MARTINGALE-BASED GRADIENT DESCENT ALGORITHM FOR ESTIMATING FREE ENERGY VALUES OF DIFFUSIONS

HAN CHENG LIE[†‡], CHRISTOF SCHÜTTE[†§], AND CARSTEN HARTMANN[†]

Abstract. Thermodynamic free energies or cumulant generating functions play a significant role in the estimation of rare event statistics of equilibrated systems because of their interpretation as normalisation constants. In this article we discuss a recently proposed method [C. Hartmann and C. Schütte, *J. Stat. Mech. Theor. Exp.*, (2012), P11004] for variance reduction of free energy estimates of reversible diffusions by minimisation of a certain control functional. Our derivation of the method using the Cameron-Martin-Girsanov formula adds a martingale term to the control functional. Using numerical examples involving the calculation of rare event probabilities, we show that the martingale-based functional exhibits smaller variance under suboptimal controls, and that minimisation of the control functional by gradient descent yields free energy function approximations that exhibit more stability.

Key words. Importance sampling, optimal control, rare events, free energy, change of measure

AMS subject classifications. 49M20, 65C20, 82C31, 90C59, 93E20

1. Introduction.

Free energy differences are important quantities for characterising the properties of physical systems. For example, in chemical physics, one may model a chemical reaction by viewing the reactive complex as a physical system. Roughly speaking, the free energy difference of the reactive complex between its educt and product states characterises the energetic barrier which must be overcome in order for the reaction to take place. Crucially, free energy differences may be used to estimate the reaction rate, via Kramers' theory.

A popular model for reactive complexes in computational physics are diffusions - i.e., Markovian dynamical systems with continuous paths. This is because diffusions on energy landscapes can exhibit the multiscale nature of molecular dynamics, from the short time scale dynamics corresponding to random fluctuations, to the long time scale dynamics corresponding to transitions between basins of low energy on the energy landscape. Furthermore, it makes sense to model chemical reactions using the inter-basin transition events, because one can also model the educts and products as regions of the system's state space corresponding to basins of low energy.

In computational molecular dynamics, the challenge in studying transition events is their rarity. The random fluctuations that drive inter-basin transitions are uncorrelated, and impart amounts of energy that are small compared to the associated free energy differences. Thus, it becomes necessary to develop methods for importance sampling of diffusions and for computing free energy differences. One important class of methods is based on the relations between nonequilibrium and equilibrium statistical mechanical systems, e.g., those of Jarzynski [14] and Crooks [3]. These nonequilibrium methods involve applying a biasing force in order to drive the system out of equilibrium, and exploiting relations between equilibrium and nonequilibrium probability densities of observables in order to estimate the equilibrium average of a given property using nonequilibrium measurements. The mathematical theory behind

[†]Institut für Mathematik, Freie Universität Berlin, Berlin, Germany
(hlie@zedat.fu-berlin.de, schuette@mi.fu-berlin.de, chartman@mi.fu-berlin.de)

[‡]This author is supported by a International Max-Planck Research School (IMPRS) scholarship from the Max Planck-Institut für Molekulare Genetik, Berlin, Germany.

[§]Zuse-Institut Berlin, Berlin, Germany

importance sampling methods for diffusions also continues apace; see [7,8,29] and the references therein for some examples.

In this paper, we derive and extend a nonequilibrium method [13] for estimating properties of diffusions on landscapes via the solution of a stochastic optimal control problem. The method expresses such properties in terms of a quantity which formally resembles the Helmholtz free energy in statistical physics. Our derivation uses the Cameron-Martin-Girsanov formula to obtain the key inequality (2.12) of the method. The derivation yields a nonequilibrium estimator of the free energy which modifies the estimator in [13] by the addition of a martingale term. The main contributions of this paper are to show how the martingale term changes the statistical properties of the nonequilibrium estimator and leads to different gradient descent dynamics. In particular, we show that the martingale-based estimator exhibits smaller variance, and that the resulting martingale-based gradient descent dynamics are more stable.

The outline of this paper is as follows. We present the theory behind the method in Section §2, the numerical discretisation scheme in Section §3, and an illustrative example in Section §4. In addition to the derivation of the method in Section §2.1, the new results in this paper are proofs of convergence of measures in Section §2.3, an analysis of the role of the martingale in the approximate gradient descent dynamics in Section §3.1, and an analysis of the sensitivity of the gradient descent algorithm to some parameters in Section §4.3. We conclude in Section §5.

2. Theory. In this section, we apply Girsanov's theorem to derive the key inequality of the method presented in [13]. The requisite theory of continuous-time stochastic processes may be found in standard texts such as [9], and that of controlled Markov processes in [10]. All stochastic integrals in this paper are of Itô type.

2.1. Changing between Equilibrium and Nonequilibrium Dynamics. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ be the filtered probability space of continuous, $\mathbb{R}^d$-valued paths, i.e., $\Omega = C([0,\infty), \mathbb{R}^d)$, and B be a d-dimensional Brownian motion with respect to P on this probability space. Let $V : \mathbb{R}^d \to \mathbb{R}$ be a differentiable potential and ε be a noise scaling factor which is small relative to the differences between the maxima and minima of the potential V. The object of interest is an ergodic, Markovian, d-dimensional dynamical system X having continuous paths, which satisfies the stochastic differential equation

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\varepsilon}dB_t, \quad t \in [0,T], \tag{2.1}$$

$$X_0 = x \in \mathbb{R}^d, \tag{2.2}$$

for an arbitrary $T \in (0,\infty)$. We define (2.1) as the *equilibrium dynamics* of the dynamical system. The equilibrium dynamics yield a map $X : \Omega \to \Omega$, where for every $\omega \in \Omega$, the function $X(\omega) : [0,T] \to \mathbb{R}^d$ satisfies

$$X_t(\omega) = X_0 + \int_0^t -\nabla V(X_s(\omega))ds + \sqrt{2\varepsilon}B_t(\omega), \ t \in [0,T].$$

We assume that the solution X of (2.1), (2.2) is pathwise unique. Recall that for a general stochastic differential equation with measurable b and Σ

$$dX_t = b(t, X_t)dt + \Sigma(X_t)dB_t, \tag{2.3}$$

there exists a one-to-one correspondence between distributions of solutions of (2.3) and solutions of the *martingale problem* $MP(b, \Sigma\Sigma^\top)$, where solutions of $MP(b, \Sigma\Sigma^\top)$

are defined to be probability measures on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$. In particular, for the special case (2.1), we can think of the solutions to the martingale problems as probability measures $\mu^x := P^x \circ X^{-1}$ that solve the martingale problem $MP(-\nabla V, 2\varepsilon)$ subject to the initial conditions $X_0 = x$ for any given $x \in \mathbb{R}^d$, where the superscript on a measure denotes conditioning on the initial condition. By the assumption that (2.1)–(2.2) has a unique strong solution, it holds that μ^x is the unique solution to $MP(-\nabla V, 2\varepsilon)$ (see Theorem 5.4.1 in [9]). Consistent with our definition of (2.1) as the equilibrium dynamics, we define μ^x to be the equilibrium measure of X.

Let $W : \Omega \to \mathbb{R}$ be a measurable random variable with the property that the scaled, conditional cumulant generating function

$$F(\sigma; x) := -\sigma^{-1} \log E^x[\exp(-\sigma W)] \tag{2.4}$$

is well-defined and finite for all $\sigma > 0$ and $x \in \mathbb{R}^d$. In this paper, we shall refer to W as 'work' and F as the 'free energy'. We consider a method for estimating the expected value of the equilibrium work done by X,

$$E^x[W] := \int_\Omega W \, d\mu^x = \int_\Omega W(X(\omega)) P^x(d\omega). \tag{2.5}$$

To estimate $E^x[W]$, we shall use another dynamical system Y, whose evolution is given by the *non-equilibrium* dynamics

$$dY_t = c(t, Y_t)dt - \nabla V(Y_t)dt + \sqrt{2\varepsilon} dB_t, \quad t \in [0, T], \tag{2.6}$$

where c is a bounded, measurable, nonvanishing control. We assume that (2.6) has a unique strong solution for every given initial condition $Y_0 = x$, so the non-equilibrium measure $\nu^x := P^x \circ Y^{-1}$ uniquely solves the martingale problem $MP(c - \nabla V, 2\varepsilon)$ for every $x \in \mathbb{R}^d$. Note that we omit the explicit dependence of Y and ν on c.

The choice of the control c is often motivated by the need to sample events with small μ-probability, i.e., by importance sampling. In importance sampling problems, there is a set $A \in \mathcal{F}$ such that $P^x(X \in A) \ll 1$ which one wishes to sample more frequently. In this case, c should be such that $P^x(Y \in A) \gg P^x(X \in A)$.

If $c \neq 0$ and $x = X_0 = Y_0$, then given $\omega \in \Omega$, the functions $X(\omega)$ and $Y(\omega)$ differ. Hence, $\nu^x \neq \mu^x$, so the equilibrium and nonequilibrium expectations of work differ:

$$E^x[W] \neq E^x_\nu[W] := \int_\Omega W \, d\nu^x = \int_\Omega W(Y(\omega)) P^x(d\omega).$$

Thus, we cannot use nonequilibrium measurements of W to estimate the equilibrium expectation of W. However, if μ is absolutely continuous with respect to ν, and if we can sample values of the Radon-Nikodym derivative $d\mu/d\nu$, then by

$$E^x[W] = E^x_\nu\left[\frac{d\mu}{d\nu} W\right], \tag{2.7}$$

we may obtain an unbiased estimate of the expected value of $W(X)$ using sample means of the product of $d\mu/d\nu$ with $W(Y)$.

Change of measure based on Girsanov's theorem. We now explain how to sample values of $d\mu/d\nu$. By our assumption of boundedness of c, we have (see Theorem 5.5.1 in [9]) that there exists a probability measure Q on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t)$ which is locally

equivalent to P, such that $Q^x \circ Y^{-1}$ solves the martingale problem $MP(-\nabla V, 2\varepsilon)$. Since $MP(-\nabla V, 2\varepsilon)$ is uniquely solved by μ, it holds that $\mu^x = Q^x \circ Y^{-1}$. Thus

$$D_t := \left(\frac{d\mu}{d\nu}\right)_t = \left.\frac{dQ}{dP}\right|_{\mathcal{F}_t}(Y). \tag{2.8}$$

By the Cameron-Martin-Girsanov formula for dQ/dP,

$$D_t = \exp\left(-\frac{1}{\sqrt{2\varepsilon}}\int_0^t c(s,Y_s)dB_s - \frac{1}{4\varepsilon}\int_0^t |c(s,Y_s)|^2 ds\right), \tag{2.9}$$

or $D_t = \exp(-M_t - A_t/2)$, where we define

$$M_t := \frac{1}{\sqrt{2\varepsilon}}\int_0^t c(s,Y_s)dBs, \tag{2.10}$$

$$A_t := \frac{1}{2\varepsilon}\int_0^t |c(s,Y_s)|^2 ds. \tag{2.11}$$

Recall that M and D are continuous local martingales with respect to P, that D is strictly positive due to the boundedness of c, and that A is the quadratic variation of $-M$. We omit the dependence of D, M and A on c.

Remark. Changing the minus sign preceding the Itô integral in (2.9) to a plus sign yields a measure $\overline{Q} \circ Y^{-1}$ that differs from the desired equilibrium measure μ. In Section §3.1 we show that changing the sign also affects the gradient descent dynamics.

THEOREM 2.1. *Let $\sigma > 0$, $x \in \mathbb{R}^d$, and $F(\sigma; x)$ be as in (2.4). Then*

$$F(\sigma;x) \le E_P^x\left[W(Y) + \frac{1}{4\varepsilon\sigma}\int_0^T |c(s,Y_s)|^2 ds\right]. \tag{2.12}$$

Proof. By using (2.7)–(2.9) and Jensen's inequality, we have

$$\begin{aligned}
-\sigma^{-1}\log E^x\left[\exp(-\sigma W)\right] &= -\sigma^{-1}\log E_\nu^x\left[D\exp(-\sigma W)\right] \\
&= -\sigma^{-1}\log E_\nu^x\left[\exp\left(-\sigma\left(W + \sigma^{-1}M_T + (2\sigma)^{-1}A_T\right)\right)\right] \\
&\le E_\nu^x\left[W + \sigma^{-1}M_T + (2\sigma)^{-1}A_T\right]
\end{aligned}$$

The conclusion follows, using (2.11) and the martingale property of M. ∎

From the proof of Theorem 2.1, we obtain two random variables which differ solely by the martingale M, which we shall write as

$$K(\alpha,\sigma) := W + (2\sigma)^{-1}A_T + \alpha\sigma^{-1}M_T, \quad \alpha \in \{0,1\}. \tag{2.13}$$

Note that equality holds in Jensen's inequality and thus in (2.12) if and only if $W + \sigma^{-1}M_T + (2\sigma)^{-1}A_T$ is ν^x-almost surely constant, since $-\sigma\log(\cdot)$ is strictly convex. In particular, $K(\alpha,\sigma)$ is in general a biased estimator of $F(\sigma;x)$ for any x.

In [13], the authors proposed a nonequilibrium estimator that equals $K(0,\varepsilon^{-1})$ (up to scaling of c), and remarked that (2.12) provides an example of Legendre-type duality relationships [2, 5, 12] between free energy and relative entropy,

$$F(\sigma;x) = \inf_{\nu \ll \mu}\left\{E_\nu^x[W] + \sigma^{-1}KL(\nu^x|\mu^x)\right\}, \tag{2.14}$$

where the relative entropy or Kullback-Leibler divergence [17] is

$$KL(\nu^x|\mu^x) := \begin{cases} E_\nu^x\left[\log\left(\frac{d\nu}{d\mu}\right)\right] & \nu \ll \mu \\ \infty & \text{otherwise.} \end{cases} \tag{2.15}$$

Remark. In statistical physics, the duality (2.14) between entropy and free energy appears in the definition of the canonical ensemble: If $W\exp(-\sigma W) \in L^1(\mu)$, then the measure ν^* for which the infimum in (2.14) is attained satisfies (cf. [5])

$$\left(\frac{d\nu^x}{d\mu^x}\right)^* = \frac{\exp(-\sigma W)}{E^x[\exp(-\sigma W)]} = \exp(-\sigma(W - F(\sigma; x))). \tag{2.16}$$

The rightmost expression can be interpreted as the Boltzmann distribution at inverse temperature $\sigma > 0$. Closely related is Jarzynski's equality [14] that states that $\exp(-\beta(W - \Delta F))$ is a probability density with respect to the probability measure of an externally driven Hamiltonian system when W is the nonequilibrium work done on the system, β is the inverse temperature, and $\Delta F = F_B - F_A$ is the free energy difference between two states A and B of the system in equilibrium.

We remark that by Theorem 2.1, the change of measure defined by (2.16) is optimal in that a Monte-Carlo estimator of the free energy F based on the optimal change of measure will have zero variance, i.e., the estimator will converge to the free energy value in one step (cf. [21]). However, this observation is of limited practical use because the normalization constant that appears in the Radon-Nikodym derivative in (2.16) is exactly the quantity that one wants to estimate, namely F.

2.2. Stochastic optimal control. We briefly review the stochastic optimal control problem presented in [12, 13]. We specify the random variable W in (2.5) as

$$W(\omega) = \int_0^\tau f(\omega_s)ds + g(\omega_\tau) =: \int_0^\tau f_s ds + g_\tau, \tag{2.17}$$

where $f, g \geq 0$ are nonnegative and bounded continuous functions, $f_s := f(\omega_s)$, $g_\tau := g(\omega_\tau)$, and τ is an almost surely finite random stopping time,

$$\tau = \inf\{t > 0 \mid X_t \notin O\}, \tag{2.18}$$

for bounded, open $O \subset \mathbb{R}^d$ with smooth boundary. Let $\psi_\sigma(x) := E^x[\exp(-\sigma W)]$ be the 'moment-generating function' of W, considered as a function of the initial data $x \in \overline{O}$. For any $w \in C^2(\mathbb{R}^d, \mathbb{R})$ define $\mathcal{A}$ to be the linear operator

$$\mathcal{A}w = \sum_{j=1}^d \varepsilon \frac{\partial^2 w}{\partial x_j^2} - \frac{\partial V}{\partial x_j}\frac{\partial w}{\partial x_j}. \tag{2.19}$$

Then by the Feynman-Kac theorem [19, pp. 201], ψ_σ satisfies the linear equation

$$(\mathcal{A} - \sigma f)\,\psi_\sigma(x) = 0\,, \quad x \in O \tag{2.20}$$

with boundary condition

$$\psi_\sigma(x) = \exp\left(-\sigma g(x)\right)\,, \quad x \in \partial O\,. \tag{2.21}$$

For smooth potentials $V \in C^\infty(\mathbb{R}^d, \mathbb{R})$ and sufficiently small $\sigma > 0$, the strong maximum principle for elliptic partial differential equations implies that (2.20)–(2.21) has a classical solution $\psi_\sigma \in C^2(O, \mathbb{R}) \cap C(\overline{O}, \mathbb{R})$. Moreover ψ_σ is bounded away from zero, which implies that the free energy $F(\sigma, x) = -\sigma^{-1}\log\psi_\sigma(x)$ is a bounded and smooth function of x whenever $\sigma > 0$ is sufficiently small.

Solving a partial differential equation to compute the free energy is clearly not recommended beyond one- or two-dimensional toy systems. An alternative, variational characterization of F that has been studied in [12, 13] is in terms of the value function of a stochastic control problem: minimise the cost functional

$$\mathfrak{I}(u) = E_P\left[\int_0^\tau f(Y_s) + \frac{1}{4\varepsilon\sigma}|u_s|^2 ds + g(Y_\tau)\right] \tag{2.22}$$

over the set U of admissible controls $u : [0,\infty) \to \mathbb{R}^d$, subject to

$$dY_t = (u_t - \nabla V(Y_t))\, dt + \sqrt{2\varepsilon} dB_t. \tag{2.23}$$

We specify U as the set of bounded controls that are adapted to the filtration generated by the Brownian motion, such that (2.23) admits a unique strong solution for every initial condition $Y_0 = x$. Note that solving the optimal control problem gives the optimal control c^* for which equality holds in (2.12).

We define $\mathfrak{I}(u;x)$ as the version of $\mathfrak{I}(u)$ conditioned on $Y_0 = x$, and call

$$v_\sigma(x) = \min_{u\in U} \mathfrak{I}(u;x)\,, \quad x \in \overline{O} \tag{2.24}$$

the 'value function' or 'optimal cost-to-go' of the optimal control problem (2.22)–(2.23). It follows from the dynamic programming principle of stochastic control [10] that the optimal controls $u^* = \operatorname{argmin}\mathfrak{I}(u)$ are Markovian feedback laws of the form

$$u_t^* = c^*(Y_t) \tag{2.25}$$

where $c^* \colon \mathbb{R}^d \to \mathbb{R}^d$ is a bounded and measurable function that depends on the control problem, through F. Since $\psi_\sigma(x) > 0$, we can rewrite (2.20) and (2.21) as

$$f(x) + \mathcal{A}F(\sigma;x) - \varepsilon\sigma|\nabla F(\sigma;x)|^2 = 0 \qquad x \in O, \tag{2.26}$$

$$F(\sigma;x) = g(x) \quad x \in \partial O. \tag{2.27}$$

Using that

$$-\varepsilon\sigma|\nabla F(\sigma;x)|^2 = \min_{c(x)\in\mathbb{R}^d}\left\{c(x)\cdot\nabla F(\sigma;x) + \frac{1}{4\varepsilon\sigma}\,|c(x)|^2\right\}, \tag{2.28}$$

it follows that (2.26) is equivalent to

$$\min_{c(x)\in\mathbb{R}^d}\left\{\mathcal{A}F(\sigma;x) - c(x)\cdot\nabla F(\sigma;x) + f(x) + \frac{1}{4\varepsilon\sigma}|c(x)|^2\right\} = 0, \quad x \in O. \tag{2.29}$$

Thus, (2.29) is the Hamilton-Jacobi-Bellman equation of the optimal control problem defined by (2.22) and (2.23), and the unique optimal control which solves this problem is the argument that minimises the quantity in the curly braces in (2.28),

$$c^*(x) = -2\varepsilon\sigma\nabla F(\sigma;x) = 2\varepsilon\frac{\nabla\psi_\sigma(x)}{\psi_\sigma(x)}\,. \tag{2.30}$$

Thus, by substituting the definition of c^* in (2.30) into the nonequilibrium dynamics (2.6), we find that the optimal nonequilibrium dynamics are of gradient form:

$$dY_t = -\nabla[2\varepsilon\sigma F + V](Y_t)dt + \sqrt{2\varepsilon}dB_t. \tag{2.31}$$

By definition of the value function and (2.17), it follows from Theorem 2.1 that $F(\sigma;x) = v_\sigma(x)$, and hence $c^*(x) = -2\varepsilon\sigma\nabla v_\sigma(x)$. Thus, one can obtain an approximation of the free energy F by quadrature if one has an approximation of the optimal control, and the optimal control by differentiation if one has the free energy. Note that, since $\psi_\sigma \in C^2(O,\mathbb{R})$, the optimal control c^* is bounded in O.

2.3. Convergence results. Recall that for every control c, equations (2.8)–(2.9) give the corresponding change of measure $D = d\mu/d\nu$, where $\nu = P \circ Y^{-1}$. Note that we suppress the explicit dependence of Y, and hence of ν, on c. As c approaches the optimal control c^*, we expect that ν converges in some sense to the optimal change of measure ν^*, i.e. the unique minimizer of (2.12) as given by (2.16).

THEOREM 2.2. *Let $x \in \overline{O} \subset \mathbb{R}^d$. For an arbitrary bounded control c, let D be the Radon-Nikodym derivative given by (2.9) and ν be the corresponding nonequilibrium measure. When $c = c^*$, then we write D^* and ν^*, where the optimal control c^* is given by (2.30). Then as $\|c - c^*\|_\infty \to 0$,*

(i) D_τ converges to D^*_τ P^x-almost surely, and

(ii) when restricted to $\mathcal{F}_t$ for any finite $t \geq 0$, the nonequilibrium measure ν converges to (ν^*) in the total variation norm, i.e.,

$$TV(\nu, \nu^*)|_t := \sup\{|\nu(A) - \nu^*(A)|,\ A \in \mathcal{F}_t\} \to 0. \tag{2.32}$$

Proof. The uniqueness of c^* follows from standard arguments and the specific form of the stochastic control problem that is 'linear-quadratic' in the control variables [13]; cf. [10, Sec. IV.2]. Moreover, the assumptions on the potential V and the boundary ∂O of the set $O \subset \mathbb{R}^d$ entail that the value function is differentiable almost everywhere. Therefore the optimal feedback control c^* is essentially bounded, and it suffices to consider controls of the form $c = c^* + \delta$ for $\delta > 0$. By (2.8) and (2.9), we obtain

$$\left(\frac{D}{D^*}\right)_\tau = \left(\frac{d\nu^*}{d\nu}\right)_\tau = \exp\left(-\frac{\delta B_\tau}{\sqrt{2\varepsilon}} - \frac{\delta}{2\varepsilon}\int_0^\tau c^*(Y_s)ds - \frac{\delta^2}{4\varepsilon}\right). \tag{2.33}$$

To prove part (i), observe that the term in the exponenti77al satisfies

$$\left|\frac{\delta B_\tau}{\sqrt{2\varepsilon}} + \frac{\delta}{2\varepsilon}\int_0^\tau c^*(Y_s)ds + \frac{\delta^2}{4\varepsilon}\right| \leq \delta\left(\frac{|B_\tau|}{\sqrt{2\varepsilon}} + \frac{\tau\|c^*\|_\infty}{2\varepsilon} + \frac{\delta}{4\varepsilon}\right). \tag{2.34}$$

We only need to consider the first two of the three terms on the right-hand side of (2.34). Since $B_\tau : \Omega \to \mathbb{R}^d$ is a Gaussian random variable, it holds that P^x-almost surely, $|B_\tau| < \infty$, and hence $\delta B_\tau \to 0$ as $\delta \to 0$. By our assumption in Section §2.2 that $\tau < \infty$ P^x-almost surely, it also holds that $\delta\tau \to 0$. By (2.34),

$$P^x\left(\lim_{\delta\to 0}\left|\frac{\delta B_\tau}{\sqrt{2\varepsilon}} + \frac{\delta}{2\varepsilon}\int_0^\tau c^*(Y_s)ds + \frac{\delta^2}{4\varepsilon}\right| = 0\right) = 1,$$

which, with (2.33), implies that D_τ converges to D^*_τ P^x-almost surely.

To prove part (ii), observe that, by (2.33) and (2.15), the Kullback-Leibler divergence of $\nu^x|_{\mathcal{F}_t}$ with respect to $(\nu^*)^x|_{\mathcal{F}_t}$ satisfies

$$KL(\nu|\nu^*) \leq \delta\left(\frac{t\|c^*\|_\infty}{2\varepsilon} + \frac{\delta}{4\varepsilon}\right). \tag{2.35}$$

Applying the Csiszar-Kullback-Pinsker inequality [4, 16, 22] to $\nu^x|_{\mathcal{F}_t}$ and $(\nu^*)^x|_{\mathcal{F}_t}$,

$$TV(\nu^x, (\nu^*)^x)|_t \leq \sqrt{2KL(\nu^x|(\nu^*)^x)|_t}.$$

Letting $\delta \to 0$, the conclusion follows. ∎

Remark. Note that we can replace t in Theorem 2.2(ii) with the bounded stopping time $\min\{\tau, T\}$, since then we can replace t with T in the right-hand side of (2.35). However, we cannot then let $T \to \infty$ (i.e., we cannot set $t = \tau$), since our assumption that τ is almost surely finite does not guarantee that $E^x_\nu[\tau]$ is finite.

The importance of Theorem 2.2 is that the nonequilibrium measures depend in a continuous way on the control. In particular, for any bounded, measurable random variable $\phi : \Omega \to \mathbb{R}$ defined on the filtered probability space, Theorem 2.2(ii) implies that the statistics of ϕ with respect to ν^x will converge to the statistics of ϕ with respect to $(\nu^*)^x$. Unfortunately, since neither $K(\alpha, \sigma)$ nor its variance are in general bounded, we cannot expect that the variance of $K(\alpha, \sigma)$ will go to zero as $\|c - c^*\|_\infty \to 0$, even though Theorem 2.1 implies that the variance of $K(\alpha, \sigma)$ is zero when $c = c^*$.

2.4. Reweighted estimation. Recall that we seek to estimate the equilibrium average $E^x[W]$ (2.5) of the path functional W defined in (2.17). In order to estimate $E^x[W]$ from $F(\sigma; x)$, one may use the property of cumulant generating functions, i.e.,

$$\left.\frac{d}{d\sigma}\right|_{\sigma=0} \sigma F(\sigma; x) = E^x[W]. \tag{2.36}$$

Or, one may use the Taylor series approximation of $\log(1 + r)$ for $|r| < 1$ to conclude that for sufficiently small σ,

$$F(\sigma; x) \approx E^x[W] - \frac{\sigma}{2}\mathrm{var}^x(W), \tag{2.37}$$

assuming that the variance $\mathrm{var}^x(W)$ of W with respect to P^x is finite. The method which we consider in this study for estimating equillibrium expectations from free energy values directly benefits from the addition of the martingale term in the nonequilibrium estimator. This procedure was applied to committor probabilities (see Section §4 below) in [12].

Observe that, by (2.13), the martingale-based estimator $K(\alpha = 1, \sigma)$ satisfies $\exp(-\sigma K(1, \sigma)) = \exp(-\sigma W)D$, where $D = d\mu/d\nu$. By (2.7), we expect that $\exp(-\sigma K(1, \sigma))$ is an unbiased estimator of $E^x[\exp(-\sigma W)]$. Suppose we seek to estimate the μ^x-expectation of an integrable path functional $\mathfrak{W}$, that is bounded from below and of the form (2.17). We may assume that $\mathfrak{W} > 0$. Define

$$W := -\sigma^{-1} \log(\mathfrak{W}), \tag{2.38}$$

so that the free energy function $F(\sigma; x)$ of W satisfies

$$E^x[\mathfrak{W}] = E^x[\exp(-\sigma W)] = \exp(-\sigma F(\sigma; x)). \tag{2.39}$$

Thus, we can estimate $E^x[\mathfrak{W}]$ by using sample means of $\exp(-\sigma K(1, \sigma))$. This method requires only one estimate of the free energy function to obtain an estimate of $E^x[\mathfrak{W}]$. However, the method applies only to functionals $\mathfrak{W}$ which are bounded from below. By Jensen's inequality and the strict concavity of the logarithm, the estimates of F that are obtained from unbiased estimates of $E^x[\mathfrak{W}]$ will be biased in general.

Remark. It is tempting to estimate $E^x[W]$ using the identity $E^x[W] = E^x_{\nu^*}[D^*W]$, where the optimal change of measure D^* for the cumulant generating function $F(\sigma; x)$ is available. It turns out, however, that when $W \geq 0$, this leads to estimators that perform worse than the equilbrium estimator for $E^x[W]$ in that they have larger

variance. The property of unbiasedness permits comparison of the second moments of the equilibrium and nonequilibrium estimators:

$$E^x_{\nu^*}\left[(D^*W)^2\right] = E^x\left[D^*W^2\right] \quad \text{versus} \quad E^x[W^2]\,.$$

It readily follows that the second moment - and hence the variance of the nonequilibrium estimator for the first moment - is reduced if and only if $E^x[D^*W^2] < E^x[W^2]$. If $W \geq 0$, then by (2.16), we observe that

$$E^x\left[D^*W^2\right] \geq E^x[D^*]\,E^x[W^2] \geq E^x[W^2]\,,$$

since W^2 and $D^* = \exp(\sigma(W - F))$ are positively correlated, and since $E^x[D^*] \geq 1$ by Jensen's inequality. Thus, the variance of the nonequilibrium estimator of the first moment will be larger than the variance of the equilibrium estimator, even though D^* admits a zero variance estimate of the cumulant generating function. In Section §4 we explore the issue of variance reduction for reweighted estimates of $E^x[\exp(-\sigma W)]$.

3. Discretisation. We follow the same numerical discretisation procedure as described in [13]. Let $(b_i)_{1\leq i\leq n}$ be a set of linearly independent, sufficiently weakly differentiable basis functions from $\mathbb{R}^d$ to $\mathbb{R}$. We project the free energy function F and control to the span of the $(b_i)_i$, so that for some constant C, and using (2.30),

$$F(\sigma; x, a) := -\frac{1}{2\varepsilon\sigma}\sum_{i=1}^{n} a_i b_i(x) + C, \tag{3.1}$$

$$c(x; a) := \sum_{i=1}^{n} a_i \nabla b_i(x). \tag{3.2}$$

In order to solve the stochastic optimal control problem of minimising (2.22) subject to the nonequilibrium dynamics (2.23), we need to find coefficients $a = (a_i)_i$ which give the projection of the optimal control c^* to the linear approximation space.

Since the optimal control is of feedback form (2.30), we replace the measurable control u_t in (2.23) accordingly. Euler-Maruyama discretisation then yields

$$\widehat{Y}_{k+1} - \widehat{Y}_k = (c(\widehat{Y}_k; a) - \nabla V(\widehat{Y}_k))\Delta t + \sqrt{2\varepsilon\Delta t}\eta_{k+1}, \quad k = 0, 1, 2, \ldots, \tag{3.3}$$

where the $(\eta_k)_k$ are i.i.d. centred Gaussian random variables in $\mathbb{R}^d$, $t_{k+1} = t_k + \Delta t$ for $t_0 = 0$, and Δt is the fixed time step of integration. With the discretisations of the random variables τ, W, M, and A given in Section §2.2,

$$\widehat{\tau} := \inf\left\{k \in \mathbb{N} \mid \widehat{Y}_k \notin O\right\}, \tag{3.4}$$

$$\widehat{W}(\widehat{Y}) := \sum_{k=0}^{\widehat{\tau}-1} f\left(\widehat{Y}_k\right)\Delta t + g\left(\widehat{Y}_{\widehat{\tau}}\right), \tag{3.5}$$

$$\widehat{M}_{\widehat{\tau}}(a) := \frac{1}{\sqrt{2\varepsilon}}\sum_{k=0}^{\widehat{\tau}-1} c\left(\widehat{Y}_k; a\right)\sqrt{\Delta t}\eta_{k+1}, \tag{3.6}$$

$$\widehat{A}_{\widehat{\tau}}(a) := \frac{1}{2\varepsilon}\sum_{k=0}^{\widehat{\tau}-1}\left|c\left(\widehat{Y}_k; a\right)\right|^2\Delta t. \tag{3.7}$$

we obtain the discretisations $\widehat{D}$ and $\widehat{K}(\alpha, \sigma; a)$ by replacing W, M and A with their discrete counterparts in D (2.9) and $K(\alpha, \sigma)$ (2.13).

The discretised stochastic control problem involves finding a control $c(\cdot, a^*)$ which minimises the objective function $\widehat{K}(\alpha, \sigma; a)$ subject to the dynamics (3.3), i.e., of finding optimal coefficients in (3.2). In the gradient descent algorithm of [13], each a_i in (3.2) changes at iteration ℓ from its current value a_i^ℓ to $a_i^{\ell+1}$ by step size Δa_i^ℓ:

$$a_i^{\ell+1} = a_i^\ell - \frac{\partial}{\partial a_i} E_\nu^x \left[\widehat{K}(\alpha, \sigma; a) \right]^\ell \Delta a_i^\ell, \tag{3.8}$$

$$\frac{\partial}{\partial a_i} E_\nu^x \left[\widehat{K}(\alpha, \sigma; a) \right] = E_\nu^x \left[\frac{\partial}{\partial a_i} \widehat{K}(\alpha, \sigma; a) \right] - \mathrm{cov}_\nu^x \left(\widehat{K}(\alpha, \sigma; a), \frac{\partial \widehat{S}}{\partial a_i} \right). \tag{3.9}$$

The derivatives of the Gaussian path density $\widehat{S}$ and $\widehat{K}(\alpha, \sigma; a)$ are

$$\frac{\partial \widehat{S}}{\partial a_i} = -\sqrt{\frac{\Delta t}{\varepsilon}} \sum_k^{\widehat{\tau}-1} \nabla b_i(\widehat{Y}_k) \cdot \eta_{k+1}, \tag{3.10}$$

$$\frac{\partial \widehat{K}(\alpha, \sigma; a)}{\partial a_i} = \frac{1}{2\sigma} \frac{\Delta t}{\varepsilon} \sum_{k=0}^{\widehat{\tau}-1} c(\widehat{Y}_k; a) \cdot \nabla b_i \left(\widehat{Y}_k \right) + \frac{\alpha}{\sigma} \sqrt{\frac{\Delta t}{2\varepsilon}} \sum_{k=0}^{\widehat{\tau}-1} \eta_{k+1} \cdot \nabla b_i \left(\widehat{Y}_k \right). \tag{3.11}$$

We emphasise that, in the partial derivatives (3.9), (3.10), and (3.11), we have made the approximation that $\widehat{\tau}$ and $\widehat{W}$ are independent of a. This can introduce significant errors in the gradient descent dynamics, as we shall see in Section §4.3.

3.1. Role of the martingale term. Observe that (3.10) and (3.11) give the discrete-time approximations of the following stochastic processes at time $t = \tau$:

$$S_t^i := -\frac{1}{\sqrt{\varepsilon}} \int_0^t \nabla b_i(Y_s) dB_s \tag{3.12}$$

$$K_t^i(\alpha, \sigma; a) := \frac{1}{2\sigma} \frac{1}{\varepsilon} \int_0^t c(Y_s; a) \cdot \nabla b_i(Y_s) ds + \frac{\alpha}{\sigma} \sqrt{\frac{1}{2\varepsilon}} \int_0^t \nabla b_i(Y_s) dB_s. \tag{3.13}$$

Since S^i is an Itô integral with respect to Brownian motion, it is a martingale. Recall that, given two continuous-time stochastic processes X and X',

$$\mathrm{cov}_\nu^x(X_t, X_t') = E_\nu^x \left[\langle X, X' \rangle_t \right], \quad t \geq 0.$$

If $X = M + N$ and $X' = M' + N'$ are real-valued continuous semimartingales, where M and M' are continuous local martingales and N and N' are processes of locally bounded variation, then the covariance process $\langle X, X' \rangle_t$ is defined by

$$\langle M + N, M' + N' \rangle_t = \langle M, M' \rangle_t := \frac{1}{4} \left(\langle M + M' \rangle_t - \langle M - M' \rangle_t \right). \tag{3.14}$$

In particular, processes of locally bounded variation do not contribute to the covariance. Since A is the quadratic variation of M, it is increasing and hence locally of bounded variation. Consider the work process defined by

$$W_t := W_t^f + W_t^g := \int_0^t f_s ds + g_\tau 1_{[\tau, \infty)}(t), \tag{3.15}$$

and note that W_τ defined by (3.15) equals the random variable W defined in (2.17). Under the assumptions that $f, g \geq 0$ are continuous and bounded, W^f is an increasing

process with continuous paths, and hence is locally of bounded variation, and W^g is locally of bounded variation since W^g has a single jump-type discontinuity at $t = \tau$. Therefore $(W_t)_t$ is locally of bounded variation. The conclusion holds if we weaken the conditions on f and g and require that f has countably many jump discontinuities and that both f and g are finite. Thus, we have

$$\mathrm{cov}_\nu^x(K(\alpha,\sigma;a), S_\tau^i) = E_\nu^x\left[\langle K(\alpha,\sigma;a), S^i\rangle_\tau\right] = \alpha\sigma^{-1}E_\nu^x\left[\langle M, S^i\rangle_\tau\right],$$

where $K(\alpha,\sigma;a)$ is the limit of $\widehat{K}(\alpha,\sigma;a)$ as $\Delta t \to 0$,

$$K_t(\alpha,\sigma;a)(\omega) := W_t(\omega) + \frac{1}{4\epsilon\sigma}\int_0^t |c(\omega_s;a)|^2 ds + \frac{\alpha}{\sigma\sqrt{2\varepsilon}}\int_0^t c(\omega_s;a)dB_s. \tag{3.16}$$

By definition of M,

$$\left\langle M, S^i\right\rangle_t = -\frac{1}{\varepsilon\sqrt{2}}\int_0^t c(Y_s;a)\cdot\nabla b_i(Y_s)ds.$$

Thus, the preceding analysis shows that in the limit as $\Delta t \to 0$, (3.9) becomes

$$\frac{\alpha\sqrt{2}+1}{2\sigma\varepsilon}E_P^x\left[\int_0^\tau c(Y_s;a)\cdot\nabla b_i(Y_s)ds\right]. \tag{3.17}$$

If $\alpha = 1$, then the martingale term $\widehat{M}$ is included in the nonequilibrium estimator $\widehat{K}$ of the free energy, and the resulting martingale-based gradient descent dynamics given by (3.11) will differ from the dynamics for $\alpha = 0$. We also expect that changing the minus sign preceding the Itô integral in (2.9) to a plus sign yields a different change of measure, since this implies $\alpha = -1$.

The expression (3.17) is important for two reasons. First, it shows that the value of the non-covariance term $E_\nu^x[\partial/\partial a_i \widehat{K}(\alpha,\sigma;a)]$ in (3.9) will increase by a factor of $1+\sqrt{2}$ if $\alpha = 1$, since $\sqrt{2} > 1$. Therefore, for any i, the contribution to (3.13) from a sample path of $\widehat{Y}$ is zero when the control $c(\cdot;a)$ is smaller by a factor of $1+\sqrt{2}$, if $\alpha = 1$, so the martingale-based gradient will not produce the correct free energy landscape. To obtain the correct free energy landscape, we only need to multiply the coefficients a^ℓ obtained from martingale-based gradient descent by $1+\sqrt{2}$. We shall use this correction procedure in Section §4.2. Second, (3.17) shows that the gradient given by (3.9) is not optimal for minimising the discretisation of the cost functional in (2.22), because in the limit of small Δt, the work W does not appear in the gradient. A better approximation of the gradient should take into account the dependence of the stopping time τ on the projected control $c(\cdot;a)$. In Section §4.2, we present results which suggest that gradient descent using the approximate gradient given by (3.9) can still yield reasonable estimates of the free energy value and the free energy surface.

4. Illustrative example. The equilibrium average we wish to estimate is the equilibrium probability of commitment to B, or simply the B-committor,

$$p_B(x) := E_P^x\left[\mathbf{1}_B(X_{\tau(X)}\right] \tag{4.1}$$

where we recall that X denotes a random solution of (2.1), and the set O in the definition of the stopping time (2.18) is $(A\cup B)^\complement$ for disjoint, closed subsets $A, B \subset \mathbb{R}^d$. The value $p_B(x)$ gives the probability that the equillibrium dynamical system $X(\omega)$ which starts at $X_0 = x$ will first reach B before A. In the physics literature,

Onsager's study [20] of ion-pair dissociation is often cited as the first to consider committor probabilities in physics. The committor is an important quantity because it describes how likely a reaction will proceed to completion (i.e., to set B), given that the reactants are in a certain 'state' x. In practice, obtaining the entire committor function poses computational difficulties, because molecules have high-dimensional state spaces. However, individual values of $p_B(x)$ are still useful, because one may use committor values to determine whether a given observable can describe the progress of a reaction [6, 11].

Since $\mathbf{1}_B \geq 0$, we may apply the derivative-free estimation method presented in Section §2.4. For fixed $\lambda > 0$, the work

$$W(\omega) := -\sigma^{-1} \log(\mathbf{1}_B(\omega_\tau) + \lambda) \tag{4.2}$$

has the free energy $F(\sigma; x) = -\sigma^{-1} \log(p_B(x) + \lambda)$. Thus, $p_B + \lambda = \psi_\sigma$ in (2.20)–(2.21), with $f \equiv 0$ and $\exp(-\sigma g(x)) = \mathbf{1}_B(x) + \lambda$. Note that g and W can assume negative values in this example. By (2.30), the optimal control c^*

$$c^*(x) := \frac{2\nabla p_B(x)}{p_B(x) + \lambda}, \tag{4.3}$$

Thus, if $\lambda > 0$, then the optimal control c^* is bounded. The example of the committor in the context of derivative-free estimation was studied in [12].

The key element of our treatment is the parameter λ, the importance of which is immediate from (4.3): as λ decreases to zero, the magnitude of the optimal control grows arbitrarily large. This is undesirable, since the boundedness of the optimal control was crucial in the proof of convergence of measures in the total variation norm (cf. Theorem 2.2). On the other hand, if λ is large, then the optimal control c^* will be small, and there may be no significant advantage to controlling the system. Hence, λ should be small enough such that the optimal control is sufficiently different from zero, while keeping the forces suitably small for the system under study. For example, in simulations of single-molecule pulling experiments, one must keep the pulling forces sufficiently small for numerical stability of the integration scheme.

4.1. Parameters and reference solutions. We considered an artificial 1-dimensional landscape shown in Figure 1, which was obtained from projecting GROMACS [24] time series data of solvated alanine dipeptide at 300K down to a dihedral angle defined by four atoms on the backbone of the molecule.

Potential and domain. The landscape is defined on the circle $[-180, 180)$ and has two minima, the smaller of which is in $A := [10, 50]$ and the other in $B := [162, -158]$ (A and B are the projections of two metastable regions to the aforementioned dihedral angle). Since $(A \cup B)^{\complement}$ is just the disjoint union of two domains, we can consider each domain separately. In this illustrative example, we set the domain of interest to be $O := (50, 162)$. The difference between the minimum in A and the maximum in $(50, 162)$ is approximately four times the value of the noise temperature $\varepsilon = 2.5$. Note that, although all domain variables are expressed in terms of degrees, the numerical calculations were performed in radians.

Finite differences approximation. Using finite differences with the regular discretisation of the domain given by

$$x_l := 50 + (l-1)\Delta x, \quad 1 \leq l \leq (162 - 50)/\Delta x + 1, \quad \Delta x = 10^{-1}, \tag{4.4}$$

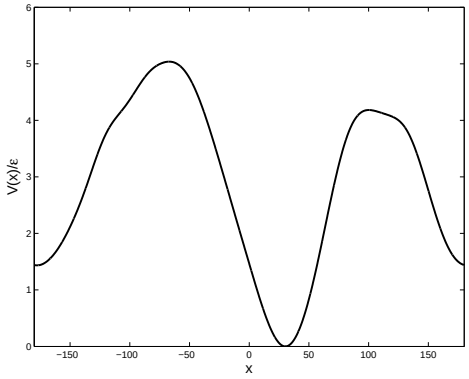

FIG. 1. *Potential V/ε over* $[-180, 180)$

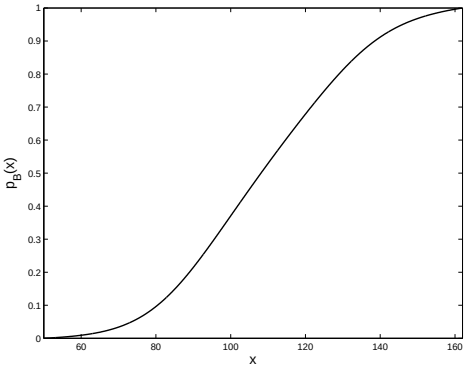

FIG. 2. *Committor p_B over* $[50, 162]$.

we computed the equilibrium committor, the free energy, and the optimal control by solving the boundary value problem (2.20), (2.21). We show the equilibrium committor over the domain of interest in Figure 2.

To measure the approximation quality of the free energy function at each iterate with respect to the finite-differences approximation, we computed $F(\sigma; \cdot, a^k)$ from the current value of the coefficients $(a_i^k)_i$ using (3.1) and set the value of C so that $F(\sigma; x = 50, a^k) := -\sigma^{-1} \log \lambda$. We measured the function approximation error with the uniformly weighted L^1 norm

$$\|F^k - F\|_1 := \sum_l \left| F(\sigma; x_l, a^k) - F(\sigma; x_l) \right| \Delta x. \tag{4.5}$$

The projection of the finite-differences approximation of the optimal control c^* to the linear subspace defined by the gradients of the basis functions is given by setting $a = a^*$ in (3.2), where the optimal values of the expansion coefficients a^* satisfy

$$Ga^* = c, \quad c_l = c^*(x_l), \quad G_{il} = \nabla b_i(x_l) \tag{4.6}$$

for x_l given in (4.4).

4.2. Sample implementation. We set $\sigma = 0.4 = \varepsilon^{-1}$ and $\lambda = 10^{-3}$ in (4.2). The difference $\rho(\lambda) := \sigma^{-1} \log(1 + \lambda^{-1})$ between the maximum and the minimum values of work W was approximately 17.3. We performed $N_{\text{iter}} = 600$ iterations of gradient descent. The coefficients were initialised at $a^0 = 0$. The descent step size $\Delta a_i = \Delta a = 5$ was constant for every iteration ℓ and coefficient a_i. For the basis functions in (3.2), we used $n = 7$ tent basis functions with equidistant centres:

$$b_i(x) := \begin{cases} h - |x - \mu_i| & 0 \le |x - \mu_i| \le h \\ 0 & \text{otherwise,} \end{cases} \tag{4.7}$$

$$\nabla b_i(x) := \begin{cases} -(x - \mu_i)(|x - \mu_i|)^{-1} & 0 < |x - \mu_i| < h \\ 0 & \text{otherwise,} \end{cases} \tag{4.8}$$

$$\mu_i := 50 + h(i - 1), \tag{4.9}$$

$$h = \frac{162 - 50}{n - 1}. \tag{4.10}$$

At each iteration, we sampled $N_{\text{traj}} = 10$ trajectories, all initialised at $\widehat{Y}_0 = x = 60$. Trajectories of (3.3) terminated when $\widehat{Y}_k \notin (50, 162)$. Since $p_B(x) \approx 0.0091$, the problem of estimating the free energy value is equivalent to the problem of estimating a rare event probability. With respect to P^x, the equilibrium estimator $\mathbf{1}_B(X_{\tau(X)})$ has the Bernoulli distribution with parameter $p = p_B(x)$, and hence its standard deviation $(p(1-p))^{1/2} \approx 0.0951$.

For a given vector v of length N_{iter}, we define the running average over the window of N_{win} iterations to be the vector r of length N_{iter} whose i-th element is the average of the $\min\{i, N_{\text{win}}\}$ elements of v up to and including v_i:

$$r_i := j^{-1}(v_{i-j+1} + \cdots + v_{i-1} + v_i), \quad j := \min\{i, N_{\text{win}}\}.$$

We chose $N_{\text{win}} = 10$ to elucidate trends in the sample statistics, while keeping the averaging of fluctuations over iterations small.

In Figure 3, the running average of the sample mean of $\widehat{K}(\alpha, \sigma; a)$ decreases in the first 300 iterations, with the running average decreasing by a larger amount when $\alpha = 0$. From the 301st iteration onwards, the running averages oscillate around values that are greater than $F(\sigma; x) \approx 11.4$. Thus, the gradient descent algorithm succeeds in minimising the objective function, but only to a limited extent. The running average of the sample mean of $\widehat{K}(\alpha, \sigma; a^\ell)$ better approximates the value of $F(\sigma; x)$ from the 150th to about the 220th iteration. The fluctuations in the running average tend to be smaller when $\alpha = 1$.

In Figure 4, the running average of the sample standard deviation of $\widehat{K}(\alpha, \sigma; a)$ increases in the first 200 iterations, with the running average increasing by a smaller amount when $\alpha = 1$. From the 201st iteration onwards, the running average oscillates between 3 and 4 (for $\alpha = 1$) and between 5 and 6 (for $\alpha = 0$). The standard deviation of $\widehat{K}(\alpha, \sigma; a)$ tends to be smaller when $\alpha = 1$.

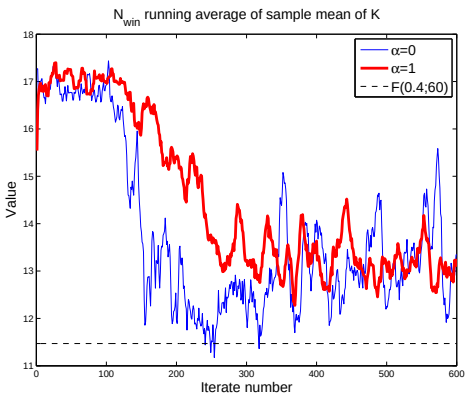

FIG. 3. *Sample mean of* $\widehat{K}(\alpha, \sigma; a^\ell)$

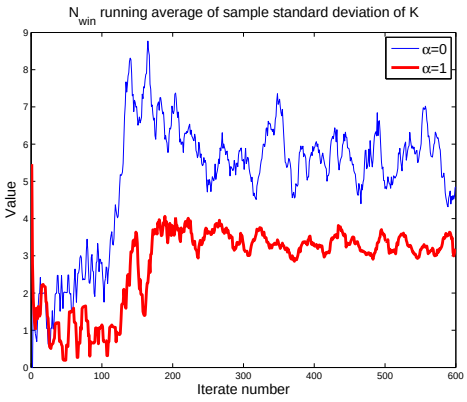

FIG. 4. *Sample std. dev. of* $\widehat{K}(\alpha, \sigma; a^\ell)$

In Figure 5, the projected control $c(x_l; a^{601})$ for $\alpha = 0$ is approximately $1 + \sqrt{2}$ times as large as the projected control for $\alpha = 1$. This observation is consistent with the analysis in Section §3.1. The projected control for $\alpha = 0$ appears closer to the finite-differences optimal control c^* when the latter is large, e.g., over (50,88). On the other hand, when the optimal control c^* is moderate, the projected control for $\alpha = 1$ appears closer.

In Figure 6, we plot the L^1 norm (4.5) with respect to the finite-differences free energy $F(\sigma;\cdot)$ of $F(\sigma;\cdot;a^\ell)$ (for $\alpha = 0$) and of $F(\sigma;\cdot,a^\ell(1+\sqrt{2}))$ (for $\alpha = 1$), as a function of ℓ. The graphs track each other reasonably well, and thus support the scaling relation described at the end of Section §3.1. The L^1 error remains close to the L^1 error of $F(\sigma;a^*)$ between the 200th and 300th iterations. On the other hand, the coefficients produced by gradient descent approach their (albeit incorrect) limiting values sooner with the martingale-based gradient descent: while the L^1 error exhibits an increasing trend from the 201st iteration onwards for $\alpha = 0$, the L^1 error begins to oscillate around a 'limiting' value within 300 iterations for $\alpha = 1$.

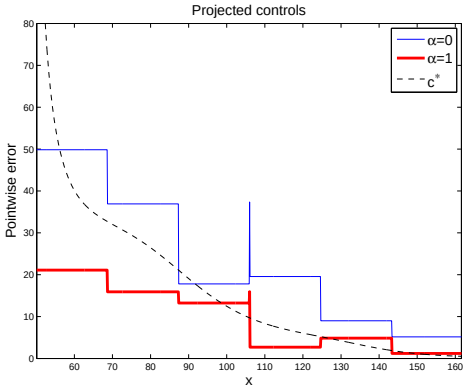

FIG. 5. *Projected controls* $c(x_l;a^{601})$

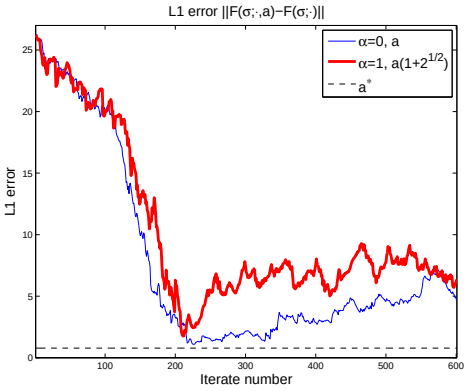

FIG. 6. L^1 *error* $\|F(\sigma;\cdot,a^\ell) - F(\sigma;\cdot)\|_1$

We now consider the derivative-free estimation presented in Section §2.4. For the work W defined by (4.2), it follows from (2.38) that $\exp(-\sigma W) = \mathfrak{W} = 1_B(Y_\tau) + \lambda$ and thus $E^x_\nu[D\mathfrak{W}] = p_B(x) + \lambda$, by (2.9) and (2.39). In Figure 7, the running average of the sample mean oscillates around the value of $p_B(x)$ obtained by finite-differences solution of (2.20)–(2.21), even though the controls are suboptimal (see Figure 6). In Figure 8, the running average of the sample standard deviation is smaller than the theoretical standard deviation.

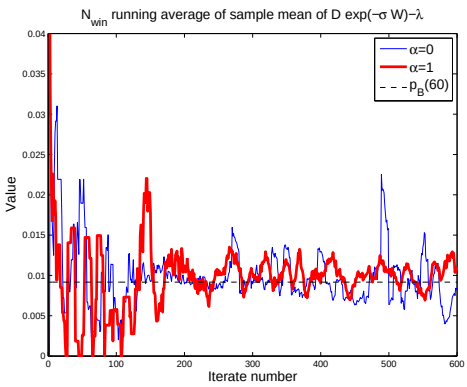

FIG. 7. *Sample mean of* $D\exp(-\sigma W) - \lambda$

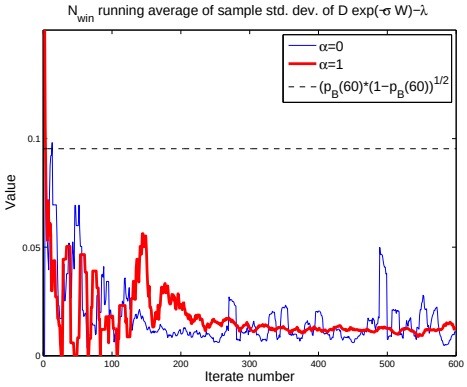

FIG. 8. *Sample std. dev. of* $D\exp(-\sigma W) - \lambda$

Figures 3, 4 and 6, suggest that, if one does not change the descent or sampling

parameters, then the method will not yield significantly better free energy value estimates or function approximations, after a certain number of iterations. At this point, one may terminate the method or change the parameters. In particular, by the 250th iteration, the sample mean, standard deviation and L^1 error appear to have attained the best observed values (for $\alpha = 0$) or to have stabilised (for $\alpha = 0$). We shall show in Section §4.3 that these quantities need not attain their best observed values simultaneously, and postpone the investigation of termination criteria to future work.

4.3. Sensitivity analysis. In order to qualitatively determine the sensitivity of the gradient descent algorithm to the gradient descent parameters, we examined the effect of changing the value of a single parameter, while keeping all other parameters fixed. We used the running average over 10 iterations of the sample mean of $\widehat{K}(\alpha, \sigma; a)$ and the L^1 error of the free energy function approximation to compare the performance of the gradient descent. The parameters we changed were: the value of λ from 10^{-3} to 10^{-2}; the initial control from $c(\cdot; a^0) \equiv 0$ to $c(\cdot; a^0) \equiv \rho(\lambda)/(162-50)$; and the basis functions, from tent functions (4.7)–(4.8) to non-normalised Gaussians

$$b_i(x) := h \exp\left(-\frac{|x-\mu_i|^2}{2h^2}\right) \tag{4.11}$$

$$\nabla b_i(x) := \frac{\mu_i - x}{h} \exp\left(-\frac{|x-\mu_i|^2}{2h^2}\right) \tag{4.12}$$

with μ_i and h as in (4.9) and (4.10). The results below are consistent with those observed in Section §4.2, in the sense that the fluctuations in the running average of the sample mean of $\widehat{K}(\alpha, \sigma; a)$ are smaller, and the function approximations of $F(\sigma; \cdot)$ are more stable, when $\alpha = 1$. The sample standard deviations of $\widehat{K}(\alpha, \sigma; a)$ are also smaller when $\alpha = 1$ (data not shown).

Changing the value λ from 10^{-3} to 10^{-2} leads to a smaller difference $\rho(\lambda)$ between the minimum and maximum values of $\widehat{W}$. This implies that the value of $\widehat{K}(\alpha, \sigma; a)$ for a sample path of $\widehat{Y}$ decreases by a smaller amount when $\widehat{Y}$ terminates at B, and in particular, the gradient (3.9) tends to assume smaller values. As a result, the L^1 error decreases more slowly in Figure 10 compared to in Figure 6, e.g., for $\alpha = 1$, the L^1 error decreases to the observed minimum within 200 (400) iterations for $\lambda = 10^{-3}$ (10^{-2}). The slower decrease in L^1 error explains the slower decrease in the sample mean of $\widehat{K}(1, \sigma; a)$ in Figure 9 compared to Figure 3.

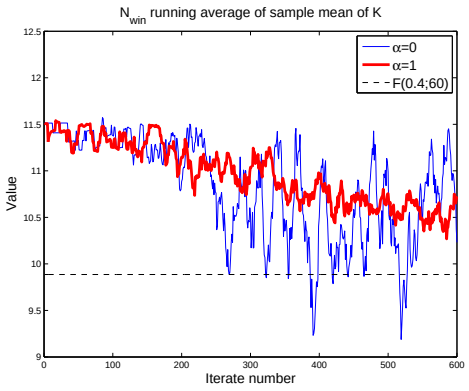

FIG. 9. *Sample mean of* $\widehat{K}(\alpha, \sigma; a)$

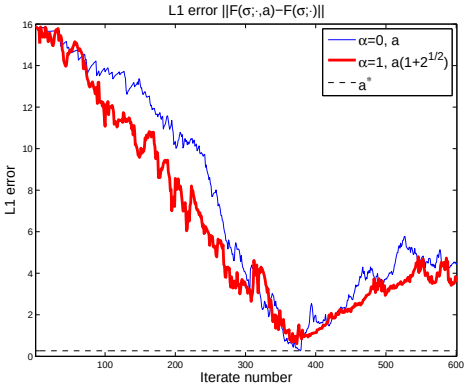

FIG. 10. L^1 *error* $\|F(\sigma; \cdot, a^\ell) - F(\sigma; \cdot)\|_1$

Note that, for $\alpha = 1$, although the L^1 error has already begun to increase from its minimum after 400 iterations in Figure 10, the running average continues to decrease towards the value of $F(\sigma; x)$ in Figure 9.

Changing the value of the initial control from 0 to $\rho(\lambda)/(162 - 50) \approx 8.8$ (recall that calculations involving domain data were performed in radians) results in running averages of the sample mean that decrease to their limiting values in fewer iterations, cf. Figure 11. Note that the limiting values around which the running averages and L^1 errors oscillate are comparable to those in Figures 3 and 6. When $\alpha = 1$, the running average of the sample mean of $\widehat{K}(\alpha, \sigma; a^\ell)$ exhibits more stable behaviour and smaller fluctuations, as was observed in Figure 3. Figure 10 shows that the L^1 error is more stable, and reaches its limiting value sooner, when $\alpha = 1$.

The preceding observations suggest that the limiting values of the coefficients $(a_i)_i$ will be the same if they are initialised in a suitable range of values, corresponding to sufficiently small initial controls. If the value of the initial control were so large that most, if not all, the sample paths had short duration and terminated at B, then $\widehat{\tau}$, $\widehat{M}$, $\widehat{A}$, $\widehat{W}$ and hence the finite sample-approximation of the gradient (3.9) would be close to zero, and the coefficients $(a_i)_i$ would remain near their initial values.

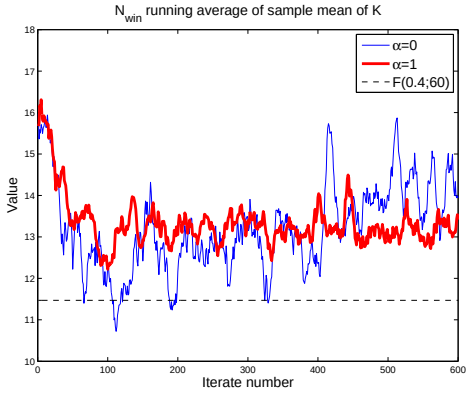

FIG. 11. *Sample mean of* $\widehat{K}(\alpha, \sigma; a)$

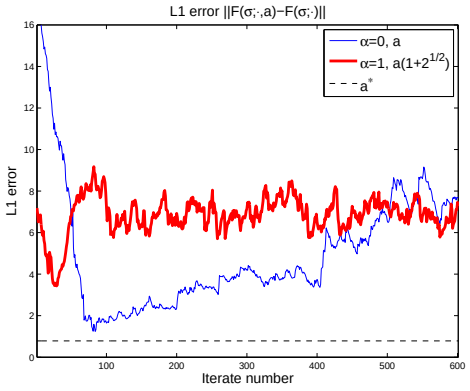

FIG. 12. L^1 *error* $\|F(\sigma; \cdot, a^\ell) - F(\sigma; \cdot)\|_1$

Changing the basis functions to non-normalised Gaussian basis functions results in running averages and L^1 errors that decrease in fewer iterations to similar values. It is reassuring that the limiting values are similar for both quantities, since this suggests that the choice of basis function does not significantly affect the limiting value of the coefficient vector a. Since the choice of basis function affects the estimator $\widehat{K}(\alpha, \sigma; a)$ only through the control, it suffices to explain the faster decrease in L^1 error in order to explain the faster decrease in the running average.

The faster reduction in L^1 error boils down to an analysis of the gradient terms (3.10) and (3.11). Fix i, and consider $\nabla b_i(\widehat{Y}_k)$ for some $k \leq \widehat{\tau}$. When ∇b_i is given by (4.8), $\nabla b_i(\widehat{Y}_k)$ will be nonzero for fewer k, compared to when ∇b_i is given by (4.12). Thus, when Gaussian basis functions are used, the gradient terms will differ from zero for more k, and thus the contribution to the gradient (3.9) from a single sample path $\widehat{Y}$ will be larger on average. As a result, the coefficients a_i will change by a larger amount after a single iteration.

5. Discussion. The sensitivity analysis in Section §4.3 emphasises the importance of the choice of basis functions: if the supports of the gradients are too small,

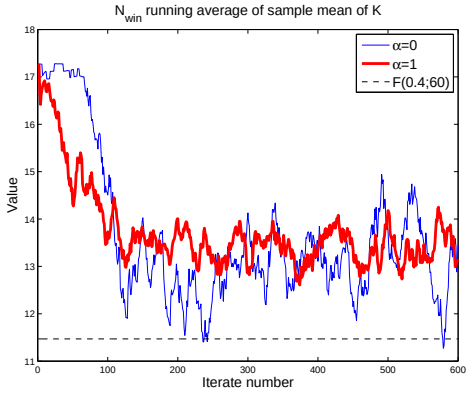

FIG. 13. *Sample mean of* $\widehat{K}(\alpha, \sigma; a)$

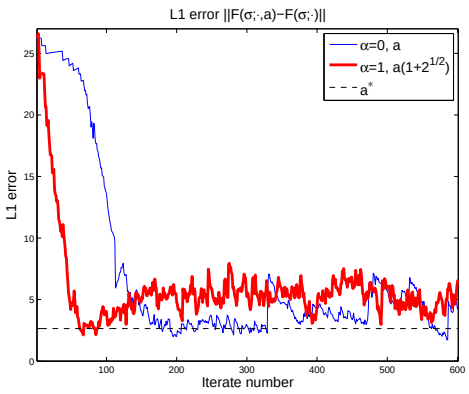

FIG. 14. L^1 *error* $\|F(\sigma; \cdot, a^\ell) - F(\sigma; \cdot)\|_1$

then the coefficients change slowly per iteration, and vice versa. If available, a priori information about the potential V should be used to devise better basis functions. When no a priori information is available, methods for exploring energy landscapes such as metadynamics [18] can be used to identify metastable sets. Although we have not done so here, setting the basis functions adaptively may be useful. For example, one criterion could be to define the basis functions so that the controlled trajectories will, on average, spend roughly the same amount of time in the supports of their gradients. When the system is high-dimensional there is no way around globally supported (radial) basis functions, so a clever choice in this case is crucial.

Many methods for stochastic approximation, e.g., [15,23,25,28] advocate schemes for reducing descent step sizes, and specify criteria for terminating the iterations. We have made no attempt to propose any such step reduction scheme in this study, because we found that common step reduction schemes reduced the descent step size too quickly. We leave an analysis of the convergence of the martingale-based method and the construction of suitable schemes for step reduction to future study.

In this paper, we derived the gradient descent algorithm for free energy value estimation in [13], using Girsanov's theorem. The derivation leads to the addition of a martingale term to the (biased) estimator of the free energy value, and to a different gradient descent dynamics, as shown in Section §3.1. The martingale term does not solve the problem caused by the omission in the gradient (3.13) of the partial derivatives of $\widehat{W}$ and $\widehat{\tau}$ with respect to the expansion coefficients $(a_i)_i$. We presented results which suggest that the martingale-based estimator has smaller variance, and that the martingale-based gradient descent dynamics are more stable.

We believe that the martingale-based gradient descent method presents an interesting addition to the existing methods for stochastic optimisation of high-dimensional systems, such as policy iteration [1] or the cross-entropy method [26, 27, 30]. This is because, in addition to features described earlier, the martingale-based method permits the simultaneous estimation of the desired equilibrium average $E^x[\mathfrak{W}]$ by free energy estimation and importance sampling.

REFERENCES

[1] D. P. BERTSEKAS, *Approximate policy iteration: a survey and some new methods*, J. Control Theor. Appl., 9 (2011), pp. 310–335.

[2] M. Boué and P. Dupuis, *A variational representation for certain functionals of brownian motion*, Ann. Prob., 26 (1998), pp. 1641–1659.
[3] G. E. Crooks, *Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems*, J. Stat. Phys., 90 (1998), pp. 1481–1487.
[4] I. Csiszar, *Information-type measures of difference of probability distributions and indirect observations*, Studia Sci. Math. Hungar., 2 (1967), pp. 299–318.
[5] P. dai Pra, L. Meneghini, and W. J. Runggaldier, *Connections between stochastic control and dynamic games*, Math. Control Signals Syst., 9 (1996), pp. 303–326.
[6] R. Du, V. S. Pande, A. Y. Grosberg, T. Tanaka, and E. S. Shakhnovich, *On the transition coordinate for protein folding*, J. Chem. Phys., 108 (1998), pp. 334–350.
[7] P. Dupuis, K. Spiliopoulos, and H. Wang, *Importance sampling for multiscale diffusions*, Multiscale Model. Simul., 10 (2012), pp. 1–27.
[8] P. Dupuis and H. Wang, *Importance sampling, large deviations, and differential games*, Stochastics Stochastic Rep., 76 (2004), pp. 481–508.
[9] R. Durrett, *Stochastic calculus : a practical introduction*, Probability and stochastics series, CRC Press, Boca Raton, 1996.
[10] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer, New York, 2006.
[11] P. L. Geissler, C. Dellago, and D. Chandler, *Kinetic pathways of ion pair dissociation in water*, J. Phys. Chem. B, 103 (1999), pp. 3706–3710.
[12] C. Hartmann, R. Banisch, M. Sarich, T. Badowski, and C. Schütte, *Characterization of rare events in molecular dynamics*, Entropy, 16 (2013), pp. 350–376.
[13] C. Hartmann and C. Schütte, *Efficient rare event simulation by optimal nonequilibrium forcing*, Journal of Statistical Mechanics: Theory and Experiment, 2012 (2012), p. P11004.
[14] C. Jarzynski, *Nonequilibrium equality for free energy differences*, Phys. Rev. Lett., 78 (1997), pp. 2690–2693.
[15] J. Kiefer and J. Wolfowitz, *Stochastic estimation of the maximum of a regression function*, Ann. Math. Stat, 23 (1952), pp. 462–466.
[16] S. Kullback, *A lower bound for discrimination information in terms of variation (corresp.)*, IEEE Trans. Inf. Theor., 13 (2006), pp. 126–127.
[17] S. Kullback and R. A. Leibler, *On information and sufficiency*, The Annals of Mathematical Statistics, 22 (1951), pp. 79–86.
[18] A. Laio and M. Parrinello, *Escaping free-energy minima*, Proc. Natl. Acad. Sci. U.S.A., 99 (2002), pp. 12562–12566.
[19] B. K. Øksendal, *Stochastic Differential Equations: An Introduction With Applications*, Springer, Berlin, 2003.
[20] L. Onsager, *Initial recombination of ions*, Phys. Rev., 54 (1938), pp. 554–557.
[21] A. Owen and Y. Zhou, *Safe and effective importance sampling*, J. Am. Statist. Assoc., 95 (2000), pp. 135–143.
[22] M.S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden-Day, San Francisco, 1964.
[23] B. Polyak and A. Juditsky, *Acceleration of stochastic approximation by averaging*, SIAM J. Control Optim., 30 (1992), pp. 838–855.
[24] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl, *Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit*, Bioinformatics, 29 (2013), pp. 845–854.
[25] H. Robbins and S. Monro, *A stochastic approximation method*, Ann. Math. Stat, 22 (1951), pp. 400–407.
[26] R. Y. Rubinstein, *Optimization of computer simulation models with rare events*, Eur. J. Oper. Res., 99 (1997), pp. 89 – 112.
[27] R. Y. Rubinstein and D. P. Kroese, *The Cross Entropy Method: A Unified Approach To Combinatorial Optimization, Monte-carlo Simulation (Information Science and Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2004.
[28] J.C. Spall, *Multivariate stochastic approximation using a simultaneous perturbation gradient approximation*, IEEE T. Automat. Control, 37 (1992), pp. 332–341.
[29] E. Vanden-Eijnden and J. Weare, *Rare event simulation of small noise diffusions*, Comm. Pure Appl. Math., 65 (2012), pp. 1770–1803.
[30] W. Zhang, H. Wang, C. Hartmann, M. Weber, and C. Schütte, *Applications of the cross-entropy method to importance sampling and optimal control of diffusions*, SIAM J. Sci. Comput., 36 (2014), pp. A2654–A2672.