Improving Policies for Hamilton-Jacobi-Bellman Equations by Postprocessing

Jan Blechschmidt* Roland Herzog†

May 30, 2017

A postprocessing strategy for optimal controls obtained via numerical solution of Hamilton-Jacobi-Bellman equations is proposed. The method can be applied to a large variety of problems for which the optimal control law can be stated explicitly in terms of the value function and its derivatives, or evaluated numerically. Our approach is based on an adaption of the postprocessing strategy introduced in [Meyer, Rösch, SIAM Journal on Control and Optimization (2004)] in the context of open-loop control. Numerical results are presented and improved convergence rates are confirmed numerically.

Keywords: HJB equation; control law; postprocessing

1. Introduction

Hamilton-Jacobi-Bellman (HJB) equations are a fundamental approach for solving dynamical closed-loop control problems both in deterministic and stochastic settings. HJB equations are stationary or time-dependent partial differential equations which are often highly nonlinear. This results in a variety of interesting and demanding problems related to their numerical solution. Beyond classical methods such as Markov-Chain approximations, see for instance Kushner, 1990, and finite difference schemes,

*Technische Universität Chemnitz, Faculty of Mathematics, Professorship Numerical Mathematics (Partial Differential Equations), 09107 Chemnitz, Germany, jan.blechschmidt@mathematik.tu-chemnitz.de, https://www.tu-chemnitz.de/mathematik/part_dgl/people/blechschmidt

†Technische Universität Chemnitz, Faculty of Mathematics, Professorship Numerical Mathematics (Partial Differential Equations), 09107 Chemnitz, Germany, roland.herzog@mathematik.tu-chemnitz.de, https://www.tu-chemnitz.de/herzog

see Kushner, Dupuis, 2001, new algorithms have been developed in recent years. Among them are finite element (FE) methods for (partial integro) HJB differential equations by Camilli, Jakobsen, 2009, the artificial diffusion FE method by Jensen, Smears, 2012, the discontinuous Galerkin FE method by Smears, Süli, 2014 and the semi-Lagrangian finite difference scheme by Chen, Forsyth, 2007.

The solution of an HJB equation is the value function v, which describes the optimal value of the objective as a function of the initial/current condition in state space. Of equal importance is the optimal control or policy as a function of the current state, denoted by u. Nevertheless, most studies concerning convergence and convergence rates focus primarily on the value function; see for instance Krylov, 1997; 2000 and more recently, Barles, Jakobsen, 2002; 2005; 2007 for finite difference approximations, and Jensen, Smears, 2013 for finite elements.

The optimal policy can usually be expressed, at least implicitly, as a pointwise function of the value function and its derivatives, such as

$$u(t,x) = c(t,x,v(t,x),Dv(t,x),D^2v(t,x)). \tag{1.1}$$

Therefore convergence results for the optimal policy could be inferred in principle from results available for the value function and the required derivatives. However in practice the admissible set for the controls is often infinite and it is discretized during the computation of the value function to facilitate the implementation. This leads to an additional discretization error, which is usually not accounted for in the analysis. Moreover, the policy computed in the process is tied to the chosen discretization.

In this paper we propose a postprocessing strategy which improves the accuracy of discretized optimal policies. Our approach does not interfere with the procedure employed to compute a discrete approximation of the value function but it comes as an additional step once the value function has been found.

The material is organized as follows: The remaining part of Section 1 introduces the general problem structure. Section 2 presents the main concepts of the proposed postprocessing method. Numerical examples which illustrate the procedure are given in Section 3. We conclude with a summary and an outlook on more complex problems to which the proposed framework can be extended.

1.1. Notation and Introductory Example

To set the stage and introduce notation we consider the following simple but typical stochastic optimal control problem. The value function v is defined by

$$v(t,x) = \inf_{U \in \mathcal{U}} \mathbb{E}_{t,x} \int_t^{T\wedge\tau} f(X_s,U_s)\,\mathrm{d}s + g(X_{T\wedge\tau}) \tag{1.2}$$

with underlying controlled stochastic dynamics

$$\begin{aligned} \mathrm{d}X_s &= b(X_s,U_s)\,\mathrm{d}s + \sigma(X_s,U_s)\,\mathrm{d}W_s, \\ X_t &= x. \end{aligned} \tag{1.3}$$

The process $X = (X_s)_{s\in[t,T]}$ describes a d-dimensional controlled Itô-diffusion with drift function b and dispersion matrix σ. The control U can be described by another process $U = (U_s)_{s\in[t,T]}$. We denote by $\mathcal{U}$ the set of all *admissible* control processes U. As usual, a control process U is admissible if it attains values in some set $A \subset \mathbb{R}^m$ and is progressively measurable with respect to the filtration $(\mathcal{F}_s)_{s\in[t,T]}$. The process W in (1.3) is a standard d-dimensional Brownian motion, and $\tau := \inf\{t > 0 \,|\, X_t \notin \Omega\}$ is a stopping time, viz. the exit time from the bounded domain $\Omega \subset \mathbb{R}^d$, which is assumed to be sufficiently regular. The integral in (1.2) extends from t to $T \wedge \tau := \min\{T, \tau\}$, which indicates that the process either stops reaching the boundary $\partial\Omega$ at time $\tau \le T$, or else the process remains in Ω for all of $[t, T]$.

Through the dynamic programming principle we obtain the following parabolic HJB equation governing the value function,

$$
\begin{aligned}
-\frac{\partial v}{\partial t}(t,x) - \inf_{\alpha\in A}\{L^\alpha v(t,x) - f(t,x,\alpha)\} &= 0 && \text{for } (t,x) \in Q = (0,T)\times\Omega,\\
v(T,x) &= g(x) && \text{for } x \in \Omega,\\
v(t,x) &= g(x) && \text{for } x \in \partial\Omega.
\end{aligned}
\tag{1.4}
$$

We refer the reader to Pham, 2009 or Fleming, Soner, 2006 for details on the derivation of HJB equations for continuous-time stochastic control problems. The differential operator L^α is given by the infinitesimal generator of the controlled diffusion process (1.3):

$$
\begin{aligned}
[L^\alpha v](t,x) &= b^T(x,\alpha)\,\nabla v(t,x) + \frac{1}{2}\,\sigma(x,\alpha)\,\sigma^T(x,\alpha) : \nabla^2 v(t,x)\\
&= \sum_{i=1}^d b_i(x,\alpha)\,\frac{\partial v}{\partial x_i}(t,x) + \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(x,\alpha)\,\frac{\partial^2 v}{\partial x_i\,\partial x_j}(t,x)
\end{aligned}
\tag{1.5}
$$

with the coefficient matrix

$$a_{i,j}(x,\alpha) = \sum_{k=1}^d \sigma_{i,k}(x,\alpha)\,\sigma_{j,k}(x,\alpha).$$

If one finds a measurable function $u : Q \to A$ such that

$$\inf_{\alpha\in A}\{L^\alpha v(t,x) - f(t,x,\alpha)\} = L^{u(t,x)} v(t,x) - f(t,x,u(t,x))$$

holds for almost all $(t,x) \in Q$, then u is an optimal control and it is implicitly defined by the value function, i.e.,

$$u(t,x) = \operatorname*{arg\,min}_{\alpha\in A} L^\alpha v(t,x) - f(t,x,\alpha). \tag{1.6}$$

2. Postprocessing of HJB controls

In this section we introduce a postprocessing strategy in order to improve the accuracy of control policies derived from discrete solutions of HJB equations. This strategy has originally been proposed and analyzed in Meyer, Rösch, 2004 in the context of open-loop optimal control of elliptic partial differential equations. It has been shown to offer superior convergence rates of optimal controls w.r.t. the mesh size compared to standard approaches in which the control is discretized but no postprocessing is applied. We make similar observations in the numerical experiments.

2.1. Main Idea

The majority of discretization schemes for HJB equations proceed by discretizing both the value function v and the (usually infinite) control set A. The discretization of the latter can be either explicit or implicit, i.e., depending on the point (t, x) in state space. This observation applies to semi-Lagrangian schemes, see Chen, Forsyth, 2007 and d'Halluin, Forsyth, Labahn, 2005, Markov chain approximations, see Kushner, 1990; Kushner, Dupuis, 2001, and generally to finite difference or finite element schemes.

The main idea of postprocessing is fairly straightforward and it can be summarized as follows. We begin with a discrete approximation $v^h : Q \to \mathbb{R}$ of the value function obtained "as always", by applying some discretization scheme. If necessary, e.g., in case of a finite difference approximation on a grid $Q_h \subset Q$, we prolong the discrete approximation v^h to all of Q by interpolation. As part of the approximation the scheme for the value function returns a discrete approximation of the control function $u^h : Q \to A$ as well. Rather than use (and perhaps interpolate when necessary) this discrete control function u^h, we instead obtain a postprocessed control $\bar{u}^h : Q \to A$ by evaluating the control law (1.6) on v^h pointwise. This amounts to considering the postprocessed control strategy

$$\bar{u}^h(t, x) = \arg\min_{\alpha \in A} L^\alpha v^h(t, x) - f(t, x, \alpha). \tag{2.1}$$

This idea is illustrated in Figure 2.1 by means of the simple, exemplary control law[1]

$$\bar{u}^h = \max\{u_a, \min\{u_b, v^h\}\} \tag{2.2}$$

in one space dimension.

It has been shown in Meyer, Rösch, 2004 for a class of open-loop optimal control problems for elliptic PDEs with right hand side controls, piecewise linear approximations of the state as well as piecewise constant approximations of the control that the postprocessed control $\bar{u}^h$ enjoys a higher convergence order than u^h. This result has been obtained by establishing certain superconvergent approximations under conditions which

[1]The control law (2.2) is used here for simplicity of illustration. It is obtained from (1.6) when $A = [u_a, u_b] \subset \mathbb{R}$, $L^\alpha v = \frac{1}{2}v^2 - \alpha\, v$ and $f(t, x, \alpha) = -\frac{1}{2}\alpha^2$ hold. Notice that this particular form of the generator is not covered by (1.5) but can be obtained for more general, discounted objectives in (1.2).

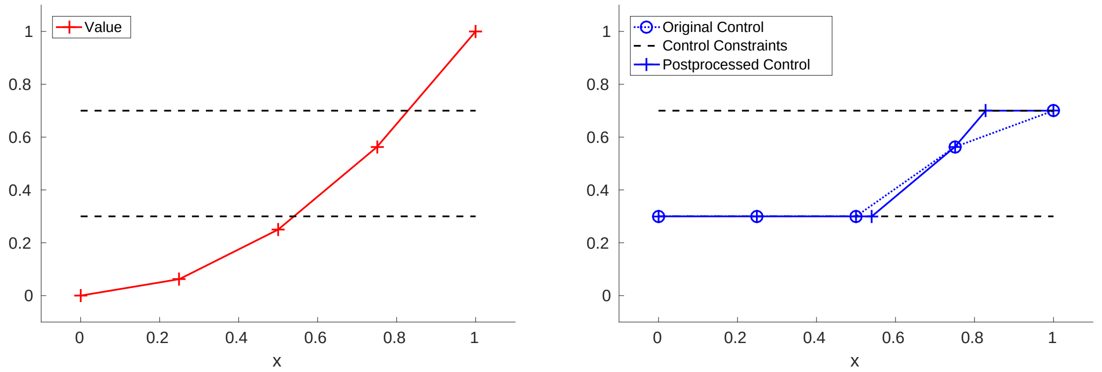

Figure 2.1: Illustration of the postprocessing strategy using the exemplary control law $\bar{u}^h = \max\{u_a, \min\{u_b, v^h\}\}$ in one space dimension.

ensure increased regularity of the adjoint state, which plays the role of the value function in (2.2) in our context.

An investigation of these techniques for the problem class at hand is significantly more involved and will not be pursued in the present paper. The increased level of difficulty stems mainly from the fact that the generator L^α depends on the control. A secondary issue is the parabolic nature of the problem. Nevertheless we remark that the main principles in the technique of proof in Meyer, Rösch, 2004 are also in place here, first and foremost the regularity gain of the control-to-value function map. We therefore expect and are able to verify numerically that the postprocessed control $\bar{u}^h$ exhibits a higher order of convergence than the discretized control u^h.

In order to obtain a high-fidelity approximation $\bar{u}^h$ of the feedback control, formula (2.1) must be evaluated potentially at many points (t, x). This in turn requires that the underlying minimization problem can be solved with high accuracy and at reasonable computational cost. Often (2.1) can even be solved explicitly. We present further examples beyond (2.2) in Section 3.

To summarize, the postprocessing strategy (2.1) can be combined with any discretization scheme yielding an approximation v^h of the value function v. No alterations are required in this part of the solution process. Postprocessing comes in as a secondary step to obtain an improved approximation of the feedback control $\bar{u}^h$ from v^h. Besides for function evaluations, (2.1) can also be used as a basis to compute, e.g., level sets of the control function, switching boundaries between optimal feedback strategies, and other quantities of interest. This will be illustrated in Section 3.

2.2. Evaluation of postprocessed control

Let us briefly address two options of comparing the usual discrete control u^h to the postprocessed control $\bar{u}^h$.

1. When the optimal "continuous" control u^* is available or at least an approximation on a very fine grid, then the *errors in the control* $\|u^h - u^*\|$ and $\|\bar{u}^h - u^*\|$ can be compared in appropriate norms.
2. Of equal interest is the comparison of the *values* incurred by operating according to the controls u^h and $\bar{u}^h$ starting from an initial state $X_t = x$ of interest. In case of u^h this value is defined by

   $$\mathbb{E}_{t,x} \int_t^{T\wedge\tau} f(X_s, u^h(s, X_s))\,\mathrm{d}s + g(X_{T\wedge\tau})$$

   with underlying controlled stochastic dynamics

   $$\begin{aligned} \mathrm{d}X_s &= b(X_s, u^h(s, X_s))\,\mathrm{d}s + \sigma(X_s, u^h(s, X_s))\,\mathrm{d}W_s, \\ X_t &= x. \end{aligned}$$

   The value can be approximated either by Monte-Carlo simulations or by solving the corresponding Fokker-Planck-Kolmogorov equation which promotes the probability density of the state variable forward in time.

Both procedures will be illustrated in the numerical examples in the following section.

3. Numerical Examples

In this section we present different examples arising in optimal stochastic control and their HJB equations. We provide a short description of each problem before discussing and evaluating the postprocessing procedure in the respective situation.

3.1. Minimum Arrival Time

We consider the problem of minimizing the expected time-to-arrival with control cost of a particle exposed to stochastic interference,

$$\inf_{U \in \mathcal{U}} \mathbb{E}_{0,x} \int_0^{T\wedge\tau} 1 + \frac{\gamma}{2} |U_s|_2^2 + \mu\, |U_s|_1 \,\mathrm{d}s. \tag{3.1}$$

Here $\gamma, \mu \geq 0$ are parameters and the underlying controlled stochastic dynamics is given by

$$\begin{aligned} \mathrm{d}X_s &= U_s\,\mathrm{d}s + \sigma\,\mathrm{d}W_s, \\ X_0 &= x. \end{aligned} \tag{3.2}$$

The term *minimum arrival time* results from the fact that the process is either stopped by reaching the boundary $\partial\Omega$ of the bounded domain $\Omega \subset \mathbb{R}^n$ (resulting in $\tau < T$), or else the process is stopped at final time T.

The vector-valued control process U_s with values in $A \subset \mathbb{R}^n$ defines the drift while the diffusion is constant over Ω. To allow for different characteristics in the control we impose in (3.1) control costs of the form

$$f_\gamma(\alpha) = \frac{\gamma}{2} |\alpha|_2^2 = \frac{\gamma}{2} \left(\alpha_1^2 + \ldots + \alpha_n^2\right)$$

and

$$f_\mu(\alpha) = \mu |\alpha|_1 = \mu \left(|\alpha_1| + \ldots + |\alpha_n|\right).$$

While the problem with $\mu = 0$ is classical, the term f_μ is known to promote sparse controls in many other situations; see for instance Stadler, 2009; Vossen, Maurer, 2006. We are not aware of the discussion of objectives involving f_μ in the *stochastic, closed-loop* situation. For the *deterministic*, infinite horizon closed-loop case we refer the reader to the recent Kalise, Kunisch, Rao, 2017.

Assuming $\gamma > 0$ and $A = \mathbb{R}^n$ we obtain from (1.6) in a straightforward way an optimal control law of the form (1.1) as a function of the *gradient* Dv of the value function, viz.

$$u_i^* = \begin{cases} -\frac{1}{\gamma}\left(\frac{\partial v}{\partial x_i} + \mu\right) & \text{where } \frac{\partial v}{\partial x_i} \leq -\mu, \\ -\frac{1}{\gamma}\left(\frac{\partial v}{\partial x_i} - \mu\right) & \text{where } \frac{\partial v}{\partial x_i} \geq \mu, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

We infer that the optimal drift $u_i^*(t,x)$ acting along the i-th coordinate direction is zero in $\left\{(t,x) \in (0,T) \times \Omega : \left|\frac{\partial v}{\partial x_i}(t,x)\right| \leq \mu\right\}$. In the case of vanishing quadratic costs ($\gamma = 0$) we have to impose a bounded control set A to maintain a well-defined problem. For this particular case we obtain a *bang-off-bang* control that attains only the extreme values in A or zero (provided that $0 \in A$). Switching for the i-th control occurs where $\left|\frac{\partial v}{\partial x_i}\right| = \mu$ holds.

The problem is considered on $\Omega = (0,1)^2$ with time horizon $T = 2$ and solved by the finite element (FE) method of artificial diffusion proposed in Jensen, Smears, 2012; 2013. The method is based on an explicit-implicit splitting of the convection and diffusion coefficients b_i and $a_{i,j}$ in (1.5) and it relies on a strictly acute triangulation of the spatial domain. The value function v^h is discretized with piecewise linear, continuous (CG1) finite elements. The monotonicity of the approximation scheme requires sufficient diffusion in the discrete operator, which is accomplished by adding *artificial diffusion*, which goes to zero for decreasing grid size. For our studies we consider an equally spaced temporal grid with $n_t = 1000$ time steps. Spatial refinements are made in a uniform way in order to maintain the strict acuteness of the grid.

In view of (3.3) the discretization of the value function by linear elements naturally suggests a discretization of the control function u^h by piecewise constant, discontinuous (DG0) finite elements.

Figure 3.1 shows the value function, the optimal policy as a vector function and the pointwise 2-norm of the optimal control policy at time $t = 0$ for certain problem parameters given in the figure caption. Since $\mu > 0$ holds, we indeed observe a sparse control, i.e., a non-trivial region in the state space Ω where a zero control is optimal.

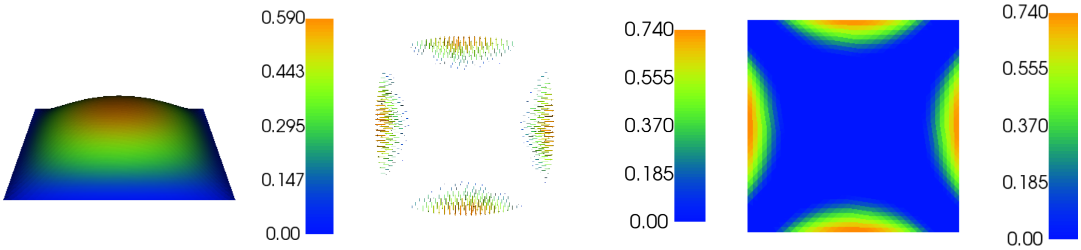

Figure 3.1: Illustration of CG1 value function v^h (left), DG0 control u^h (middle) and $|u^h|_2$ (*without postprocessing*) for the minimum arrival time problem (3.1)–(3.2) with cost parameters $\gamma = \mu = 1.5$ and diffusion $\sigma = 0.5\, I_{2\times 2}$ on a relatively fine grid (3584 cells). All functions are shown at time $t = 0$.

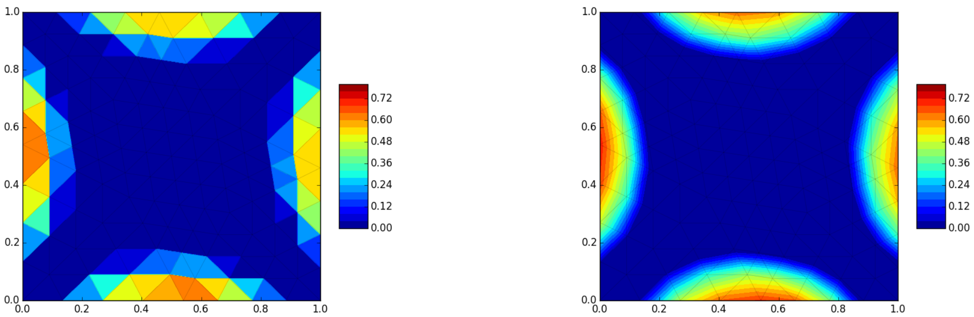

Figure 3.2: Comparison between DG0 control $|u^h|_2$ (*without postprocessing*, left) and $|\bar{u}^h|_2$ (*with postprocessing*, right) for the minimum arrival time problem (3.1)–(3.2) at time $t = 0$ on a coarse mesh with 224 cells for better discernability. Data as in Figure 3.1.

The increased accuracy by postprocessing is clearly visible in Figure 3.2. Let us explain how the postprocessed control $\bar{u}^h$ was obtained. Notice that contrary to our initial illustrative example control law (2.2) and contrary to the problem considered in Meyer, Rösch, 2004, the control now depends on the *gradient* of the value function. A straightforward evaluation of (3.3) (with the CG1 function v^h in place of the continuous value function v) would therefore not lead to an improvement in the postprocessed control. Therefore we first employ a "gradient recovery" procedure. This technique is well known from a posteriori error estimation; see Zienkiewicz, Zhu, 1987. To be more precise we project the DG0 gradient ∇v^h into the vector-valued CG1 space by solving the following L^2 projection problem for the recovered gradient $g^h := R^h(\nabla v^h)$ in $[CG1(\Omega)]^2$,

$$\int_\Omega \nabla v^h \cdot w \, \mathrm{d}x = \int_\Omega g^h \cdot w \, \mathrm{d}x \quad \text{for all } w \in [CG1(\Omega)]^2.$$

Subsequently the postprocessed control is obtained from the control law (3.3), i.e.,

$$\bar{u}_i^h = \begin{cases} -\frac{1}{\gamma}\big([g^h]_i + \mu\big) & \text{where } [g^h]_i \le -\mu, \\ -\frac{1}{\gamma}\big([g^h]_i - \mu\big) & \text{where } [g^h]_i \ge \mu, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that the postprocessed control $\bar{u}^h$ is not a FE function but possesses sub-grid resolution.

To confirm the improvement of the postprocessed control $\bar{u}^h$ compared to u^h also numerically we evaluate the errors $\|u^h - u^*\|_{L^2(\Omega)}$ and $\|u^h - u^*\|_{L^\infty(\Omega)}$ at time $t = 0$; see Figure 3.3. Due to the true solution being unknown we use an (*unpostprocessed*) fine grid solution u^* obtained on a grid with 229 376 cells and 115 201 vertices. Our study shows that the postprocessing procedure indeed improves the convergence order as well as the error constant, as can be seen in Figure 3.3.

Next we consider the case $\gamma = 0$ with control constraints described by the set of admissible controls $A = [-0.1, 0.1]$. In this case the optimal policy is of type *bang-off-bang* and the switching boundaries are sufficient to determine the control law. We illustrate how to obtain the switching boundaries with sub-grid resolution using the postprocessing strategy.

As before we use the discretization method of artificial diffusion and compute a CG1 approximation of the value function v^h. The recovered gradient with components in CG1 is denoted again by g^h.

The two switching boundaries for the i-th component of the control can be found using

$$\begin{aligned} \pi_i^+(t) &:= \big\{x \in \Omega : -[g^h(t,x)]_i + \mu = 0\big\}, \\ \pi_i^-(t) &:= \big\{x \in \Omega : -[g^h(t,x)]_i - \mu = 0\big\}. \end{aligned} \tag{3.4}$$

At a given time t from the time grid, these sets are computed by looping over all triangles T. Whether or not a triangle T intersects $\pi_i^\pm(t)$ can be tested by considering the

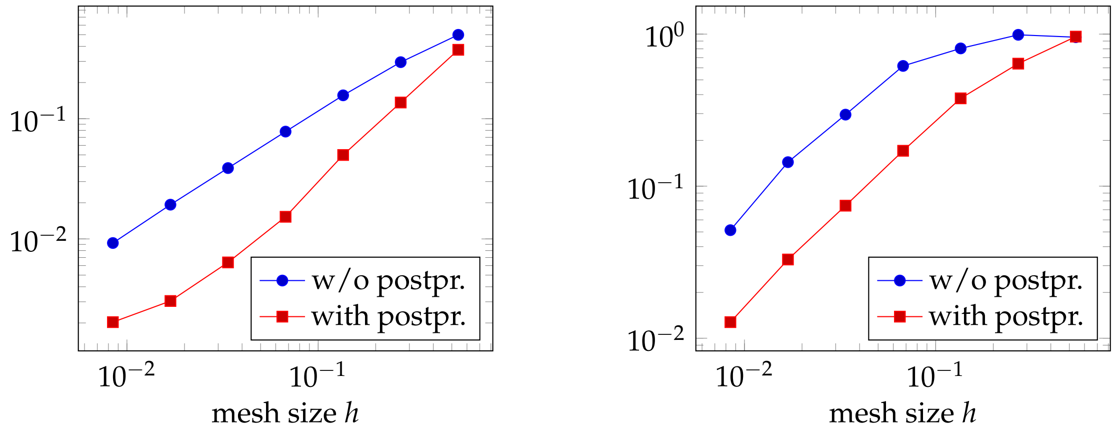

Figure 3.3: Convergence results for the minimum arrival time problem (3.1)–(3.2) with constant isotropic diffusion $\sigma = 0.125\, I_{2\times 2}$, cost parameters $\gamma = \mu = 0.5$. The plots show the L^2-error (left) and L^∞-error (right) of $u^h - u^*$ and $\bar{u}^h - u^*$ at time $t = 0$, respectively.

signs of the values of the linear function $[g^h(t,\cdot)]_i \pm \mu$ in the vertices of T. If T intersects, then it is easy to determine $\pi_i^\pm(t) \cap T$ by finding the two points on the boundary ∂T where $[g^h(t,\cdot)]_i \pm \mu = 0$ holds and connecting them linearly. This procedure is illustrated in Figure 3.4, and it applies not only to two dimensional state spaces but—with appropriate modifications—also in higher dimensions; see Section 3.2.

3.2. Optimal Energy Storage

While the minimum arrival time problem considered in Section 3.1 presents the main ideas of the postprocessing procedure in a simple setting, the following example highlights a more complex application. In contrast to Section 3.1 the state space is three-dimensional and therefore a high spatial resolution incurs considerable numerical effort. Postprocessing can help mitigate the computational cost of obtaining control laws of sufficient accuracy, especially in the case of finite time horizons, which leads to time-dependent HJB equations.

The problem we are considering is the optimal control of a gas storage facility. The model extends previous work by Chen, Forsyth, 2007 and is described and analyzed in detail in the forthcoming Blechschmidt et al., 2017. While the gas storage process is deterministic, the price and consumption are stochastic mean-reverting processes with seasonalities. This setup leads to a singular diffusion coefficient matrix $(a_{i,j})$ due to vanishing viscosity in one spatial dimension. Moreover the problem features time-dependent convection coefficients as well as a cost functional f which is discontinuous due to the structure of the control costs. Although the problem fits well in the framework introduced in Section 1.1 there a multiple sources of potential issues which we will not discuss in detail here. We only mention that standard methods may fail solv-

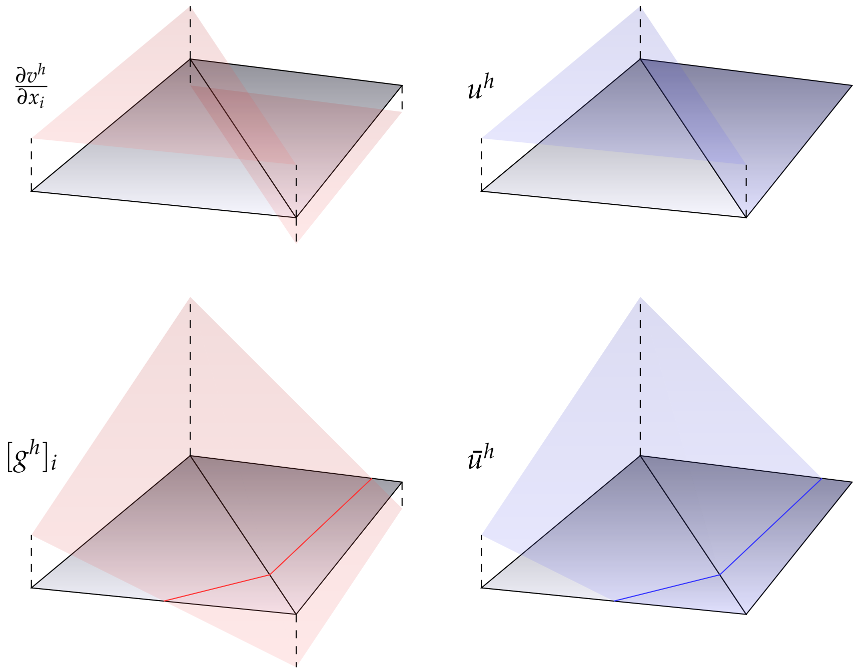

Figure 3.4: Illustration of determination of switching curves

ing mixed deterministic/stochastic optimal control problems due to the degeneracy of the diffusion.

The stochastic dynamics for $X = (P, Q, C)$ for the price process $P_s = p_s$, the storage level $Q_s = q_s$ and the consumption process $C_s = c_s$, together with their initial values, are given by

$$\mathrm{d}X_s = \begin{pmatrix} \kappa_p\,(\mu_p(s) - P_s) \\ U_s - C_s \\ \kappa_c\,(\mu_c(s) - C_s) \end{pmatrix} \mathrm{d}s + \begin{pmatrix} \sigma_p & 0 & 0 \\ 0 & 0 & 0 \\ \rho\,\sigma_c & 0 & \sqrt{1-\rho^2}\,\sigma_c \end{pmatrix} \mathrm{d}W_s$$

where $\kappa_p, \kappa_c, \sigma_c, \sigma_p > 0$ are constants and $\rho \in [0, 1]$ is a correlation coefficient.[2] The functions μ_p and μ_c are time-dependent and they describe seasonalities in the price and consumption dynamics.

The value function for this problem is defined by

$$v(t, x) = \sup_{U \in \mathcal{U}} \mathbb{E}_{t,x} \int_t^T f(X_s, U_s)\,\mathrm{d}s + g(X_T), \tag{3.5}$$

with the supremum ranging over all admissible control processes $U_s \in \mathcal{U}$. Instead of considering merely a maximization of the return we impose different kinds of transac-

[2] The complete set of parameters for our numerical study is given in Appendix A.

tion costs as well as a storage cost. The costs considered are of the type

$$f(x,\alpha) = -d_q^p\, q - \begin{cases} \alpha\left((1+d_+^p)\, p + d_+^f\right) - d_+^c, & \text{where } \alpha > 0, \\ \alpha\left((1-d_-^p)\, p - d_-^f\right) - d_-^c, & \text{where } \alpha < 0, \\ 0 & \text{otherwise.} \end{cases}$$

The constants $d_q^p, d_+^p, d_-^p, d_+^f, d_-^f, d_+^c, d_-^c \in \mathbb{R}_{\geq 0}$ represent storage costs and proportional, fixed and constant transaction costs, respectively. The terminal condition is determined by a penalty payment in case one falls short of meeting the required minimal terminal fill level $\bar{q}_T$:

$$g(x) = \begin{cases} -p\,(1+d_T^p)\,(\bar{q}_T - q_T) & \text{where } \bar{q}_T > q_T, \\ 0, & \text{otherwise.} \end{cases}$$

where $d_T^p > 0$ is another constant. The rigorous derivation of the corresponding HJB equation and proper boundary conditions is a challenging problem in its own right. We state it here without proof and refer to Blechschmidt et al., 2017 for details:

$$\begin{aligned} -\frac{\partial v}{\partial t}(t,x) - \sup_{\alpha\in A}\left\{L^\alpha v(t,x) + f(x,\alpha)\right\} &= 0, \\ v(T,x) &= g(x) \end{aligned} \tag{3.6}$$

with

$$\begin{aligned} L^\alpha v = {} & \kappa_p\,(\mu_p(t) - p)\,\frac{\partial v}{\partial p} + (\alpha - c)\,\frac{\partial v}{\partial q} + \kappa_c\,(\mu_c(t) - c)\,\frac{\partial v}{\partial c} \\ & + \frac{1}{2}\left[\sigma_P^2\,\frac{\partial^2 v}{\partial p^2} + 2\,\rho\,\sigma_P\,\sigma_C\,\frac{\partial^2 v}{\partial p\partial c} + \sigma_C^2\,\frac{\partial^2 v}{\partial c^2}\right]. \end{aligned} \tag{3.7}$$

The optimal control is of bang-off-bang type and it can be expressed explicitly in the form (1.1) in terms of the first derivative of v with respect to the storage direction q:

$$u^*(t,x) = \begin{cases} \zeta(t,x), & \text{where } p(t,x) \geq \frac{1}{1-d_-^p}\left(\frac{\partial v}{\partial q}(t,x) + d_-^f - \frac{d_-^c}{\zeta(t,x)}\right), \\ \eta(t,x), & \text{where } p(t,x) < \frac{1}{1+d_+^p}\left(\frac{\partial v}{\partial q}(t,x) - d_+^f - \frac{d_+^c}{\eta(t,x)}\right), \\ \max\{\zeta(t,x), 0\}, & \text{otherwise.} \end{cases} \tag{3.8}$$

The functions $\zeta, \eta : Q \to \mathbb{R}$ denote the state-dependent maximum inflow and outflow rates and they determine the feasible region A for the control. We use $\Omega = (0,80) \times (0,550\,000) \times (0,16\,000)$ as our computational domain and the considered time horizon is $T = 365$ days to capture one full cycle of seasonality.

We solve the problem using a semi-Lagrangian finite difference approach which exploits the non-existing viscosity in the storage direction q, which would otherwise cause numerical difficulties if not handled properly. This amounts to integrating the PDE (3.6) along Lagrangian trajectories for fixed consumption and price, which allows us to substitute the term

$$\frac{\partial v}{\partial t} + (\alpha - c)\,\frac{\partial v}{\partial q}$$

in (3.6)–(3.7) by a Lagrangian directional derivative

$$\frac{Dv}{Dt} = \frac{\partial v}{\partial t} + (\alpha - c)\,\frac{\partial v}{\partial q}.$$

The resulting system is solved by a fully implicit time-stepping scheme. An important advantage of the semi-Lagrangian scheme, described in detail in Chen, Forsyth, 2007 and d'Halluin, Forsyth, Labahn, 2005, lies in the reduction of the problem size. Instead of solving a linear system of size $n_p \cdot n_q \cdot n_c$ in each time step the method requires the solution of n_q systems of size $n_p \cdot n_c$. In addition, one circumvents a full policy iteration without losing convergence to the unique viscosity solution. For a detailed discussion of the problem including the semi-Lagrangian discretization scheme, existence and uniqueness of a viscosity solution to problem (3.5)–(3.8) and a comparison with previous studies of energy and gas storage problems we refer the reader to Blechschmidt et al., 2017.

We compare the discrete control u^h (without postprocessing) to the postprocessed control $\bar{u}^h$ by means of Monte-Carlo simulations as described at the end of Section 2.2. We ran $m = 10\,000$ samples with either control policy to obtain the estimates $v^h_{\text{sim}}(0, x_0)$ (implementing the policy u^h) and $\bar{v}^h_{\text{sim}}(0, x_0)$ (corresponding to $\bar{u}^h$) by averaging over all samples. All sample paths start at $x_0 = (40, 110\,000, 8000)$ at time $t = 0$. We consider an equally spaced grid in time with $n_t = 365$ time steps and two types of spatial grids, one with uniform distributions of cells in each dimension and one with non-uniform distributions. In the latter, approximately one third of the intervals in each coordinate direction are once refined around the mean reverting levels. Figure 3.5 shows the value function and the optimal discrete control (*without postprocessing*) for a fixed consumption level $c = 7200$ at time $t = 0$.

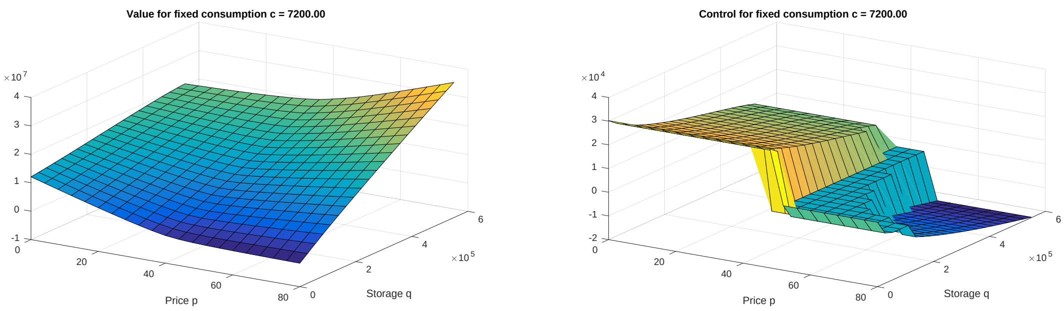

Figure 3.5: Value function (left) and optimal control (right) for energy storage problem with at time $t = 0$ on a coarse (equally spaced) grid with $n_{\text{dofs}} = 9261$ degrees of freedom for better discernability. Shown are slice plots of v^h and u^h for a fixed consumption $c = 7200$.

The results of this study are presented in Table 3.1. The first five rows give the results for equally spaced grids in space, while the remaining rows pertain to non-uniform grids as described above. Column $v^h(0, x_0)$ gives the value as computed by the semi-

Table 3.1: Improvement from postprocessing measured in terms of the incurred estimated values at $(0, x_0)$. Comparison of the values incurred by operating according to the controls u^h and $\bar{u}^h$ starting from an initial state $x_0 = (40, 110\,000, 8000)$ at time $t = 0$.

$n_p = n_q = n_c$	n_{dofs}	$v^h(0, x_0)$	$v^h_{\text{sim}}(0, x_0)$	$\bar{v}^h_{\text{sim}}(0, x_0)$	$s(0, x_0)$
10	1,331	$5.38 \cdot 10^6$	$1.65 \cdot 10^6$	$3.72 \cdot 10^6$	2.3785
20	9,261	$4.42 \cdot 10^6$	$2.09 \cdot 10^6$	$3.98 \cdot 10^6$	1.2408
40	68,921	$4.36 \cdot 10^6$	$2.66 \cdot 10^6$	$4.02 \cdot 10^6$	0.6241
80	$5.31 \cdot 10^5$	$4.30 \cdot 10^6$	$3.45 \cdot 10^6$	$4.07 \cdot 10^6$	0.2003
160	$4.17 \cdot 10^6$	$4.26 \cdot 10^6$	$3.85 \cdot 10^6$	$4.08 \cdot 10^6$	0.0670
15	4,096	$4.39 \cdot 10^6$	$2.03 \cdot 10^6$	$3.91 \cdot 10^6$	1.1894
27	21,952	$4.30 \cdot 10^6$	$2.55 \cdot 10^6$	$4.01 \cdot 10^6$	0.7553
56	$1.76 \cdot 10^5$	$4.27 \cdot 10^6$	$3.03 \cdot 10^6$	$4.04 \cdot 10^6$	0.4239
108	$1.26 \cdot 10^6$	$4.25 \cdot 10^6$	$3.73 \cdot 10^6$	$4.07 \cdot 10^6$	0.1056

Lagrangian method, in comparison to the averages

$$v^h_{\text{sim}}(0, x_0) = \frac{1}{m} \sum_{i=1}^{m} v^h_{\text{sim},i}(0, x_0) \quad \text{and} \quad \bar{v}^h_{\text{sim}}(0, x_0) = \frac{1}{m} \sum_{i=1}^{m} \bar{v}^h_{\text{sim},i}(0, x_0)$$

over the simulated values. The quantity $s(0, x_0)$ denotes the *mean improvement through postprocessing*

$$s(0, x_0) = \frac{1}{m} \sum_{i=1}^{m} \frac{\bar{v}^h_{\text{sim},i}(0, x_0) - v^h_{\text{sim},i}(0, x_0)}{v^h_{\text{sim},i}(0, x_0)}. \tag{3.9}$$

To ensure a meaningful comparison across all discretization levels an identical collection of $m = 10\,000$ samples was used for all discretization levels so that the sample paths for the price and consumption processes are the same on each level. The simulation of the (correlated) price and consumption processes is straightforward. On each level the control process U_t *without postprocessing* is computed by an interpolation of the discrete control function u^h while the control process $\bar{U}_t$ *with postprocessing* is determined by a pointwise evaluation of (3.8). Instead of recovering the gradient as in Section 3.1 we use a first-order central difference in the storage direction q to obtain $\frac{\partial v^h}{\partial q}(t, x)$. This takes into account that the implemented semi-Lagrangian method is based on a finite difference approach.

Combining this with the pointwise evaluation of (3.8) we obtain a significant improvement by postprocessing in terms of the simulated values. For the coarsest spatial discretization with $n_p = n_q = n_c = 10$ cells, we obtain a relative improvement of $s(0, x_0) = 238\%$ in terms of the value, and even for the finest discretization with $n_{\text{dofs}} = 4\,173\,281$ spatial degrees of freedom we still obtain an improvement of about 7%. This amounts to an absolute surplus of 2.29×10^5 in terms of the objective value. In addition we wish to highlight that the postprocessed control for a grid size of $n_p = n_q = n_c = 20$ cells already leads to a (simulated) value which exceeds the value on the finest level without postprocessing. This is an astonishing result, recalling that the only difference between the control without postprocessing and the postprocessed one lies in the pointwise evaluation of the control law (3.8) instead of a plain interpolation of u^h.

A further application of the proposed postprocessing idea lies in the possibility to compute the switching boundaries as explained for the two-dimensional case in Section 3.1. The procedure is similar in three dimensions. Each cuboid is divided into six simplices, for which we compute the zero-level sets of the different cases in (3.8) by linear interpolation. The resulting boundaries are shown in Figure 3.6 at three different time stages. Notice that without preprocessing these boundaries would consist of piecewise axis-parallel segments of significantly coarser resolution.

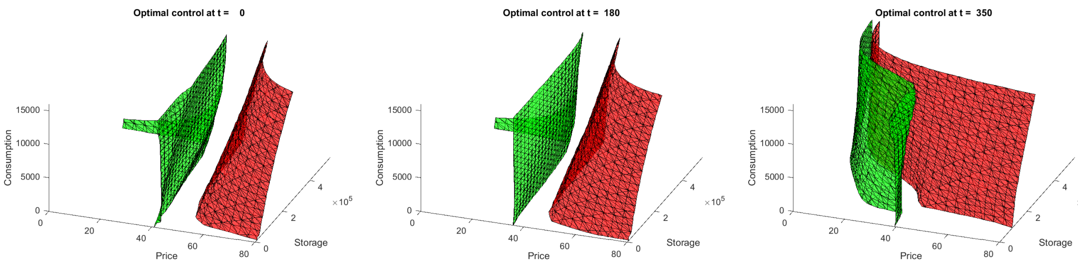

Figure 3.6: Illustration of postprocessed switching boundaries for the energy storage problem at time $t = 0$, $t = 180$ and $t = 350$, respectively. The region to the left of the green surfaces indicates that "buy at maximum rate" is optimal, while the region to the right of the red surface corresponds "sell at maximum rate". The region between these surfaces correponds to the third case in (3.8).

4. Conclusion and Outlook

In this paper we proposed a postprocessing strategy for the policy obtained from numerical solutions of Hamilton-Jacobi-Bellman equations. Our approach has the advantage of improving the accuracy of policies without the necessity of changing the procedure to compute an approximation of the value function. We presented the main ideas on the basis of two examples of different complexities. These examples also show that the approach can be used with finite element and finite difference approximations of the value function. We verified numerically that postprocessing can lead to significant improvements in the accuracy of control laws and their incurred values. A theoretical confirmation of improved convergence rates for the policy w.r.t. the mesh size remains a topic for future research. Another interesting topic is the extension of the approach to HJB variational inequalities (HJB VIs) and quasi-variational inequalities (HJB QVIs).

ACKNOWLEDGMENTS

This work was partially supported by DFG grant HE 6077/7–1 *Impulse Control Problems and Adaptive Numerical Solution of Quasi-Variational Inequalities in Markovian Factor Models*. Funding is gratefully acknowledged. The authors wish to thank Yves Läßig (University of Freiburg) for fruitful discussions about the energy storage problem.

A. PARAMETERS FOR THE ENERGY STORAGE PROBLEM

Parameter	Value	Description
κ_p	0.02	rate of mean-reversion of price process
κ_c	0.007	rate of mean-reversion of consumption process
σ_p	0.50	diffusion coefficient of price process
σ_c	100.00	diffusion coefficient of consumption process
ρ	0.30	correlation coefficient between price and consumption
d_q^p	0.001	proportional storage cost
d_+^p	0.05	proportional cost of buying
d_-^p	0.05	proportional cost of selling
d_+^f	0.00	fixed cost of buying
d_-^f	0.00	fixed cost of selling
d_+^c	10 000.00	constant cost of buying
d_-^c	10 000.00	constant cost of selling
$\bar{q}_T$	110 000	terminal fill level
d_T^p	0.20	proportional cost for penalty payment

MEAN PRICE LEVEL

$$\mu_p(t) = 40 - 3\,\frac{t}{365} + 2.5\,\cos\Big(2\,\pi\,(1 + \frac{t}{365})\Big)$$

MEAN CONSUMPTION LEVEL

$$\mu_c(t) = 8000 - 100\,\frac{t}{365} + 1400\,\cos\Big(5.93 + 2\,\pi\,\frac{t}{365}\Big) + 100\,\cos\Big(5.52 + 4\,\pi\,\frac{t}{365}\Big)$$

MAXIMUM INFLOW RATE

$$\zeta(t,x) = -36.41\,\sqrt{q(t,x)} + c(t,x)$$

Maximum outflow rate

$$\eta(t,x) = 9\,535\,300\sqrt{\frac{1}{q(t,x)+137\,500} - \frac{1}{687\,500}} + c(t,x)$$

References

Barles, Guy, Jakobsen, Espen R. (2002). "On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations". *M2AN. Mathematical Modelling and Numerical Analysis* 36.1, pp. 33–54. ISSN: 0764-583X. DOI: 10.1051/m2an:2002002.

Barles, Guy, Jakobsen, Espen R. (2005). "Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations". *SIAM Journal on Numerical Analysis* 43.2, pp. 540–558. ISSN: 0036-1429. DOI: 10.1137/S003614290343815X.

Barles, Guy, Jakobsen, Espen R. (2007). "Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations". *Mathematics of Computation* 76.260, pp. 1861–1893. ISSN: 0025-5718. DOI: 10.1090/S0025-5718-07-02000-5.

Blechschmidt, J., Herzog, R., Laessig, Y., Schmidt, T. (2017). *Optimal Control and Valuation of Energy Storage with Stochastic Consumption*. Tech. rep. in preparation. Technische Universität Chemnitz, Albert-Ludwigs-Universität Freiburg.

Camilli, Fabio, Jakobsen, Espen R. (2009). "A finite element like scheme for integro-partial differential Hamilton-Jacobi-Bellman equations". *SIAM Journal on Numerical Analysis* 47.4, pp. 2407–2431. ISSN: 0036-1429. DOI: 10.1137/080723144.

Chen, Zhuliang, Forsyth, Peter A. (2007). "A semi-Lagrangian approach for natural gas storage valuation and optimal operation". *SIAM Journal on Scientific Computing* 30.1, pp. 339–368.

d'Halluin, Y., Forsyth, Peter A., Labahn, G. (2005). "A semi-Lagrangian approach for American Asian options under jump diffusion". *SIAM Journal on Scientific Computing* 27.1, pp. 315–345. ISSN: 1064-8275. DOI: 10.1137/030602630.

Fleming, Wendell H., Soner, H. Mete (2006). *Controlled Markov processes and viscosity solutions*. Second. Vol. 25. Stochastic Modelling and Applied Probability. New York: Springer, pp. xviii+429. ISBN: 978-0387-260457; 0-387-26045-5.

Jensen, Max, Smears, Iain (2012). "Finite Element Methods with Artificial Diffusion for Hamilton-Jacobi-Bellman Equations". *Numerical Mathematics and Advanced Applications 2011*. Springer Berlin Heidelberg, pp. 267–274. DOI: 10.1007/978-3-642-33134-3_29.

Jensen, Max, Smears, Iain (2013). "On the convergence of finite element methods for Hamilton-Jacobi-Bellman equations". *SIAM Journal on Numerical Analysis* 51.1, pp. 137–162. DOI: 10.1137/110856198.

Kalise, Dante, Kunisch, Karl, Rao, Zhiping (2017). "Infinite horizon sparse optimal control". *Journal of Optimization Theory and Applications* 172.2, pp. 481–517. ISSN: 0022-3239. DOI: 10.1007/s10957-016-1016-9.

Krylov, N. V. (1997). "On the rate of convergence of finite-difference approximations for Bellman's equations". *Rossiĭskaya Akademiya Nauk. Algebra i Analiz* 9.3, pp. 245–256. ISSN: 0234-0852.

Krylov, N. V. (2000). "On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients". *Probability Theory and Related Fields* 117.1, pp. 1–16. ISSN: 0178-8051. DOI: 10.1007/s004400050264.

Kushner, Harold J. (1990). "Numerical methods for stochastic control problems in continuous time". *SIAM Journal on Control and Optimization* 28.5, pp. 999–1048. ISSN: 0363-0129. DOI: 10.1137/0328056.

Kushner, Harold J., Dupuis, Paul (2001). *Numerical methods for stochastic control problems in continuous time*. Second. Vol. 24. Applications of Mathematics (New York). Stochastic Modelling and Applied Probability. New York: Springer-Verlag, pp. xii+475. ISBN: 0-387-95139-3.

Meyer, C., Rösch, A. (2004). "Superconvergence Properties of Optimal Control Problems". *SIAM Journal on Control and Optimization* 43.3, pp. 970–985. ISSN: 0363-0129. DOI: 10.1137/S0363012903431608.

Pham, Huyên (2009). *Continuous-time stochastic control and optimization with financial applications*. Vol. 61. Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin, pp. xviii+232. ISBN: 978-3-540-89499-5. DOI: 10.1007/978-3-540-89500-8.

Smears, Iain, Süli, Endre (2014). "Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients". *SIAM Journal on Numerical Analysis* 52.2, pp. 993–1016.

Stadler, G. (2009). "Elliptic Optimal Control Problems with L^1-Control Cost and Applications for the Placement of Control Devices". *Computational Optimization and Applications* 44.2, pp. 159–181. DOI: 10.1007/s10589-007-9150-9.

Vossen, G., Maurer, H. (2006). "On L^1-Minimization in Optimal Control and Applications to Robotics". *Optimal Control Applications and Methods* 27.6, pp. 301–321.

Zienkiewicz, O. C., Zhu, J. Z. (1987). "A simple error estimator and adaptive procedure for practical engineering analysis". *International Journal for Numerical Methods in Engineering* 24.2, pp. 337–357. ISSN: 0029-5981. DOI: 10.1002/nme.1620240206.