &\lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{\tau=0}^{\tau(T_{\text{sim}})} f(\mathbf{X}_\tau) \frac{dt}{d\tau} d\tau \\ &= \lim_{T_{\text{sim}} \rightarrow \infty} \frac{1}{T_{\text{sim}}} \int_{\tau=0}^{\tau(T_{\text{sim}})} f(\mathbf{X}_\tau) g(\mathbf{X}_\tau) d\tau. \end{aligned}$$

Writing $\mathcal{T}_{\text{sim}} := \tau(T_{\text{sim}})$, this can alternatively be written as

$$\lim_{\mathcal{T}_{\text{sim}} \rightarrow \infty} \frac{1}{t(\mathcal{T}_{\text{sim}})} \int_{\tau=0}^{\mathcal{T}_{\text{sim}}} f(\mathbf{X}_\tau) g(\mathbf{X}_\tau) d\tau, \quad (\text{A10})$$

where $g(\mathbf{X}_\tau) = 1/D^2(\mathbf{X})$ by the definition of time rescaling. Next, we integrate $\frac{dt}{d\tau} = g(\mathbf{X})$ from 0 to \mathcal{T}_{sim} to obtain an

expression for $t(\mathcal{T}_{\text{sim}})$,

$$t(\mathcal{T}_{\text{sim}}) = \int_{\tau=0}^{\mathcal{T}_{\text{sim}}} g(\mathbf{X}_\tau) d\tau, \quad (\text{A11})$$

Substituting Equation (A11) and the relation $\mathbf{X}_\tau = \mathbf{R}\mathbf{Y}_\tau$ into Equation (A10), we have:

$$\int_{\mathbb{R}^n} f(\mathbf{X}) \rho(\mathbf{X}) d\mathbf{X} = \lim_{\mathcal{T}_{\text{sim}} \rightarrow \infty} \frac{\int_{\tau=0}^{\mathcal{T}_{\text{sim}}} f(\mathbf{R}\mathbf{Y}_\tau) g(\mathbf{R}\mathbf{Y}_\tau) d\tau}{\int_{\tau=0}^{\mathcal{T}_{\text{sim}}} g(\mathbf{R}\mathbf{Y}_\tau) d\tau},$$

as required. \blacksquare

Theorem A.10: *Performing a time rescaling followed by a Lamperti transform to a multivariate Brownian dynamics process with \mathbf{D} matrix $\mathbf{D}(\mathbf{X}) = \mathbf{D}^{(1)}(\mathbf{X})\mathbf{R}\mathbf{D}^{(2)}(\mathbf{X})$, where \mathbf{R} is not diagonal, results in a constant-diffusion process that is not Brownian dynamics.*

Proof: First, consider the time-rescaled process \mathbf{X}_τ , where $\frac{dt}{d\tau} = 1/D^2(\mathbf{X})$, which obeys the dynamics

$$d\mathbf{X}_\tau = -\frac{\mathbf{D}\mathbf{D}^T \nabla_{\mathbf{X}} V - kT \text{div}(\mathbf{D}\mathbf{D}^T)}{D^2} dt + \sqrt{2kT}\mathbf{R}\mathbf{D}^{(2)} d\mathbf{W}_\tau.$$

Defining a transformed process $\mathbf{Y}_\tau = \mathbf{R}^{-1}\mathbf{X}_\tau$, then applying the multidimensional Itô formula gives

$$d\mathbf{Y}_\tau = -\mathbf{R}^{-1} \left(\frac{\mathbf{D}\mathbf{D}^T \nabla_{\mathbf{X}} V|_{\mathbf{R}\mathbf{Y}_\tau} - kT \text{div}(\mathbf{D}\mathbf{D}^T)|_{\mathbf{R}\mathbf{Y}_\tau}}{D^2} \right) dt + \sqrt{2kT}\mathbf{D}^{(2)} d\mathbf{W}_\tau,$$

where

$$V = V(\mathbf{R}\mathbf{Y}_\tau), \quad \mathbf{D} = \mathbf{D}(\mathbf{R}\mathbf{Y}_\tau), \\ \mathbf{D}^{(2)} = \mathbf{D}^{(2)}(\mathbf{R}\mathbf{Y}_\tau), \quad D = D(\mathbf{R}\mathbf{Y}_\tau).$$

Finally, we apply a Lamperti transform to remove the noise dependence on $\mathbf{D}^{(2)}$. The transformed process $\mathbf{Z}_\tau = \phi(\mathbf{Y}_\tau)$ then satisfies (using Einstein summation convention for sums over repeated indices)

$$d\mathbf{Z}_{i,\tau} \\ = -R_{ij}^{-1} \left(\frac{D_{jk}D_{lk}\partial_l V|_{\mathbf{R}\phi^{-1}(\mathbf{Z}_\tau)} - kT\partial_l(D_{jk}D_{lk})|_{\mathbf{R}\phi^{-1}(\mathbf{Z}_\tau)}}{D^2 D_i} \right) dt + \sqrt{2kT} dW_{i,\tau}.$$

Changing variables,

$$\frac{\partial}{\partial X_k} = \sum_{l=1}^n \frac{\partial Z_l}{\partial X_k} \frac{\partial}{\partial Z_l} = \sum_{l=1}^n \frac{R_{lk}^{-1}}{D_k(X_k)} \frac{\partial}{\partial Z_l}$$

and substituting $D_{ij} = \delta_{ik}\delta_{jl}R_{kl}DD_l$ this becomes

$$d\mathbf{Z}_{i,\tau} = -R_{ij}^{-1} \left(\frac{\delta_{jm}\delta_{nk}\delta_{lp}\delta_{qk}R_{mn}R_{pq}D^2 D_n D_q R_{rl}^{-1} \nabla_{Y_r} V}{D^2 D_i D_l} \right. \\ \left. \times \frac{-kTR_{rl}^{-1} \nabla_{Y_r} (\delta_{jm}\delta_{nk}\delta_{lp}\delta_{qk}R_{mn}R_{pq}D^2 D_n D_q)}{D^2 D_i D_l} \right) dt + \sqrt{2kT} dW_{i,\tau},$$

which simplifies to

$$d\mathbf{Z}_{i,\tau} \\ = -R_{ij}^{-1} \left(\frac{R_{jk}R_{lk}D^2 D_k^{-1} \nabla_{Y_r} V}{-kTR_{rl}^{-1} \nabla_{Y_r} (R_{jk}R_{lk}D^2 D_k^2)} \right) dt + \sqrt{2kT} dW_{i,\tau}, \\ d\mathbf{Z}_{i,\tau} \stackrel{(\text{nosumi})}{=} \\ - \left(\frac{R_{li}D^2 D_i^2 R_{rl}^{-1} \nabla_{Y_r} V - kTR_{rl}^{-1} R_{li} \nabla_{Y_r} (D^2 D_i^2)}{D^2 D_i D_l} \right) dt + \sqrt{2kT} dW_{i,\tau}.$$

Expanding this expression does not lead to a great simplification of terms. In particular, the non-vanishing of the \mathbf{R} matrix elements in the dt term means that the drift term cannot be written as a gradient of a potential energy, hence this is not Brownian dynamics. \blacksquare

Theorem A.11: *Consider a multivariate Brownian dynamics process \mathbf{X}_t following (2), where the diffusion tensor \mathbf{D} is defined as*

$$\mathbf{D}(\mathbf{X}) = \mathbf{D}^{(1)}(\mathbf{X})\mathbf{D}^{(2)}(\mathbf{X}),$$

where $\mathbf{D}^{(1)}$ and $\mathbf{D}^{(2)}$ are given by (25). Then the transformed process $\mathbf{Y}_\tau = \phi(\mathbf{X}_\tau)$, resulting from a time rescaling where $\frac{dt}{d\tau} = g(\mathbf{X}) := 1/D^2(\mathbf{X})$, followed by a Lamperti transform given by

$$Y_{i,\tau} = \sqrt{2kT} \int_{x_0}^{X_{i,\tau}} \frac{1}{D_i(x)} dx := \sqrt{2kT} \phi_i(X_{i,\tau}),$$

satisfies the constant-diffusion Brownian dynamics process:

$$dY_{i,\tau} = -\nabla_{Y_i} \hat{V}(\mathbf{Y}) dt + \sqrt{2kT} dW_i,$$

where $\hat{V}(\mathbf{Y})$ is the effective potential defined as

$$\hat{V}(\mathbf{Y}) = V(\phi^{-1}(\mathbf{Y})) - 2kT \ln D(\mathbf{Y}) \\ - kT \sum_{i=1}^n \ln D_k(\phi_k^{-1}(Y_{k,\tau})).$$

Proof: The time-rescaling transformation gives

$$d\mathbf{X}_\tau = -\frac{\mathbf{D}\mathbf{D}^T \nabla_{\mathbf{X}} V - kT \text{div}(\mathbf{D}\mathbf{D}^T)}{D^2} dt + \sqrt{2kT}\mathbf{D}^{(2)}(\mathbf{X}) d\mathbf{W}_\tau. \quad (\text{A12})$$

Applying the Lamperti transform then gives

$$dY_{i,\tau} = \sqrt{2kT} \left(-\frac{D_{ij}D_{kj}\partial_k V - kT\partial_k(D_{ij}D_{kj})}{\sqrt{2kT}D^2 D_i} \right. \\ \left. - \frac{1}{2} \sqrt{2kT} \partial_i D_i \right) dt + \sqrt{2kT} dW_i,$$

where $V(\mathbf{X}_\tau)$, $\mathbf{D}(\mathbf{X}_\tau)$, $D(\mathbf{X}_\tau)$ and $\mathbf{D}_i(\mathbf{X}_\tau)$ are functions of \mathbf{Y}_τ through the relation $\mathbf{X}_\tau = \phi^{-1}(\mathbf{Y}_\tau)$. Since

$$D_{ij}(\mathbf{X}) = \sum_{k=1}^n D_{ik}^{(1)}(\mathbf{X})D_{kj}^{(2)}(\mathbf{X})$$

$$= \sum_{k=1}^n \delta_{ik} \delta_{kj} D(\mathbf{X}) D_k(X_k) = D(\mathbf{X}) D_i(X_i),$$

this becomes

$$dY_{i,\tau} = -D_i \partial_i V dt + kT \frac{\partial_i (D^2 D_i^2)}{D^2 D_i} dt - kT \partial_i D_i dt + \sqrt{2kT} dW_i$$

which simplifies to

$$dY_{i,\tau} = -D_i \partial_i V dt + 2kT \frac{D_i}{D} \partial_i D dt + kT \partial_i D_i dt + \sqrt{2kT} dW_i.$$

Changing variables so that the derivatives are with respect to \mathbf{Y} we get

$$\frac{\partial}{\partial X_k} = \sum_{l=1}^n \frac{\partial Y_l}{\partial X_k} \frac{\partial}{\partial Y_l} = \sum_{l=1}^n \frac{\delta_{lk}}{D_k(X_k)} \frac{\partial}{\partial Y_l} = \frac{1}{D_k} \frac{\partial}{\partial Y_k},$$

and the transformed equation becomes

$$dY_{i,\tau} = -\nabla_{Y_i} V dt + 2kT \nabla_{Y_i} \ln D dt + kT \nabla_{Y_i} \ln D_i dt + \sqrt{2kT} dW_i,$$

which we identify as Brownian motion in an effective potential

$$\hat{V}(\mathbf{Y}) = V(\phi^{-1}(\mathbf{Y})) - 2kT \ln D(\phi^{-1}(\mathbf{Y})) - kT \sum_{i=1}^n \ln D_i(\phi_i^{-1}(Y_i)),$$

as required. ■