Projection-based algorithms for optimal control and importance sampling of diffusions

Carsten Hartmann[1], **Christof Schütte**[1],[2] **and Wei Zhang**[3]

[1] Institute of Mathematics, Freie Universität Berlin, Arnimallee 6, 14195 Berlin, Germany

[2] Zuse Institute Berlin, Takustrasse 7, 14195 Berlin, Germany

E-mail: chartman@mi.fu-berlin.de, schuette@zib.de, wei.zhang@fu-berlin.de

Abstract. We propose numerical algorithms for solving complex, high-dimensional control and importance sampling problems based on reduced-order models. The algorithms approach the "curse of dimensionality" by a combination of model reduction techniques for multiscale diffusions and stochastic optimization tools, with the aim of reducing the original, possibly high-dimensional problem to a lower dimensional representation of the dynamics, in which only few relevant degrees of freedom are controlled or biased. Specifically, we study situations in which either an suitable set of slow collective variables onto which the dynamics can be projected is known, or situations in which the dynamics shows strongly localized behaviour in the small noise regime. The idea is to use the solution of the reduced-order model as a predictor of the exact solution that, in a corrector step, can be used together with the original dynamics, where no explicit assumptions about small parameters or scale separation have to be made. We illustrate the approach with simple, but paradigmatic numerical examples.

AMS classification scheme numbers: 65C20, 93E20, 82C31

Keywords: Stochastic optimal control, importance sampling, collective variables, conditional expectation, transition path, cross entropy method.

1. Introduction

Stochastic differential equation models are ubiquitous in real world applications from various different disciplines, such as molecular dynamics [1], economy [2], or climate science [3]. As a typical feature, these often high-dimensional models display dynamics on vastly different time and length-scales, which makes their analysis and numerical simulation a nontrivial task. In most situations of practical relevance, however, these different scales are either not well resolved or the equations of motion are too complicated to explicitly identify scaling parameters that would make the equations amenable to asymptotic methods.

In this work we focus on stochastic dynamics with multiple time scales, without assuming that the system admits an explicit separation into slow and fast variables. Specifically, we are interesting in solving stochastic control problems under these dynamics that may arise in an optimal control context or in the estimation of rare events based on variance minimizing Monte Carlo schemes. There are various model reduction techniques that approximate a complex (i.e. large-scale or multiscale) stochastic dynamics without assuming explicit scale separation or resolution of the dynamics on all scales. Typically, these methods assume that a suitable set of collective variables or macroscopic observables that describe the essential behaviour of the system is known and then seek a statistical closure by projecting the dynamics onto these variables [4, 5, 6, 7, 8]; cf. also [9, 10] and the references therein. If the collective variables are chosen appropriately, then the (forward) dynamics as a function of the control is well approximated by the projected equation (cf. [6]). We shall refer to this as *forward stability* of the approximation scheme. However, here the situation is that the controls will be functions of the dynamics in the sense of a feedback law, and the question is whether the optimal controls for the projected equations—that are easily computable—yield reasonable approximations of the optimal control for the original equation and therefore can be used to control the original dynamics. Approximations with this property are referred to as *backward stable*. The question of backward stability of the projected dynamics has been recently analyzed in the context of averaging and homogenization of multiscale diffusions [11, 12]. (Roughly speaking, the loss of backward stability of some standard approximation schemes for optimal control has to do with the fact that the associated dynamic programming or *Hamilton-Jacobi-Bellman* equations are nonlinear, whereas the approximation is based on the linear evolution equations of the forward dynamics.)

The main purpose of this work is to design backward stable numerical algorithms for solving certain types of optimal control and sampling problems. For both optimal control and sampling, we devise algorithms to compute suboptimal, but nearly optimal control policies based on reduced-order models. Our approximation schemes, the associated algorithms and the terminology used are partially based on ideas from dynamical systems, in particular molecular dynamics. One concept borrowed from molecular dynamics is the concept of *transition pathways* that play a key role in the study of transition events between metastable sets of a stochastic dynamical system. Transition pathways are curves in state space along which transition events are likely to occur and along which typical trajectories concentrate (see, e.g., [13, 14, 15, 16]; cf. also [17, 18]). Once a transition pathway is known, it is possible to design locally optimal control policies based on the information that the dynamics concentrate in a certain region of phase space, using any stochastic optimization technique such as the

cross-entropy method (see [19] or [20] and the references therein) with suitably placed ansatz functions. A related notion that we will adopt is known by the name of *reaction coordinate*, a generalized coordinate by which the dynamics along the transition pathway can be parametrized [21, 22]. Putting this into a broader perspective, one may replace the term *reaction coordinate* by *collective variable* or any other macroscopic observable that describes the essential behaviour of the system. By projecting the dynamics onto these variables, one then obtains a low-dimensional equation for the essential dynamics (after applying an appropriate closure, e.g., conditional expectations [5, 6]), and, as the projected equations are low-dimensional, it is possible to solve the corresponding optimal control by means of traditional PDE-based discretization techniques that would be out of scope otherwise.

The article is organized as follows: In Section 2, we introduce the indefinite time horizon stochastic control problem and its dual sampling problem that represent the two basic computational problems to be solved by the methods considered herein. Section 3 is devoted to projected dynamics based on conditional expectations with respect to a given set of collective variables, a method also called *first-order optimal prediction*. We formulate the projected dynamic programming equations for the underlying control and sampling problems and sketch a numerical discretization for computing suboptimal control policies that are active only in the direction of the collective variables. In Section 4 we recall the string method for computing transition pathways and explain how it can be used together with the cross-entropy method to solve stochastic control and importance sampling problems. Numerical examples are presented in Section 5. The article concludes with a brief discussion of possible issues and future directions in Section 6.

2. Two related stochastic control problems

In this paper, we consider diffusion processes governed by stochastic differential equations (SDE) of the form

$$dx_s = -\nabla V(x_s)ds + \sqrt{2\beta^{-1}}dw_s\,, \tag{1}$$

where $x_s \in \mathbb{R}^n$ is the system state at time $s \geq 0$, $V\colon \mathbb{R}^n \to \mathbb{R}$ is a sufficiently smooth (e.g. C^∞) potential energy function, $\beta > 0$ is an arbitrary scaling factor for the noise (inverse temperature), and w_s is an n-dimensional standard Wiener process.‡ We moreover consider the process to be confined to some open and bounded set $\Omega \subset \mathbb{R}^n$ with sufficiently smooth boundary, and define the stopping time $\tau = \inf\{s > 0\colon x_s \notin \Omega\}$ to be the first exit time from Ω.

As the next step, we will introduce two apparently unrelated functionals of the dynamics (1), namely, an exponential expectation over realizations of (1) and a closely related stochastic optimal control problem. The two formulations turn out to be related by a Legendre-type duality principle that will be important in the course of the article. See [23, 12, 19, 11] for further background.

‡ It is possible to generalize most of the considerations in this article to the case of general, non-degenerate diffusions of the form $dx_s = b(x_s)dt + \sigma(x_s)dw_s$, but for the sake of clarity, we refrain from doing so.

2.1. Stochastic optimal control of diffusions

To begin with, we consider the following optimal control problem: minimize

$$J(u;x) = \mathbf{E}_{\mu_u}\Big[\int_0^\tau \Big(G(x_s^u) + \frac{1}{2}|u_s|^2\Big)ds + H(x_\tau^u) \;\Big|\; x_0^u = x\Big]\,, \tag{2}$$

over all admissible controls $u \in \mathcal{A}$ and subject to the dynamics

$$dx_s^u = \Big(\sqrt{2}u_s - \nabla V(x_s^u)\Big)\,ds + \sqrt{2\beta^{-1}}dw_s^u\,, \quad x_0^u = x\,, \tag{3}$$

that is is a controlled version of (1). Here $G, H\colon \mathbb{R}^n \to \mathbb{R}^+$ are non-negative and continuous running and terminal cost functions, and the set $\mathcal{A}$ of admissible controls consists of all bounded and measurable functions $u\colon [0,\infty) \to A \subset \mathbb{R}^n$, such that (3) admits a unique strong solution. The expectation in (2) is taken over all realizations of the controlled dynamics (3), which we denote by an expectation with respect to a probability measure μ_u on the space $C([0,\infty),\mathbb{R}^n)$ of continuous trajectories. In the controlled SDE, w_s^u is again an n-dimensional standard Wiener process where the superscript "u" is only used to distinguish it from w_s in (1), when it comes to the application of the Girsanov's theorem in Subsection 2.2. It will be dropped later on when it is convenient. (The scaling factor $\sqrt{2}$ in front of the control is for notational convenience only.)

Now let

$$U(x) = \inf_{u\in\mathcal{A}} J(u;x) \tag{4}$$

be the value function or *optimal cost-to-go* associated with (2)–(3). Intuitively speaking, computing the control value U means to drive the dynamics (3) starting from $x_0^u = x \in \Omega$ until it reaches the set boundary $\partial\Omega$, in such a way that the cost (2) is minimized. Under certain regularity conditions on the coefficients V, G, H and the boundary of the domain Ω, it can be shown (see, e.g., [23]) that U satisfies a Hamilton-Jacobi-Bellman type boundary value PDE:

$$\begin{aligned} 0 &= \min_{\alpha\in\mathbb{R}^n}\Big\{\beta^{-1}\Delta U + (\sqrt{2}\alpha - \nabla V)\cdot\nabla U + G + \frac{1}{2}|\alpha|^2\Big\} \\ H &= U|_{\partial\Omega}\,. \end{aligned} \tag{5}$$

or, equivalently,

$$\begin{aligned} 0 &= \beta^{-1}\Delta U - \nabla V\cdot\nabla U - |\nabla U|^2 + G \\ H &= U|_{\partial\Omega}\,. \end{aligned} \tag{6}$$

For this type of optimal control problem, it is known that the infimum of $J(u)$ is attained, with a unique minimizer $\hat{u}$ that is given by (see, e.g., [23])

$$\hat{u}_s = -\sqrt{2}\nabla U(x_s^u), \tag{7}$$

where x_s^u is system's state. Note that the optimal control $\hat{u}$ is a Markovian feedback control, i.e. the action u_t at time t depends only on the system state x_t^u at time t.

2.2. Adaptive importance sampling

The stochastic control problem (2)–(3) has a well-known dual formulation in terms of a sampling problem [24, 25, 26, 27, 28]. Consider again the uncontrolled dynamics (1) and define the function ψ by

$$\psi(x) = \mathbf{E}\left[\exp\left(-\beta\int_0^\tau G(x_s)ds - \beta H(x_\tau)\right)\,\middle|\, x_0 = x\right]. \tag{8}$$

Here the expectation is taken over all realizations of (1) with initial condition $x_0 = x$. By the Feymann-Kac formula (e.g., [23, App. D]), $\psi(x)$ satisfies the PDE

$$\begin{aligned} 0 \quad &= \beta^{-1}\Delta\psi - \nabla V\cdot\nabla\psi - \beta G\psi\,, \quad x\in\Omega\,,\\ e^{-\beta H} &= \psi|_{\partial\Omega}\,. \end{aligned} \tag{9}$$

Clearly, computing $\psi(x)$ by solving the above boundary value problem is numerically infeasible if the state space is high-dimensional. We could compute (8) by Monte Carlo, simulating a sufficiently large number of independent realizations of (1) until they leave the domain Ω. However, in most of the relevant cases, this straightforward method suffers from poor convergence of the corresponding Monte Carlo estimator, especially when τ is large, i.e. when escaping from Ω is a rare event.

As a remedy, we could try to generate realizations from the controlled SDE, with a control such that reaching the set boundary $\partial\Omega$ is no longer a rare event, and then account for this change of the underlying probability measure by reweighting the Monte Carlo estimator with the likelihood ratio between the two distributions. This is the idea behind importance sampling, and, in our case estimators, it can be based on the reweighted expectation

$$\psi(x) = \mathbf{E}_{\mu_u}\left[\exp\left(-\beta\int_0^\tau G(x_s^u)ds - \beta H(x_\tau^u)\right)Z_\tau^{-1}\,\middle|\, x_0 = x\right], \tag{10}$$

where x_s^u is the solution of the controlled SDE (3) with—not necessarily optimal—control u. The random variable $Z = d\mu_u/d\mu$ is the Radon-Nikodym derivative (likelihood ratio) between the probability measures of the controlled and the uncontrolled dynamics:

$$\begin{aligned} Z_\tau &= \exp\left(\beta^{1/2}\int_0^\tau u_s\cdot dw_s - \frac{\beta}{2}\int_0^\tau |u_s|^2 ds\right)\\ &= \exp\left(\beta^{1/2}\int_0^\tau u_s\cdot dw_s^u + \frac{\beta}{2}\int_0^\tau |u_s|^2 ds\right), \end{aligned} \tag{11}$$

with the shorthand $dw_s^u = dw_s - \beta^{\frac{1}{2}}u_s ds$. By Girsanov's theorem [29, Thm. 8.6.8], the μ^u-law of the transformed Brownian motion w_τ^u is the same as the μ-law of the standard Brownian motion w_τ, where $d\mu^u|_{\mathcal{F}_\tau} = Z_\tau d\mu|_{\mathcal{F}_\tau}$ where $\mathcal{F}_\tau$ is the filtration generated by w_s up to time $s = \tau$, hence (8) and (10) agree (see Remark 1 below)

Note that even though a standard Monte-Carlo estimator based on (10) is an unbiased estimator of ψ for any admissible control u, the variance of the estimator, and hence its convergence, heavily depends on the choice of u. As a consequence, a careful design of the importance sampling estimator (i.e., a clever choice of u) is crucial for obtaining efficient Monte Carlo estimators.

Efficiency of importance sampling In term of computational cost, the efficiency of importance sampling will depend on both the average length τ of simulated trajectories and the variance of the estimator. To simplify notation, call $\mathcal{I}$ the user-specified statistical estimator of (10) with $\psi(x) = \mathbf{E}_{\mu_u}[\mathcal{I}]$. The standard deviation of $\mathcal{I}$ is

$$\sqrt{\mathrm{Var}_{\mu_u}[\mathcal{I}]} = \sqrt{\mathbf{E}_{\mu_u}[\mathcal{I}^2] - (\mathbf{E}_{\mu_u}[\mathcal{I}])^2}\,.$$

Now suppose that we have generated N independent realizations of x_s^u, and let $\mathcal{I}_i$ denote the estimator based on the i-th realization where $1 \le i \le N$. Then we can approximate $\psi(x)$ by

$$\hat{\mathcal{I}}_N = \frac{1}{N}\sum_{i=1}^{N} \mathcal{I}_i\,,$$

with standard deviation

$$\sigma_{\hat{\mathcal{I}}_N} = \sqrt{\frac{\mathrm{Var}_{\mu_u}[\mathcal{I}_1]}{N}}\,.$$

Given a tolerance $\epsilon > 0$, such that $\sigma_{\hat{\mathcal{I}}_N} \le \epsilon$, it follows that we have to use $N \ge \mathrm{Var}[\mathcal{I}_1]/\epsilon^2$ realizations to approximate ψ within the prescribed accuracy. To quantify the efficiency of the importance sampling scheme, let Δt_u be the constant time step (depending on u) used to numerically compute the realization of the controlled SDE. The computational cost for a single trajectory then is proportional to $\mathbf{E}_{\mu_u}[\tau]/\Delta t_u$, which results in a total computational cost of

$$\mathcal{C}(\epsilon, u) = \frac{\mathrm{Var}_{\mu_u}[\mathcal{I}]}{\epsilon^2} \cdot \frac{\mathbf{E}_{\mu_u}[\tau]}{\Delta t_u}. \tag{12}$$

to achieve the given tolerance ϵ. We call the quantity

$$\mathrm{Accel}(u) = \frac{\mathcal{C}(\epsilon, 0)}{\mathcal{C}(\epsilon, u)} \tag{13}$$

the *acceleration index* of the importance sampling estimator, comparing standard Monte Carlo based on the uncontrolled SDE and importance sampling based on the control u. Note that the acceleration index is independent of ϵ. (When Δt is independent of u, the acceleration index coincides with the definition in [19].) Further note that for

$$\hat{u}_s = \sqrt{2}\beta^{-1}\frac{\nabla\psi(x_s^u)}{\psi(x_s^u)}\,, \tag{14}$$

the variance of the estimator is zero [25], hence $\mathrm{Accel}(\hat{u}) = \infty$.

Remark 1 *A brief comment on the application of the Girsanov's theorem with a random terminal time τ is in order. Let $W = C([0,\infty), \mathbb{R}^n)$ be the space of continuous paths (i.e. realizations of our SDE) of arbitrary length, equipped with the Borel σ-algebra $\sigma(W)$ that is generated by all cylinder sets of the form $\{f \in W : f(t_1) \in E_1, f(t_2) \in E_2, \ldots, f(t_k) \in E_k\}$ where $k \in \mathbb{N}$, $E_i \in \mathcal{B}(\mathbb{R}^n)$ and $0 \le t_1 < t_2 < \ldots < t_k < \infty$. Further let $\mathcal{F}_t = \sigma(\{w_s : s \le t\})$ denote the σ-algebra generated by the Brownian motion up to time $t < \infty$. Then, Girsanov's theorem holds on the measurable space $(W, \sigma(W))$ as long as the family $(Z_t)_{t\ge 0}$ of random variables*

$$Z_t = \exp\Big(\beta^{1/2}\int_0^t u_s \cdot dw_s - \frac{\beta}{2}\int_0^t |u_s|^2 ds\Big) \tag{15}$$

is a uniformly integrable martingale. By Ito's formula [29], $(Z_t)_{t\geq 0}$ is a nonnegative local martingale, which is uniformly integrable if Novikov's condition holds:

$$\mathbf{E}\left[\exp\left(\frac{\beta}{2}\int_0^\infty |u_s|^2 ds\right)\right] < \infty$$

Now given an admissible control u that is defined up to a random stopping time τ, we define a stopped version of the control by

$$\tilde{u}_s = \begin{cases} u_s & s \leq \tau\,, \\ 0 & s > \tau. \end{cases}$$

The stopped control $\tilde{u}$ satisfies Novikov's condition and thus the change of measure that is the basis of (10)–(11) is well-defined on $\mathcal{F}_\tau$ by Doob's optional stopping theorem.

2.3. Duality between optimal control and importance sampling

From the previous consideration, it follows that both the optimal control problem (4) and the importance sampling problem (10)–(11) boil down to finding an optimal feedback control, either (7) or (14). It turns out that the sought optimal controls are the same and that the functions ψ and U are related by

$$\psi = \exp(-\beta U)\,, \tag{16}$$

as can be readily seen by checking that $\exp(-\beta U(x))$ solves linear PDE (9); (16) is called a *logrithmic transformation* [23]. By a slight variation of Theorem 3.1 in [23, Sec. VI.3], the solution to (9) is $C^2(\Omega)\cap C(\overline{\Omega})$, under mild additional regularity and growth conditions on the coefficients and the domain, hence the assertion follows from chain rule [12, 19]. This proves the important duality relation (cf. [30, 31]):

Lemma 1 *Let x_s and x_s^u denote the solutions of (1) and (3), respectively. Further let $G, H \geq 0$ be bounded. Then*

$$\begin{aligned} -\beta^{-1}\ln \mathbf{E}\Big[\exp\Big(-\beta\Big(\int_0^\tau G(x_s)ds + H(x_\tau)\Big)\Big)\;\Big|\; x_0 = x\Big] \\ = \min_{u\in\mathcal{A}} \mathbf{E}_{\mu_u}\Big[\int_0^\tau \Big(G(x_s^u) + \frac{1}{2}|u_s|^2\Big)ds + H(x_\tau^u)\;\Big|\; x_0^u = x\Big]\,. \end{aligned} \tag{17}$$

We conclude that the optimal control (7) that is the unique minimizer of (2)–(3) coincides with the optimal change of measure based on (14) that leads to a zero-variance importance sampling estimator of (8). This clearly implies that computing the optimal change of measure is as difficult as solving the dynamic programming equation (5) of the optimal control problem. On the other hand, it is neither possible (in high dimension) nor neccessary to find the *optimal* control to compute either (8) or the value function (4) exactly, since we can always sample the exact solution using the importance sampling identity (10)–(11). The idea is that, even with a suboptimal control policy based on, e.g., an approximation of the SDE, we can use reweighting to compute an unbiased estimate of ψ and hence a (biased) estimate of $U = -\beta^{-1}\ln\psi$. By continuity of the cost functional J as a function of the control, suboptimal controls that are close to the optimal control (in some norm) will lead to efficient estimators, in that they have small variance and require only short trajectories.

Motivated by these considerations, our purpose is to design numerical algorithms for computing cheap suboptimal controls u, without solving a PDE like (5) or (9).

3. Projection onto essential coordinates

In this section, we derive approximations of the optimal control problem (4) and its dual sampling problem (8) under the assumption that a collective variable is known that describes the essential dynamics of the system. We will first recall some facts on model reduction using collective variables (first-order optimal prediction in the sense of [4]), and then discuss algorithms for solving (4) or (8) in this situation.

3.1. First-order optimal prediction of diffusions

We suppose that there is a smooth function $\xi\colon \mathbb{R}^n \to \mathbb{R}^k$, $1 \le k < n$, such that the level sets $\xi^{-1}(z)$ are smooth submanifolds of codimension k in $\mathbb{R}^n$ for all values $z \in \mathbb{R}^k$ of ξ, in other words, the Jacobian $\nabla\xi(x)$ has constant rank k for all $x \in \mathbb{R}^n$.

Now let $\Sigma_z = \xi^{-1}(z)$ denote the isosurface of ξ for any regular value $z \in \mathbb{R}^k$. We borrow some terminology from molecular dynamics and call ξ the *essential coordinate*, *reaction coordinate* or *collective variable*. The rationale behind projecting the SDE (1) onto the collective variable ξ is that the value $z_s = \xi(x_s)$ varies slowly, whereas the orthogonal dynamics on Σ_{z_s} is fast and therefore relaxes to its (conditional) invariant measure while the slow variable $z_s \approx z$ stays approximately constant—provided that ξ is carefully chosen. As a consequence, the influence of the fast variables on the essential dynamics is only through their average, where averages are understood with respect to the invariant measure of the fast variables condition on the slow variables, and thus it makes sense to construct a closure of the ξ-dynamics by taking conditional expectations of drift and diffusion coefficients in (1).

Closure scheme We shall briefly recall the first-order optimal prediction approach using conditional expectations that goes back to [6, 32]; cf. also [4, 5]. To this end, let x_s be the solution of (1). By Ito's formula, $\xi(x_s)$ satisfies the SDE

$$d\xi = \big(\beta^{-1}\Delta\xi(x_s) - \nabla\xi(x_s)^T\nabla V(x_s)\big)\,ds + \sqrt{2\beta^{-1}}\nabla\xi(x_s)^T\,dw_s \quad (18)$$

where we use the convention that the Jacobian matrix $\nabla\xi$ of $\xi\colon \mathbb{R}^n \to \mathbb{R}^k$ is an $(n\times k)$-matrix. (Gradients are represented by column vectors.) Note that (18) is not a closed equation for ξ, but rather depends on the solution of (1). To obtain a differential equation for ξ only, we need an appropriate closure scheme for (18). One such closure is obtained by taking conditional expectations with respect to the invariant probability distribution of (1) with density

$$\rho(x) = \exp(-\beta V(x))\,, \qquad \int_{\mathbb{R}^n} \rho(x)\,dx = 1\,, \quad (19)$$

where we assume without loss of generality that $\exp(-\beta V)$ is normalized to have total probability mass one. The corresponding conditional expectation, given that $\xi(x) = z$, reads

$$\mathbf{E}_\rho[f|\xi = z] = \frac{1}{Q(z)}\int_{\Sigma_z} f\rho\,(\mathrm{vol}(\nabla\xi))^{-1}\,d\sigma_z\,, \quad (20)$$

with $\mathrm{vol}(A) = \sqrt{\det(A^TA)}$ denoting the matrix volume of a (not necessarily square) matrix, σ_z the Riemannian volume element of Σ_z, and Q the normalization constant

$$Q(z) = \int_{\Sigma_z} \rho\,(\mathrm{vol}(\nabla\xi))^{-1}\,d\sigma_z\,. \quad (21)$$

Taking conditional expectations of the right hand side (18), i.e., computing the best approximation with respect to the ρ-weighted L^2-norm, we obtain a closed equation for the essential variable:

$$dz_s = b(z_s)ds + \sqrt{2\beta^{-1}}\sigma(z_s)\, dw_s, \quad z_0 = \xi(x_0)\,, \tag{22}$$

Here $z_s \in \mathbb{R}^k$, w_s is k-dimensional Wiener process and the coefficients b, σ are given by

$$\begin{aligned} b(z) &= \mathbf{E}\left[\beta^{-1}\Delta\xi - \nabla\xi^T\nabla V \,|\, \xi = z\right]\,, \\ a(z) &= \mathbf{E}\left[\nabla\xi^T\nabla\xi \,|\, \xi = z\right]\,, \end{aligned} \tag{23}$$

where $a = \sigma\sigma^T$, and $\Delta\xi$ in the first equation is understood component-wise. We call (22)–(23) the *effective dynamics* of along ξ. For the case of a scalar essential variable, the approximation error of the effective dynamics has been analyzed in [6]. Numerical schemes to compute (23) by constrained simulations have been studied in [33, 34].

For later use, we also define the thermodynamic free energy with respect to ξ as

$$F(z) = -\beta^{-1}\ln Q(z)\,, \tag{24}$$

with Q being the normalization constant from (20)–(21).

3.2. Dynamic programming based on first-order optimal prediction

The aim of this paragraph is to derive an effective equation for the control problem (2)–(3) in terms of the nonlinear dynamic programming equation (5). By the duality principle, Lemma 1, this will give us an analogous effective equation for the corresponding sampling problem (8). We will use the duality principle in reverse order and start from the linear boundary value problem (9).

Further assumptions The following notation and additional assumptions will be used throughout this section.

(a) Let $\Omega \subset \mathbb{R}^n$ be a bounded open set, in which the dynamics takes place, and assume that there exists an open and bounded set $S \subset \mathbb{R}^k$ such that $\Omega = \xi^{-1}(S) \subset \mathbb{R}^n$.

(b) The random variable

$$\tau = \inf\{s > 0 \colon \xi(x_s) \notin S\}$$

is a stopping time with respect to the filtration generated by $\xi(x_s)$.

(c) There exists a continuous and bounded function h such that

$$H(x) = h(\xi(x))\,,$$

i.e., the terminal cost H depends only on ξ.

Petrov-Galerkin method The starting point of our formal derivation is the integral form of the elliptic PDE (9) based on test functions that are constant on the level sets Σ_z for every $z \in \mathbb{R}^k$. To this end, we call $L^2(\Omega, \rho)$ the space of measurable functions on Ω that are square integrable with respect to the Boltzmann density ρ and consider N smooth functions $\psi_i \in L^2(\Omega, \rho)$ of the form $\psi_i(x) = f_i(\xi(x))$ for $x \in \Omega$, where the $f_i \colon S \to \mathbb{R}$, $1 \le i \le N$ are smooth functions that vanish on the boundary

∂S. Further let ψ be the unique strictly positive solution of (9). We suppose that $\psi \in \operatorname{span}\{\psi_1, \psi_2, \cdots, \psi_N\} + \{e^{-\beta H}\} \subset L^2(\Omega, \rho)$, in other words,

$$\psi = \sum_{j=1}^{N} a_j \psi_j + e^{-\beta H}\,, \tag{25}$$

in accordance with the boundary condition in (9). Multiplying (9) by a test function $\phi(x) = \widetilde{\phi}(\xi(x))$, $\phi \in L^2(\Omega, \rho)$ and integrating over the domain Ω with respect to ρ, we obtain

$$\begin{aligned} &\sum_{j=1}^{N} a_j \int_\Omega \left(\beta^{-1}\Delta\psi_j - \nabla V \cdot \nabla\psi_j\right) \phi\, \rho\, dx \\ &\qquad = \beta \sum_{j=1}^{N} a_j \int_\Omega G\psi_j \phi\, \rho\, dx + \beta \int_\Omega G e^{-\beta H} \phi\, \rho\, dx\,. \end{aligned} \tag{26}$$

To derive an equation for ξ, we call ∇_ξ the gradient with respect to ξ. By chain rule,

$$\begin{aligned} \nabla_\xi \psi_j|_{\xi=z} &= (\nabla\xi|_{\xi=z}) \nabla f_j(z)\,, \\ \Delta_\xi \psi_j|_{\xi=z} &= (\nabla\xi^T \nabla\xi|_{\xi=z})\colon \nabla^2 f_j(z) + \nabla f_j(z) \cdot \Delta\xi|_{\xi=z}\,, \end{aligned}$$

for all functions $\psi_j(x) = f_j(\xi(x))$ where $\Delta\xi$ in the second equation should be interpreted as a column vector with components $\Delta\xi_i$. Since the test functions are constant on the level sets Σ_z, the total law of expectation (equivalently: the co-area formula) and the stability of the conditional expectation $\mathbf{E}_\rho[\cdot|\xi = z]$ imply that (26) is equivalent to

$$\begin{aligned} &\sum_{j=1}^{N} a_j \int_S \left(\beta^{-1} a\colon \nabla^2 f_j + b \cdot \nabla f_j\right) \widetilde{\phi}\, e^{-\beta F} dz \\ &= \beta \sum_{j=1}^{N} a_j \int_S g f_j\, \widetilde{\phi}\, e^{-\beta F} dz + \beta \int_S g e^{-\beta h}\, \widetilde{\phi}\, e^{-\beta F} dz \end{aligned}$$

where F is the free energy (24), $g(z) = \mathbf{E}_\rho[G|\xi = z]$ is the average running cost, and all functions such as f_j, $\widetilde{\phi}$ etc. are evaluated at $\xi = z$. The last equation can be viewed as the Galerkin discretization of the linear boundary value problem

$$\begin{aligned} 0 \quad &= b \cdot \nabla\varphi + \beta^{-1} a\colon \nabla^2\varphi - \beta g\varphi, \quad z \in S \\ e^{-\beta h} &= \varphi|_{\partial S}\,. \end{aligned} \tag{27}$$

for $\varphi = \varphi(z)$, which is a discretization of (27) on $\operatorname{span}\{f_1, f_2, \cdots, f_N\} + \{e^{-\beta h}\}$, considered as an affine subspace of $L^2(S, \bar{\rho})$ with $\bar{\rho}(z) \propto \exp(-\beta F(z))$.

Equation (27) is the best approximation of (9) in $L^2(\Omega, \rho)$ as a function of the essential variable ξ. The next statement is a direct consequence of Lemma 1, the dynamic programming principle (5) and the Feynman-Kac formula (cf. [12, 35]):

Theorem 1 *Let z_s^v be the solution of (22) and the controlled projected SDE*

$$dz_s^v = b(z_s^v)ds + \sqrt{2}\sigma(z_s^v)v_s\, ds + \sqrt{2\beta^{-1}}\sigma(z_s^v)\, dw_s\,. \tag{28}$$

with w_s standard k-dimensional Brownian motion and coefficients b, σ as given by (23). Further let φ be the solution of the projected boundary value problem (27). Then

$$-\beta^{-1} \ln \varphi(z) = \inf_{v \in \mathcal{A}_\xi} \mathbf{E}_{\nu_v} \left[\int_0^\tau g(z_s^v) + \frac{1}{2}|v_s|^2 ds + h(z_\tau^v) \,\middle|\, z_0^v = z \right] \tag{29}$$

where $\tau = \inf\{s > 0 \,|\, z_s^v \notin S\}$, ν_v *is the path probability measure associated with (28), and* $\mathcal{A}_\xi$ *is the set of all bounded measurable controls adapted to* w_s*, such that (28) has a strong solution. The optimal control, for which the minimum is attained, is unique and is given by*

$$\hat{v}_s = \hat{v}(z_s^v) = \sqrt{2}\beta^{-1}\frac{\nabla\varphi(z_s^v)}{\varphi(z_s^v)}\,. \tag{30}$$

From the above reasoning and Theorem 1 we expect that, if $\psi(x) \approx \varphi(\xi(x))$, then the solution to the projected boundary value problem (27) provides a good approximation to the original optimal control or sampling problems (7) or (14), since

$$\hat{u}_s = \sqrt{2}\beta^{-1}\frac{\nabla\psi(x_s)}{\psi(x_s)} \approx \sqrt{2}\beta^{-1}\frac{\nabla\xi(x_s)\nabla_\xi\varphi(\xi(x_s))}{\varphi(\xi(x_s))} = \nabla\xi(x_s)\hat{v}_s\,, \tag{31}$$

with $\hat{v}_s = \hat{v}(\xi(x_s))$ as defined by (30). We refrain from making precise statements about the "≈" sign in the last equation, which would involve L^2 estimates of the control and instead refer to Remark 2 below (cf. also [12, 19]).

This suggests the following algorithm for approximating the control policies (7) or (14) and solving the corresponding control or sampling problems:

Algorithm 1 Control along essential coordinates.

1: Compute the coefficients (23) of the effective SDE (22).
2: Compute the solution $\hat{v}$ of the low-dimensional optimal control problem (28)–(29).
3: Set $u_s = \nabla\xi(x_s)\hat{v}_s$ in the original dynamics (3) where $\hat{v}_s = \hat{v}(\xi(x_s))$ as defined by (30).

The low-dimensional optimal control problem in Step 2 of Algorithm 1 can be solved either directly or by solving the PDE (27) the equivalent dynamic programming equation (32) by one's favourite discretization scheme, which is typically feasible up to dimension $k = 3$ provided that the averaged coefficients b, σ, g can be efficiently computed (cf. [33, 34]).

Remark 2 *It is easy to see that the value function* $\vartheta = -\beta^{-1}\ln\varphi$ *of the simplified optimal control problem (29) solves the projected dynamic programming equation*

$$\begin{aligned} &\beta^{-1}a\colon\nabla^2\vartheta + b\cdot\nabla\vartheta - |\nabla\vartheta|^2 + g = 0 \\ &\vartheta|_{\partial S} = h\,. \end{aligned} \tag{32}$$

Nevertheless, the best approximation property of (27) does not imply that ϑ *is a best approximation of the original value function* U *in any reasonable sense, because* φ *and* ϑ *are related by a logarithmic transformation, and the nonlinear transformation destroys the linear structure of the approximating* L^2*-subspace.*

One may argue, however, that the logarithm is locally Lipschitz when its argument is bounded away from zero, which restores the best approximation property in some sense. Specifically, if we assume that ψ *and* φ *are classical solutions of (9) and (27) that are bounded away from zero, which is guaranteed if the coefficients* $\nabla V, G$ *and* b, σ, g *satisfy suitable regularity conditions (see, e.g., [23, Ch. VI.3]), and if we let* $L > 0$ *denote the Lipschitz constant of the logarithm, then*

$$\mathbf{E}_\rho\big[|\ln\psi - \ln\varphi|^2\big] \le L^2\mathbf{E}_\rho\big[|\psi - \varphi|^2\big]$$

where the right hand side of the inequality cannot be made smaller by approximating the coefficients in (9) by functions of ξ (if we keep L fixed). Therefore we may assume that good approximations of ψ will lead to good approximations of U and thus of the optimal control (assuming that U is a classical solution and thus differentiable). For a rigorous approximation result for systems with time scale separation, see [12, 19].

4. Localization of the dynamics near optimal trajectories

The previous approach is useful if appropriate (e.g. problem specific) collective variables are known beforehand. But even then, the numerical algorithm requires that one can efficiently evaluate the coefficients (23) of the projected equation, which may be numerically challenging if the space orthogonal to the collective variables contains slowly mixing components that impede the convergence of the corresponding conditional expectations. Moreover, first and second derivatives of ξ need to be explicitly available in Algorithm 1, which may be prohibitive in some cases.

In this section, we consider the situation that collective variables or their effective equations are either not available or are difficult to find, even though they may exist. The key idea then is to first numerically identify a most likely pathway, along which the typical realizations of the dynamics concentrate, and then place basis functions along this pathway and iteratively compute an approximation to the optimal control using the cross-entropy method [19].

4.1. Finding typical pathways in metastable systems

In many situations of practical relevance, e.g., in molecular dynamics, modelling of materials or chemical reactions, it is very likely that transition events of a stochastic dynamical system concentrate along an isolated pathway, often called *most probable transition path (MPTP)*. One such example is diffusions in multi-well energy landscapes at low temperature, and large deviations theory states that, with high probability, the typical transition follows the MPTP that minimizes the Freidlin-Wentzell rate functional [36]. For gradient-type diffusions of the form (1), the MPTP coincides with the minimum energy pathway γ that is characterized by [13]

$$\nabla V(\gamma(s)) \parallel \dot\gamma(s) \iff \nabla^\perp V(\gamma(s)) = 0\,, \quad s \geq 0\,, \tag{33}$$

where $\dot\gamma(s)$ is the tangent vector of the curve $\gamma\colon [0,\infty) \to \mathbb{R}^n$ at $\gamma(s)$, and $\nabla^\perp$ denotes the component of the gradient that is perpendicular to the curve.

The string method The minimum energy pathway can be numerically identified using the zero temperature string (ZTS) method [13, 37], extensions of which include the finite temperature string (FTS) method [14, 38] that allows for computinmg approximations to transition paths in rugged potential energy surfaces or at finite temperature. It is convenient to parametrize the curve γ by arc length, i.e., we consider γ as a differentiable map on $[0,1]$ rather than the real half line $[0,\infty)$, with boundary conditions $\gamma(0) = c_0$ and $\gamma(1) = c_1$ that satisfy $\nabla V(c_i) = 0$ and represent the relevant metastable sets of the dynamics. Both ZTS and FTS methods compute an approximation of the MPTP based on a discretization $(x_l)_{l=0,1,\dots}$ of γ and a constrained gradient descent algorithm, where the main difference between the two methods is that in the ZTS method each discretization node follows the gradient direction $d_l = -\nabla V(x_l)$, while in the FTS method a descent direction d_l based on an

ensemble average of the local gradients in the Voronoi cell corresponding to x_l is used. The key steps of ZTS and FTS methods are summarized in Algorithm 2 below; for further details we refer to the original papers [13, 37, 14, 38].

Algorithm 2 String method

1: **Initialization.** At step $k = 0$, set $x_0^0 = c_0$, $x_N^0 = c_1$. Compute $x_l^0, l = 1, 2, \cdots, N-1$ by linearly interpolating x_0^0 and x_N^0.
2: **Update.** At the k-th step, for each discretized nodes x_l^k compute the descent direction d_l to obtain $x_l^{k+\frac{1}{2}}, l = 0, 2, \cdots, N$ by
$$x_l^{k+\frac{1}{2}} = x_l^k + \Delta t\, d_l + r_l^k$$
where $r_0^k = r_N^k = 0$, and $r_l^k = \kappa(x_{l+1}^k + x_{l-1}^k - 2x_l^k)$, $1 \le l \le N-1$.
3: **Reparameterization.** (Linearly) interpolate the updated nodes $x_l^{k+\frac{1}{2}}$ and obtain x_l^{k+1} at $k+1$ step, $l = 0, 2, \cdots, N$.
4: **Termination.** Iterate Steps 2 and 3 until convergence.

4.2. Cross-entropy minimization based on typical pathways

The cross-entropy (CE) method is a numerical method, with a wide range of applications in combinatorial optimization and statistics [20, 39]. The idea is as follows: Given a target probability measure μ^* and a parameterized subset of probability measures $\mathcal{D} = \{\mu(\boldsymbol{\omega}) \mid \boldsymbol{\omega} \in \Theta\}$, where Θ is any parameter set, the CE method seeks an optimal probability measure $\mu(\omega^*)$ in the parametric family $\mathcal{D}$ that minimizes the Kullback-Leibler divergence with respect to μ^*. For the reader's convenience, we briefly recall the basics of the cross-entropy method for diffusions [19], and discuss its application to the numerical approximation of optimal controls of the form (7) and (14); since the optimal control policy provides a zero-variance estimator for the sampling problem (8) and vice versa, we take the optimal control problem as a starting point. To this end, we observe that (2) has the following entropy representation (see [19]):

$$J(u) = J(\hat{u}) + \beta^{-1} D(\mu_u \,|\, \hat{\mu}) \tag{34}$$

where $\hat{u}$ is the unique minimizer of $J(u)$, and $\hat{\mu} = \mu_{\hat{u}}$ is the corresponding probability measure on the space of SDE realizations. Here $D(\cdot \,|\, \cdot)$ denotes the Kullback-Leibler (KL) divergence

$$D(\mu_1 \,|\, \mu_2) = \begin{cases} \mathbf{E}_{\mu_1}\left[\ln \dfrac{d\mu_1}{d\mu_2}\right] & \mu_1 \ll \mu_2\,, \\ \infty & \text{otherwise} \end{cases} \tag{35}$$

between two probability measures μ_1 and μ_2. Notice that $D(\cdot \,|\, \cdot)$ is not symmetric, i.e., in general, $D(\mu_1 \,|\, \mu_2) \neq D(\mu_2 \,|\, \mu_1)$. Nevertheless the KL divergence can be regarded as a measure of similarity between two probability measures μ_1 and μ_2 in that $D(\mu_1 \,|\, \mu_2) \ge 0$ and equality holds if and only if $\mu_1 = \mu_2$. As a consequence, minimizing $J(u)$ in (34) is equivalent to minimizing $D(\mu_u \,|\, \hat{\mu})$. However, minimizing $D(\mu_u \,|\, \hat{\mu})$ is not feasible because $\hat{\mu}$ is not explicitly available.

The key idea of the CE method is to solve a relaxed minimization problem

$$\min_{\mu_u \in \mathcal{D}} D(\hat{\mu} \,|\, \mu_u)\,, \tag{36}$$

which should yield a good approximation of $\hat{\mu}$ provided that $\hat{\mu}$ can be accurately approximated within the parametric family of candidate probability measures. It turns out that for solving (36), it is only necessary to know $\hat{\mu}$ up to a normalization constant (which depends on the value function and therefore is not known).

Specifically, we suppose that there are m candidate control policies $u_i : \mathbb{R}^n \to \mathbb{R}^n, i = 1, 2, \cdots, m$, such that $\hat{u}$ can be approximated by feedback controls of the form

$$u(x\,;\boldsymbol{\omega}) = \sum_{i=1}^{m} \omega_i u_i(x)\,, \quad x \in \mathbb{R}^n, \quad \boldsymbol{\omega} = (\omega_1, \omega_2, \cdots, \omega_m)^T \in \mathbb{R}^m\,.$$

In terms of the candidate policies, (36) takes the form

$$\min_{\boldsymbol{\omega} \in \mathbb{R}^m} D(\hat{\mu} \,|\, \mu_{\boldsymbol{\omega}})\,, \tag{37}$$

where $\mu_{\boldsymbol{\omega}}$ coresponds to the probability measure in path space related to dynamics

$$dx_s^u = -\nabla V(x_s^u)ds + \sqrt{2}u(x_s^u\,;\boldsymbol{\omega})\,ds + \sqrt{2\beta^{-1}}dw_s\,, \quad x_0^u = x\,. \tag{38}$$

Now let ν be the scaled Wiener measure (corresponding to $\sqrt{2\beta^{-1}}w_s$). By Girsanov's theorem, it follows that there exists a measurable function f such that $\mu_{\boldsymbol{\omega}}(dz) = f(z;\boldsymbol{\omega})\nu(dz)$, with

$$f(z;\boldsymbol{\omega}) = \exp\Big(\frac{\beta}{2}\int_0^\tau b(z_s)\cdot dz_s - \frac{\beta}{4}\int_0^\tau \big|b(z_s)\big|^2 ds\Big) \tag{39}$$

with

$$b(x) = \sqrt{2}\sum_{i=1}^{m}\omega_i u_i(x) - \nabla V(x) \tag{40}$$

and the shorthands z and z_s for the random variables $(x_s^u)_s$ and x_s^u. On the other hand, the dual relationship (17) and Jensen's inequality imply that $\hat{\mu}(dz) = g(z)\nu(dz)$ where

$$g(z) \propto \exp\Big[-\beta\Big(\int_0^\tau G(z_s)ds + H(z_\tau)\Big)\Big] f(z\,;\mathbf{0})\,. \tag{41}$$

To exploit this relation, note that, for two probability measures μ_1, μ_2 that are both absolutely continuous with respect to ν such that $d\mu_i = g_i d\nu$, we have

$$D(\mu_1 \,|\, \mu_2) = \int g_1 \ln g_1\, d\nu - \int g_1 \ln g_2\, d\nu\,. \tag{42}$$

Hence minimizing $D(\mu_1 \,|\, \mu_2)$ with respect to the second argument is equivalent to maximizing the cross-entropy

$$CE_\nu(g_1 \,|\, g_2) = \int g_1 \ln g_2\; d\nu\,, \tag{43}$$

between the probability densities g_1 and g_2. Hence the name *cross-entropy method*.

Combining (39)–(43), it turns out that the target function in (37) is quadratic and strictly convex in $\boldsymbol{\omega}$, as a consequence of which a necessary and sufficient condition for optimality is that the coefficient vector $\boldsymbol{\omega}$ satisfies the linear system of equations

$$A\boldsymbol{\omega} = \boldsymbol{r} \tag{44}$$

with

$$\begin{aligned} A_{ij} =& \mathbf{E}_{\mu_{\boldsymbol{v}}}\left[B(z)\int_0^\tau u_i(z_s)\cdot u_j(z_s)ds\,\eta(z\,;\,\boldsymbol{v})\right] \\ r_i =& \ \frac{1}{\sqrt{2}}\mathbf{E}_{\mu_{\boldsymbol{v}}}\Bigg[[B(z)\left(\int_0^\tau u_i(z_s)\cdot dz_s\right. \\ &\left.+\int_0^\tau u_i(z_s)\cdot\nabla V(z_s)ds\right)\eta(z\,;\,\boldsymbol{v})\Bigg] \end{aligned} \tag{45}$$

for $1\le i,j\le m$ and

$$B(z)=\exp\left(-\beta\left(\int_0^\tau G(z_s)ds+H(z_\tau)\right)\right),\ \eta(z\,;\,\boldsymbol{v})=\frac{f(z\,;\,\mathbf{0})}{f(z\,;\,\boldsymbol{v})}\,. \tag{46}$$

Notice that (44) can be calculated under any reference probability measure $\mu_{\boldsymbol{v}}$, $\boldsymbol{v}\in\mathbb{R}^m$. Although the linear system (44) can be solved directly, in practice it is advisable to solve the CE minimization problem (37) iteratively, with an increasing sequence of inverse temperatures β_j, to improve the overall convergence of the method. The complete CE algorithm is summarized in Algorithm 3, and readers are referred to [19, 20, 39] for further details.

Algorithm 3 Cross entropy method.

1: Define $0<\beta_0<\beta_1<\cdots<\beta_k=\beta$, set $\boldsymbol{v}^{(0)}=0$, or some nonzero initial guess.
2: **for** $j=0$ to k **do**
3: generate N_j trajectories $z_i, i=1,2,\cdots,N_j$ from dynamics (38), with $\boldsymbol{\omega}=\boldsymbol{v}^{(j)}$, $\beta=\beta_j$.
4: compute the coefficients of $A^{(j)},\boldsymbol{r}^{(j)}$ from (45) with $\boldsymbol{v}=\boldsymbol{v}^{(j)}$, and solve the linear equations $A^{(j)}\boldsymbol{v}^{(j+1)}=\boldsymbol{r}^{(j)}$.
5: **end for**

Choice of ansatz functions We shall now discuss how to combine the string method and the cross-entropy method, so as to efficiently approximate the control forces (7) and (14) in the optimal control and the dual sampling problem. Although the CE method has no specific requirements regarding the choice of ansatz functions, attention should be paid to their placement in order to avoid combinatorial explosion in the number of basis functions. Our key idea is to utilize the information about the MPTP from the string method to optimize the design of the ansatz functions in the cross-entropy method.

Following the consideration in Subsection 4.1, we assume that we have computed a discretization $x_0,\ldots,x_N$ of the transition path γ using Algorithm 2. The unit tangent direction t_l along γ at x_l can then be approximated by finite differences:

$$t_l=\frac{x_{l+1}-x_l}{|x_{l+1}-x_l|}\,,\quad 0\le l\le N-1 \tag{47}$$

where we set $t_N=t_{N-1}$. Now consider a Voronoi partition of the state space associated with the nodes $\{x_l\colon 0\le l\le N\}$ and denote by C_l the Voronoi cell corresponding to the l-th node. Further let $l\colon\mathbb{R}^n\to\{0,\ldots,N\}$ be the membership function that is defined by $l(x)=i$ for $x\in C_i^\circ$ and every $0\le i\le N$, where C_i° denotes the interior of the Voronoi cell C_i.

We consider the following types of ansatz policies for the control.

(i) Gradients of Gaussian basis functions:

$$u_i(x) = \nabla\varphi_i(x)\,, \quad \text{with}\varphi_i(x) = \exp\big(-\frac{|x-x_{k_i}|^2}{2r^2}\big), \tag{48}$$

where $1 \le i \le m$ and $0 \le k_1 < k_2 < \cdots < k_{m-1} < k_m \le N$. Here $r > 0$ is an adjustable parameter that may be chosen globally or for each φ_i.

(ii) Piecewise linear functions with Chebyshev coefficients:

$$u_i(x) = T_{i-1}\left(\frac{2l(x)}{N} - 1\right)\phi\left(\frac{|x-x_{l(x)}|}{R_{l(x)}}\right)t_{l(x)}\,, \quad 1 \le i \le m \tag{49}$$

where the n-th Chebychev polynomial T_n is defined as

$$T_n(x) = \cos(n\cos^{-1}(x))\,, \quad x \in [-1,1]\,, \quad n \in \mathbb{N}_0 \tag{50}$$

See [40] for details. The function ϕ is a nonnegative smooth cutoff function (for example, take $\phi(r) = 1, |r| \le 1$ and $\phi(r) = 0, |r| > 2$ and interpolate smoothly in between), and R_l is a fitting parameter, depending on the diffusivity of the dynamics within cell C_l (or width of the "transition tube", which can be inferred by the string method). Intuitively, the control force $u_i(x)$ in the Voronoi cell $C_{l(x)}$ is along the tangent direction $t_{l(x)}$ of the transition path and the magnitude is proportional to the $(i-1)$-th Chebyshev polynomial T_{i-1}. Furthermore, if x is far away from the centre $x_{l(x)}$ of $C_{l(x)}$, then $u_i(x) = 0$, i.e., the control force is only active inside the transition tube.

(iii) Localized piecewise functions: Similarly to the previous case, set $m = N+1$ and

$$u_i(x) = \phi\left(\frac{|x-x_i|}{R_i}\right)\mathbf{1}_{C_i}(x)t_i\,, \quad 1 \le i \le N+1\,, \tag{51}$$

where ϕ is a cutoff function as above. In this case, we have $N+1$ basis control functions which have disjoint supports and are parallel to t_i inside the Voronoi cell C_i.

Algorithm 4 Control along the transition path

1: Compute the discretized transition path using the string method (Algorithm 2).
2: Interpolate the unit tangent directions t_l for each $x_l, 0 \le l \le N$ from (47). Construct m basis functions from (48), (49) or (51).
3: Solve for the optimal candidate policy $u(x, \boldsymbol{\omega})$ using the CE method (Algorithm 3).

Computational cost We summarize the MPTP-based approach for computing the control policies (7) or (14) in Algorithm 4 and conclude this section with a brief discussion of the computational complexity of calculating (44). For simplicity we consider a single temperature level $\beta_0 = \beta$ in Algorithm 3. Further assume that the cost, M_{int}, of generating and computing the path integral along each trajectory is constant and that M trajectories are generated to evaluate the expectation in (44). Depending on the type of basis functions used the computational costs are as follows, where we assume throughout that m functions are used:

(i) The total computational cost for evaluating the coefficient matrix A in (44) with a Gaussian basis is of the order

$$\mathcal{O}\left(\frac{m(m+1)MM_{int}}{2}\right),$$

taking into account that A is symmetric.

(ii) As for the piecewise linear basis with Chebyshev coefficients, notice that the i-th Chebyshev function is a polynomial of order i, i.e. $T_i(z) = \sum_{j=0}^{i} T_{i,j} z^j$ for some $T_{i,j} \in \mathbb{R}$. For each trajectory z, instead of computing A_{ij} directly, we first calculate

$$\int_0^\tau \left(\frac{2l(z_s)}{N} - 1\right)^k \phi^2\left(\frac{|x - x_{l(z_s)}|}{R_{l(z_s)}}\right) ds, \quad 0 \le k \le 2m-2\,, \tag{52}$$

and then, using (49),

$$\begin{aligned}&\int_0^\tau u_i(z_s)u_j(z_s)ds\\ &= \sum_{\substack{0\le k_1\le i-1\\ 0\le k_2\le j-1}} T_{i,k_1}T_{j,k_2}\int_0^\tau \left(\frac{2l(z_s)}{N} - 1\right)^{k_1+k_2} \phi^2\left(\frac{|x - x_{l(z_s)}|}{R_{l(z_s)}}\right) ds.\end{aligned}$$

The overall computational cost for evaluating A in this case is $\mathcal{O}\left((2m-1)MM_{int}\right)$.

(iii) Similarly, localized piecewise linear control basis functions have the property that any two of the $m = N+1$ basis functions in (51) have disjoint support, which entails that A is a diagonal matrix. The total computational cost therefore is $\mathcal{O}\left((N+1)MM_{int}\right)$.

Remark 3 *A remark on the interpretation of the MPTP and the string method in the context of optimal control is in order. Let us suppose that $\Omega = \mathbb{R}^n$ and replace the first exit time τ by a finite, but sufficiently large deterministic stopping time $T < \infty$, in which case the value function U in (4) becomes a function of both x and t via the initial condition $x_t^u = x$. Further let $G = 0$ and consider the limit $\beta \to \infty$ in the dynamic programming equation (5). If the potential energy function V and the terminal cost H satisfy mild regularity conditions, then the limit of (5) has a unique solution in the viscosity sense that is given by*

$$U_\infty(t,x) = \inf_{\substack{\gamma\in\mathcal{AC}(t,T)\\ \gamma(t)=x}} \left\{\frac{1}{2}\int_t^T \left\|\gamma'(s) + \nabla V(\gamma(s))\right\|^2 ds + H(\gamma(T))\right\},$$

where $\mathcal{AC}(t,T)$ is the space of absolutely continuous functions $\varphi\colon [t,T] \to \mathbb{R}^n$. The string method basically approximates the minimizer γ^* that attains the control value $U_\infty(0,c_0)$, with the auxiliary boundary condition $\gamma(T) = c_1$ that can be enforced by choosing the terminal cost H such that it penalizes deviations from hitting the endpoint c_1 at time T. Hence the string method provides a reasonable initial guess for the optimal trajectories of the stochastic control problem (4).

5. Numerical examples

In the following, we study several examples to demonstrate and compare the different algorithmic approaches. For simplicity, we choose $G(x) = \lambda > 0$ and $H = 0$. In this

case, the value function (4) becomes

$$U(x) = \inf_{u_s} \mathbf{E}_{\mu_u}\left[\lambda\tau + \frac{1}{2}\int_0^\tau |u_s|^2 ds\right] \tag{53}$$

and the objective function (8) in the associated sampling problem is $\psi(x) = \mathbf{E}\left[e^{-\beta\lambda\tau}\right]$. For notational simplicity we have dropped the conditioning on the initial value and consider only trajectories with fixed initial condition $x_0 = x$ or $x_0^u = x$.

5.1. Two-dimensional diffusion in a stiff potential

The following example is taken from [6]. Here, the dynamics is given by (1) with the stiff potential

$$V_\epsilon(x_1, x_2) = \left(x_1^2 - 1\right)^2 + \frac{1}{\epsilon}\left(x_1^2 + x_2 - 1\right)^2 . \tag{54}$$

The potential V_ϵ has two local minima $c_1 = (-1, 0)$ to $c_2 = (1, 0)$, and two possible candidates for the essential variables to describe transitions between the wells are

$$\xi_1(x_1, x_2) = x_1 \exp(-2x_2) \quad \text{and} \quad \xi_2(x_1, x_2) = x_1 . \tag{55}$$

For details, see [6].

As a first test, we solve the control task (2) using Algorithm 1 with ξ_1 and ξ_2 as essential coordinates. The simulation parameters are set to $\beta = 3.0$, $\epsilon = 0.02$ and $G(x) = 1.0$, and the stopping time τ is defined as the first hitting time of the basin of attraction of c_2 under the condition that the process is initialized at $x = c_1$:

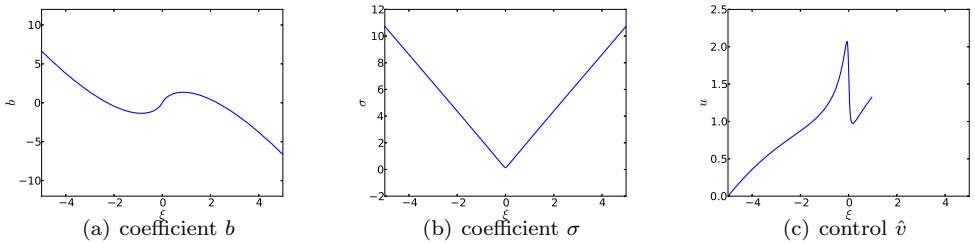
$$\tau = \inf_{s>0}\{|x_s^u - c_2| < 0.02 \colon x_0 = c_1\} ,$$

For the projected dynamics, we define

$$\tau^\xi = \inf_{s>0}\{|z_s^u - \xi^{end}| < 0.001 \colon z_0^u = \xi^{start}\}$$

where z_s^u is the solution of (28), and we set $\xi^{start} = -1.0$ and $\xi^{end} = 1.0$. (Note that the definitions of the stopping sets and τ^ξ, τ differ slightly from the definition on page 9.)

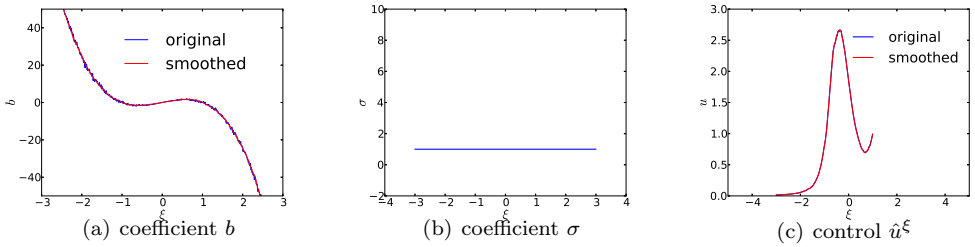
Projection onto ξ_1 We compute the effective dynamics (22) for $\xi = \xi_1$ on the interval $[-5.0, 5.0]$, with mesh size 3.0×10^{-2} on $[-5.0, -0.5]$ and $[0.5, 5.0]$, and 5.0×10^{-3} on $[-0.5, 0.5]$. The coefficients of the effective dynamics b, σ were computed from (23) using constrained dynamics [33]. The constrained simulation step size was $\mathrm{dt} = 10^{-5}$ where $N = 10^6$ time steps were generated to compute the expectations in (23) at each grid point $\xi \in [-5.0, 5.0]$. The computed profiles of b and σ are shown in Figure 1(a,b). The optimal control $\hat{v}$ in (29)–(28), wich has been computed by a finite differences discretization of (32), is shown in Figure 1(c).



(a) coefficient b (b) coefficient σ (c) control $\hat{v}$

Figure 1. The coefficients b, σ for the 1D observable $\xi = \xi_1$ and the resulting optimal control policy $\hat{v}$ as function of $\xi = \xi_1$ (with a slight abuse of notation).

Projection onto ξ_2 For comparison we consider the effective dynamics (22) for $\xi = \xi_2$, computed in the interval $[-3.0, 3.0]$, with uniform mesh size 0.01 and all other simulation parameters as before. The resulting coefficients b, σ and the optimal control policy as a function of the coordinate ξ are shown in Figure 2. Notice that, unlike in the previous case, the effective drift b is oscillatory, which indicates that the coordinate ξ_2 is not orthogonal to the fast dynamics and involves slow or metastable components, hence it is not a good essential coordinate for this problem. As a test, we have computed the optimal control using the oscillatory drift coefficient and a smoother approximation (see Figure 2(a)). The resulting one-dimensional control policies are shown in the rightmost panel of Figure 2, and they indicate that the result is not at all sensitive to the sampling error that produced the random oscillations in the effective drift b. This somewhat surprising outcome may be explained by the smoothing effect that the elliptic operator has in the one-dimensional dynamic programming equation (32).



(a) coefficient b (b) coefficient σ (c) control $\hat{u}^{\xi}$

Figure 2. Coefficients b, σ as functions of $\xi = \xi_2$ and the corresponding optimal control $\hat{v}$.

Now, to test the validity of the approach, we approximate the optimal control for the original 2D dynamics (1) by the 1D control policies $\hat{v}$ for both ξ_1 and ξ_2, following the route of Algorithm 1. By construction, the approximation to the full-dimensional optimal control is always acting in the direction of the respective essential coordinates. We have then simulated the original systems with the approximating control policies and computed the control value (53) and the mean first hitting time (MFHT). The results that are shown in Table 1, and we observe that ξ_1 gives a much better estimate of the optimal cost value than ξ_2. We further observe that the ξ_1 estimate is closer

	ξ_1		ξ_2	
	cost	MFHT	cost	MFHT
reduced	3.02	–	2.20	–
full	3.13	1.5	4.53	2.3

Table 1. Approximations of the optimal control problem based on coordinates ξ_1, ξ_2. The row with the label "reduced" is the cost value of the projected dynamics (28)–(29). The row with the label "full" records the cost value and the mean value of τ, obtained from a simulation of 10^5 trajectories of the controlled 2D dynamics, with the approximate controls.

	MC	ξ_1	ξ_2
$\mathbf{E}\left[e^{-\beta\tau}\right]$	1.2×10^{-4}	1.1×10^{-4}	1.1×10^{-4}
$\sqrt{\mathrm{Var}}$	1.9×10^{-3}	2.2×10^{-5}	1.2×10^{-3}
MFHT	38.3	1.5	2.3
Δt	10^{-4}	10^{-5}	10^{-5}
Accel(u)	1.0	1.9×10^4	4.2

Table 2. Importance sampling estimates. Here, $\sqrt{\mathrm{Var}}$ is the standard deviation of the estimators, "MFHT" records the average length of the trajectories, and 'Accel(u)" is the acceleration index as defined in (13). The columns "MC", "ξ_1", "ξ_2" show the results from standard Monte Carlo and importance sampling with biasing force u in the directions of ξ_1 or ξ_2.

to the cost value of the reduced optimal control problem, which indicates that ξ_2 is not a good (i.e. slow) variable and which is in agreement with the asymptotic results obtained in [12] for systems with slow and fast variables.

As an additional test, we have solved the dual sampling problem $\mathbf{E}\left[e^{-\beta\tau}\right]$ by importance sampling using the approximate controls obtained from the one-dimensional approximation. The results are shown in Table 2, along with a comparison to standard Monte Carlo simulation. In accordance with the observations from the optimal control problem, we see again that the importance sampling estimator based on controlling ξ_1 clearly outperforms the estimator based on controlling ξ_2 in terms of sample variance and expected trajectory length (cf. the discussion on page 6). Nevertheless, the second scheme that biases the dynamics only in the direction of ξ_2 appears to be more efficient than brute-force Monte Carlo.

Remark 4 *By Lemma 1, the importance sampling estimator yields an accurate (though biased) estimator of the true optimal control value, even when a suboptimal control is used.*

5.2. One-dimensional bistable dynamics: choice of ansatz functions

The purpose of this example is to study the effect of the control basis functions on the accuracy of the cross-entropy (CE) minimization (see Subsection 4.2). As reference dynamics, we consider a one-dimensional diffusion ($n = 1$) in a double-well potential

$$V(x) = \frac{1}{2}\left(x^2 - 1\right)^2$$

because no approximation error is introduced by reducing the state space to the neighbourhood of the transition path, and thus the system allows for a systematic comparison of different basis functions. The potential $V(x)$ has two local minima at $x = \pm 1$. We set $x_0^u = -1$ and $G(x) = 1$, which, together with the stopping time definition

$$\tau = \inf\{s \geq 0 \colon x_s^u \geq 1.0\}$$

leads to a cost functional of the form (53). We seek to minimize (2) subject to the dynamics (3) by forcing the dynamics to make a quick transition from one local minimum to the other. The related sampling task then is to compute $\mathbf{E}\big[e^{-\beta\tau}\big]$. With the choice $\beta = 3.0$ we obtain the estimate $\mathbf{E}[\tau] = 12.5$ based on 10^5 independent realizations of the uncontrolled reference dynamics, with a time step $\Delta t = 10^{-4}$ in the Euler discretized SDE.

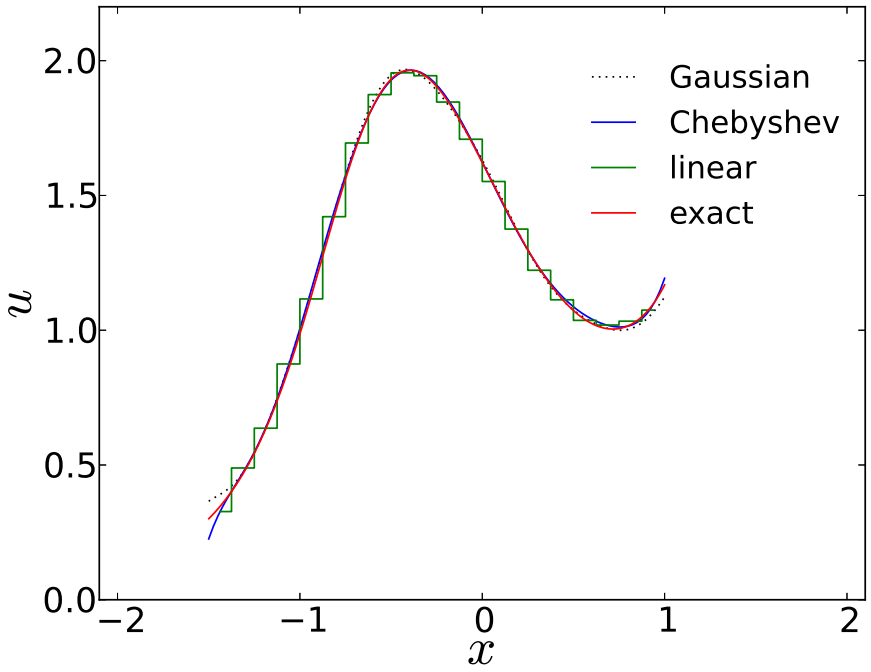


Figure 3. CE approximations of the reference ("exact") optimal control for the dynamics in a 1D double-well potential based on different ansatz functions.

Based on this prior information, we solve the CE minimization problem with the following choices of basis functions: For the Gaussian basis, $m = 20$ basis functions are used with centers equally distributed in $[-1.25, 1.0]$ and $r = 0.5$ as in (48). For the Chebyshev basis, $m = 9$ basis functions are used. For the localized piecewise linear basis, the interval $[-1.5, 1.0]$ is divided evenly into 20 subintervals (Voronoi cells), leading to $m = 20$ basis functions. Figure 3 displays the resulting control approximations together with the reference solution that has been obtained by directly solving the boundary value PDE (9) using a finite difference scheme. In each case the optimal control policy is quite well approximated by the CE method.

	Gaussian	Chebyshev	linear	exact
$J(u)$	2.04	2.04	2.04	2.03
MFHT	1.01	1.01	1.01	1.01

Table 3. Approximations of the optimal cost value $J(\hat{u})$ using the CE-based controls as in Figure 3. For comparison, the row "MFHT" records the mean trajectory length under the controlled dynamics, averaged over 10^5 independent realizations.

	MC	Gaussian	Chebyshev	linear
$\mathbf{E}\left[e^{-\beta\tau}\right]$	2.2×10^{-3}	2.2×10^{-3}	2.2×10^{-3}	2.2×10^{-3}
$\sqrt{\mathrm{Var}}$	1.3×10^{-2}	6.9×10^{-5}	5.9×10^{-5}	2.0×10^{-4}
MFHT	12.5	1.01	1.01	1.01
Accel(u)	1.0	4.4×10^5	6.0×10^5	5.2×10^4

Table 4. Importance sampling of $\mathbf{E}\left[-e^{\beta\tau}\right]$ with CE-based controls as in Figure 3. Here "MC" denotes standard Monte Carlo, $\sqrt{\mathrm{Var}}$ is the standard deviation of the estimator and "Accel(u)" is the acceleration index defined in (13). All estimates were computed using 10^5 trajectories with a time step $\Delta t = 10^{-4}$.

As a next step, we plugged the approximations of the optimal control into the controlled SDE and generated 10^5 independent realizations, from which we estimated the cost value $J(u)$ and the average stopping time $\mathbf{E}_{\mu_u}[\tau]$. The simulation results are shown in Table 3 and demonstrate that all three types of basis functions lead to very accurate approximations of $J(\hat{u}) = U(-1)$ and $\mathbf{E}_{\mu_u}[\tau]$. We then repeated the calculation for the computation of $\mathbf{E}\left[e^{-\beta\tau}\right]$, using importance sampling with different control bases. For a better comparison, average trajectory length and standard deviation were recorded for each estimator and compared to standard Monte Carlo (see Table 4). The importance sampling estimators differ slightly in the achieved variance reduction and thus lead to slightly different acceleration indices that, however, should be contrasted with the computational complexity of each method (see page 16). But even then, a tremendous acceleration is obtained over standard Monte Carlo for every choice of ansatz functions.

5.3. Two-dimensional dynamics with Müller-Brown potential

As a nontrivial test case for the transition path based CE minimization, we study the two-dimensional dynamics with Müller-Brown potential

$$V(x_1, x_2) = \sum_{i=1}^{4} \alpha_i \, e^{a_i(x_1 - x_i^0)^2 + b_i(x_1 - x_i^0)(x_2 - y_i^0) + c_i(x_2 - y_i^0)^2}$$

with $\alpha_i, x_i^0, y_i^0, a_i, b_i, c_i \in \mathbb{R}$, $1 \le i \le 4$, taken from [41, 42], where all the α_i have been rescaled with a factor 0.1, so as to yield a realistic average potential energy and realistic transition probabilities for $\beta = 0.8$. With this choice of parameters, the system exhibits two metastable regions centred around $c_0 = (-0.5, 1.5)$ and $c_1 = (0.5, 0)$ that are connected by a third, less metastable region close to the point $c_{1/2} = (-0.1, 0.5)$; see Figure 4.

As a first step, we have computed the MPTPs using both ZTS and FTS methods. The MPTPs are shown in Figure 4; cf. [38]. Notice even for this relatively smooth

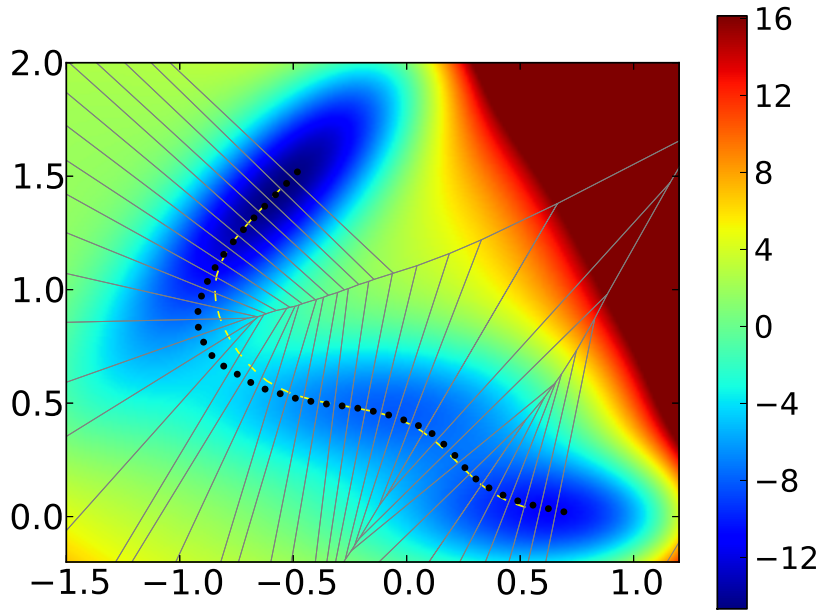


Figure 4. Scaled Müller-Brown potential with Voronoi cells. Yellow dashed line and black dotted line are the MPTPs computed using ZTS method and FTS method respectively. The Voronoi cell C_i corresponds to node x_i are shown.

potential, the MPTP obtained from ZTS method is slightly different from the one obtained from FTS method due to temperature-related entropic effects. (In the FTS case, we have decreased the temperature to $\beta = 3.0$ in order to reduce the diffusivity in the system after having observed that the FTS method becomes rather sensitive to the initialization of the string and may converge to an unphysical transition path.)

As the next step, we have solved the CE minimization problem (53) and its dual sampling problem for $G(x) = \lambda$, with $\lambda \in \{0.5, 1.0, 2.0\}$. To this end, we have placed basis functions along the discrete nodes x_l, $0 \leq l \leq N = 39$ of the discrete representation of the FTS-MPTP (black dots in Figure 4), where we define $x_2 = (-0.53, 1.47)$ to be the initial point as it has lower potential energy than x_0 and x_1. Denoting by $\{C_l : 0 \leq l \leq N\}$ the associated Voronoi tessellation, we define τ to be the first hitting time of C_{39}, which is centred at the last node x_{39} on the MPTP.

Before going into the comparison of the results, we have generated 10^5 independent realizations of the uncontrolled dynamics ($u = 0$, $\Delta t = 10^{-4}$) and estimated the mean first hitting time $\mathbf{E}[\tau] \approx 258.6$, with standard deviation $\sqrt{\mathrm{Var}[\tau]} \approx 258.3$. The large mean value and the fact that the standard deviation is of the same order of magnitude indicate that the typical transition paths from C_2 to C_{39}, given our choice of simulation parameters for the uncontrolled dynamics, are typically long and diffusive.

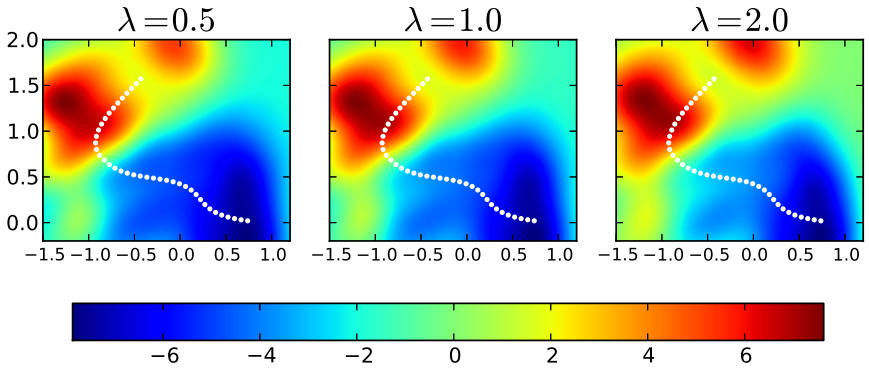
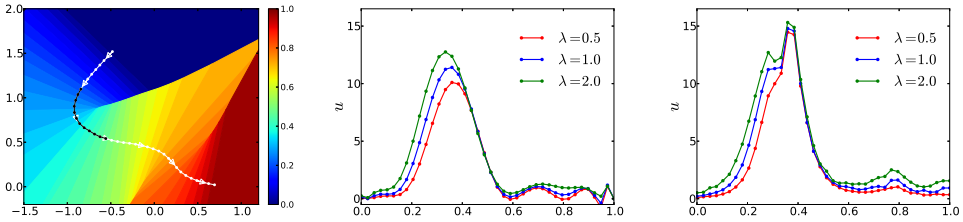


Figure 5. The added potential $-\sqrt{2}\sum_i \omega_i \varphi_i(x)$ corresponding to the control force computed with Gaussian basis functions. The white dots are the discrete nodes of the approximation to the MPTP and are chosen to be centers of the Gaussians.



(a) Voronoi cells (b) Chebyshev basis (c) Localized basis

Figure 6. (a) Voronoi cells C_l associated with the nodes x_l, $l = 0, 1, \cdots, N$; the MPTP is in white, the black segment corresponds to the region, in which the control is large. (b) Average control force per Voronoi cell with Chebyshev basis. (c) Average control force with localized piecewise linear basis functions.

We have then computed the control force using Algorithm 4 with $N + 1 = 40$ Gaussians $\varphi_i = \mathcal{N}(x_i, r_i^2)$ as basis functions centered at each node x_i and radii r_i that were chosen to be 5 times the mean Euclidean distance of x_i to its nearest neighbours $x_{i\pm1}$. Algorithm 3 was first used to solve the optimal control problem for a 5-step temperature sequence $\beta_j \in \{0.5, 0.6, 0.7, 0.8, 0.8\}$ with $N_j \in \{10^5, 10^5, 10^5, 10^6, 10^6\}$ realizations. The calculation was repeated with Chebyshev basis functions (49) and localized piecewise linear basis functions (51), with a Chebychev basis up to order 12; for simplicity we choose the cutoff function to be $\phi(r) \equiv 1$ (no cut-off).

For Gaussian type basis functions, it is clear that adding a control force to the dynamics is equivalent to modifying the original potential $V(x)$ by adding a linear combination of Gaussians, and Figure 5 visualizes the added potential. For the other two types of non-integrable basis functions, Figure 6 displays the 2-norm of the control forces along the MPTP. Once the CE-based control forces have been computed, we sample the corresponding cost values $J(u)$ and the average trajectory length are computed from 10^6 independent realizations of the controlled dynamics. The results are shown in Table 5. As in the previous examples, and in agreement with our expectations, the control renders the trajectories to become much shorter compared to the average trajectory length of the original dynamics for $u = 0$.

	$\lambda = 0.5$	$\lambda = 1.0$	$\lambda = 2.0$
Gaussian	5.3 / 7.9	2.6 / 8.8	1.3 / 9.8
Chebyshev	5.2 / 8.0	2.6 / 9.1	1.4 / 10.3
Localized	2.6 / 7.5	1.4 / 8.8	0.9 / 10.2
MC	- / 6.0	- / 7.1	- / 8.4
PDE	- / 6.0	- / 7.0	- / 8.3
PDE-MC	2.6 / 6.1	1.5 / 7.2	0.9 / 8.5

Table 5. Cost value $J(u)$ and average trajectory length for various λ and different basis functions (Gaussian, Chebyshev, localized). The rows "MC" and "PDE" contain the optimal cost from the left hand side of dual relation (17), as obtained by sampling of the uncontrolled dynamics or by solving PDE (9). The row "PDE-MC" records estimates of $\mathbf{E}_{\mu_u}[\tau]$ and $J(u)$ based on sampling the dynamics with the optimal PDE-based control. 10^6 trajectories were simulated in each case, except for "MC", for which 10^5 trajectories were used.

In order to validate our numerical results for the control problem, we also calculate the optimal cost by means of the dual relation (17), whose left hand side can be computed either by directly generating trajectories and computing the expectation or by solving the PDE (9). The results of the brute-force Monte Carlo calculation based on 10^5 trajectories with time step $\Delta t = 10^{-4}$ starting from x_2 is shown in the row "MC" of Table 5. For a thorough comparison, we moreover discretized (9) using the finite volume method (FVM) introduced in [43] on the domain $[-1.5, 1.0] \times [-0.2, 2.0]$ with a regular 1000×1000 grid; the Dirichlet boundary condition with value one (corresponds to $H = 0$) was applied to all discretization boxes inside the Voronoi cell C_{39}. The latter approach leads to a (non-symmetric) linear system of size 10^6, that was solved in parallel using the numerical software package PETSs [44]. We found that the convergence is fast when the BiCGSTAB method (stabilized version of BiConjugate Gradient Squared method) was used [45]. The resulting value function U is plotted in Figure 7. Notice that from (7) and (3), the added potential corresponding to value function U is $2U$, i.e. Figure 7 should be rescaled by a factor 2.0 if one want to compare it with Figure 5. In Table 5, the cost value $U(x_2)$ and the cost obtained by simulating the controlled dynamics with the PDE-based optimal control are recorded (see rows "PDE" and "PDE-MC"). We conclude that the controls obtained from any of the above basis functions leads to a cost value $J(u)$ which is close to the optimal cost. Another general observation is that the control forces increase with λ (see Table 5 or Figures 5 and 6) becomes larger, which is not surprising if one bears in mind that the larger λ, the more weight is put on the minimization of $\mathbf{E}_{\mu_u}[\tau]$, but less on the control penalization.

λ		MC	Gaussian	Chebyshev	Localized	PDE-MC
0.5	$\mathbf{E}\left[e^{-\beta\lambda\tau}\right]$	8.0×10^{-3}	8.0×10^{-3}	8.0×10^{-3}	8.0×10^{-3}	8.0×10^{-3}
	$\sqrt{\mathrm{Var}}$	5.8×10^{-2}	1.4×10^{-2}	2.7×10^{-2}	3.6×10^{-2}	3.5×10^{-3}
	MFHT	258.6	5.3	5.2	2.6	2.6
	Accel(u)	1.0	4.2×10^{2}	1.1×10^{2}	1.3×10^{2}	1.4×10^{4}
1.0	$\mathbf{E}\left[e^{-\beta\lambda\tau}\right]$	3.5×10^{-3}	3.4×10^{-3}	3.4×10^{-3}	3.4×10^{-3}	3.4×10^{-3}
	$\sqrt{\mathrm{Var}}$	3.3×10^{-2}	6.4×10^{-3}	1.4×10^{-2}	2.2×10^{-2}	1.6×10^{-3}
	MFHT	258.9	2.6	2.6	1.4	1.5
	Accel(u)	1.0	1.3×10^{3}	2.8×10^{2}	2.1×10^{2}	3.7×10^{4}
2.0	$\mathbf{E}\left[e^{-\beta\lambda\tau}\right]$	1.2×10^{-3}	1.3×10^{-3}	1.2×10^{-3}	1.3×10^{-3}	1.3×10^{-3}
	$\sqrt{\mathrm{Var}}$	1.9×10^{-2}	2.8×10^{-3}	7.0×10^{-3}	9.8×10^{-3}	6.2×10^{-4}
	MFHT	257.6	1.3	1.4	0.9	0.9
	Accel(u)	1.0	4.6×10^{3}	6.8×10^{2}	5.4×10^{2}	1.3×10^{5}

Table 6. Importance sampling of $\mathbf{E}\left[e^{-\beta\lambda\tau}\right]$ and comparison with standard Monte Carlo. "PDE-MC" records the result of importance sampling based on the PDE solution. For standard Monte Carlo, a time step $\Delta t = 10^{-4}$ and 10^5 trajectories were used, otherwise 10^6 trajectories were simulated at time step $\Delta t = 5 \times 10^{-5}$.

The second numerical test with the Müller-Brown potential involved the sampling problem (8). With G as before, we were aiming at computing $\mathbf{E}\left[e^{-\beta\lambda\tau}\right]$, which can be interpreted as the moment-generating of the random variable $-\beta\tau$ with $\beta > 0$. As before, we compared importance sampling with the CE-based controls and brute-force Monte Carlo and generally observed large gains in the computational efficiency, independently of the control bases used (see Table 6). Upon closer inspection, we observed that the average length of the trajectories was shortest for the localized linear basis set, while the Gaussian basis provided the smallest standard deviations and, consequently, the highest acceleration index. This superiority of the Gaussian basis in this case was clear in despite of the fact that the computational cost of the Gaussian CE minimization is larger than for the other bases sets (due to the global support of the Gaussians). Finally, we repeated the importance sampling calculation with the control that had been obtained from solving the boundary value problem (9). As was to be expected, the variance was smallest and so the largest computational gain was achieved. Note that even with the PDE reference solution at hand, a zero variance importance sampling estimator is out of reach, because of inevitable space and time discretization errors.

6. Conclusions

In this article, we have studied optimal control and adaptive importance sampling from the perspective of reduced-order models of diffusions. The idea is to replace a possibly high-dimensional dynamical system by a simpler (i.e. lower dimensional) dynamics, for which the underlying control problem can be easily solved, where in contrast to our previous work [12, 35] no explicit assumptions on scale separation have been made. Specifically, we have discussed two different scenarios in which reduced-order models provide sufficiently accurate solutions that can be used to control the original (high-dimensional) dynamics close to optimality: Situations in which a suitable collective

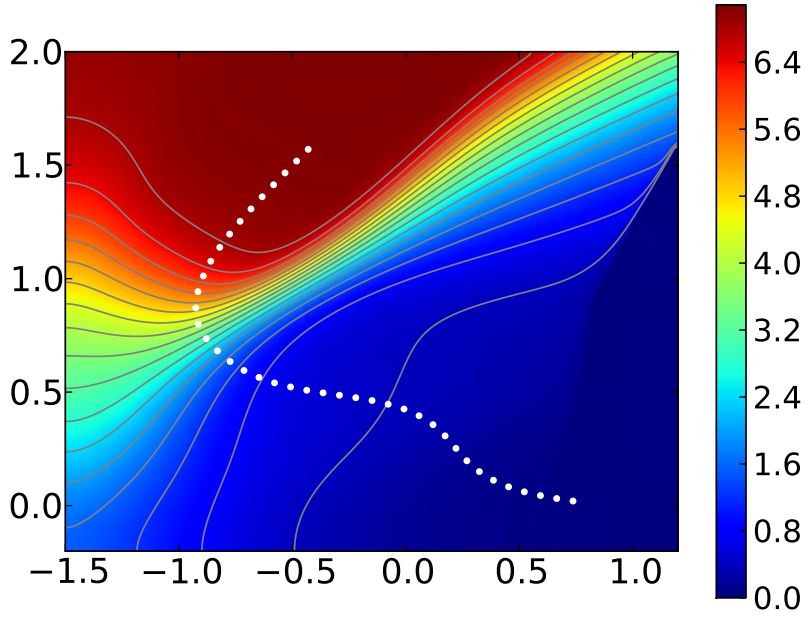


Figure 7. Reference value function U for $\lambda = 1$ based on the finite-volume discretization of (9) on a 1000×1000 grid. For comparison the 40 nodes of the discretized string are shown.

variable that describes the essential dynamics is known, and situations in which the system either exhibits domimant pathways or is close to its viscosity limit. We have proposed numerical algorithms based on first-order optimal prediction [4] and a combination of string method [13] and cross-entropy minimization [19] that can cope with either situation and that are flexible enough to be combined with almost any other available numerical algorithm for solving optimal control and importance sampling problems. All algorithms have been validated and systematically compared to each other on the basis of extensive numerical tests.

This is clearly not the end of the story. Firstly, a thorough test of the algorithms should involve high-dimensional problems, such as a realistic model of a small protein or a peptide, which we have refrained from looking at because the purpose of this article was on systematic tests and comparisons of the algorithms with regard to their accuracy rather than efficiency and implementation issues. Secondly, it would be desirable to derive computable error bounds (e.g. in the spirit of [6, 46]) and better understand the convergence of and the approximations invoked by the cross-entropy method. Thirdly, systematic algorithms how to find good collective variables are still missing and there is room for hope that the duality between sampling and control that was used here can be also used to identify the essential dynamics and suitable coordinates in an iterative manner by a clever combination of free and targeted state space exploration, i.e., by iterating sampling and control in the right way.

Acknowledgement

The authors acknowledge financial support by the Einstein Center of Mathematics (ECMath) and the DFG-CRC 1114 "Scaling Cascades in Complex Systems".

References

[1] B. Leimkuhler and Ch. Matthews. *Molecular Dynamics: With Deterministic and Stochastic Numerical Methods*. Springer, 2015.

[2] J. Di Nunno and B. Øksendal. *Advanced Mathematical Methods for Finance*. Springer, 2011.

[3] A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden. A mathematical framework for stochastic climate models. *Commun. Pure Appl. Math.*, 54:891–974, 2001.

[4] A.J. Chorin, O.H. Hald, and R. Kupferman. Optimal prediction and the Mori-Zwanzig representation of irreversible processes. *Proc. Natl. Acad. Sci. USA*, 97(7):2968–2973, 2000.

[5] A.J. Chorin, O.H. Hald, and R. Kupferman. Optimal prediction with memory. *Physica D*, 166:239–257, 2002.

[6] F. Legoll and T. Lelièvre. Effective dynamics using conditional expectations. *Nonlinearity*, 23(9):2131, 2010.

[7] C. Hartmann. Balanced model reduction of partially observed Langevin equations: an averaging principle. *Math. Comput. Model. Dyn. Syst.*, 17(5):463–490, 2011.

[8] B. Turkington. An optimization principle for deriving nonequilibrium statistical models of Hamiltonian dynamics. *J. Stat. Phys.*, 152:569–597, 2013.

[9] A. J. Majda. Challenges in climate science and contemporary applied mathematics. *Comm. Pure Appl. Math.*, 65(7):920–948, 2012.

[10] D. Givon, R. Kupferman, and A. M. Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):R55–R127, 2004.

[11] C. Schütte, S. Winkelmann, and C. Hartmann. Optimal control of molecular dynamics using markov state models. *Math. Program. Ser. B*, 134(1):259–282, 2012.

[12] C. Hartmann, J.C. Latorre, W. Zhang, and G.A. Pavliotis. Optimal control of multiscale systems using reduced-order models. *J. Comput. Dyn.*, 1(2):279–306, 2014.

[13] W. E, W. Ren, and E. Vanden-Eijnden. String method for the study of rare events. *Phys. Rev. B*, 66:052301, 2002.

[14] W. E, W. Ren, and E. Vanden-Eijnden. Finite temperature string method for the study of rare events. *J. Phys. Chem. B*, 109(14):6688–6693, 2005.

[15] P. Metzner, C. Schütte, and E. Vanden-Eijnden. Transition path theory for markov jump processes. *Multiscale Model Simul.*, 7(3):1192–1219, 2009.

[16] R. Zhao, J. Shen, and R. D. Skeel. Maximum flux transition paths of conformational change. *J. Chem. Theory and Comput.*, 6(8):2411–2423, 2010.

[17] E. Vanden-Eijnden. Transition path theory. In M. Ferrario, G. Ciccotti, and K. Binder, editors, *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology Volume 1*, volume 703 of *Lecture Notes in Physics*, pages 453–493. Springer Berlin Heidelberg, 2006.

[18] W. E and E Vanden-Eijnden. Transition-path theory and path-finding algorithms for the study of rare events. *Annu. Rev. Phys. Chem.*, 61(1):391–420, 2010.

[19] 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(6):A2654–A2672, 2014.

[20] P. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein. A tutorial on the cross-entropy method. *Annals of Operations Research*, 134(1):19–67, 2005.

[21] M. Nic, J. Jirat, and B. Kosata. IUPAC compendium of chemical terminology (gold book), online version, 2012.

[22] S. Huo and J. E. Straub. The MaxFlux algorithm for calculating variationally optimized reaction paths for conformational transitions in many body systems at finite temperature. *J. Chem. Phys.*, 107(13):5000–5006, 1997.

[23] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer, 2006.

[24] P. Dupuis and H. Wang. Importance sampling, large deviations, and differential games. *Stochastics and Stochastic Rep.*, 76(6):481–508, 2004.

[25] C. Hartmann and Ch. Schütte. Efficient rare event simulation by optimal nonequilibrium forcing. *J. Stat. Mech. Theor. Exp.*, 2012:P11004, 2012.

[26] K. Spiliopoulos. Large deviations and importance sampling for systems of slow-fast motion. *Appl. Math. Optim.*, 67:123–161, 2013.

[27] K. Spiliopoulos, P. Dupuis, and X. Zhou. Escaping from an attractor: Importance sampling and rest points, part I. *Ann. Appl. Probab.*, 1303.0450v1, 2014.

[28] E. Vanden-Eijnden and J. Weare. Rare event simulation of small noise diffusions. *Comm. Pure Appl. Math.*, 65(12):1770–1803, 2012.

[29] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 5th edition, 2000.

[30] M. Boué and P. Dupuis. A variational representation for certain functionals of Brownian motion. *Ann. Probab.*, 26(4):1641–1659, 1998.

[31] P. Dai Pra., L. Meneghini, and W. J. Runggaldier. Connections between stochastic control and dynamic games. *Math. Control Signals Systems*, 9:303–326, 1996.

[32] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an ito differential. *Probab. Th. Rel. Fields*, 71(4):501–516, 1986.

[33] G. Ciccotti, T. Lelivre, and E. Vanden-Eijnden. Projection of diffusions on submanifolds: Application to mean force computation. *Comm. Pure Appl. Math.*, 61(3):371–408, 2008.

[34] C. Hartmann. *Model reduction in classical molecular dynamics*. PhD dissertation, Fachbereich Mathematik und Informatik, Freie Universität Berlin, 2007.

[35] C. Hartmann, Ch. Schütte, M. Weber, and W. Zhang. Importance sampling in path space for diffusion processes. *Submitted*, 2015. ArXiv Preprint1502.07899.

[36] M. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*. Springer, 1998.

[37] W. E, W. Ren, and E. Vanden-Eijnden. Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *J. Chem. Phys.*, 126:164103, 2007.

[38] E. Vanden-Eijnden and M. Venturoli. Revisiting the finite temperature string method for the calculation of reaction tubes and free energies. *J. Chem. Phys.*, 130(19):194103, 2009.

[39] R. Y. Rubinstein and D. P. Kroese. *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning (Information Science and Statistics)*. Springer, 1 edition, 2004.

[40] A. Gil, J. Segura, and N. M. Temme. *Numerical Methods for Special Functions*. Society for Industrial and Applied Mathematics, 1 edition, 2007.

[41] K. Müller. Reaction paths on multidimensional energy hypersurfaces. *Angew. Chem. Int. Ed. Engl.*, 19(1):1–13, 1980.

[42] K. Müller and L. Brown. Location of saddle points and minimum energy paths by a constrained simplex optimization procedure. *Theor. Chem. Acc. (Theor. Chim. Acta)*, 53(1):75–93, 1979.

[43] J. C. Latorre, Ph. Metzner, C. Hartmann, and C. Schütte. A structure-preserving numerical discretization of reversible diffusions. *Commun. Math. Sci.*, 9(4):1051 – 1072, 2011.

[44] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, K. Rupp, B. Smith, and H. Zhang. PETSc Web page. `http://www.mcs.anl.gov/petsc`, 2015.

[45] H. A. van der Vorst. Bi-CGSTAB : A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 13(2):631–644, 1992.

[46] K. Spiliopoulos. Non-asymptotic performance analysis of importance sampling schemes for small noise diffusions. *J. Appl. Probab. (to appear)*, 2015.