

# Efficient Rare Event Simulation by Optimal Nonequilibrium Forcing

Carsten Hartmann, Christof Schütte

E-mail: [chartman@mi.fu-berlin.de](mailto:chartman@mi.fu-berlin.de), [schuette@mi.fu-berlin.de](mailto:schuette@mi.fu-berlin.de)

**Abstract.** Rare event simulation and estimation for systems in equilibrium are among the most challenging topics in molecular dynamics. As was shown by Jarzynski and others, nonequilibrium forcing can theoretically be used to obtain equilibrium rare event statistics. The advantage seems to be that the external force can speed up the sampling of the rare events by biasing the equilibrium distribution towards a distribution under which the rare events is no longer rare. Yet algorithmic methods based on Jarzynski's and related results often fail to be efficient because they are based on sampling in path space. We present a new method that replaces the path sampling problem by minimization of a cross-entropy-like functional which boils down to finding the optimal nonequilibrium forcing. We show how to solve the related optimization problem in an efficient way by using an iterative strategy based on milestoning.

## 1. Introduction

Molecular dynamics (MD) simulations allow for analysis and understanding of the dynamical behaviour of molecular systems. However realistic simulations on timescales beyond microseconds are still infeasible even on the most powerful general purpose computers, which renders the MD-based analysis of many biological equilibrium processes, that are often rare compared to the characteristic time scale of the system and hence require prohibitively long simulations, impossible. The hallmark of these rare events is that the average waiting time between the events is orders of magnitude longer than the timescale of the switching event itself. Thus rare event simulation and estimation are among the most challenging topics in molecular dynamics.

The molecular dynamics literature on rare event simulations is rich. Since direct numerical equilibrium simulation is infeasible, all available techniques try to sample from the rare event statistics by biasing the system in one or the other way. Roughly speaking, we can distinguish between two major classes of sampling techniques: class *A* consists of splitting methods that decompose state space, but are still essentially based on an equilibrium distribution, whereas methods from class *B* proceed by driving the system under consideration into a nonequilibrium regime that changes the rare events statistics. For a general overview of Monte-Carlo methods for rare events in other application fields, we refer to the textbook [2].

The list of methods in class *A* range from reaction-coordinate based techniques via path-space oriented techniques to approaches based on interface sampling or generalized dynamics. Reaction-coordinate based techniques consider the marginal of the equilibrium distribution in some low-dimensional collective variables like in

direct free energy calculations [4]; they suffer from the fact that appropriate reaction coordinates are often not available. Path-space oriented techniques approximate the most important reaction paths that govern the rare event statistics either by sampling distribution of reactive paths like in transition path sampling (TPS) [9, 3] or by optimizing an appropriate path functional like in the string method [13]; they become problematic if the path space distribution is multi-modal or generally too complex (e.g., involving bifurcations). Interface sampling techniques like milestoning [14] or forward flux sampling (FFS) [1] place a set of suitably chosen interfaces in state space between the initial and final state and use them to follow the transition of the system in an iterative manner using equilibrium trajectories that connect neighbouring interfaces. The idea of *generalized dynamics* such as hyperdynamics [39], metadynamics [24], conformational flooding [17], or the adaptive biasing force (ABF) method [7] is to bias the system on-the-fly (e.g., by filling in certain energy wells in which the system got trapped during a simulation) so as to enhance rare transitions between metastable states. Although seemingly different, generalized dynamics belong to class *A*, in that they only alter the underlying equilibrium distribution along a predefined set of low-dimensional collective variables. Although these methods have proven to be very efficient, they require that the interesting processes can be described by a few collective coordinates that have to be known in advance.

Class *B* consists of methods based on the Jarzynski and Crooks formulae [21, 5] that relate the equilibrium Helmholtz free energy to the nonequilibrium work exerted under external forcing. Instances of nonequilibrium simulations that mimic experiments on controlling and manipulating single molecules (see, e.g., [33, 28]) are single-molecule pulling [19], steered molecular dynamics [36] or bridge sampling [29], to mention just a few. The corresponding path functionals have the form of cumulant-generating functions for the exerted work [23, 26] which poses immense challenges to Monte-Carlo simulations and limits the usability of the formulae in practice. Roughly speaking, the usability is limited by the fact that the likelihood ratio between equilibrium and nonequilibrium trajectories is highly degenerate, for the overwhelming majority of nonequilibrium forcings generate trajectories that have almost zero weight with respect to the equilibrium distribution that is relevant for the rare event; cf. also the discussion in [27]. Nevertheless the underlying idea is appealing and a cleverly designed external force may speed up the sampling of the rare events by biasing the equilibrium distribution of the system towards a distribution under which the rare events is no longer rare, while giving numerical estimators that are useful in terms of variance and convergence properties.

The method presented in this article belongs to the latter class, but shares some ideas with ideas from class *A*. It takes up the idea that external forcings can speed up the rare event but avoids sampling issues related to nonequilibrium processes. Instead it uses optimal nonequilibrium forcing in connection with splitting methods such as FFS or milestoning, in the sense that the new method uses interfaces to follow the transition of an *optimally driven* system where the external forcing that drives the system from one interface to the next results in a considerable speed-up compared to FFS or milestoning. Specifically, the new method replaces the path sampling problem using an exponential change of measure that can be explicitly computed by minimizing a cross-entropy-like functional, which then yields the optimal forcing. Although the minimization involves solving an optimal control problem, the numerical effort can be drastically reduced when the minimization is done in a clever way; one reason is that the path functional becomes linear after the change of measure whereas it was

exponential in the original cumulant-generating function.

Transformations based on exponential change of measures have a rich tradition in the (risk-sensitive) optimal control literature [20, 6, 16] and the theory of large deviations [15, 40], and are regularly rediscovered—mostly aiming at turning certain optimal control problems into linearly solvable sampling problems [22, 38, 12]; cf. also [37, 32]. Here we pursue the reversed strategy and turn a difficult rare event estimation problem into an optimal control problem that can be solved by minimizing a suitable functional. Thus the basic outline of the new method is: iteratively determine the optimal nonequilibrium forcing by an optimization procedure based on milestone ideas that avoid path-space sampling and compute the equilibrium rare event statistics from the optimal nonequilibrium forcing.

Besides introducing the new method the purpose of this article is to explain the basic ideas of how to use optimal control for the estimation and simulation of rare events. Therefore we present only the simplest possible scenario (a particle following an overdamped Langevin dynamics in a conservative force field), without paying too much attention to complete generality or mathematical rigour. The first issue in Section 2 then is to introduce the variational characterization of (generalized) free energy and the exponential change of measures that are the basis of our optimal control approach. The precise formulation of the optimal control problem, a stochastic control problem with quadratic control costs and an indefinite time horizon, is given in Section 3. In Section 4 we describe the numerical method for computing the optimal control, based on an inexact gradient descent in connection with a milestone algorithm, and apply it to the controlled first passage between metastable sets. We briefly summarize the results in Section 5 and sketch possible generalization that have been omitted for the sake of brevity.

## 2. A variational characterization of free energy

We consider a particle with position  $X_t \in \mathbb{R}^n$  at time  $t > 0$  which moves in an energy landscape  $V: \mathbb{R}^n \rightarrow \mathbb{R}$  according to the equation

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\epsilon} dB_t, \quad X_0 = x. \quad (2.1)$$

Here  $B_t$  denotes standard  $n$ -dimensional Brownian motion, and  $\epsilon > 0$  is the temperature of the system. Under mild conditions on the energy landscape function  $V$  we have ergodicity, and the law of  $X_t$  converges to a unique equilibrium distribution with density

$$\rho(x) = Z^{-1} \exp(-\epsilon^{-1}V(x)), \quad Z = \int_{\mathbb{R}^n} \exp(-\epsilon^{-1}V(x)) dx.$$

We assume throughout that the temperature is small, relative to the largest energy barriers, i.e.,  $\epsilon \ll \Delta V_{\max}$ . As a consequence, the relaxation of the dynamics towards equilibrium is dominated by the rare transitions over the largest energy barriers.

Let  $W$  be a random variable that depends on the sample paths  $(X_t)_{0 \leq t \leq \tau}$  up to a stopping time  $\tau$ . We will call  $W$  *work* in the following. Given some continuous function  $f: \mathbb{R}^n \times [0, \infty) \rightarrow \mathbb{R}$ , we suppose that it can be expressed as<sup>‡</sup>

$$W = \int_0^\tau f(X_t) dt. \quad (2.2)$$

<sup>‡</sup> The following considerations below are not at all limited to systems of the form (2.1) and path functionals like (2.2) and can be easily generalized to, e.g., non-gradient systems with multiplicative and/or degenerate noise or observables  $f$  that are explicitly time-dependent.

Let us further denote by  $P$  the probability measure on the space of continuous trajectories that is generated by the Brownian motion in (2.1), and let  $\mathbf{E}^x[\cdot] = \mathbf{E}[\cdot | X_0 = x]$  be the expectation with respect to  $P$ , i.e., the average over all realizations of  $X_t$  starting at  $X_0 = x$ . We call the quantity

$$F(x) = -\epsilon \log \mathbf{E}^x[\exp(-W/\epsilon)]. \quad (2.3)$$

the (conditional) free energy of  $W$  with respect to  $P$ .

**Remark 1.** Clearly, the functions and the expectation on the right hand side of (2.3) do not commute, and it follows by Jensen's inequality that  $F(x) \leq \mathbf{E}^x[W]$ , in accordance with the second law of thermodynamics. But  $F$  encodes information about the cumulants of the work  $W$  (assuming they exist), namely,

$$F(x) = \mathbf{E}^x[W] + \frac{1}{2\epsilon} \mathbf{E}^x[(W - \mathbf{E}^x[W])^2] + \dots$$

**Remark 2.** The similarity between (2.3) and Jarzynski's formula [21] is no coincidence. If  $\tau = T$  is a deterministic stopping time and  $W$  is the nonequilibrium work done on a system during a transition between two equilibrium states  $E_1$  and  $E_2$ , then  $F(E_1)$  equals the equilibrium free energy difference between  $E_1$  and  $E_2$ .

The phrases "work" for the quantity  $W$  defined in (2.2) and "free energy" for  $F$  as of (2.3) are just used to relate to Jarzynski's formula. The framework is much more general as the following example will show.

*Guiding example.* One example, of which we will consider variants below, is the first hitting time of a subset of state space. To this end let  $S \subset \mathbb{R}^n$  a set and define

$$\tau = \inf\{t > 0: X_t \in S\}.$$

to be the first time at which  $X_t$  hits  $S$ . Choosing the constant function  $f = \sigma$  in (2.2), the free energy

$$F_\sigma(x) = -\epsilon \log \mathbf{E}^x[\exp(-\sigma\tau/\epsilon)].$$

considered as a function of  $\sigma$  is the scaled cumulant-generating function of  $\tau$  when  $X_t$  is started at  $X_0 = x$ . In particular, we can compute the mean first hitting time by

$$\epsilon \left. \frac{dF_\sigma}{d\sigma} \right|_{\sigma=0} = \mathbf{E}^x[\tau].$$

### 2.1. Relative entropy and change of measures

The strict convexity of the exponential function implies that equality  $F(x) = \mathbf{E}^x[W]$  is only attained if  $W$  is  $P$ -almost surely constant; one such case is the adiabatic limit

$$W = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X_t) dt.$$

We will restore (2.3) to an expression that becomes linear in  $W$  after a suitable change of measure. To this end let  $Q$  denote a probability measure on the space of continuous trajectories that is absolutely continuous with respect to  $P$  (i.e.,  $\varphi = dQ/dP$  exists). We define the relative entropy of  $Q$  with respect to  $P$  as

$$I(Q\|P) = \int_{\mathbb{R}^n} \log \left( \frac{dQ}{dP} \right) dQ. \quad (2.4)$$

(This is also called the *Kullback-Leibler divergence*.) We declare that  $I(Q\|P) = \infty$  if  $Q$  is not absolutely continuous with respect to  $P$ . Then, by Jensen's inequality,

$$\begin{aligned} F(x) &= -\epsilon \log \mathbf{E}^x[\exp(-W/\epsilon)] \\ &= -\epsilon \log \mathbf{E}_Q^x[\exp(-W/\epsilon - \log \varphi)] \\ &\leq \mathbf{E}_Q^x[W] + \epsilon I(Q\|P), \end{aligned} \tag{2.5}$$

where we have used the notation  $\mathbf{E}_Q[\cdot]$  to denote the expectation with respect to  $Q$ . The last inequality that appears in the literature in various forms as *second-law-like identity* or *generalized Jarzynski inequality* (cf. [35, 18]) suggests that the free energy and the relative entropy are related by a Legendre-type transformation, viz.,

$$F(x) = \inf_Q \{ \mathbf{E}_Q^x[W] + \epsilon I(Q\|P) \},$$

and a result in [6] implies that the infimum exists and is attained when  $Q$  runs over all path measures that are absolutely continuous with respect to  $P$ . By the strict convexity of the exponential function, the latter implies that  $W + \epsilon \log \varphi$  is  $Q$ -almost surely constant. Note that this is equivalent to saying that estimators based on the new random variable  $W + \epsilon \log \varphi$  has zero variance with respect to  $Q$ . The idea of the approach sketched below then is to represent  $Q$  in terms of suitable (parametric) control variables and minimize the right hand side of (2.5) over all admissible controls.

### 3. An optimal control problem

The aim of this section is to derive necessary and sufficient conditions for the optimal change of measure that turns (2.5) into an equality. To this end we follow ideas by Fleming and co-workers [15, 10] and consider the exponential cost functional:

$$\psi(x) = \mathbf{E}^x \left[ \exp \left( -\epsilon^{-1} \int_0^\tau f(X_s) ds \right) \right]. \tag{3.1}$$

For a stopping time  $\tau$  that is the first hitting time of a set  $S \subset \mathbb{R}^n$ , the Feynman-Kac formula [31] implies that  $\psi$  solves the elliptic boundary value problem

$$\epsilon L\psi = f\psi, \quad \psi|_{\partial S} = 1, \tag{3.2}$$

where

$$L = \epsilon \nabla^2 + \nabla V \cdot \nabla. \tag{3.3}$$

is the infinitesimal generator of  $X_t$ , defined on a suitable subspace of  $L^2(\mathbb{R}^n)$ . We want to transform the boundary value problem (3.2) into an equation for the unknown control variable in (2.5). For this we proceed in two steps.

*Step 1:* We can safely assume that  $\tau$  is almost surely finite. As a consequence, the function  $\psi$  in (3.1) admits a formal representation of the form

$$\psi = \exp(-F/\epsilon).$$

We seek an equation for the free-energy  $F$ . By chain rule, it follows that

$$\epsilon \exp(F/\epsilon) L \exp(-F/\epsilon) = -LF + |\nabla F|^2,$$

which entails that (3.2) is equivalent to

$$LF - |\nabla F|^2 + f = 0, \quad F|_{\partial S} = 0, \tag{3.4}$$

The last equation is known as the Hamilton-Jacobi-Bellmann (HJB) equation of optimal control [16]; its solution is called *value function* or *optimal cost-to-go*.

*Step 2:* To reveal the stochastic optimal control problem that corresponds to the HJB equation (3.4), we first note that

$$-|\nabla F|^2 = \min_{c \in \mathbb{R}^n} \left\{ \sqrt{2}c \cdot \nabla F + \frac{1}{2}|c|^2 \right\},$$

from which we recognize that (3.4) is equivalent to

$$\min_{c \in \mathbb{R}^n} \{L(c)F + g(x, c)\} = 0, \quad F|_{\partial S} = 0, \quad (3.5)$$

with the shorthands

$$g(x, c) = f(x) + \frac{1}{2}|c|^2$$

and

$$L(c) = \epsilon \nabla^2 + (\sqrt{2}c - \nabla V) \cdot \nabla.$$

Equation (3.5) is the Hamilton-Jacobi-Bellman equation of the following optimal control problem that should be compared to the right hand side of (2.5): minimize

$$I(u) = \mathbf{E} \left[ \int_0^\tau g(X_t, u_t) dt \right] \quad (3.6)$$

over an admissible set  $U$  of control laws  $u$  with values in  $\mathbb{R}^n$  and subject to the tilted dynamics

$$dX_t = \left( \sqrt{2}u_t - \nabla V(X_t) \right) dt + \sqrt{2\epsilon} dB_t. \quad (3.7)$$

That is, the expectation in (3.6) has to be taken wrt the path measure  $Q$  generated by the dynamics given by (3.7).

**Remark 3.** *The dynamics that generates the new path measure  $Q$  is again of gradient form if  $u = u^*$  is the optimal Markovian feedback control, i.e. when  $Q = Q(u^*)$ . As a consequence, the optimally controlled process satisfies detailed balance [26]. Indeed, since (3.6) is quadratic and (3.7) is affine in the control, the minimizer*

$$c^*(x) = \underset{c}{\operatorname{argmin}} \{L(c)F + g(x, c)\},$$

*in (3.5) is unique (provided that  $F$  is sufficiently smooth). The optimal feedback law is then given by  $u_t^* = -\sqrt{2}\nabla F(X_t)$  and gives rise to the tilted dynamics*

$$dX_t = -\nabla G(X_t)dt + \sqrt{2\epsilon} dB_t, \quad X_t \in \mathbb{R}^n \setminus S,$$

*with the tilted potential*

$$G(x) = V(x) + 2F(x).$$

*Guiding example, cont'd.* In some cases it is helpful to pursue a reverse strategy and transform the nonlinear HJB equations of an optimal control problem into a linear equation that may be easier to solve (cf. [22, 38]).

Consider a Brownian particle under a microscope with a moveable object holder. Let  $D \subset \mathbb{R}^2$  denote the microscope's focal disc,  $X_t \in \mathbb{R}^2$  the particle position at time  $t > 0$ , relative to the position of the object holder, and  $u_t$  the motor force. The control task is to move the object holder such that the particle stays in the focus as long as possible. Hence the control objective is the maximization of the mean first exit time from  $D$  which amounts to minimizing the cost functional

$$I(u) = \mathbf{E} \left[ -\tau + \frac{1}{2} \int_0^\tau |u_t|^2 dt \right],$$

subject to

$$dX_t = \sqrt{2}u_t + \sqrt{2\epsilon} dB_t.$$

Let

$$F(x) = \min_{u \in U} \mathbf{E}^x \left[ -\tau + \frac{1}{2} \int_0^\tau |u_t|^2 dt \right],$$

be the value function (free energy) of the problem and

$$\psi(x) = \mathbf{E}^x[\exp(\tau/\epsilon)].$$

Then the linear boundary value problem for  $\psi = \exp(-F/\epsilon)$  is a Helmholtz equation with Dirichlet boundary conditions,

$$\epsilon^2 \nabla^2 \psi + \psi = 0, \quad \psi|_{\partial D} = 1,$$

which can be solved by standard means.

#### 4. Greedy milestoning algorithm

At first sight it seems that we have not gained much, for we have transformed the original path sampling problem into a complicated nonlinear optimal control problem. However the optimal control formulation opens up other options for the numerical treatment of the rare event sampling in terms of a minimization problem. Another advantage is that it is relatively easy to construct unbiased estimators of the control functional, avoiding both bias and variance issues when estimating exponential observables such as (2.3).

*Discretization* Together with the information that the optimal Markov control is of feedback form our minimization problem (3.6)–(3.7) takes the form

$$F(x) = \min_{u_t=c(X_t)} \mathbf{E}_Q^x \left[ \int_0^\tau g(X_t, u_t) dt \right]$$

with  $Q$  denoting the path measure generated by the dynamics given by (3.7). We discretize this optimization problem by choosing a finite dimensional ansatz space for the space of admissible feedback functions  $c$ : We choose sufficiently smooth and integrable vector fields  $b_j: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $j = 1, \dots, m$ , so that

$$c(x) = \sum_{j=1}^m a_j b_j(x), \quad a_j \in \mathbb{R},$$

or, respectively, we choose scalar ansatz function  $v_j: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $j = 1, \dots, m$ , so that

$$F(x) = \sum_{j=1}^m a_j v_j(x), \quad b_j = -\sqrt{2} \nabla v_j.$$

The minimization problem then amounts to minimizing the cost functional

$$\tilde{I}(a) = \mathbf{E}_Q \left[ \int_0^\tau \left( f(X_s) + \frac{1}{2} \left| \sum_j a_j(s) b_j(X_s) \right|^2 \right) ds \right] \quad (4.1)$$

over the unknown coefficients  $a = (a_1, \dots, a_m)$  where  $Q = Q(a)$ , the path measure of the controlled diffusion (3.7) also depends on the coefficients; for the moment we remain with the imprecise statement that the measure  $Q$  has a density  $\varphi(\cdot; a)$  with

respect to a (fictitious) uniform measure on the space of all continuous paths in  $\mathbb{R}^n$ , which is a function of the unknown coefficients.§

*Gradient descent* We minimize the cost functional  $\tilde{I}(a)$  by a doing a gradient descent in the coefficient vector  $a = (a_1, \dots, a_m)$ . Specifically, we iterate the map

$$a^{(i+1)} = a^{(i)} - \alpha_i \nabla \tilde{I}(a^{(i)}),$$

where  $i$  is the iteration index and  $(\alpha_i)_{i \geq 1}$  is a bounded sequence of stepsizes for the gradient search. For instance, we can do a line search in the descent direction and determine  $\alpha_i$  so that it satisfies the Wolfe condition [30]. Details of the iteration that is based on an Euler-Maruyama discretization of the path measure  $Q$  will be given below in the appendix. The overall algorithm thus has the following steps:

- Choose scalar-valued ansatz functions  $v_j$  with support in the interesting region of state space and related vector fields  $b_j = -\sqrt{2}\nabla v_j$ .
- Choose initial coefficients  $a^{(0)} = (a_j^{(0)})$  such that the free energy or value function  $\sum_{j=1}^m a_j v_j(x)$  fills up the main wells in the energy landscape  $V$ .
- Iterate the following steps in  $i$ , starting with  $i = 0$ , until a prescribed termination criterium is satisfied:
  - (i) Sample the path measure  $Q = Q(a^{(i)})$  and evaluate  $\nabla \tilde{I}(a^{(i)})$  (see formula (A.4) in the appendix).
  - (ii) Perform a gradient descent  $a^{(i+1)} = a^{(i)} - \alpha_i \nabla \tilde{I}(a^{(i)})$ .

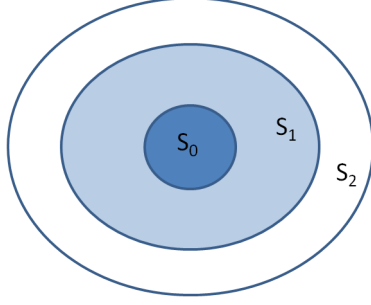
**Remark 4.** *The gradient search algorithm can be regarded as a variant of the cross-entropy method that is a relatively new Monte-Carlo technique for the sampling of rare events which goes back to Rubinstein and others [34]. It is based on the idea that an optimal change of measure can be found by minimizing the Kullback-Leibler divergence (2.4) over a family of probability measures  $Q$  in terms of the tilting parameter  $c$ . Compared to equilibrium rare event simulation algorithms used in molecular dynamics using the optimal change of measure has the advantage that the likelihood ratio  $dQ/dP$  stays of order one, while rare events under the original dynamics (here: diffusion in an energy landscape  $V$ ) are no longer rare under the forced dynamics (3.7). As a consequence, sampling the path measure  $Q$  is significantly more efficient than sampling the original path measure  $P$  since the trajectories to be sampled from  $Q$  are much shorter on average (i.e., the expected hitting time is considerably shorter).*

*Milestoning algorithm* For problems with a large state space or for strongly metastable systems, the above algorithm may still be inefficient since sampling the path measure  $Q$  may involve many rather long trajectories. In this case the computation can be broken down to transitions between neighbouring interfaces as in milestoning [14] or in FFS [1]. We explain the basic steps of this procedure: Let

$$\tilde{F}(x) = \min_a \mathbf{E}_Q^x \left[ \int_0^T \tilde{g}(X_s, c(X_s)) ds \right]$$

§ More precisely,  $Q = Q_x^\delta$  is the probability to find paths  $(X_s)_{0 \leq s \leq T}$  in a small tube around a smooth curve  $\gamma: [0, T] \rightarrow \mathbb{R}^n$ , i.e.,  $Q_x^\delta(\gamma) = P(\|X_s - \gamma(s)\| \leq \delta | X_0 = x)$ . By the Girsanov theorem,  $Q_x = \lim_{\delta \rightarrow 0} Q_x^\delta$  has a density  $\varphi = \exp(-S(\gamma))$  with respect to the Gaussian measure induced by the Brownian motion  $\tilde{B}_s = x + \sqrt{2\epsilon}B_s$ , where  $S(\gamma)$  is the Onsager-Machlup functional [11].





**Figure 1.** Illustration of nesting of sets for the milestoning iteration.

denote the semi-discretized value function of the problem, with the shorthand

$$\tilde{g}(x, c(x)) = \sigma f(x) + \frac{1}{2} \left| \sum_j a_j b_j(x) \right|^2.$$

Suppose that  $S = S_0$  is the set of interest and  $\tau = \tau_0$  is the first hitting time of  $S_0$ ; we now choose nested sets or *milestones*  $S_0 \subset S_1 \subset S_2 \subset \dots$  (cf. Figure 1). We first compute  $\tilde{F}$  in  $S_1 \setminus S_0$  by finding the optimal control policy  $c$  in  $S_1 \setminus S_0$ . That is, our ansatz functions in the above gradient descent algorithm only have to be non-vanishing in  $S_1 \setminus S_0$ . In particular this gives  $\tilde{F}$  on  $\partial S_1$ , the outer boundary of  $S_1 \setminus S_0$ . We can repeat the same algorithm in the set  $S_2 \setminus S_1$ ; then letting  $x \in S_2 \setminus S_1$  and letting  $\tau_1$  denote the first entry time into  $S_1$ , we have

$$\tilde{F}(x) = \min_a \mathbf{E}_Q^x \left[ \int_0^{\tau_1} \tilde{g}(X_s, c(X_s)) ds + \tilde{F}(X_{\tau_1}) \right],$$

where  $X_{\tau_1} \in \partial S_1 \subset S_1 \setminus S_0$  for which  $\tilde{F}$  has been computed in the previous step. By iterating the algorithm we eventually obtain  $\tilde{F}$  on all set boundaries  $\partial S_i$ ,  $i = 0, 1, 2, \dots$ . Thus, the milestoning iteration can be implemented as an outer loop which contains the above gradient descent algorithm in every of its iterations.

**Remark 5.** *The milestoning variant of the gradient descent algorithm only requires the computation of an ensemble of short trajectories of the controlled system (3.7). Here "short" means that they are orders of magnitude shorter than those in typical path-space sampling algorithms like TPS, and equilibrium milestoning or FFS.*

#### 4.1. Guiding example: computing the mean first passage time

We consider the uncontrolled dynamics (2.1) with the one-dimensional potential shown in Figure 2. Suppose we are interested in computing the mean first passage time to the set  $S = [-1.1, -1]$  in terms of the free energy (2.3). Let

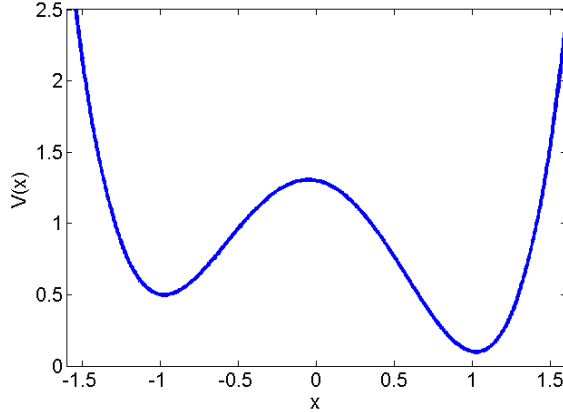
$$\tau = \inf\{t > 0: X_t \in \partial S\}.$$

be the first hitting time of  $S$ , consider the constant function  $f = \sigma$ , and the scaled moment generating function

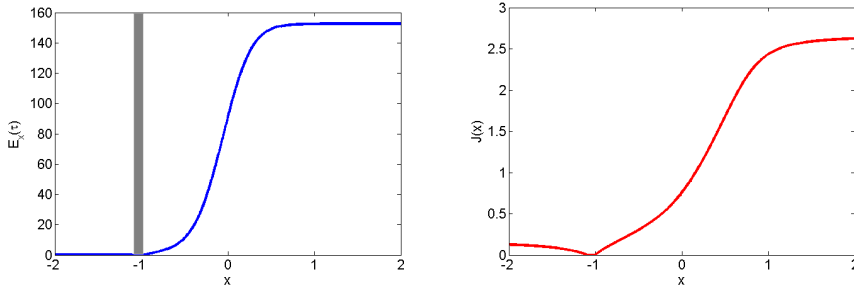
$$\psi_\sigma(x) = \mathbf{E}^x[\exp(-\sigma\tau/\epsilon)],$$

considered as a function of  $\sigma$ . The quantity of interest is the mean first passage time of the uncontrolled dynamics,

$$-\epsilon \left. \frac{d\psi_\sigma}{d\sigma} \right|_{\sigma=0} = \mathbf{E}^x[\tau],$$



**Figure 2.** Skew double-well potential  $V$ .



**Figure 3.** Reference solution for the uncontrolled mean first passage time (left panel) and the related free energy  $F_\sigma$  for  $\sigma = 1$  (right panel). Results based on finite element discretization of (3.2) with high precision.

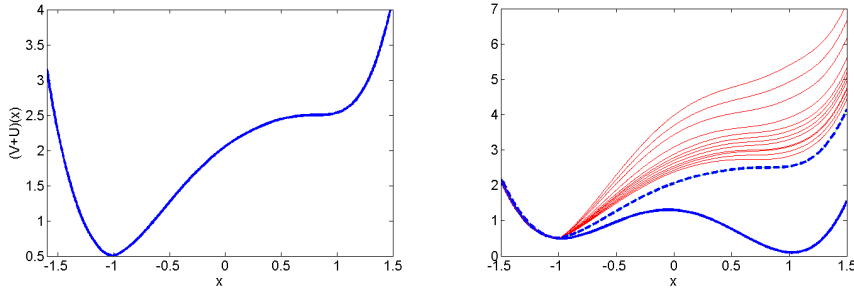
for  $\epsilon = 0.5$ .

In order to obtain a reference solution with high accuracy we first compute  $\psi_\sigma$  by discretizing the elliptic boundary value problem (3.2) based on a standard finite element discretization on a fine grid. This is possible because the state space dimension in this guiding example is small but will not be possible in realistically high dimensions. The resulting reference solution for  $\mathbf{E}^x[\tau]$  is shown in the left panel of Figure 3 below, along with the associated free energy  $F_\sigma(x) = -\epsilon \log \psi_\sigma(x)$  in the right panel.

An approximation of the free energy was then computed by the gradient-descent-milestoning algorithm described above that minimizes the cost functional (4.1) in the coefficients  $a = (a_1, \dots, a_m)$ . As scalar ansatz functions  $v_j$  we chose  $m = 10$  Gaussians with width 0.1 whose centers were uniformly spaced in the complement of  $S$ . Once the minimization had been converged, the value function (free energy) and the resulting optimal control law were given by

$$\tilde{F} = \sum_{j=1}^m a_j v_j(x), \quad c^*(x) = \sum_{j=1}^m a_j b_j(x),$$

with  $b_j = -\sqrt{2}\nabla v_j$ . The result agrees with the reference solution shown in Fig. 3 (deviations are of the order of the accuracy threshold used in the gradient descent



**Figure 4.** Optimally tilted potential potential (left panel) and the first 11 iterates of the gradient descent (right panel).

algorithm). Figure 4 shows the resulting optimally tilted potential  $G = V + 2\tilde{F}$ , together with first few iteration steps of the gradient search. The mean first passage time of the tilted system

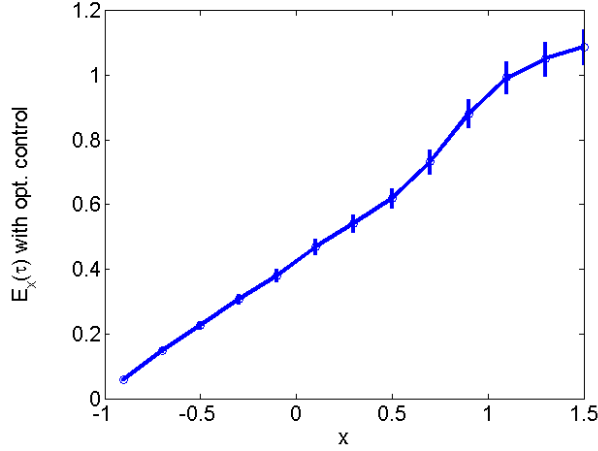
$$dX_t = -\nabla G(X_s)dt + \sqrt{2\epsilon}dB_s, \quad (4.2)$$

i.e., with  $V$  in (2.1) replaced by the new potential  $G$ , is shown in Figure 5.

As has been outlined above the algorithm only requires the computation of rather short trajectories since for all iterative potentials the mean first passage time is orders of magnitude smaller than for the original dynamics; the mean first passage time of the optimally tilted potential, e.g., is around 100 times smaller than originally.

## 5. Conclusions and outlook

We have developed a simulation scheme for rare events that is based on an optimal change of measure that boils down to a logarithmic transformation of the path functional under consideration. The measure transformation turns the original exponential path functional into the functional of an optimal control problem that is linear in the observable and quadratic in the control variables. Although analytic solutions to the optimal control problem are available only in simple situations and computing the optimal change of measure may require to solve a possibly high-dimensional optimal control problem numerically, there is a considerable speed-up coming from (a) the fact that the functional is linear-quadratic and allows for the design of robust unbiased Monte-Carlo estimators and (b) the fact that events that were rare originally are no longer rare under the new probability measure. The gain in the numerical complexity requires that the optimal control problem can be solved efficiently, and, with the equivalence between path sampling and optimal control in hand, we have sketched a numerical algorithm for computing the optimal control that is based on an easy-to-implement inexact gradient descent that can be solved rather efficiently using milestoning. The algorithm was tested, computing the optimal feedback for the controlled passage between metastable sets in a double-well potential. Even though the numerical example that we presented is tiny on the scale of typical molecular dynamics applications, we emphasize that the minimization algorithm is independent of the dimension of the system and hence admits an easy generalization to more complicated systems; we refer to the rich literature on machine learning and queuing networks where various strategies for treating high dimensional systems



**Figure 5.** Unbiased estimate of the first mean passage time, based on 2000 realizations of (4.2) after Euler-discretization. The error bars indicate 95% confidence intervals that were computed from the estimator’s standard deviation.

have been developed (e.g., see [8]). Finally we note that all ideas presented in this article can be readily extended to more complicated dynamics (e.g., degenerate diffusions with dissipation) and time-dependent path functionals (e.g., to simulate single-molecule experiments); it is even possible to consider situations where the exponential path functional involves additional control variables, in which case a logarithmic transformation leads to a game rather than an optimal control problem (cf. [25]). Further open issues are the deterministic limit of the stochastic control problem, the convergence analysis of the gradient descent and the rigorous analysis of fluctuations in systems under feedback control (cf. [35]).

## Appendix A. Computational aspects

In order to compute the gradient of (4.1) with respect to the unknown coefficients  $a = (a_1, \dots, a_m)$ , it is convenient to discretize the path measure  $Q = Q(a)$ . To this end, let  $0 = t_0 < t_1 < \dots < t_N = \tau$  be a set of time nodes with  $h = t_{k+1} - t_k$  where we assume for the moment that  $\tau < \infty$  is deterministic. Euler’s method applied to

$$dX_t = \left( \sqrt{2}c(X_t) - \nabla V(X_t) \right) dt + \sqrt{2\epsilon} dB_t.$$

gives

$$\tilde{X}_{k+1} = \tilde{X}_k + h \left( \sqrt{2}c(\tilde{X}_k) - \nabla V(\tilde{X}_k) \right) + \sqrt{2h\epsilon} \eta_{k+1}$$

where the  $\eta_k$  are i.i.d. random variables that are normally distributed with mean zero and unit covariance. Since the  $\eta_k$  are Gaussian, the density of the distribution  $Q_h(a)$  of discrete paths  $(\tilde{X}_0, \dots, \tilde{X}_N) \subset \mathbb{R}^n$  conditional on  $\tilde{X}_0 = x_0$  is readily shown to be

$$\varphi_h(x_0, \dots, x_N; a) = (Z_h(a))^{-1} \exp(-S_h(x_0, \dots, x_N; a)) \quad (\text{A.1})$$

with the discrete action

$$S_h = \frac{h}{4\epsilon} \sum_{k=0}^{N-1} \left| \frac{x_{k+1} - x_k}{h} + \nabla V(x_k) - \sqrt{2}c(x_k) \right|^2 \quad (\text{A.2})$$

and the normalization constant

$$Z_h = \int_{\mathbb{R}^n \times \dots \times \mathbb{R}^n} \exp(-S_h(x_0, \dots, x_N; a)) dx_1 \dots dx_N. \quad (\text{A.3})$$

Computing the gradient of the discretized functional

$$\tilde{I}_h(a; \tilde{X}_0) = \mathbf{E}_{Q_h}^{\tilde{X}_0} \left[ \sum_{k=0}^{N-1} \tilde{g}_h(\tilde{X}_k, c(\tilde{X}_k)) \right]$$

with

$$\tilde{g}_h(x, c(x)) = h \left( f(x) + \frac{1}{2} \left| \sum_j a_j b_j(x) \right|^2 \right)$$

is now straightforward. Assuming that  $\tilde{X}_0$  is independent of the control, we have

$$\frac{\partial \tilde{I}_h}{\partial a_j} = \mathbf{E}_{Q_h}^{\tilde{X}_0} \left[ \sum_{k=0}^{N-1} \frac{\partial \tilde{g}_h}{\partial a_j} - \tilde{g}_h \left( \frac{\partial S_h}{\partial a_j} + \frac{1}{Z_h} \frac{\partial Z_h}{\partial a_j} \right) \right],$$

where both  $\tilde{g}_h$  and  $\partial \tilde{g}_h / \partial a_j$  are evaluated at  $(x_k, c(x_k))$ . Specifically,

$$\begin{aligned} \frac{\partial \tilde{g}_h}{\partial a_j} &= h c(x_k, t_k) b_j(x_k) \\ \frac{\partial S_h}{\partial a_j} &= -\frac{h}{\epsilon \sqrt{2}} \sum_{k=0}^{N-1} \left( \frac{x_{k+1} - x_k}{h} + \nabla V(x_k) - \sqrt{2} c(x_k) \right) b_j(x_k) \\ \frac{\partial Z_h}{\partial a_j} &= -Z_h \mathbf{E}_{Q_h}^{\tilde{X}_0} \left[ \frac{\partial S_h}{\partial a_j} \right]. \end{aligned}$$

Together with the projection property of the conditional expectation this gives

$$\frac{\partial \tilde{I}_h}{\partial a_j} = \mathbf{E}_{Q_h}^{\tilde{X}_0} \left[ \frac{\partial \tilde{g}_h}{\partial a_j} \right] + \mathbf{C}_{Q_h}^{\tilde{X}_0} \left[ \tilde{g}_h, \frac{\partial S_h}{\partial a_j} \right], \quad (\text{A.4})$$

where  $\mathbf{C}_{Q_h}$  denotes the covariance operator

$$\mathbf{C}_{Q_h}[u, v] = \mathbf{E}_{Q_h}[uv] - \mathbf{E}_{Q_h}[u] \mathbf{E}_{Q_h}[v].$$

*Inexact gradient* We are interested in the situation when  $\tau$  in (3.6) is a random stopping time rather than a fixed time; otherwise the optimal control policy would be a function of time, i.e.,  $u_t = c(X_t, t)$ . But in case that  $\tau$  is a first entry time of a set  $S \subset \mathbb{R}^n$ , this stopping time  $\tau = \tau(c)$  will be a function of the control. Hence the derivative of the cost functional with respect to the unknown control coefficients  $a_j$  would involve additional derivatives of  $\tau$  or its time-discrete counterpart  $N_\tau$ ; for example, for the discretized running cost this would result in an expression like

$$\frac{\partial}{\partial a_j} \sum_{k=0}^{N_\tau-1} \tilde{g}_h(x_k, c(x_k)) = \tilde{g}_h(x_{N_\tau-1}, c(x_{N_\tau-1})) \frac{\partial N_\tau}{\partial a_j} + \sum_{k=0}^{N_\tau-1} \frac{\partial \tilde{g}_h}{\partial a_j}$$

In principle the dependence of the stopping time on the control variable can be made explicit in terms of the solution to an elliptic boundary value problem for  $\tau$ , yet it is unclear how terms such as  $\partial N_\tau / \partial a_j$  can be handled numerically efficiently.

In many cases the gradient descent will also converge even though the gradient  $\nabla \tilde{I}$  is not exact, and it turns out that the boundary cost in the last equations is

typically small compared to the accumulated cost. Ignoring the contribution from the boundary terms in the derivatives hence gives a gradient descent method with inaccurate gradient. In our numerical example where  $f = \sigma$  is constant, the inexact gradient reads

$$\begin{aligned} \frac{\partial \tilde{I}_h}{\partial a_j} &= h \mathbf{E}_{Q_h^{\tilde{X}_0}} \left[ \sum_{k=0}^{N_\tau-1} c(x_k) b_j(x_k) \right] \\ &\quad - \frac{h^{3/2}}{\epsilon \sqrt{2}} \mathbf{C}_{Q_h^{\tilde{X}_0}} \left[ \sum_{k=0}^{N_\tau-1} \left( \sigma + \frac{1}{2} |c(X_k)|^2 \right), \sum_{k=0}^{N_\tau-1} \eta_{k+1} b_j(x_k) \right] \\ &= h \mathbf{E}_{Q_h^{\tilde{X}_0}} \left[ \sum_{k=0}^{N_\tau-1} c(x_k) b_j(x_k) \right] \\ &\quad - \frac{h^{3/2}}{\epsilon \sqrt{2}} \mathbf{E}_{Q_h^{\tilde{X}_0}} \left[ \left( \sum_{k=0}^{N_\tau-1} \left( \sigma + \frac{1}{2} |c(X_k)|^2 \right) \right) \sum_{k=0}^{N_\tau-1} \eta_{k+1} b_j(x_k) \right], \end{aligned}$$

where  $N_\tau = \lceil \tau/h \rceil$  is the discrete analog of the first hitting time (here  $\lceil x \rceil$  is the nearest integer larger than  $x$ ), and we used the fact that  $b_j(X_k)$  and  $\eta_{k+1}$  are independent.

## References

- [1] R. J. Allen, P. B. Warren, and P. R. ten Wolde. Sampling rare switching events in biochemical networks. *Physical Review Letters*, 94:018104, 2005.
- [2] J. Bucklew. *Introduction to Rare Event Simulation*. Springer, New York, 2004.
- [3] D. Chandler. Finding transition pathways: Throwing ropes over rough mountain passes, in the dark. In G. Ciccotti B. J. Berne and D. F. Coker, editors, *Computer Simulation of Rare Events and Dynamics of Classical and Quantum Condensed-Phase Systems – Classical and Quantum Dynamics in Condensed Phase Simulations*, pages 51–66. World Scientific, 1998.
- [4] C. Chipot and A. Pohorille. *Free Energy Calculations: Theory and Applications in Chemistry and Biology*. Springer Series in Chemical Physics, Vol. 86. Springer, Berlin, 2007.
- [5] G.E. Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. *J. Stat. Phys.*, 90:1481–1487, 1998.
- [6] P. Dai Pra, L. Meneghini, and W. Runggaldier. Connections between stochastic control and dynamic games. *Math. Control Signals Systems*, 9:303–326, 1996.
- [7] E. Darve, D. Rodriguez-Gomez, and A. Pohorille. Adaptive biasing force method for scalar and vector free energy calculations. *J. Chem. Phys.*, 128:144120, 2008.
- [8] P.-T. De Boer, D. Kroese, S. Mannor, and R. Rubinstein. A tutorial on the cross-entropy method. *Ann. Oper. Res.*, 134:19–67, 2005.
- [9] Ch. Dellago, P. Bolhuis, and D. Chandler. Efficient transition path sampling: Application to lennard-jones cluster rearrangements. *The Journal of Chemical Physics*, 108(22):9236, 1998.
- [10] P. Dupuis and W.M. McEneaney. Risk-sensitive and robust escape criteria. *SIAM J. Control Optim.*, 35:2021–2049, 1997.
- [11] D. Dürr and A. Bach. The Onsager-Machlup function as Lagrangian for the most probable path of a diffusion process. *Comm. Math. Phys.*, 60:153–170, 1978.
- [12] K. Dvijotham and E. Todorov. Linearly-solvable optimal control. In F.L. Lewis and Liu D., editors, *To appear in: Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*, chapter 6. Wiley & Sons, 2012.
- [13] W. E, W. Ren, and E. Vanden-Eijnden. Finite temperature string method for the study of rare events. *J. Phys. Chem. B*, 109:6688–6693, 2005.
- [14] A.K. Faradjian and R. Elber. Computing time scales from reaction coordinates by milestoning. *J. Chem. Phys.*, 120:10880–10889, 2004.
- [15] W.H. Fleming. Exit probabilities and optimal stochastic control. *Appl. Math. Optim.*, 4:329–346, 1977.
- [16] W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer, 2006.

- [17] H. Grubmüller. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E*, 52:2893–2906, 1995.
- [18] J.M. Horowitz and S. Vaikuntanathan. Nonequilibrium detailed fluctuation theorem for repeated discrete feedback. *Phys. Rev. E*, 82:061120, 2010.
- [19] G. Hummer and A. Szabo. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl. Acad. Sci. USA*, 98:3658–3661, 2001.
- [20] M. James. Asymptotic analysis of nonlinear stochastic risk-sensitive control and differential games. *Math. Control Signals Systems*, 5:401–417, 1992.
- [21] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78:2690–2693, 1997.
- [22] H.J. Kappen. Path integrals and symmetry breaking for optimal control theory. *J. Stat. Mech. Theor. Exp.*, 2005(11):P11011, 2005.
- [23] J. Kurchan. Fluctuation theorem for stochastic dynamics. *J. Phys. A: Math. Gen.*, 31:3719–3729, 1998.
- [24] A. Laio and M. Parrinello. Escaping free-energy minima. *PNAS*, 99:12562–12566, 2002.
- [25] J. Latorre, C. Hartmann, and Ch. Schütte. Free energy computation by controlled Langevin processes. *Procedia Computer Science*, 1:1591–1600, 2010.
- [26] J.L. Lebowitz and H. Spohn. A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics. *J. Stat. Phys.*, 95:333–365, 1999.
- [27] T. Lelièvre, G. Stoltz, and M. Rousset. *Free Energy Computations: A Mathematical Perspective*. Imperial College Press, 2010.
- [28] R. Merkel, P. Nassoy, A. Leung, K. Ritchie, and E. Evans. Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature*, 397:50–53, 1999.
- [29] D.L.D. Minh and J.D. Chodera. Optimal estimators and asymptotic variances for nonequilibrium path-ensemble averages. *J. Chem. Phys.*, 131:134110, 2009.
- [30] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, New York, 1999.
- [31] B.K. Øksendal. *Stochastic Differential Equations: An Introduction With Applications*. Springer, 2003.
- [32] K. Rawlik, M. Toussaint, and S. Vijayakumar. On stochastic optimal control and reinforcement learning by approximate inference. In *Proc. Robotics: Science and Systems Conference (RSS '12)*, 2012 (in press).
- [33] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, and H.E. Gaub. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science*, 276:1109–1112, 1997.
- [34] R.Y. Rubinstein and D.P. Kroese. *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning*. Springer, New York, 2004.
- [35] T. Sagawa and M. Ueda. Generalized Jarzynski equality under nonequilibrium feedback control. *Phys. Rev. Lett.*, 104:090602, 2010.
- [36] K. Schulten and S. Park. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.*, 120:5946–5961, 2004.
- [37] Ch. Schütte, S. Winkelmann, and C. Hartmann. Optimal control of molecular dynamics using markov state models. *Math. Program. Series B*, 134:259–282, 2012.
- [38] E. Todorov. Efficient computation of optimal actions. *Proc. Natl. Acad. Sci. USA*, 106(28):11478–11483, 2009.
- [39] A.F. Voter. Hyperdynamics: Accelerated molecular dynamics of infrequent events. *Phys. Rev. Lett.*, 78:3908–3911, 1997.
- [40] P. Whittle. Risk-sensitivity, large deviations and stochastic control. *Eur. J. Oper. Res.*, 73:295–303, 1994.