

# A SPARSE MARKOV CHAIN APPROXIMATION OF LQ-TYPE STOCHASTIC CONTROL PROBLEMS

RALF BANISCH\* AND CARSTEN HARTMANN†

**Abstract.** We propose a novel Galerkin discretization scheme for stochastic optimal control problems on an indefinite time horizon. The control problems are linear-quadratic in the controls, but possibly nonlinear in the state variables, and the discretization is based on the fact that problems of this kind admit a dual formulation in terms of linear boundary value problems. We show that the discretized linear problem is dual to a Markov decision problem, prove an  $L^2$  error bound for the general scheme and discuss the sparse discretization using a basis of so-called committor functions as a special case; the latter is particularly suited when the dynamics are metastable, e.g., when controlling biomolecular systems. We illustrate the method with several numerical examples, one being the optimal control of Alanine dipeptide to its helical conformation.

**1. Introduction.** A large body of literature, going back to the seminal work by Kramers [25] in the late 1930ies, is concerned with the question: How well can a continuous diffusion in a multi-well energy landscape be approximated by a Markov jump process (MJP) in the regime of low temperatures?. Qualitatively, the approximation should be good if the system under consideration is metastable, in which case the process stays in the neighbourhood of the potential energy minima for a long time and occasionally makes rapid transitions (jumps) between the wells. These metastable regions then become the states of the MJP, and the jump rates are determined by the frequency of the transitions (see, e.g., [32, 35]).

In this article we consider the approximation of optimal control problems with nonlinear diffusive dynamics by discrete Markov decision problems. This situation is more complicated than in Kramers' case, for one has to approximate the dynamics as well as the corresponding cost functional and the resulting control forces. The difficulty comes from the fact that dynamics enter as a constraint in the optimization of the cost functional, which makes the problem nonlinear, because the controls become *a priori* unknown functions of the state variables. Specifically, we consider reversible diffusions on an unbounded domain, with a cost functional that is linear-quadratic in the controls, but possibly nonlinear in the state variables, and that is defined up to some random stopping time. The control problem admits a dual formulation in form of linear boundary value problem that is amenable to a discretization by standard means (cf. [37, 48]). Problems of this kind appear relevant in various applications, including molecular dynamics [41], material science [42] or quantum computing [33], to mention just a few. We propose a Markov chain approximation (MCA) that is based on a Galerkin discretization the underlying dynamic programming equations, with basis functions that cover the metastable sets of the dynamics. The discretized problem can be interpreted as a Markov decision problem where the dynamics is given by a continuous-time Markov process on the metastable sets (so-called *core sets*). As these core sets are not assumed to cover the state space, the method is sparse and meshless and hence can be applied to large-scale problems.

MCA are a versatile tool to approximate stochastic control problems. Going back to Kushner [26], the idea of MCA is to approximate the spatially discretized dynamics by a Markov chain and reformulate the underlying continuous control problem as

---

\*Freie Universität Berlin, Institut für Mathematik, Arnimallee 6, 14195 Berlin, Germany (ralf.banisch@fu-berlin.de). This author holds a scholarship from the Berlin Mathematical School.

†Corresponding author. Freie Universität Berlin, Institut für Mathematik, Arnimallee 6, 14195 Berlin, Germany (chartman@mi.fu-berlin.de).

a control problem for the approximating chain. For a spatial discretization with uniform grid, the MCA amounts to a finite-difference discretization of the dynamic programming equation [28, 30]. Recent advances in MCA include collocation-based schemes with radial basis functions [24], finite volume approximations [46], switched systems and jump diffusions [40], or multi-player differential games [39].

The MCA that we propose here for the indefinite time-horizon case is based on a meshless Galerkin discretization of the dual problem. This discretization is standard in that it shares features with related Galerkin methods for linear elliptic equations (see, e.g., [1, 47]), however it is novel in that it preserves the particular structure of the reversible dynamics and the duality of the underlying Markov decision problem when the basis functions form a partition of unity. Related problems with indefinite time-horizon using grid-based techniques have been studied in [22, 43, 27]; cf. also [4] for a (non-Markovian) finite element method. For meshless discretizations of the parabolic dynamic programming equation of finite time-horizon problems we refer to, e.g., [9, 24]; related ideas for the first-order dynamic programming equations of deterministic control have been used in [3, 8].

**A simple paradigm.** As an introductory example consider the one-dimensional diffusion process  $(X_t)_{t \geq 0}$  satisfying the Itô stochastic differential equation

$$dX_t^u = (u_t - \nabla V(X_t^u))dt + \sqrt{2\epsilon}dB_t, t \geq 0 \quad (1.1)$$

where  $B_t$  is standard Brownian motion,  $\epsilon > 0$  is noise intensity, called *temperature* in the following,  $V$  is the bistable potential energy shown in Figure 1.1, and  $u$  is a bounded measurable function, the *control*. Let us suppose that the control task is to force the particle in the left well to reach the right well in minimum time  $\tau$ . When  $u = 0$  and noise is low, the average transition time  $\mathbf{E}[\tau]$  is exponentially large in the energy barrier  $\Delta V$  between the wells, and we can decrease the barrier by tilting the potential according to  $V(x) \mapsto V(x) - ux$ . A controller will then seek to minimize the average transition time by tilting the potential without applying too much force, which leads to a quadratic cost functional of the form [37]

$$J(u) = \mathbf{E} \left[ \tau + \gamma \int_0^\tau |u_t|^2 dt \right], \quad \gamma > 0,$$

that must be minimized over a set of admissible (e.g. adapted) control strategies, subject to the stochastic dynamics (1.1). (Here  $\gamma > 0$  is an adjustable parameter.)

In this paper we deal with the question how to solve optimal control problems of the above form beyond simple one-dimensional examples. The typical application that we have in mind is molecular dynamics that is both high-dimensional and displays vastly different time scales. This defines the basic requirements of the numerical method: it must handle problems with large state space dimension and it must be able to capture the relevant processes of the dynamics, typically the slowest degrees of freedom in the system. For moderate controls, and if the temperature is small compared to the energy barrier, the dynamics in the above example basically consists of rare jumps between the potential wells, with the jump rate being controlled by  $u$ . An efficient discretization would be one that resolves only the jumps between the wells by a 2-state Markov jump process with adjustable jump rates, according to the value of the control (cf. [12, 35]). It is known (see [16]) that control problems of the above form can be transformed into a dual linear boundary value problem that can be approximated by an MJP on the metastable sets. The discretized linear

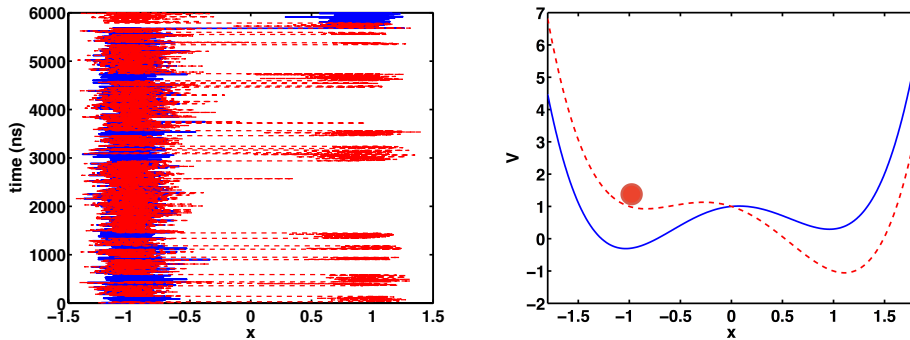


Fig. 1.1: Two typical realizations of the bistable system (1.1), with and without tilting (left panel). The corresponding potential energies are shown in the right panel.

problem turns out to be dual to a Markov decision problem and thus represents the natural Markovian discretization of the original stochastic control problem. The discretization is *meshless*, in that the number of states of the Markov model does not scale exponentially with the dimension of the continuous state space, hence the method avoids the curse of dimensionality of most grid-based schemes.

**Organization of the article.** The rest of the paper is organised as follows: In section 2 we introduce the class of optimal control problems studied and state the duality between optimal control and sampling for both continuous SDEs and MJPs. In section 3, the Galerkin projection method is introduced, and some results about the approximation error are discussed. We also give a stochastic interpretation of the discretized linear equation in terms of Elber’s milestoning process [14]. Finally, we construct sampling estimators. Section 3 is the core parts of the paper and contains new results, including the meshless Galerkin algorithm and an estimate of the discretization error in  $L^2$ . In section 4, we discuss numerical examples.

**1.1. Elementary notation and assumptions.** We implement the following notation and standing assumptions that will be used throughout the paper.

**Dynamics.** Let  $\mathcal{S} = \mathbb{R}^d$  and consider the potential energy function  $V: \mathcal{S} \rightarrow \mathbb{R}$ , that we assume to be two times continuously differentiable and bounded from below. Further assume that  $V(x)$  is polynomially growing at infinity like  $|x|^{2k}$  for some positive integer  $k$ . We consider the process  $X_t \in \mathcal{S}$  solving

$$dX_t^u = (u_t - \nabla V(X_t^u))dt + \sqrt{2\epsilon}dB_t, \quad t \geq 0, \quad (1.2)$$

where  $B_t \in \mathbb{R}^d$  is  $d$ -dimensional Brownian motion under a probability measure  $P$ , and  $u: [0, \infty) \rightarrow U \subset \mathbb{R}^d$  is a time dependent measurable and bounded function.

**Reversibility and invariant measure.** For test functions  $\varphi: \mathcal{S} \rightarrow \mathbb{R}$  that are two times continuously differentiable, the infinitesimal generator of the uncontrolled process  $X_t = X_t^0$  is defined as the second-order differential operator

$$L\varphi = \epsilon\Delta\varphi - \nabla V \cdot \nabla\varphi.$$

Define

$$d\mu(x) = \exp(-\epsilon^{-1}V(x))dx$$

to be the Boltzmann measure at temperature  $\epsilon > 0$ . Without loss of generality, we assume that  $\mu$  is normalized, so that  $\mu(\mathcal{S}) = 1$ . For the subsequent analysis it will be convenient to think of  $L$  as an operator acting on a suitable subspace of

$$L^2(\mathcal{S}, \mu) = \left\{ \phi: \mathcal{S} \rightarrow \mathbb{R}: \int_{\mathcal{S}} |\phi(x)|^2 d\mu(x) < \infty \right\},$$

that is a weighted Hilbert space equipped with the scalar product

$$\langle v, w \rangle_{\mu} = \int_{\mathcal{S}} v(x)w(x)d\mu(x).$$

It can be readily seen that  $L$  is symmetric with respect to the weighted scalar product,

$$\langle Lv, w \rangle_{\mu} = \langle v, Lw \rangle_{\mu},$$

which implies that  $X_t$  is reversible with respect to the Boltzmann measure  $\mu$ . Moreover, by the above assumptions on the potential energy function,  $\mu$  is the unique invariant measure of the process  $X_t$  and satisfies

$$\int_{\mathcal{S}} (L\psi)d\mu = \int_{\mathcal{S}} \psi(L\mathbf{1})d\mu = 0$$

for all test functions  $\psi \in L^2(\mathcal{S}, \mu)$ ; see [29] for details.

**Quadratic cost criterion.** We now introduce the cost criterion that the controller choosing  $u$  in (1.2) seeks to minimize. To this end let  $A \subset \mathcal{S}$  be a closed bounded subset of positive measure  $\mu(A) > 0$  with smooth (at least  $C^3$ ) boundary  $\partial A$  and call  $\tau_A < \infty$  the random stopping time

$$\tau_A = \inf\{t > 0: X_t \in \partial A\}.$$

We define the cost functional

$$J(u) = \mathbf{E} \left[ \int_0^{\tau_A} \left\{ f(X_t^u) + \frac{1}{4}|u_t|^2 \right\} dt \right], \quad (1.3)$$

where  $f: \mathcal{S} \rightarrow \mathbb{R}$ , called *running cost*, is any bounded nonnegative function with bounded first derivative; the factor 1/4 in the penalization term is merely conventional. Cost functionals of this form are called *indefinite time horizon cost*, because the terminal time  $\tau_A$  is random. We will sometimes need the conditioned variant of the above cost functional:

$$J(u; x) = \mathbf{E}_x \left[ \int_0^{\tau_A} \left\{ f(X_t^u) + \frac{1}{4}|u_t|^2 \right\} dt \right]. \quad (1.4)$$

Here  $\mathbf{E}_x[\cdot] = \mathbf{E}[\cdot | X_0 = x]$  is a shorthand for the expectation over all realizations of  $X_t$  starting at  $X_0 = x$ , i.e. the expectation with respect to  $P$  conditional on  $X_0 = x$ . We use  $J$  to denote both unconditional and conditional cost functionals; it should always be clear from the context, which one is meant.

**Admissible control strategies.** We call a control strategy  $u = (u_t)_{t \geq 0}$  *admissible* if it is adapted to the filtration generated by  $B_t$ , i.e., if  $u_t$  depends only on the history of the Brownian motion up to time  $t$ , and if the equation for  $X_t^u$  has a unique strong solution. The set of admissible strategies is denoted by  $\mathcal{A}$ .

Even though  $u_t$  may depend on the entire past history of the process up to time  $t$ , it turns out that optimal strategies are Markovian, i.e., they depend only on the current state of the system at time  $t$ . In our case, in which the costs are accumulated up to a random stopping time  $\tau_A$ , the optimal strategies are of the form

$$u_t = \alpha(X_t^u)$$

for some function  $\alpha: \mathcal{S} \rightarrow \mathbb{R}^d$ . Hence the optimal controls are time-homogeneous feedback policies, depending only on the current state  $X_t^u$ , but not on  $t$ .

**2. Optimal control and logarithmic transformation.** In this section we establish a connection between controlled diffusions and certain path sampling problems, the latter are associated with a linear boundary value partial differential equation (PDE) that can be discretized by standard numerical techniques for PDEs or Monte-Carlo. The duality between optimal control and path sampling goes back to Wendel Fleming and co-workers (see, e.g, [15]) and is based on a logarithmic transformation of the value function (see also [16, Sec. VI] and the references therein)

$$W(x) = \min_{u \in \mathcal{A}} J(u; x). \quad (2.1)$$

**2.1. Duality between control and path sampling for diffusions.** Our simple derivation of the duality between path sampling optimal control will be based on the Hamilton-Jacobi-Bellman equations of optimal control. To this end, we recall the dynamic programming principle for optimal control problems of the form (1.2)–(1.3) that we adapt from [16, Secs. VI.3–5] and that we state without proof.

**THEOREM 2.1.** *Let  $W \in C^2(\mathcal{S} \setminus A) \cap C(\partial A)$  be the solution of*

$$\begin{aligned} \min_{c \in \mathbb{R}^d} \left\{ (L + c \cdot \nabla)W(x) + f + \frac{1}{4}|c|^2 \right\} &= 0, \quad x \in \mathcal{S} \setminus A \\ W(x) &= 0, \quad x \in \partial A. \end{aligned} \quad (2.2)$$

Then

$$W(x) = \min_{u \in \mathcal{A}} J(u; x)$$

where the minimizer  $u^* = \operatorname{argmin} J(u)$  is unique and given by the feedback law

$$u_t = -2\nabla W(X_t^u). \quad (2.3)$$

Before we proceed with the derivation of the dual sampling problem, we shall briefly discuss some of the consequences of the dynamic programming approach. Equation (2.2) is the Hamilton-Jacobi-Bellman (HJB) equation, also called *dynamic programming equation* associated with the following optimal control task:

$$\min_{u \in \mathcal{A}} J(u) \quad \text{s.t.} \quad dX_t^u = (u_t - \nabla V(X_t^u))dt + \sqrt{2\epsilon}dB_t.$$

The function  $W(x)$  is called *value function* or *optimal cost-to-go*. Using the fact that optimal control is the gradient of two times the value function, the optimally controlled process  $X_t^*$  solves the SDE

$$dX_t^* = -\nabla U(X_t^*)dt + \sqrt{2\epsilon}dB_t. \quad (2.4)$$

with the new potential

$$U(x) = V(x) + 2W(x).$$

Note that  $X_t^*$  is reversible with respect to a tilted Boltzmann distribution having the density  $\rho^* = \exp(-U/\epsilon)$ . The reversibility follows from the fact that the value function does not depend on  $t$ , which would not be the case if the terminal time  $\tau_A$  were a deterministic stopping time rather than a first exit time.<sup>1</sup>

**Logarithmic transformation and Feynman-Kac formula (part I).** The approach that is pursued in this article is to discretize the HJB equation by first removing the nonlinearity by a logarithmic transformation of the value function. Let

$$\phi(x) = \exp(-\epsilon^{-1}W(x)). \quad (2.5)$$

It follows by chain rule that

$$\epsilon \frac{L\phi}{\phi} = -LW + |\nabla W|^2, \quad \phi \neq 0, \quad (2.6)$$

which, together with the relation

$$-|\nabla W|^2 = \min_{c \in \mathbb{R}^m} \left\{ c \cdot \nabla W + \frac{1}{4}|c|^2 \right\},$$

implies that (2.2) is equivalent to the linear boundary value problem

$$\begin{aligned} (L - \epsilon^{-1}f)\phi(x) &= 0, & x \in \mathcal{S} \setminus A \\ \phi(x) &= 1, & x \in \partial A. \end{aligned} \quad (2.7)$$

By the above assumptions and the strong maximum principle for elliptic PDEs it follows that (2.7) has a classical solution  $\phi \in C^2(\mathcal{S} \setminus A) \cap C(\partial A)$  that is uniformly bounded away from zero. The latter, together with (2.5)–(2.6), implies existence and uniqueness of classical solutions of that the dynamic programming equation (2.2) and hence smoothness of the value function.

Now, by the Feynman-Kac theorem [31, Thm. 8.2.1], the linear boundary value problem has an interpretation in terms of a sampling problem. The solution (2.7) can be expressed as the conditional expectation

$$\phi(x) = \mathbf{E}_x \left[ \exp \left( -\frac{1}{\epsilon} \int_0^{\tau_A} f(X_t) ds \right) \right] \quad (2.8)$$

over all realizations of the following SDE on  $\mathcal{S}$ :

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\epsilon}dB_t, \quad X_0 = x. \quad (2.9)$$

---

<sup>1</sup>For finite time-horizon control problems the value function depends on the time  $\tau_A - t$  remaining until the terminal time  $\tau_A$ .

**2.2. Duality between control and path sampling for jump processes.** In the last section, we have established a connection between an optimal control problem and sampling of a continuous path observable  $\phi(x)$ . In this section, we will repeat the same construction for MJP, however, in reverse order: starting from a path observable for a MJP, we derive the dual optimal control using a logarithmic transformation.

Let  $(\hat{X}_t)_{t \geq 0}$  be a MJP on the discrete state space  $\hat{\mathcal{S}} = \{1, \dots, n\}$  with infinitesimal generator  $G \in \mathbb{R}^{n \times n}$ . (For simplicity, we assume that  $\hat{\mathcal{S}}$  is finite.) The entries of the generator matrix  $G$  satisfy

$$G_{ij} \geq 0 \text{ for } i \neq j \text{ and } G_{ii} = -\sum_{j \neq i} G_{ij},$$

where the off-diagonal entries of  $G$  are the jump rates between the states  $i$  and  $j$ .

**Logarithmic transformation and Feynman-Kac formula (part II).** In accordance with the previous subsection let  $\hat{f}: \hat{\mathcal{S}} \rightarrow \mathbb{R}$  be nonnegative and define the stopping time  $\tau_A = \inf\{t > 0: \hat{X}_t \in A\}$  to be the first hitting time of a subset  $A \subset \hat{\mathcal{S}}$ . As before we introduce a function

$$\hat{\phi}(i) = \mathbf{E}_i \left[ \exp \left( -\frac{1}{\epsilon} \int_0^{\tau_A} \hat{f}(\hat{X}_s) ds \right) \right],$$

with  $\mathbf{E}_i[\cdot] = \mathbf{E}[\cdot | \hat{X}_0 = i]$  being the conditional expectation over the realizations of  $\hat{X}_t$  starting at  $\hat{X}_0 = i$ . We have the following lemma that is the exact analogue of the Feynman-Kac formula for diffusions for the case of an MJP (see [17]).

LEMMA 2.2. *The function  $\hat{\phi}(i)$  solves the linear boundary value problem*

$$\begin{aligned} \sum_{j \in \hat{\mathcal{S}}} G_{ij} \hat{\phi}(j) - \epsilon^{-1} \hat{f}(i) \hat{\phi}(i) &= 0, \quad i \in \hat{\mathcal{S}} \setminus A \\ \hat{\phi}(i) &= 1, \quad i \in A. \end{aligned} \tag{2.10}$$

Now, in one-to-one correspondence with the log transformation procedure in the diffusion case, the function

$$\hat{W} = -\epsilon \log \hat{\phi}$$

can be interpreted as the value function of an optimal control problem for the MJP  $(\hat{X}_t)_{t \geq 0}$ . The derivation of the dual optimal control problem goes back to [38], and we repeat it here in condensed form for the reader's convenience (see also [16, Sec. VI.9]): First of all note that  $\hat{W}$  satisfies the equation

$$\begin{aligned} \exp(\hat{W}/\epsilon) G \exp(-\hat{W}/\epsilon) - \epsilon^{-1} \hat{f} &= 0, \quad i \in \hat{\mathcal{S}} \setminus A \\ \hat{W}(i) &= 0, \quad i \in A. \end{aligned}$$

and define a new generator matrix by

$$G^v = (G_{ij}^v)_{i,j \in \hat{\mathcal{S}}}, \quad G_{ij}^v = \frac{G_{ij} v(j)}{v(i)}, \tag{2.11}$$

with  $v(i) > 0$  for all  $i \in \hat{\mathcal{S}}$ . The exponential term in the above equation for  $\hat{W}$  can be recast as

$$\frac{(G \hat{\phi})(i)}{\hat{\phi}(i)} = \epsilon^{-1} \min_{v > 0} \{-(G^v \hat{W})(i) + k^v(i)\}$$

where we have introduced the shorthand

$$k^v(i) = \epsilon(G^v(\log v))(i) - \epsilon \frac{(Gv)(i)}{v(i)},$$

and used the identity  $\min_{y \in \mathbb{R}} \{e^{-y} + ay\} = a - a \log a$  for  $a > 0$ . As a consequence, (2.10) is equivalent (i.e. dual) to

$$\begin{aligned} \min_{v > 0} \left\{ (G^v \hat{W})(i) + k^v(i) + \hat{f}(i) \right\} &= 0, \quad i \in \hat{\mathcal{S}} \setminus A \\ \hat{W}(i) &= 0, \quad i \in A. \end{aligned} \quad (2.12)$$

which is the dynamic programming equation of a Markov decision problem, i.e. an optimal control problem for an MJP (e.g. see [16, Sec. VI.9]): Minimize

$$\hat{J}(v) = \mathbf{E} \left[ \int_0^{\tau_A} \left\{ \hat{f}(\hat{X}_s^v) + k^v(\hat{X}_s^v) \right\} ds \right] \quad (2.13)$$

over all controls  $v > 0$  and subject to the constraint that the process  $(\hat{X}_t^v)_{t \geq 0}$  is generated by  $G^v$ . It readily follows from the derivation of (2.12) that the minimizer exists and is given by  $v^*(i) = \hat{\phi}(i)$ . The next lemma records some important properties of the controlled Markov jump process with generator  $G^v$  and the corresponding cost functional (2.13).

LEMMA 2.3. *Let  $G^v$  and  $k^v$  be defined as above.*

- (i) *Let  $G$  be reversible with unique stationary distribution  $\pi$ . Then  $\pi^v(i) = Z_v^{-1} v^2(i) \pi(i)$ , with  $Z_v$  an appropriate normalization constant, is the unique probability distribution such that  $G^v$  is reversible with stationary distribution  $\pi^v$ .*
- (ii) *Let  $\hat{P}$  denote the probability measure on the space of trajectories generated by  $\hat{X}_t$  with initial condition  $\hat{X}_0 = i$ , and let  $\hat{Q}$  be the corresponding probability measure generated by  $\hat{X}_t^v$  with the same initial condition  $\hat{X}_0^v = i$ . Then  $\hat{Q}$  is absolutely continuous with respect to  $\hat{P}$  and the expected value of the running cost  $k^v$  is the Kullback-Leibler (KL) divergence between  $\hat{Q}$  and  $\hat{P}$ , i.e.,*

$$\mathbf{E}_{\hat{Q}} \left[ \int_0^{\tau_A} k^v(\hat{X}_s^v) ds \right] = \int \log \frac{d\hat{Q}}{d\hat{P}} d\hat{Q}$$

where  $\mathbf{E}_{\hat{Q}}[\dots]$  is the expectation over all realizations of  $\hat{X}_t^v$  starting at  $\hat{X}_0^v = i$ .

*Proof.* We first show (i). By assumption we have  $\pi(i)G_{ij} = \pi(j)G_{ji}$ . Now, let  $\pi^v$  be such that  $\pi^v(i)G_{ij}^v = \pi^v(j)G_{ji}^v$ . We will show that  $\pi^v$  has the proposed form:

$$\pi^v(i)G_{ij}^v = \frac{v(j)}{v(i)} \frac{\pi^v(i)}{\pi(i)} \pi(i)G_{ij} = \frac{v(j)}{v(i)} \frac{\pi^v(i)}{\pi(i)} \pi(j)G_{ji} = \frac{v^2(j)}{v^2(i)} \frac{\pi(j)}{\pi(i)} \frac{\pi^v(i)}{\pi^v(j)} \pi^v(j)G_{ji}^v$$

But since  $\pi^v(i)G_{ij}^v = \pi^v(j)G_{ji}^v$ , we must have

$$\frac{\pi^v(j)}{\pi(j)v^2(j)} = \frac{\pi^v(i)}{\pi(i)v^2(i)} \quad \forall i \neq j.$$

This can only be true if the quantity  $Z_v^{-1} = \frac{\pi^v(i)}{\pi(i)v^2(i)}$  is independent of  $i$ . This gives  $\pi^v(j) = Z_v^{-1} v^2(j) \pi(j)$  as desired. The constant  $Z_v$  is uniquely determined by the



requirement that  $\pi^v$  be normalized. Finally, from reversibility it follows directly that  $\pi^v$  is also a stationary distribution of  $G^v$ .

To show (ii), note that the running cost  $k^v(i)$  can be written as

$$k^v(i) = \epsilon \sum_{j \neq i} G_{ij} \left\{ \frac{v(j)}{v(i)} \left[ \log \frac{v(j)}{v(i)} - 1 \right] + 1 \right\}, \quad (2.14)$$

which is the KL divergence between  $\hat{Q}$  and  $\hat{P}$  (see [10, Sec. 3.1.4]). The absolute continuity between  $\hat{Q}$  and  $\hat{P}$  simply follows from the fact that  $v$  in the definition of  $G^v$  was required to be component-wise strictly positive.  $\square$

REMARK 1. *To reveal further similarities between (1.2)–(1.3) and the corresponding Markov decision problem, note that the quadratic penalization term in (1.3) equals the KL divergence between the reference measure  $P$  of the uncontrolled diffusion (2.9) and the probability measure  $Q$  of the controlled process (1.2), as can be shown using Girsanov’s theorem [31, Thm. 8.6.8]. It holds that (cf. [20, 18]):*

$$\mathbf{E}_Q \left[ \frac{1}{4} \int_0^{\tau_A} |u_s|^2 ds \right] = \epsilon \int \log \frac{dQ}{dP} dQ.$$

**3. Discretization: Galerkin projection point of view.** In this section we will develop a discretization for optimal control problems of the type discussed in Section 2 using the method of logarithmic transformations. The discretization will approximate the continuous control problem with a control problem for a Markov jump process on finite state space. Our strategy is outlined in Figure 3.1.

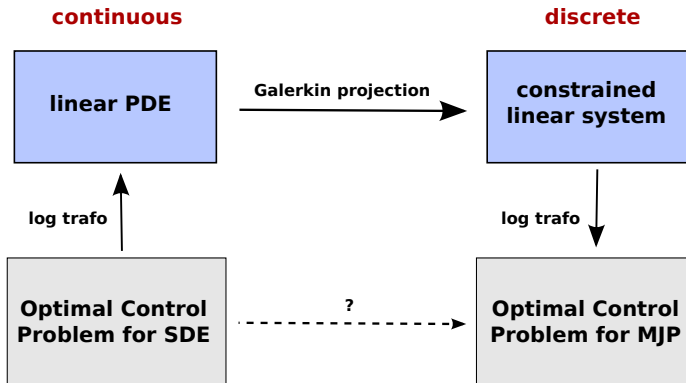


Fig. 3.1: Discretization of continuous control problems via a logarithmic transform.

In the first part of this section, we will develop the Galerkin projection for general subspaces and obtain some control of the discretization error. To refine this control, we specify the subspace  $D$  we project onto. As the state space is unbounded and possibly high-dimensional, a grid-based discretization is prohibitive. Here we suggest a meshless discretization based on an incomplete partition of state space into so called *core sets*, that are the metastable regions of the uncontrolled dynamics. We will prove an error bound which gives us detailed control over the discretization error, even if very few basis functions are used. We should mention that clearly other choices

are possible, such as radial basis functions [47] or moving least-squares [1], but for metastable systems like in molecular dynamics or chemical reaction kinetics using core sets and the associated basis of committor functions is beneficial.

In the second part of this section, we will develop the stochastic interpretation of the resulting matrix equation as the backward Kolmogorov equation of a MJP, which enables us to identify the discrete control problem for the MJP, as it was developed in Section 2. We will study the resulting discrete control problem and make a connection to Transition Path Theory [45] and core set Markov state models (MSM) [36].

**3.1. Galerkin projection of the Dirichlet problem.** As discussed above, we consider the boundary value problem

$$\begin{aligned} (L - \epsilon^{-1}f)\phi(x) &= 0, & x \in \mathcal{S} \setminus A \\ \phi(x) &= 1, & x \in \partial A. \end{aligned} \tag{3.1}$$

with  $L$  and  $f$  as given above. We declare that  $\phi|_A = 1$ , so that the domain of  $\phi$  is  $\mathcal{S}$ .

Following standard references (e.g. [6]) we construct a Galerkin projection of (3.1). For this purpose, we introduce the  $L^2$ -based Sobolev space  $H^1$  with norm  $\|\phi\|_{H^1} = \|\nabla u\|_\mu^2 + \|u\|_\mu^2$  and the Hilbert spaces  $V = \{\psi \in L^2(\mathcal{S}, \mu), \|\psi\|_{H^1} < \infty\}$  and  $V_0 = \{\psi \in V, \psi|_{\partial A} = 0\}$ . We further define the symmetric and positive bilinear form

$$\mathcal{B} : V \times V \rightarrow \mathbb{R}, \quad \mathcal{B}(\phi, \psi) = \epsilon^{-1}\langle f\phi, \psi \rangle_\mu + \epsilon\langle \nabla\phi, \nabla\psi \rangle_\mu.$$

Now if  $\phi$  is a solution of (3.1), then it also solves the weak problem

$$\mathcal{B}(\phi, \psi) = 0 \quad \forall \psi \in V_0. \tag{3.2}$$

A Galerkin solution  $\hat{\phi}$  is now any function satisfying

$$\mathcal{B}(\hat{\phi}, \hat{\psi}) = 0 \quad \forall \hat{\psi} \in D_0, \tag{3.3}$$

with  $D_0$  being a suitable finite dimensional subspace  $V_0$ . Specifically, we choose basis functions  $\chi_1, \dots, \chi_{n+1}$  with the following properties:

- (S1) The functions  $\chi_i : \mathcal{S} \rightarrow \mathbb{R}$  are in  $V$ .
- (S2) The  $\chi_i$  form a partition of unity, that is  $\sum_{i=1}^{n+1} \chi_i = \mathbb{1}$ .
- (S3) The  $\chi_i$  satisfy  $\chi_{n+1}|_A = 1$  and  $\chi_i|_A = 0$  for  $i \in \{1, \dots, n\}$ .

All elements of  $D_0 := \text{lin}\{\chi_1, \dots, \chi_n\}$  will satisfy homogeneous Dirichlet boundary conditions in (3.1), and we will sometimes write  $D := \chi_{n+1} \oplus D_0$  and think of the Galerkin solution  $\hat{\phi}$  as an element in  $D$ . Now define the matrices

$$F_{ij} = \frac{\langle \chi_i, f\chi_j \rangle_\mu}{\langle \chi_i, \mathbb{1} \rangle_\mu}, \quad K_{ij} = -\epsilon \frac{\langle \nabla\chi_i, \nabla\chi_j \rangle_\mu}{\langle \chi_i, \mathbb{1} \rangle_\mu}.$$

Setting  $\hat{\phi} = \sum_i \hat{\phi}_i \chi_i$ , the weak form (3.3) becomes a matrix equation for the unknown coefficients  $\hat{\phi}_i$ :

$$\sum_{j=1}^{n+1} (K_{ij} - \epsilon^{-1}F_{ij}) \hat{\phi}_j = 0, \quad i \in \{1, \dots, n\} \tag{3.4}$$

$$\hat{\phi}_{n+1} = 1,$$

which is the discretization of (3.1).

In order to control the discretization error of the Galerkin method, we choose a norm  $\|\cdot\|$  on  $V$  and introduce the two error measures:

1. The *Galerkin error*  $\varepsilon = \|\phi - \hat{\phi}\|$ , i.e. the difference between original and Galerkin solution measured in  $\|\cdot\|$ .
2. The *best approximation error*  $\varepsilon_0 = \inf_{\hat{\psi} \in D} \|\phi - \hat{\psi}\|$ , i.e. the minimal difference between the solution  $\phi$  and any element  $\hat{\psi} \in D$ .

In order to obtain full control over the discretization error, we need bounds on  $\varepsilon$ , and we will get them by first obtaining a bound on the performance  $\mathbf{p} := \varepsilon/\varepsilon_0$  and then a bound on  $\varepsilon_0$ . The latter will depend on the choice of subspace  $D$ . For the former, standard estimates assume the following  $\|\cdot\|$ -dependent properties of  $A$ :

- (i) Boundedness:  $\mathcal{B}(\phi, \psi) \leq \alpha_1 \|\phi\| \|\psi\|$  for some  $\alpha_1 > 0$
- (ii) Ellipticity: for all  $\phi \in V$  holds  $\mathcal{B}(\phi, \phi) \geq \alpha_2 \|\phi\|^2$  for some  $\alpha_2 > 0$ .

If both (i) and (ii) hold, Céa's lemma states that  $\mathbf{p} \leq \frac{\alpha_1}{\alpha_2}$ , see e.g. [6]. For the *energy norm*  $\|\phi\|_B^2 := \mathcal{B}(\phi, \phi)$  we have  $\alpha_1 = \alpha_2 = 1$  and therefore  $\mathbf{p} = 1$ , thus the Galerkin solution  $\hat{\phi}$  is the best-approximation to  $\phi$  in the energy norm.

**Performance bound.** The next two statements give a bound on  $\mathbf{p}$  if errors are measured in the  $L^2$ -norm. In this case,  $\mathcal{B}(\cdot, \cdot)$  is still elliptic but possibly unbounded. Later in this section, we will specify the bound on  $\varepsilon_0$  for a specific Galerkin basis.

**THEOREM 3.1.** *Let  $\mathcal{B}$  be elliptic. Further let*

$$Q: L^2(\mathcal{S}, \mu) \rightarrow D_0 \subset L^2(\mathcal{S}, \mu), \quad Qw = \sum_{i=1}^n \langle \chi_i, w \rangle \chi_i$$

be the orthogonal projection onto  $D_0$ . Then

$$\mathbf{p}^2 = \left( \frac{\varepsilon}{\varepsilon_0} \right)^2 \leq 1 + \frac{1}{\alpha_2^2} \sup_{v \in V} \frac{\|QB(\mathbb{1} - Q)v\|_\mu^2}{\|v\|_\mu^2}$$

where  $B = \varepsilon^{-1}f - L$  is the linear operator associated with  $\phi \mapsto \mathcal{B}(\cdot, \phi)$ .

*Proof.* In Appendix A.  $\square$

**REMARK 2.** *Note that  $\|QB(\mathbb{1} - Q)v\|_\mu \leq \|QBv\|_\mu$  is always finite even though  $B$  is possibly unbounded since  $v \in V \subset L^2(\mathcal{S}, \mu)$  and  $Q$  is the projection onto a finite-dimensional subspace of  $L^2(\mathcal{S}, \mu)$ .*

The bottom line of Theorem 3.1 is that if  $B$  leaves the subspace  $D$  almost invariant, then  $\hat{\phi}$  is almost the best-approximation of  $\phi$  in  $\|\cdot\|_\mu$ . The following lemma gives a more detailed description. In the following, we will write  $\|\cdot\| = \|\cdot\|_\mu$  for convenience.

**LEMMA 3.2.** *Let  $Q^\perp = \mathbb{1} - Q$  and define*

$$\delta_L := \max_k \|Q^\perp L \chi_k\|, \quad \delta_f := \max_k \|Q^\perp \varepsilon^{-1} f \chi_k\|$$

to be the maximal projection error of the images of the  $\chi_k$ 's under  $L$  and  $f$ . Then

$$\|QBQ^\perp\| = \|Q^\perp BQ\| \leq (\delta_L + \delta_f) \sqrt{\frac{n}{m}}$$

where  $m$  is the smallest eigenvalue of  $\hat{M}$ .

*Proof.* The first statement is true since  $A$  is essentially self-adjoint. For the second statement, first of all

$$\|Q^\perp BQ\| = \|Q^\perp(\epsilon^{-1}f - L)Q\| \leq \|Q^\perp \epsilon^{-1}fQ\| + \|Q^\perp LQ\|$$

holds from the triangle inequality. We now bound the term involving  $L$ . Notice that for  $\hat{\phi} = \sum_i \hat{\phi}_i \chi_i \in D$ :

$$\|Q^\perp L\hat{\phi}\| = \left\| \sum_i \hat{\phi}_i Q^\perp L\chi_i \right\| \leq \delta_L \sum_i |\hat{\phi}_i| = \delta_L \|\hat{\phi}\|_1.$$

Then, with  $\hat{M} := (\langle \chi_i, \chi_j \rangle_\mu)_{ij}$ :

$$\|Q^\perp LQ\| = \sup_{\phi = \phi_{||} + \phi_\perp \in V} \frac{\|Q^\perp L\phi_{||}\|}{\|\phi\|} \leq \sup_{\phi_{||} \in D} \frac{\|Q^\perp L\phi_{||}\|}{\|\phi_{||}\|} \leq \delta_L \sup_{\hat{\phi} \in \mathbb{R}^n} \frac{\|\hat{\phi}\|_1}{\sqrt{\langle \hat{\phi}, \hat{\phi} \rangle_M}}$$

A similar result holds for the term involving  $f$ . The statement now follows from a standard equivalence between finite-dimensional norms,  $\|\hat{\phi}\|_1 \leq \sqrt{n}\|\hat{\phi}\|_2$ , and the fact that  $\hat{M}$  is symmetric, which implies that  $\langle \hat{\phi}, \hat{\phi} \rangle_M = \hat{\phi}^T \hat{M} \hat{\phi} \geq m \hat{\phi}^T \hat{\phi} = m \|\hat{\phi}\|_2^2$ .  $\square$

To summarize, Theorem 3.1 and Lemma 3.2 give us a formula for the projection performance  $p$  which states that

$$p^2 \leq 1 + \frac{n}{m} \frac{(\delta_L + \delta_f)^2}{\alpha_2^2}.$$

How large or small  $\delta_f$  is will depend on the behaviour of  $f$ , e.g., if  $f = \text{const}$  then  $\delta_f = 0$ . Both  $\delta_f$  and  $\delta_L$  are always finite even though  $L$  is possibly unbounded.

**Best-approximation error bound.** We now generalize results [12] on the approximation quality of MSMs for reversible equilibrium diffusions and estimate the best-approximation error  $\epsilon_0$  for the case that the subspace  $D$  is spanned by committor functions associated with the metastable sets of the dynamics. To this end suppose that the potential  $V(x)$  has  $n + 1$  deep minima  $x_1, \dots, x_{n+1}$ . Let  $C_1, \dots, C_{n+1}$  be convex *core* sets around  $x_1, \dots, x_{n+1}$  and such that  $A = C_{n+1}$ . We write  $C = \cup_{i=1}^{n+1} C_i$  and  $T = \mathcal{S} \setminus C$  and introduce  $\tau_C = \inf\{t \geq 0 : X_t \in C\}$ . We take  $\chi_i$  to be the committor function associated with the set  $C_i$ , that is

$$\chi_i(x) = \mathbf{P}(X_{\tau_C} \in C_i | X_0 = x). \quad (3.5)$$

These  $\chi_i$  satisfy the assumptions (S2)–(S3) and (S1) expect on the core set boundaries, which is a set of measure zero. Since we do not have a grid parameter, by which the approximation error can be controlled, standard PDE techniques for bounding  $\epsilon_0$  fail. Indeed, typically we will have very few basis functions compared to a grid-like discretization. The following theorem gives a bound on  $\epsilon_0$ .

**THEOREM 3.3.** *Let  $Q$  be the orthogonal projection onto the subspace  $D$  spanned by the committor functions (3.5), and let  $\phi$  be the solution of (3.1). Then we have*

$$\epsilon_0 = \|Q^\perp \phi\|_\mu \leq \|P^\perp \phi\|_\mu + \mu(T)^{1/2} [\kappa \|f\|_\infty + 2\|P^\perp \phi\|_\infty]$$

where  $\|\cdot\| = \|\cdot\|_\mu$ ,  $\kappa = \sup_{x \in T} \mathbf{E}_x[\tau_{\mathcal{S} \setminus T}]$ , and  $P$  is the orthogonal projection onto  $V_c = \{v \in L^2(\mathcal{S}, \mu), v|_{C_i} = \text{const on every } C_i\} \subset L^2(\mathcal{S}, \mu)$ , with  $P^\perp = \mathbf{1} - P$ .

*Proof.* In Appendix B  $\square$

In Theorem 3.3,  $\kappa$  is the maximum expected time of hitting the metastable set from outside (which is short). Note further that  $P^\perp \phi = 0$  on  $T$ . The errors  $\|P^\perp \phi\|_\mu$  and  $\|P^\perp \phi\|_\infty$  measure how constant the solution  $\phi$  is on the core sets. Theorem 3.3 suggest the following strategy to minimize  $\varepsilon_0$ : (i) Place a core set  $C_i$  in every metastable region where  $\phi$  is expected to be almost constant, (ii) place core sets in regions with high invariant density  $\mu$  in order to minimize  $\mu(T)$ . This strategy requires knowledge of the invariant density  $\mu$ . Identifying the metastable regions requires additional dynamical information. If this is not available, then a good guess is usually to use the deepest wells of  $\mu$ .

REMARK 3. *Theorem 3.3 together with Theorem 3.1 gives us full control over the discretization error  $\varepsilon$ . These error bounds generalize recent results [35, 12] on the approximation quality of the dominant eigenvalues of a reversible diffusion by an MSM to general (bounded, nonnegative) observables. It would of course be nice to have an error estimate also for the value function. In general such an estimate is difficult to get, because of the nonlinear logarithmic transformation  $W = -\varepsilon \log \phi$  involved. However we know that  $\phi$  and its discrete approximation are both uniformly bounded and bounded away from zero. Hence the logarithmic transformation is uniformly Lipschitz continuous on its domain, which implies that the  $L^2$  error bounds holds for the value function with an additional prefactor given by the Lipschitz constant squared; for a related argument see [19]*

**3.2. Interpretation in terms of a Markov decision problem.** We derive an interpretation of the discretized equation (3.4) in terms of a MJP. We introduce the diagonal matrix  $\Lambda$  with entries  $\Lambda_{ii} = \sum_j F_{ij}$  (zero otherwise) and the full matrix  $G = K - \varepsilon^{-1}(F - \Lambda)$ , and rearrange (3.4) as follows:

$$\begin{aligned} \sum_{j=1}^{n+1} (G_{ij} - \varepsilon^{-1} \Lambda_{ij}) \hat{\phi}_j &= 0, \quad i \in \{1, \dots, n\} \\ \hat{\phi}_{n+1} &= 1, \end{aligned} \tag{3.6}$$

This equation can be given a stochastic interpretation. To this end let us introduce the vector  $\pi \in \mathbb{R}^{n+1}$  with nonnegative entries  $\pi_i = \langle \chi_i, \mathbf{1} \rangle$  and notice that  $\sum_i \pi_i = 1$  follows immediately from the fact that the basis functions  $\chi_i$  form a partition of unity, i.e.  $\sum_i \chi_i = \mathbf{1}$ . This implies that  $\pi$  is a probability distribution on the discrete state space  $\hat{\mathcal{S}} = \{1, \dots, n+1\}$ . We summarize properties of the matrices  $K$ ,  $F$  and  $G$ :

LEMMA 3.4. *Let  $K$ ,  $G$ ,  $F$  and  $\pi$  be as above.*

(i)  *$K$  is a generator matrix (i.e.  $K$  is a real-valued square matrix with row sum zero and positive off-diagonal entries) with stationary distribution  $\pi$  that satisfies detailed balance*

$$\pi_i K_{ij} = \pi_j K_{ji}, \quad i, j \in \hat{\mathcal{S}}$$

(ii)  *$F \geq 0$  (entry-wise) with  $\pi_i F_{ij} = \pi_j F_{ji}$  for all  $i, j \in \hat{\mathcal{S}}$ .*

(iii)  *$G$  has row sum zero and satisfies  $\pi^T G = 0$  and  $\pi_i G_{ij} = \pi_j G_{ji}$  for all  $i, j \in \hat{\mathcal{S}}$ .*

(iv) There exists a (possibly  $\epsilon$ -dependent) constant  $0 < C < \infty$  such that  $G_{ij} \geq 0$  for all  $i \neq j$  if  $\|f\|_\infty \leq C$ . In this case equation (3.6) admits a unique and strictly positive solution  $\hat{\phi} > 0$ .

*Proof.* (i) follows from  $\sum_i \chi_i(x) = \mathbf{1}$  and reversibility of  $L$ : We have  $\sum_i \pi(i)K_{ij} = \sum_i \langle \chi_i, L\chi_j \rangle_\mu = \langle L\mathbf{1}, \chi_j \rangle_\mu = 0$  and  $\pi(i)K_{ij} = \langle \chi_i, L\chi_j \rangle_\mu = \langle L\chi_i, \chi_j \rangle_\mu = \pi(j)K_{ji}$ . (ii) follows from  $f(x)$  being real and positive for all  $x$ . As for (iii),  $G$  has row sum zero by (i) and the definition of  $\Lambda$ .  $\pi(i)G_{ij} = \pi(j)G_{ji}$  follows from (i), (ii) and the fact that  $\Lambda$  is diagonal, and  $\pi^T G = 0$  follows directly. For (iv), rewrite (3.6) as the  $n \times n$ -system  $\bar{G}_\lambda \bar{\phi} = g$  where  $\bar{G}_\lambda$  is the first  $n$  rows and columns of  $G_\lambda := -G + \epsilon^{-1}\Lambda$ ,  $\bar{\phi} = (\bar{\phi}, 1)^T$  and  $-g$  is the vector of the first  $n$  entries of the  $(n+1)$ st row of  $G_\lambda$ . Choose  $C$  such that  $\epsilon^{-1}\langle \chi_i, f\chi_j \rangle_\mu \leq \langle \chi_i, L\chi_j \rangle_\mu$  for all  $i \neq j$ . Then  $g > 0$  and  $\bar{G}_\lambda$  is a non-singular  $M$ -matrix and thus inverse monotone [2], that is from  $\bar{G}_\lambda \bar{\phi} = g$  and  $g > 0$  follows  $\bar{\phi} > 0$ .  $\square$

It follows that if the running costs  $f$  are such that (iv) in Lemma 3.4 holds, then  $G$  is a generator matrix of a MJP that we shall denote by  $(\hat{X}_t)_{t \geq 0}$ , and by lemma 2.2, (3.6) has a unique and positive solution of the form

$$\hat{\phi}(i) = \mathbf{E} \left[ \exp \left( -\epsilon^{-1} \int_0^{\tau_A} \hat{f}(\hat{X}_s) ds \right) \middle| \hat{X}_0 = i \right]$$

with  $\hat{f}(i) = \Lambda_{ii}$  and  $\tau_A = \inf\{t \geq 0 | \hat{X}_t = i+1\}$ . In fact (3.6) can be interpreted as the backward Kolmogorov equation for  $\hat{\phi}$ . Moreover, the logarithmic transformation  $\hat{W} = -\epsilon \log \hat{\phi}$  is well-defined and can be interpreted as the value function of the Markov decision problem (2.12)–(2.13), that is, we seek to minimize

$$\hat{J}(v; i) = \mathbf{E} \left[ \int_0^{\tau_A} \left( \hat{f}(\hat{X}_s^v) + k^v(\hat{X}_s^v) \right) ds \middle| \hat{X}_0^v = i \right]$$

over Markov control strategies  $v : \hat{\mathcal{S}} \rightarrow (0, \infty)$  with the costs

$$\hat{f}(i) = \Lambda_{ii}, \quad k^v(i) = \epsilon \sum_{j \neq i} G_{ij} \left\{ \frac{v(j)}{v(i)} \left[ \log \frac{v(j)}{v(i)} - 1 \right] + 1 \right\}.$$

This completes the construction of the discrete control problem. Note that in general<sup>2</sup>  $G \neq K$ , but both  $K$  and  $G$  are reversible with stationary distribution  $\pi$ .

**Elber's milestoning process.** The discretized equation admits a useful stochastic representation, by which its coefficients can be computed without knowing the committor functions. Define the *forward milestoning process*  $\tilde{X}_t^+$  to be in state  $\tilde{X}_t^+ = i$  if  $X_t$  visits core set  $C_i$  next, and the *backward milestoning process*  $\tilde{X}_t^-$  to be in state  $\tilde{X}_t^- = i$  if  $X_t$  came from  $C_i$  last. Then the discrete costs can be written as

$$\hat{f}(i) = \frac{1}{\pi_i} \langle \chi_i, f \sum_j \chi_j \rangle = \int \nu_i(x) f(x) dx = \mathbf{E}_\mu \left[ f(X_t) \middle| \tilde{X}_t^- = i \right] \quad (3.7)$$

where  $\nu_i(x) = \pi_i^{-1} \chi_i(x) \mu(x) = \mathbf{P}(X_t = x | \tilde{X}_t^- = i)$  is the probability density of finding the system in state  $x$  given that it came last from  $i$ . Hence  $\hat{f}(i)$  is the average costs

<sup>2</sup>In case of a full partition,  $\cup_i C_i = S$ , the  $\chi_i$  become stepfunctions and  $K = G$  is the generator of a full partition MSM. Our method then becomes a finite volume method. Stepfunctions are not regular enough to be in  $V$  however.

conditioned on the information  $\tilde{X}_t^- = i$ , i.e.  $X_t$  came last from  $A_i$ , which is the natural extension to the full partition case where  $\hat{f}(i)$  was the average costs conditioned on the information that  $X_t \in A_i$ .

The matrix  $K_{ij} = \pi_i^{-1} \langle \chi_i, L\chi_j \rangle$  is reversible with stationary distribution

$$\pi_i = \langle \chi_i, \mathbf{1} \rangle = \mathbf{P}_\mu(\tilde{X}_t^- = i)$$

and is related to so called *core MSMs*. To see this, define the core MSM transition matrix  $P^\tau$  with components  $P_{ij}^\tau = \mathbf{P}(\tilde{X}_{t+\tau}^+ = j | \tilde{X}_t^- = i)$ , and the mass matrix  $M$  with components  $M_{ij} = \mathbf{P}(\tilde{X}_t^+ = j | \tilde{X}_t^- = i)$ . Then, it is not hard to show that for reversible processes we have  $P_{ij}^\tau = \pi_i^{-1} \langle \chi_i, T^\tau \chi_j \rangle_\mu$  and  $M_{ij} = \pi_i^{-1} \langle \chi_i, \chi_j \rangle_\mu$  so that

$$K = \frac{1}{\pi_i} \langle \chi_i, L\chi_j \rangle_\mu = \lim_{\tau \rightarrow 0} \frac{1}{\tau} (P^\tau - M).$$

Thus  $K$  is formally<sup>3</sup> the generator of the  $P^\tau$ . If the core sets are chosen as the metastable states of the system,  $K$  can be sampled directly from  $\tilde{X}_t^\pm$ . See [35, 36] for more details on the construction and sampling of core MSMs.  $F$  can also be sampled using

$$F_{ij} = \mathbf{E}_\mu \left[ f(X_t) \chi_{\{\tilde{X}_t^+ = j\}} \middle| \tilde{X}_t^- = i \right] \quad (3.8)$$

Therefore, as in the construction of core MSMs, we do not need to compute committor functions explicitly.

**4. Numerical Results.** We will present two examples to illustrate the approximation of LQ-type stochastic control problems based on a sparse Galerkin approximation using MSMs.

**4.1. 1D triple well potential.** To begin with we study diffusion in the triple well potential which is presented in Figure 4.2a. This potential has three minima at  $x_{0/1} = \mp 3.4$  and  $x_2 = 0$ . We choose  $A = [x_0 - \delta, x_0 + \delta]$  with  $\delta = 0.2$  as the target set and the running cost  $f = \sigma = \text{const}$ , such that the control goal is to steer the particle into  $C_0$  in minimum time. In Figure 4.2a the potential  $V$  and effective potential  $U$  are shown for  $\epsilon = 0.5$  and  $\sigma = 0.08$  (solid lines), cf. equation (2.4). One can observe that the optimal control lifts the second and third well up such that the system is driven into  $C_0$  quickly.

First we validate our method with a convergence test using linear finite elements as basis functions  $\chi_i$ . To do so, we compute a reference solution  $\hat{\phi}$  of (3.4) using linear finite elements on a uniform grid with spacing  $h_r = 10^{-4}$ . The resulting interpolation  $\phi_I = \sum_i \chi_i \hat{\phi}(i)$  is very close to the true solution  $\phi$  of (3.1). We also compute a reference solution for the value function  $W_I = \sum_i \chi_i \hat{W}(i)$  with  $\hat{W} = -\epsilon \log \hat{\phi}$  and for the optimal control  $u_I = -2\nabla W_I$ . Then we compute coarser solutions  $\phi_{I,h}$  using various grid spacings  $1 \geq h \geq 10^{-3}$  and compute the  $L^2$  error  $\|\phi_{I,h} - \phi_I\|_\mu$ , and  $L^2$  errors for  $W_I$  and  $u_I$  similarly. The result is shown in Figure 4.1. The  $L^2$  error of

<sup>3</sup>The  $P^\tau$  do not form a semigroup since  $M \neq \mathbf{1}$ , thus  $K$  cannot be interpreted as i.e. the generator of  $\tilde{X}_t^-$ . However, the entries of  $K$  are the transition rates between the core sets as defined in transition path theory [45].

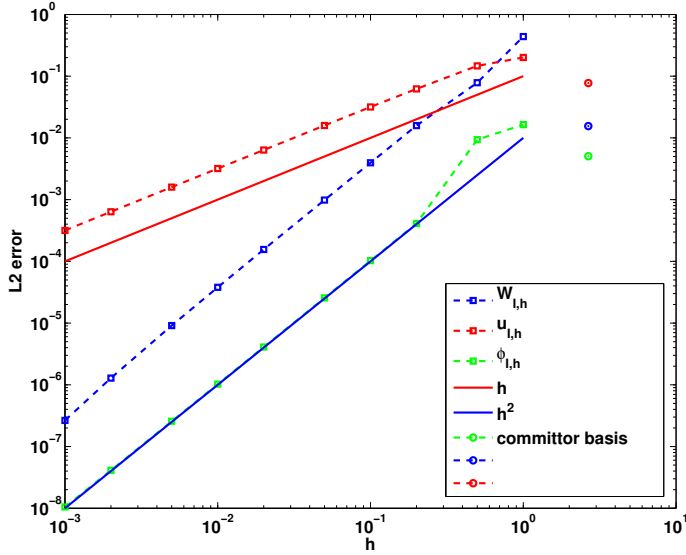


Fig. 4.1:  $L^2$  error of  $\phi_{I,h}$ ,  $W_{I,h}$  and  $u_{I,h}$  using linear finite elements (dashed lines) and using the committor basis (circles).

$\phi_{I,h}$  is quadratic in  $h$ , as expected from the theory. Additionally, the  $L^2$  error of  $W_{I,h}$  is also quadratic in  $h$  which, given that the transformation between  $\phi$  and  $W$  is nonlinear, is surprising. The error of  $u_{I,h}$  is only linear in  $h$ ; as expected one order of convergence is lost due to the fact that  $u_I$  is the gradient of  $W_I$ .

Next, we use a committor basis. In accordance with the strategy to minimize  $\epsilon_0$  in Theorem 3.3, we placed core sets  $C_i = [x_i - \delta, x_i + \delta]$  in each of the three wells of the potential shown in Figure 4.2a, resulting in the set of three basis functions shown in Figure 4.2b. The  $L^2$  errors achieved by solving (3.4) in this basis are shown as circles in Figure 4.1. We observe that the 3 committor functions achieve the same performance as linear finite elements with grid spacing  $h \approx 0.2$ , which corresponds to  $\approx 40$  basis functions. Theorem 3.3 gives  $\epsilon_0 \leq 0.08$ , while the actual error is one order of magnitude smaller. The dashed line in Figure 4.2a gives the approximation to  $U$  calculated in the committor basis, which is in good agreement to the reference solution. In Figure 4.2c the optimal control  $u$  (solid line) and its approximation  $u_I$  (dashed line) are shown. The core sets are shown in blue. The jumps in  $u_I$  at the left boundaries of the core sets are due to the fact that the committor functions are only piecewise differentiable.

The computations so far require explicit knowledge of the basis functions  $\chi_i$  to compute the matrices  $K$  and  $F$ . For high-dimensional systems the committor basis is usually not explicitly known. To mimic this situation, we construct a core MSM to sample the matrices  $K$  and  $F$ . 100 trajectories of length  $T = 20,000$  were used to build the MSM. In Figure 4.2d, the optimal cost starting from the rightmost well  $W(x_1)$  and its estimate using the core MSM are shown for  $\epsilon = 0.5$  and different values of  $\sigma$ . Each of the 100 trajectories has seen about four transitions. For comparison, a direct sampling estimate of  $W(x_1)$  using the same data is shown (green). The



direct sampling estimate suffers from a large bias and variance. In contrast, the MSM estimator for  $W(x_1)$  performs well for all considered values of  $\sigma$ . The constant  $C$  which ensures  $\hat{\phi} > 0$  when  $\sigma \leq C$  is approximately 0.2 in this case. This seems restrictive but still allows to capture all interesting information about  $\phi$  and  $W(x_1)$ .

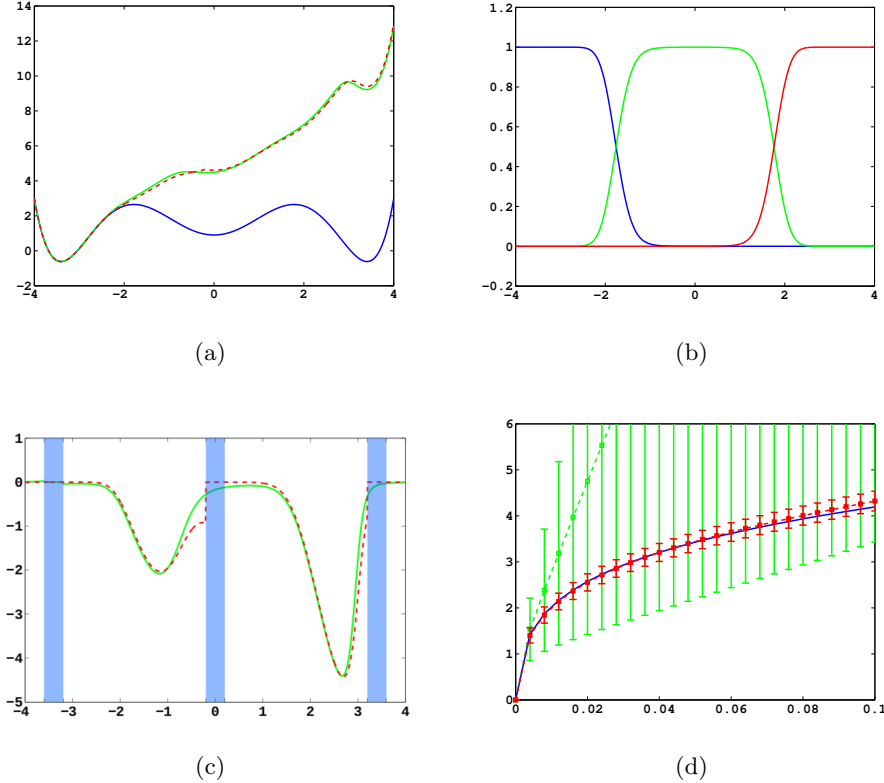


Fig. 4.2: Three well potential example for  $\epsilon = 0.5$  and  $\sigma = 0.08$ . (a) Potential  $V(x)$  (blue), effective potential  $U = V + 2W$  (green) and approximation of  $U$  with committors (dashed red). (b) The three committors. (c) The optimal control  $u(x)$  (solid line) and its approximation (dashed line). Core sets are shown in blue. (d) Optimal cost  $W(x_1)$  for  $\epsilon = 0.5$  as a function of  $\sigma$ . Blue: Exact solution. Red: Core MSM estimate. Green: Direct sampling estimate.

**4.2. Alanine dipeptide.** As a second, non-trivial example we study conformational transitions in Alanine dipeptide (ADP), a well-studied test system in molecular dynamics. We performed an all-atom simulation of ADP in explicit water (TIP3P) with the Amber FF99SB force field [23] using the GROMACS 4.5.5 simulation package [44]. The simulations were performed in the NVT ensemble, where the temperature was restrained to 300 K using the V-Rescale thermostat [7]. 20 trajectories of 200ns with 100ps equilibration runs were simulated. Covalent bonds to hydrogen atoms were constrained using the LINCS algorithm11 [21] (lincs iter = 1, lincs order = 4), allowing for an integration timestep of 2 fs. The leap-frog integrator was used. Lennard-Jones interactions were cut off at 1 nm. Electrostatic interactions were treated by the Particle-Mesh Ewald (PME) algorithm12 [11] with a real space cut-off

of 1 nm, a grid spacing of 0.15 nm, and an interpolation order of 4. Periodic boundary conditions were applied in the x, y, and z-direction.

In Figure 4.3a, a cartoon of the molecule is shown. The full system including the water molecules has about 4000 degrees of freedom. However, it is well known that the conformational dynamics, which are the slowest dynamical processes in the system, can be monitored via the backbone dihedral angles  $\phi$  and  $\psi$ . The dynamics along the other degrees of freedom happens on much faster timescales. For this reason, the value function will essentially be a function of  $\phi$  and  $\psi$ ; see [19] for precise statements. We will use this to build a Markov State Model which partitions the  $\phi - \psi$ -plane.

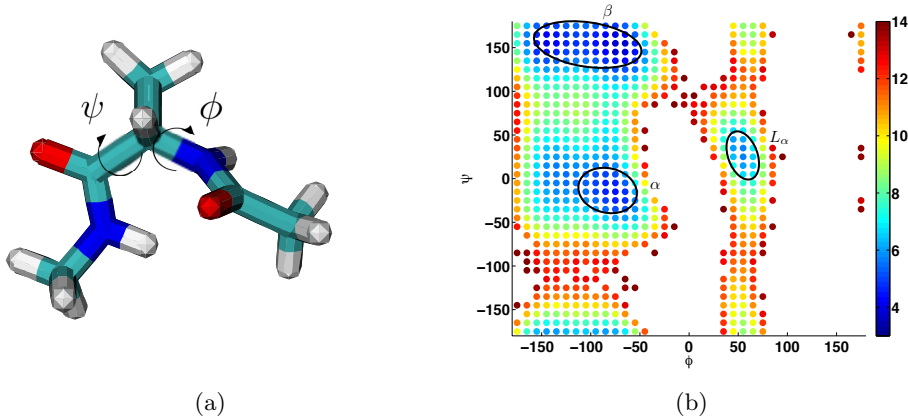


Fig. 4.3: (a) Alanine dipeptide. (b) Free energy  $g_i = -\log \pi_i$ .

**Validation of the MSM approximation.** We construct a full partition MSM using a uniform clustering into  $36 \times 36$  boxes  $A_i$  of size  $10^\circ \times 10^\circ$  in the  $\phi - \psi$ -plane, and we use characteristic basis functions  $\chi_i(x) = \mathbf{1}_{A_i}$  for the discretization<sup>4</sup>. Figure 4.3b shows the free energy  $g_i = -\log \pi_i = -\log \mathbf{P}[X_t \in A_i]$  together with the three largest molecular conformations  $\alpha$ ,  $\beta$  and  $L_\alpha$ . The missing boxes have not seen any data. The slowest dynamical process is the switching between the left-handed  $L_\alpha$  structure and the right-handed  $\alpha$  and  $\beta$  sheet structures. As is customary in MSM theory [35], we estimate the slowest *implied timescale* (ITS) as follows: For different lagtimes  $\tau$ , we construct the MSM transfer operator  $P_{ij}^\tau = \pi_i^{-1} \langle \chi_i, T^\tau \chi_j \rangle_\mu$  and compute

$$t_1(\tau) = -\frac{\tau}{\log \lambda_1(\tau)}$$

where  $\lambda_1(\tau)$  is the 2nd largest eigenvalue of  $P^\tau$ . The result is shown in Figure 4.4a. We observe a plateau for  $6\text{ps} \leq \tau \leq 30\text{ps}$  which indicates that the time-discrete snapshots  $\hat{X}_n^\tau := \hat{X}_{n\tau}$  are well described by a Markov chain with transition matrix  $P^\tau$  for  $\tau \geq 6\text{ps}$ . The plateau is used to compute  $t_1 = 1560\text{ps} \pm 6\text{ps}$ . For smaller

<sup>4</sup>Let  $C_i \subset A_i$  be a core set in cell  $A_i$ . The characteristic basis functions can be obtained from the committors by expanding  $C_i$  to fill out all of  $A_i$ . This basis is used to construct full partition MSMs and produces particularly simple sampling formulas. The characteristic functions are not in  $V$ , but the sampling results are similar to what one would obtain using committors with cores  $C_i$  which fill out most of  $A_i$ .

values of  $\tau$ ,  $\hat{X}_n^\tau$  is not Markovian due to recrossing effects. For this reason we also cannot sample  $K$  directly and have to work with the finite-time transfer operator  $P^\tau$  instead. Before proceeding to the optimal control problem, we study the effect of this time discretization on the mean first passage time (MFPT)  $t(x) = \mathbf{E}_x[\tau_{\alpha\cup\beta}]$  where  $\tau_{\alpha\cup\beta}$  is the first hitting time of  $\alpha \cup \beta$ . Since the  $\alpha$ ,  $\beta$  and  $L_\alpha$  conformations are very metastable,  $t(x)$  is almost constant on  $L_\alpha$  and  $t_{L_\alpha} = \mathbf{E}[t(x)|x \in L_\alpha]$  can be computed from  $t_1$  via  $t_{L_\alpha} = t_1/\pi_{\alpha\cup\beta}$  where  $\pi_{\alpha\cup\beta} = 0.96$  is the invariant measure of the  $\alpha$  and  $\beta$  conformation combined [36]. This gives  $t_{L_\alpha} = 1626\text{ps} \pm 6\text{ps}$ .

On the other hand, let  $N_{\alpha\cup\beta}^\tau = \inf\{n > 0 : \hat{X}_n^\tau \in \alpha \cup \beta\}$ . If the chain  $(\hat{X}_n^\tau)_n$  is Markovian, then the time-discrete MFPT  $\hat{t}^\tau(x) = \mathbf{E}_x[\tau N_{\alpha\cup\beta}^\tau]$  satisfies the matrix equation

$$(P^\tau - I)\hat{t}^\tau = -\tau \text{ outside } \alpha \cup \beta, \quad \hat{t}^\tau = 0 \text{ in } \alpha \cup \beta. \quad (4.1)$$

Additionally, since  $\tau_{\alpha\cup\beta} \in (\tau(N_{\alpha\cup\beta}^\tau - 1), \tau N_{\alpha\cup\beta}^\tau]$ , we should expect that

$$\hat{t}_{L_\alpha}^\tau := \mathbf{E}[\hat{t}^\tau(x)|x \in L_\alpha] = t_{L_\alpha} + c\tau, \quad c \leq 1. \quad (4.2)$$

In Figure 4.4b,  $\hat{t}_{L_\alpha}^\tau$  obtained by solving (4.1) is shown as a function of  $\tau$  together with  $t_{L_\alpha}$ . A linear interpolation using the values for  $6\text{ps} \leq \tau \leq 30\text{ps}$  where the Markov assumption holds gives  $\hat{t}^\tau = \hat{t}^0 + 0.8\tau$  with  $\hat{t}^0 = 1628\text{ps}$ , which is consistent with (4.2). This shows that the time discretization introduces only small, controllable errors for  $\tau \geq 6\text{ps}$ . In the following, we will work with  $\tau = 10\text{ps}$ . Notice that the time discretization amounts to setting  $K = \tau^{-1}(P^\tau - I)$  in (3.4).

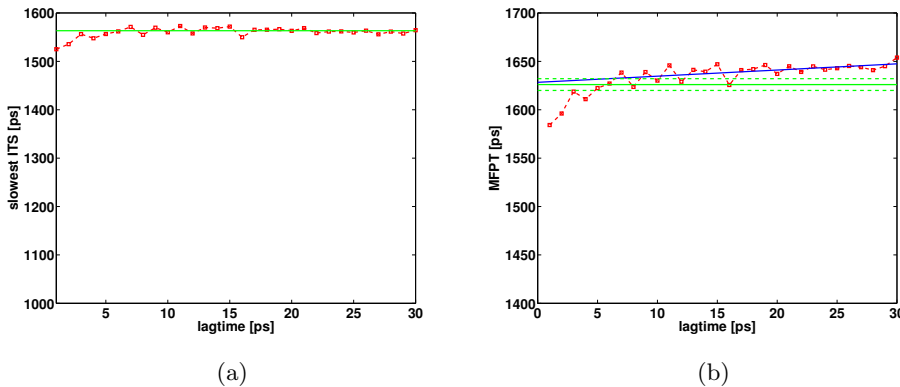


Fig. 4.4: (a) Slowest implied timescale  $t_1(\tau)$  (red) and average using the values for  $6\text{ps} \leq \tau \leq 30\text{ps}$  (green). (b) Time-discrete MFPT  $\hat{t}_{L_\alpha}^\tau$  as a function of  $\tau$  (red), linear interpolation of  $\hat{t}_{L_\alpha}^\tau$  (blue) and the reference value  $t_{L_\alpha}$  (green). The confidence interval of  $t_{L_\alpha}$  is shown as dashed lines.

**Controlled transition to the  $\alpha$ -helical structure.** Next we consider an optimal control problem for steering the molecule into the  $\alpha$ -structure. We choose as the target region  $A = \alpha$  and define running costs in the  $(\phi, \psi)$  variables as  $f(\phi, \psi) = f_0 + f_1\|\psi - \psi_\alpha\|^2$  where  $\|\cdot\|$  is a metric on the torus, and we choose  $f_0 = 0.01$  and  $f_1 = 0.001$  representing a mild penalty for being away from the target region in the  $\psi$ -direction. We discretize this control problem using the same partition

and time discretization as for the MSM construction in section 4.2 and sample  $P^\tau$  and  $F$  from the MSM data. The resulting value function  $\hat{W} = \log \hat{\phi}$  is shown in Figure 4.5a. Since the basis functions  $\chi_i$  are not differentiable and some data is missing in  $\hat{W}$ , we have to construct an interpolation  $W_I(\phi, \psi)$  from the point data  $\hat{W}$  to obtain an estimate for the optimal control force  $u(\phi, \psi) = -\sigma \nabla W_I(\phi, \psi)$ . An interpolation based on a Delauney triangulation which is  $C^1$  everywhere except at the data points is shown in Figure 4.5b.

To demonstrate that adding the control force  $u(\phi, \psi)$  has the effect of speeding up the transition from  $L_\alpha$  to  $\alpha$ , we would have to implement it in the MD simulation software. We leave that for future work. We can make a prediction of the anticipated effect within the MSM framework: In accordance with (2.11), we compute the transition matrix  $P_{v^*}^\tau$  of the optimally controlled process by  $P_{v^*}^\tau(i, j) = P^\tau(i, j) \frac{v^*(j)}{v^*(i)}$  with  $v^* = \hat{\phi}$  for  $i \neq j$ . The discretized MFPT vector  $\hat{t}^*$  of the optimally controlled process can be computed from the Matrix equation

$$(P_{v^*}^\tau - I) \hat{t}^* = -\tau \text{ outside } \alpha, \quad \hat{t}^* = 0 \text{ in } \alpha.$$

The result is shown in Figure 4.5c and gives a speed up compared to  $\hat{t}_0$  of one order of magnitude. A larger speed up could easily be achieved by increasing  $f$ . In Figure 4.5d we show the free energy of the controlled process in log scale, which according to Lemma 2.3 is given by  $g^{v^*} = -\log \pi^{v^*} = \log Z_{v^*} - 2\hat{W} - \log \pi$ . Observe that the  $L_\alpha$  and  $\beta$  conformations are now much less populated compared to the equilibrium distribution in Figure 4.3b: As in the 1D example, the control mainly has the effect of lifting the wells which are not in the target region up such that they become less metastable.

**5. Conclusions.** We have developed a Galerkin projection method that leads to an approximation of certain optimal control problems for reversible diffusions by Markov decision problems. The approach is based on the dual formulation of the optimal control problem in terms of a linear boundary value problem that can be discretized in a straightforward way. In this article we propose a discretization that preserves reversibility and the generator form of the linear equations, i.e., the discretization of the infinitesimal generator of the original diffusion process can be interpreted as the infinitesimal generator of a reversible Markov jump process (MJP). The discretized linear boundary value problem admits again a dual formulation in terms of a Markov decision problem. A sparse approximation that uses the basis of committor functions of metastable sets of the dynamics was discussed in detail: The discretization using committor functions does not require that the metastable sets partition the state space, hence the method can be applied to high-dimensional problems as they appear, e.g., in molecular dynamics. The committor functions in this case need not be known explicitly, as it is possible to sample the generator matrices and the discrete cost functions by a Monte-Carlo method, similarly to what is done in the Markov state modelling approach to protein folding. We could prove an  $L^2$  error estimate for the Galerkin scheme, moreover the discretization was shown to preserve basic structural elements of the continuous problem, such as duality, reversibility or properties of the invariant measure. Our numerical results showed very good performance of the incomplete partition discretization on a simple toy example and a high-dimensional molecular dynamics problem, even with only a few basis functions, which is in line with the theoretical error bounds presented in this paper.

While we addressed the discretization error in this paper in great detail, we did not address the sampling error. In particular, for large systems our construction requires

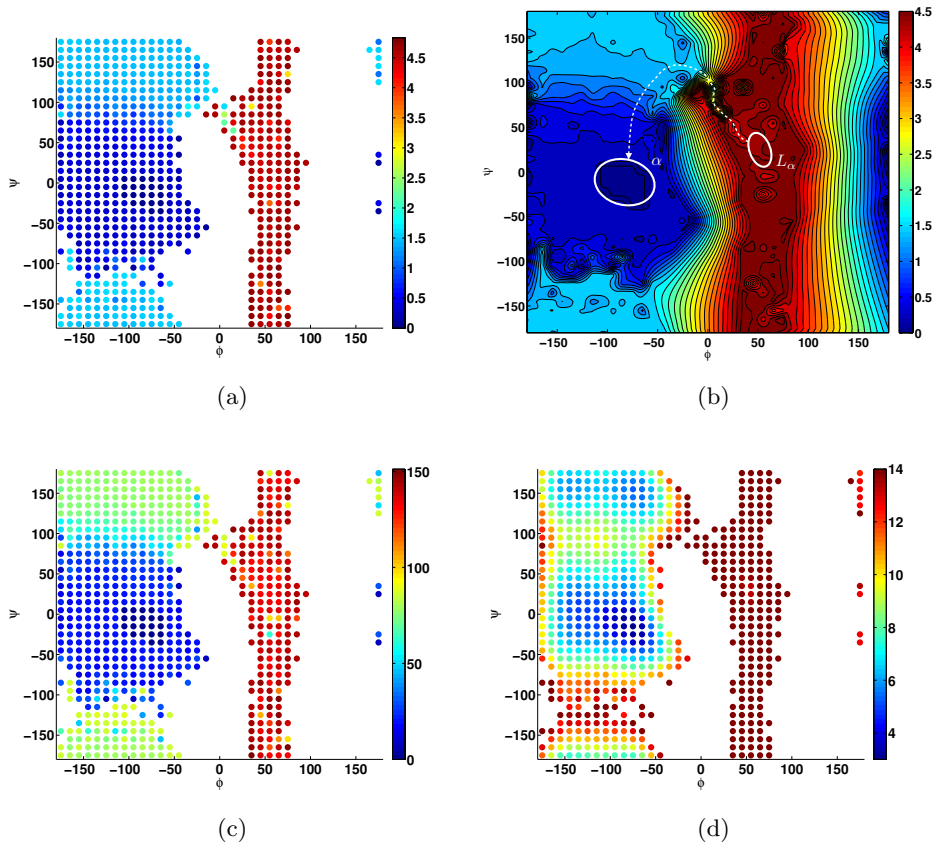


Fig. 4.5: (a) Optimal control  $\hat{W}$  for steering into the  $\alpha$ -structure. (b) Interpolation  $W_I(\phi, \psi)$  obtained from  $\hat{W}$  via a Delaunay triangulation, and a steepest descent path from  $L_\alpha$  to  $\alpha$ . (c) MFPT to the  $\alpha$  conformation for the optimally controlled process. (d) Free energy  $g^{v^*} = -\log \pi^{v^*}$  of the optimally controlled process.

the coefficients of the MJP and therefore the transition rates between all metastable states as an input. This is not fully satisfactory. We believe that the optimal control framework presented here should be linked with Monte-Carlo methods for rare events, e.g., [20, 13], that exploit the same duality between optimal control and sampling to devise efficient importance sampling strategies as we did so as to reduce the sampling error. We leave the analysis of the sampling error to future work.

**Acknowledgements.** The authors thank Marco Sarich and Christof Schütte for helpful discussions and Antonia Mey and Francesca Vitalini for providing the Alanine dipeptide data. The research was funded by the DFG Research Centre Matheon.

### Appendix A. Proof of Theorem 3.1.

Here we give the proof of Theorem 3.1 from Section 3.1. For ease of notation, let  $\|\cdot\| = \|\cdot\|_\mu$ .

*Proof.* Let  $\phi$  be the solution to (3.2), and write  $\phi = Q\phi + \phi_\perp = \phi_\parallel + \phi_\perp$  with  $\phi_\perp \in D^\perp$ . The first step is to show that  $\|\phi - \phi_\parallel\| = \inf_{\psi \in D} \|\phi - \psi\|$ , i.e. the infimum

in the definition of  $\varepsilon_0$  is attained at  $\phi_{||}$ . But this is clear since for any  $\psi \in D$ , by orthogonality we have

$$\|\phi - \psi\|^2 = \|\phi_{||} - \psi + \phi_{\perp}\|^2 = \|\phi_{||} - \psi\|^2 + \|\phi_{\perp}\|^2$$

which attains its minimum of  $\varepsilon_0^2 = \|\phi_{\perp}\|^2$  for  $\psi = \phi_{||}$ . By (3.2),  $\phi_{||}$  solves the equation

$$\mathcal{B}(\phi, \psi) = \mathcal{B}(\phi_{||}, \psi) + \mathcal{B}(\phi_{\perp}, \psi) = 0 \quad \forall \psi \in D,$$

and if we write  $\phi_{||} = \sum_{i=1}^n \hat{\phi}_i^* \chi_i + 1\chi_{n+1}$  with  $n$  unknown coefficients  $\hat{\phi}_i^*$  (note that a general element of  $D$  is of this form), this takes the matrix form

$$\hat{B}\hat{\phi}^* - c = F,$$

where in components we have  $\hat{B}_{ij} = \mathcal{B}(\chi_i, \chi_j)$ ,  $c_i = -\mathcal{B}(\phi_{\perp}, \chi_i) = -\langle \phi_{\perp}, B\chi_i \rangle_{\mu}$  and  $F_i = -\langle \chi_i, B\chi_{n+1} \rangle_{\mu}$ . On the other hand, the Galerkin solution  $\hat{\phi} = \sum_i \hat{\phi}_i \chi_i$  satisfies  $\hat{B}\hat{\phi} = F$  by 3.3, hence we obtain

$$\hat{B}(\hat{\phi}^* - \hat{\phi}) = c. \tag{A.1}$$

Now we can write

$$\begin{aligned} \varepsilon^2 &= \|\phi_{||} + \phi_{\perp} - \hat{\phi}\|^2 = \|\phi_{||} - \hat{\phi}\|^2 + \|\phi_{\perp}\|^2 \\ &= \left\langle \sum_i (\hat{\phi}_i^* - \hat{\phi}_i) \chi_i, \sum_j (\hat{\phi}_j^* - \hat{\phi}_j) \chi_j \right\rangle_{\mu} + \varepsilon_0^2 \\ &= (\hat{\phi}^* - \hat{\phi})^T \hat{M} (\hat{\phi}^* - \hat{\phi}) + \varepsilon_0^2 \end{aligned}$$

where  $\hat{M}_{ij} = \langle \chi_i, \chi_j \rangle_{\mu}$ . The scalar product  $\langle \cdot, \cdot \rangle_{\mu}$  on  $D_0 \subset V$  induces a natural scalar product on  $\mathbb{R}^n$  by the isomorphism  $\hat{\phi} \mapsto \sum_i \hat{\phi}_i \chi_i$ :

$$\left\langle \sum_i \hat{\phi}_i \chi_i, \sum_j \hat{\phi}'_j \chi_j \right\rangle_{\mu} = \hat{\phi}^T \hat{M} \hat{\phi}' =: \langle \hat{\phi}, \hat{\phi}' \rangle_M$$

The error  $\varepsilon^2$  is exactly  $\varepsilon_0^2$  plus the distance between Galerkin solution and best approximation measured in this scalar product. There is also a natural bilinear form inherited from  $\mathcal{B}$  on  $\mathbb{R}^n$ :

$$\mathcal{B} \left( \sum_i \hat{\phi}_i \chi_i, \sum_j \hat{\phi}'_j \chi_j \right) = \hat{\phi}^T \hat{B} \hat{\phi}' = \langle \hat{\phi}, \hat{M}^{-1} \hat{B} \hat{\phi}' \rangle_M$$

The Matrix  $\hat{M}^{-1} \hat{B}$  is symmetric since  $\mathcal{B}(\cdot, \cdot)$  is symmetric. Moreover, since  $\mathcal{B}(\cdot, \cdot)$  is elliptic,

$$\langle \hat{\phi}, \hat{M}^{-1} \hat{B} \hat{\phi} \rangle_M = A \left( \sum_i \hat{\phi}_i \chi_i, \sum_j \hat{\phi}_j \chi_j \right) \geq \alpha_2 \left\langle \sum_i \hat{\phi}_i \chi_i, \sum_j \hat{\phi}_j \chi_j \right\rangle_\mu = \alpha_2 \langle \hat{\phi}, \hat{\phi} \rangle_M \quad (\text{A.2})$$

In particular,  $\hat{M}^{-1} \hat{B}$  is positive, hence it has a positive and symmetric square root  $\hat{S}^2 = \hat{M}^{-1} \hat{B}$ . Now, for any  $\hat{\phi} \in \mathbb{R}^n$  it holds by virtue of (A.2),

$$\begin{aligned} \langle \hat{\phi}, \hat{\phi} \rangle_M &\leq \frac{1}{\alpha_2} \langle \hat{\phi}, \hat{M}^{-1} \hat{B} \hat{\phi} \rangle_M = \frac{1}{\alpha_2} \langle \hat{S} \hat{\phi}, \hat{S} \hat{\phi} \rangle_M \\ &\leq \frac{1}{\alpha_2^2} \langle \hat{S} \hat{\phi}, \hat{M}^{-1} \hat{B} \hat{S} \hat{\phi} \rangle_M = \frac{1}{\alpha_2^2} \langle \hat{M}^{-1} \hat{B} \hat{\phi}, \hat{M}^{-1} \hat{B} \hat{\phi} \rangle_M. \end{aligned} \quad (\text{A.3})$$

Now we apply the inequality (A.3) to  $\hat{\phi}^* - \hat{\phi}$  and use (A.1):

$$\varepsilon^2 \leq \varepsilon_0^2 + \frac{1}{\alpha_2^2} \langle \hat{M}^{-1} c, \hat{M}^{-1} c \rangle_M. \quad (\text{A.4})$$

Now for some final simplifications, note that the orthogonal projection  $Q$  onto  $D_0$  can be written as

$$Q\psi = \sum_{i,j=1}^n \hat{M}_{ij}^{-1} \langle \chi_j, \psi \rangle_\mu \chi_i.$$

Using this we can write

$$\begin{aligned} \langle \hat{M}^{-1} c, \hat{M}^{-1} c \rangle_M &= \sum_{ij} c_i \hat{M}_{ij}^{-1} c_j = \sum_{ij} \langle \chi_i, B\phi_\perp \rangle_\mu \hat{M}_{ij}^{-1} \langle \chi_j, B\phi_\perp \rangle_\mu \\ &= \left\langle \sum_{ij} \hat{M}_{ij}^{-1} \langle \chi_j, B\phi_\perp \rangle_\mu \chi_i, B\phi_\perp \right\rangle_\mu = \langle QB\phi_\perp, B\phi_\perp \rangle_\mu \\ &= \langle QB\phi_\perp, QB\phi_\perp \rangle_\mu \end{aligned}$$

To arrive at the final result, notice that

$$\begin{aligned} \langle QB\phi_\perp, QB\phi_\perp \rangle_\mu &\leq \left( \sup_{\phi'_\perp \in D^\perp} \frac{\langle QB\phi'_\perp, QB\phi'_\perp \rangle_\mu}{\langle \phi'_\perp, \phi'_\perp \rangle_\mu} \right) \cdot \langle \phi_\perp, \phi_\perp \rangle_\mu \\ &= \left( \sup_{\phi'_\perp \in D^\perp} \frac{\langle QBQ^\perp \phi'_\perp, QBQ^\perp \phi'_\perp \rangle_\mu}{\langle \phi'_\perp, \phi'_\perp \rangle_\mu} \right) \cdot \langle \phi_\perp, \phi_\perp \rangle_\mu \\ &\leq \left( \sup_{\phi' \in V} \frac{\langle QBQ^\perp \phi', QBQ^\perp \phi' \rangle_\mu}{\langle \phi', \phi' \rangle_\mu} \right) \cdot \langle \phi_\perp, \phi_\perp \rangle_\mu \\ &= \|QBQ^\perp\|^2 \langle \phi_\perp, \phi_\perp \rangle_\mu \end{aligned}$$

Plugging these inequalities into (A.4) and dividing by  $\varepsilon_0^2$  completes the proof.  $\square$

## Appendix B. Best-approximation error bound.

In this appendix, we prove lemma 3.3:

$$\varepsilon_0 = \|Q^\perp \phi\|_\mu \leq \|P^\perp \phi\|_\mu + \mu(T)^{1/2} [\kappa \|f\|_\infty + 2\|P^\perp \phi\|_\infty].$$

Recall that  $\kappa = \sup_{x \in T} \mathbf{E}_x[\tau_{S \setminus T}]$  and  $P$  is the orthogonal projection onto the subspace  $V_c = \{v \in L^2(\mathcal{S}, \mu), v = \text{const on every } C_i\} \subset L^2(\mathcal{S}, \mu)$ . Note that  $P^\perp \phi = 0$  on  $C$ . The errors  $\|P^\perp \phi\|$  and  $\|P^\perp \phi\|_\infty$  measure how constant the solution  $\phi$  is on the core sets. We write  $\|\cdot\| = \|\cdot\|_\mu$  throughout the proof for convenience.

*Proof.* The proof closely follows the proof of theorem (12) in [34]. The first step of the proof is to realize that the committor subspace  $D$  where  $Q$  projects onto can be written as  $D = \{v \in L^2(\mathcal{S}, \mu), v = \text{const on every } C_i, Lv = 0 \text{ on } C\}$ . To see this, note that the values  $v$  takes on the  $C_i$  can be used as boundary values for the Dirichlet problem  $Lv = 0$  on  $T$ . A linear combination of committor functions is obviously a solution to this problem. But the solution to the Dirichlet problem must be unique, otherwise one can construct a contradiction to the uniqueness of the invariant distribution, see [34].

By definition we have  $\|Q^\perp \phi\| \leq \|\phi - I\phi\|$  for every interpolation  $I\phi \in D$  of  $\phi$ . With the definition of  $P$  from above, we will take  $q = I\phi$  such that

$$Lq = 0 \text{ on } T, \quad q = P\phi \text{ on } \mathcal{S} \setminus T. \quad (\text{B.1})$$

Now  $D \subset V$ , therefore  $q \in V_c$  and  $Pq = q$ . Therefore (B.1) is equivalent to

$$PLPq = 0 \text{ on } T, \quad q = P\phi \text{ on } \mathcal{S} \setminus T. \quad (\text{B.2})$$

Now define  $e := P\phi - q$ . Then we have

$$PLPe = PLP(P\phi - q) = PLP\phi - PLPq = PL\phi - PLP^\perp \phi - PLPq$$

and by (B.2) and since  $L\phi = f\phi$  on  $\mathcal{S} \setminus A \supset T$ , we have

$$PLPe = Pf\phi - PLP^\perp \phi \text{ on } T, \quad e = 0 \text{ on } \mathcal{S} \setminus T. \quad (\text{B.3})$$

Therefore,  $e \in E_\Theta = \{v \in L^2(\mathcal{S}, \mu), v = 0 \text{ on } \mathcal{S} \setminus T\}$  and with  $\Theta$  being the orthogonal projection onto  $E_\Theta$ ,  $e$  has to fulfil

$$\Theta PLP\Theta e = \Theta Pf\phi - \Theta PLP^\perp \phi.$$

Since  $\Theta P = P\Theta = \Theta$ , this can be written as

$$Re := \Theta L\Theta e = \Theta f\phi - \Theta LP^\perp \phi.$$

The operator  $R = \Theta L\Theta$  is invertible on  $E_\Theta$ : If this wasn't the case, there would be a nontrivial solution  $v$  to



$$Lv = 0 \text{ on } T, \quad v = 0 \text{ on } \mathcal{S} \setminus T.$$

But the solution to this boundary value problem is again unique, and hence there is only the trivial solution. This gives

$$e = R^{-1}\Theta f\phi - R^{-1}\Theta LP^\perp\phi, \quad (\text{B.4})$$

and  $\|R^{-1}\| = \frac{1}{|\lambda_0|}$  where  $\lambda_0$  is the principal eigenvalue of  $R$ . Due to an estimate by Varadhan we have

$$\frac{1}{|\lambda_0|} \leq \sup_{x \in T} \mathbf{E}_x[\tau_{\mathcal{S} \setminus T}] =: \kappa,$$

see e.g. [5]. To complete the derivation we need to focus on the second term in (B.4). Since  $R^{-1}$  is an operator on  $E_\Theta$ , we can write it as  $R^{-1}\Theta LP^\perp\phi =: \Theta g$ , where the function  $\Theta g$  solves

$$\Theta L\Theta g = R\Theta g = \Theta LP^\perp\phi \Leftrightarrow \Theta L[\Theta g - P^\perp\phi] = 0$$

by the definition of  $R$  and  $\Theta g$ . Therefore  $w := \Theta g - P^\perp\phi$  solves the boundary value problem

$$Lw = 0 \text{ on } T, \quad w = -P^\perp\phi \text{ on } \mathcal{S} \setminus T \quad (\text{B.5})$$

which implies that  $\|w\|_\infty \leq \|P^\perp\phi\|_\infty$ , this follows from Dynkin's formula or Lemma 3 in [34]. Finally,

$$\|\Theta g\| \leq \mu(T)^{1/2} \|\Theta g\|_\infty \leq \mu(T)^{1/2} (\|P^\perp\phi\|_\infty + \|w\|_\infty) \leq 2\mu(T)^{1/2} \|P^\perp\phi\|_\infty$$

holds by the triangle inequality and the above considerations. Now focus on the first term in (B.4). Note that by the maximum principle,  $\phi$  achieves its maximum of 1 on the boundary of  $\mathcal{S} \setminus A \supset T$ , therefore  $\max_{x \in T} |\phi(x)| \leq 1$ . Then we have

$$\|\Theta f\phi\| \leq \mu(T)^{1/2} \|f\|_\infty \max_{x \in T} |\phi(x)| \leq \mu(T)^{1/2} \|f\|_\infty.$$

Now putting everything together, we arrive at

$$\begin{aligned} \|e\| &\leq \|R^{-1}\| \|\Theta f\phi\| + \|R^{-1}\Theta LP^\perp\phi\| \\ &\leq \kappa \|\Theta f\phi\| + \|\Theta g\| \\ &\leq \mu(T)^{1/2} [\kappa \|f\|_\infty + 2\|P^\perp\phi\|_\infty]. \end{aligned}$$

Finally, note that by the triangle inequality

$$\|Q^\perp\phi\| \leq \|\phi - q\| \leq \|\phi - P\phi\| + \|P\phi - q\| = \|P^\perp\phi\| + \|e\|$$

which completes the proof.  $\square$

## REFERENCES

- [1] T. Belytschko, Y. Krongauz, D. Organ, M. Fleming, and P. Krysl, *Meshless methods: An overview and recent developments*, Comput. Methods Appl. M. **139** (1996), 3 – 47.
- [2] A. Berman and R.J. Plemmons, *Nonnegative matrices in the mathematical sciences*, Academic Press, 1979.
- [3] O. Bokanowski, J. Garcke, M. Griebel, and I. Klompaker, *An adaptive sparse grid semi-lagrangian scheme for first order Hamilton–Jacobi–Bellman equations*, J. Sci. Comput. **55** (2013), 575–605.
- [4] M. Boulbrachene and M. Haiour, *The finite element approximation of Hamilton–Jacobi–Bellman equations*, Comput. Math. Appl. **41** (2001), 993 – 1007.
- [5] A. Bovier, *Methods of contemporary statistical mechanics*, Metastability, Springer, 2009.
- [6] D. Braess, *Finite elements: Theory, fast solvers, and applications in solid mechanics*, Cambridge University Press, 2007.
- [7] G. Bussi, D. Donadio, and M. Parrinello, *Canonical sampling through velocity rescaling*, J. Chem. Phys. **126** (2007), 014101.
- [8] E. Carlini, M. Falcone, and R. Ferretti, *An efficient algorithm for Hamilton–Jacobi equations in high dimension*, Comput. Visual. Sci. **7** (2004), 15–29.
- [9] T. Cecil, J. Qian, and S. Osher, *Numerical methods for high dimensional Hamilton–Jacobi equations using radial basis functions*, J. Comput. Phys. **196** (2004), 327 – 347.
- [10] P. Dai Pra, L. Meneghini, and W. Runggaldier, *Connections between stochastic control and dynamic games*, Math. Control Signals Systems **9** (1996), 303–326.
- [11] T. Darden, D. York, and L. Pedersen, *Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems*, J. Chem. Phys. **98** (1993), 10089–10092.
- [12] N. Djurdjevic, M. Sarich, and Ch. Schütte, *Estimating the eigenvalue error of Markov state models*, Multiscale Model. Simul. **10** (1) (2012), 61–81.
- [13] P. Dupuis, K. Spiliopoulos, and H. Wang, *Importance sampling for multiscale diffusions*, Multiscale Model. Simul. **10** (2012), 1–27.
- [14] A.K. Faradjian and R. Elber, *Computing time scales from reaction coordinates by milestoning*, J. Chem. Phys. **120** (2004), 10880–10889.
- [15] W.H. Fleming, *Exit probabilities and optimal stochastic control*, Appl. Math. Optim. **4** (1977), 329–346.
- [16] W.H. Fleming and H.M. Soner, *Controlled Markov processes and viscosity solutions*, Springer, 2006.
- [17] I.I. Gikhman and A.V. Skorokhod, *The theory of stochastic processes II*, Springer-Verlag, New York, 1975.
- [18] C. Hartmann, R. Banisch, M. Sarich, T. Badowski, and Ch. Schütte, *Characterization of rare events in molecular dynamics*, Entropy **16** (2014), 350–376.
- [19] C. Hartmann, J.C. Latorre, G.A. Pavliotis, and W. Zhang, *Optimal control of multiscale systems using reduced-order models*, J. Comput. Dynamics **1** (2014), 279–306.
- [20] C. Hartmann and C. Schütte, *Efficient rare event simulation by optimal nonequilibrium forcing*, J. Stat. Mech. Theor. Exp. **11** (2012), 4.
- [21] B. Hess, H. Bekker, H.J.C. Berendsen, and J. Fraaije, *Lincs: a linear constraint solver for molecular simulations*, J. Comp. Chem. **18** (1997), 1463–1472.
- [22] Ronald H.W. Hoppe, *Multi-grid methods for Hamilton–Jacobi–Bellman equations*, Numer. Math. **49** (1986), 239–254.
- [23] V. Hornak, R. Abel, . Okur, B. Strockbine, A. Roitberg, and C. Simmerling, *Comparison of multiple amber force fields and development of improved protein backbone parameters*, Proteins **65** (2006), 712–725.
- [24] C.-S. Huang, S. Wang, C.S. Chen, and Z.-C. Li, *A radial basis collocation method for Hamilton–Jacobi–Bellman equations*, Automatica **42** (2006), 2201 – 2207.
- [25] H.A. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, Physica **7** (1940), 284 – 304.
- [26] Harold J. Kushner, *A survey of some applications of probability and stochastic control theory to finite difference methods for degenerate elliptic and parabolic equations*, SIAM Review **18** (1976), 545–577.
- [27] Harold J. Kushner, *Numerical methods for stochastic control problems in finance*, Mathematics of Derivative Securities, Cambridge University Press, 1997, pp. 504–527.
- [28] H.J. Kushner and P.G. Dupuis, *Numerical methods for stochastic control problems in continuous time*, Springer Verlag, 1992.
- [29] J.C. Mattingly, A.M. Stuart, and D.J. Higham, *Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise*, Stochastic Process. Appl. **101** (2002),

- 185–232.
- [30] J. Menaldi, *Some estimates for finite difference approximations*, SIAM J. Control Optim. **27** (1989), 579–607.
  - [31] B.K. Øksendal, *Stochastic differential equations: An introduction with applications*, Springer, 2003.
  - [32] M. Peletier, G. Savaré, and M. Veneroni, *Chemical reactions as  $\Gamma$ -limit of diffusion*, SIAM Review **54** (2012), 327–352.
  - [33] H. Rabitz, R. de Vivie-Riedle, M. Motzkus, and K. Kompa, *Whither the future of controlling quantum phenomena?*, Science **288** (2000), 824–828.
  - [34] M. Sarich, *Projected transfer operators*, Ph.D. thesis, FU Berlin, 2011.
  - [35] M. Sarich, F. Noe, and Ch. Schütte, *On the approximation quality of Markov state models*, Multiscale Model. Simul. **8** (2010), 1154–1177.
  - [36] Ch. Schütte, F. Noe, J. Lu, M. Sarich, and E. Vanden-Eijnden, *Markov state models based on milestoning*, J. Chem. Phys. **134** (2011), 204105.
  - [37] Ch. Schütte, S. Winkelmann, and C. Hartmann, *Optimal control of molecular dynamics using Markov state models*, Math. Program. (Series B) **134** (2012), 259–282.
  - [38] S. Sheu, *Stochastic control and exit probabilities of jump processes*, J. Control Optim. **23** (1985), 306–328.
  - [39] Q.S. Song, *Convergence of Markov chain approximation on generalized HJB equation and its applications*, Automatica **44** (2008), 761 – 766.
  - [40] Q.S. Song, G. Yin, and Z. Zhang, *Numerical methods for controlled regime-switching diffusions and regime-switching jump diffusions*, Automatica **42** (2006), 1147 – 1157.
  - [41] H. Stapelfeldt, *Laser aligned molecules: Applications in physics and chemistry*, Physica Scripta **2004** (2004), 132–136.
  - [42] A. Steinbrecher, *Optimal control of robot guided laser material treatment*, Progress in Industrial Mathematics at ECMI 2008 (A.D. Fitt, J. Norbury, H. Ockendon, and E. Wilson, eds.), Springer Berlin Heidelberg, 2010, pp. 505–511.
  - [43] M. Sun, *Domain decomposition algorithms for solving Hamilton-Jacobi-Bellman equations*, Numer. Func. Anal. Optim. **14** (1993), 145–166.
  - [44] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H.J. C. Berendsen, *Gromacs: Fast, flexible, and free*, J. Comp. Chem. **26** (2005), 1701–1718.
  - [45] E. Vanden-Eijnden, *Transition path theory*, Lect. Notes Phys. **703** (2006), 439–478.
  - [46] S. Wang, L.S. Jennings, and K.L. Teo, *Numerical solution of Hamilton-Jacobi-Bellman equations by an upwind finite volume method*, J. Global Optim. **27** (2003), 177–192.
  - [47] H. Wendland, *Meshless galerkin methods using radial basis functions*, Math. Comput. **68** (1999), 1521–1531.
  - [48] M. Zhong and E. Todorov, *Moving least-squares approximations for linearly-solvable stochastic optimal control problems*, J. Control Theory Appl. **9** (2011), 451–463.