

# A Simulation Approach to Optimal Stopping Under Partial Information

Mike Ludkovski

*Department of Statistics and Applied Probability  
University of California  
Santa Barbara, CA 93106 USA*

---

## Abstract

We study the numerical solution of nonlinear partially observed optimal stopping problems. The system state is taken to be a multi-dimensional diffusion and drives the drift of the observation process, which is another multi-dimensional diffusion with correlated noise. Such models where the controller is not fully aware of her environment are of interest in applied probability and financial mathematics. We propose a new approximate numerical algorithm based on the particle filtering and regression Monte Carlo methods. The algorithm maintains a continuous state-space and yields an integrated approach to the filtering and control sub-problems. Our approach is entirely simulation-based and therefore allows for a robust implementation with respect to model specification. We carry out the error analysis of our scheme and illustrate with several computational examples. An extension to discretely observed stochastic volatility models is also considered.

MSC2000 Codes: 60G35, 60G40, 93E35

*Key words:* optimal stopping, nonlinear filtering, particle filters, Snell envelope, regression Monte Carlo

---

## 1. Introduction

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$  be a filtered probability space and consider a  $d$ -dimensional process  $X = (X_t)$  satisfying an Itô stochastic differential equation (SDE) of the form

$$dX_t = b(X_t) dt + \alpha(X_t) dU_t + \sigma(X_t) dW_t, \quad (1)$$

where  $U$  and  $W$  are two independent  $(\mathcal{F}_t)$ -adapted Wiener processes of dimension  $d_U$  and  $d_W$  respectively. Let  $Y$  be a  $d_Y \equiv d_U$ -dimensional diffusion given by

$$dY_t = h(X_t) dt + dU_t. \quad (2)$$

Assumptions about the coefficients of (1)-(2) will be given later. Denote by  $\mathcal{F}_t^Y = \sigma(Y_s : 0 \leq s \leq t)$  the filtration generated by  $Y$ . We study the partially observed finite horizon optimal stopping

---

*Email address:* ludkovski@pstat.ucsb.edu (Mike Ludkovski)

*Preprint submitted to Stochastic Processes and Applications*

*September 13, 2009*

problem

$$\sup_{\tau \leq T, \mathcal{F}^Y\text{-adapted}} \mathbb{E} [g(\tau, X_\tau, Y_\tau)], \quad (3)$$

where  $g : [0, T] \times \mathbb{R}^d \times \mathbb{R}^{d_Y} \rightarrow \mathbb{R}$  is the reward functional.

The probabilistic interpretation of (3) is as follows. A controller wishes to maximize expected reward  $g(t, x, y)$  by selecting an optimal stopping time  $\tau$ . Unfortunately, she only has access to the observation process  $Y$ ; the state  $X$  is not revealed and can be only partially inferred through its impact on the drift of  $Y$ . Thus,  $\tau$  must be based on the information contained solely in  $Y$ . Recall that even when  $Y$  is observed continuously, its drift is never known with certainty; in contrast the instantaneous volatility of  $Y$  can be obtained from the corresponding quadratic variation.

Such partially observed problems arise frequently in financial mathematics and applied probability where the agent is not fully aware of her environment, see Section 1.1 below. One of their interesting features is the interaction between learning and optimization. Namely, the observation process  $Y$  plays a *dual* role as a source of information about the system state  $X$ , and as a reward ingredient. Consequently, the agent has to consider the trade-off between further monitoring of  $Y$  in order to obtain a more accurate inference of  $X$ , vis-a-vis stopping early in case the state of the world is unfavorable. This tension between exploration and maximization is even more accentuated when time-discounting is present. Compared to the fully observed setting, we therefore expect that partial information would postpone decisions due to the demand for learning.

In the given form the problem (3) is non-standard because the payoff  $g(t, X_t, Y_t)$  is not adapted to the observed filtration  $(\mathcal{F}_t^Y)$  and, moreover,  $Y$  is not Markovian with respect to  $(\mathcal{F}_t^Y)$ . This difficulty is resolved by a two-step inference/optimization approach. Namely, the first filtering step transforms (3) into an equivalent fully-observed formulation using the Markov conditional distribution  $\pi_t$  of  $X_t$  given  $\mathcal{F}_t^Y$ . In the second step, the resulting standard optimal stopping problem with the Markovian state  $(\pi_t, Y_t)$  is solved.

Each of the two sub-problems above are covered by an extensive literature. The filtering problem with diffusion observations was first studied by Kalman and Bucy [21] and we refer to the excellent texts [2, 20] for the general theory of nonlinear stochastic filtering. The original linear model of [21] had a key advantage in the availability of sufficient statistics and subsequent closed-form filtering formulas for  $\pi_t$ . Other special cases where the filter was explicitly computable were obtained by [1] and [4]. However, in the general setup of (1)-(2), the conditional distribution  $\pi_t$  of  $X_t$  is measure-valued, i.e. an infinite-dimensional object. This precludes consideration of explicit solutions and poses severe computational challenges.

To address such nonlinear models, a variety of approximation tools have been proposed. First, one may linearize the system (1)-(2) by applying (A) the *extended Kalman filter* [18, 23]. Thus, the conditional distribution of  $X$  is summarized by its conditional mean  $m_t = \mathbb{E}[X_t | \mathcal{F}_t^Y]$  and conditional variance  $P_t = \mathbb{E}[(X_t - m_t)^2 | \mathcal{F}_t^Y]$ . One then derives (approximate) evolution equations for  $(m_t, P_t)$  given observations  $Y$ . More generally,  $\pi_t$  can be parameterized by a given family of probability densities, yielding the (B) *projection filter*. Let us especially single out the exponential projection methods studied by Brigo et al. [4, 5]. Third, the state space of  $\pi_t$  can be discretized through (C) *optimal quantization* methods [35, 36]. This replaces  $(\pi_t)$  by a non-Markovian approximation  $(\tilde{\pi}_t)$  whose transition probabilities are pre-processed via Monte Carlo simulation. Fourth, one may apply (D) *Wiener chaos expansion* methods [28, 27, 32] that reduce computation of  $\pi_t$  to a solution of SDE's plus ordinary differential equation systems. Finally,

(E) *interacting particle systems* have been considered to approximate  $\pi_t$  non-parametrically via simulation tools [7, 8, 9, 12].

The optimal stopping sub-problem of the second step can again be tackled within several frameworks. When the transition density of the state variables is known, classical (a) *dynamic programming* computations are possible, see e.g. [38]. If the problem state is low-dimensional and Markov, one may alternatively use the quasi-variational formulation to obtain a free-boundary partial differential equation (pde) and then implement a (b) *numerical pde solver* for an efficient solution. Thirdly, (c) *simulation-based methods* [13, 26, 40] that rely on probabilistic Snell envelope techniques can be applied.

The joint problem of optimal stopping with partial observations was treated in [16], [17], [29], [36] and [33]. All these models can be viewed as a combination of the listed approaches to the two filtering/optimization sub-problems. For example, [29] proposes to use the assumed density filter for the filtering step, followed by a pde solver for the optimization. This can be summarized as algorithm (B)/(b) in our notation. Meanwhile, [36] use (C)/(a), i.e. optimal quantization for the filter and then dynamic programming to find optimal stopping times. Methodologically, two ideas have been studied. First, using filtering techniques (A) or (B), one may replace  $\pi_t$  by a low-dimensional Markovian approximation  $\tilde{\pi}_t$ . Depending on the complexity of the model, algorithms (a) or (b) can then be applied in the second step. Unfortunately, the resulting filtering equations are inconsistent with the true dynamics of  $\pi_t$ , and require a lot of computations to derive them for each considered model. The other alternative is to use the quantization technique (C) which is robust and produces a consistent (but non-Markovian) approximation  $\tilde{\pi}_t$ . Since the state space of  $\tilde{\pi}_t$  is fully discretized, the resulting optimal stopping problem can be solved exactly using dynamic programming algorithm (a). Moreover, tight error bounds are available. The shortcomings of this approach are the need to discretize the state space of  $X$  and the requirement of offline pre-processing to compute the transition density of  $\tilde{\pi}_t$ .

In this paper we propose a new approach of type (E)/(c) that uses a *particle filter* for the inference step and a *simulation tool* in the optimization step. Our method is attractive based on three accounts. Firstly, being entirely simulation-based it can be generically applied to a wide variety of models, with only minor modifications. In particular, the implementation is robust and requires only the ability to simulate  $(X_t, Y_t)$ . For comparison, free boundary pde solvers of type (b) often use advanced numerical techniques for stability and accuracy purposes and must be re-programmed for each class of models. Also, in contrast to optimal quantization, no pre-processing is needed. Moreover, the interacting particle system approach to filtering is also robust with respect to different observation schemes. In the original system (1)-(2) it is assumed that  $Y$  is observed continuously. It is straightforward to switch our algorithm to discrete regularly-spaced observations of  $Y$  that may be more natural in some contexts.

Secondly, our approach maintains a continuous state space throughout all computations. In particular, the computed optimal stopping rule  $\tau^*$  is continuous, eliminating that source of error and leading to a more natural decision criteria for the controller. Thus, compared to optimal quantization, our approach is expected to produce more “smooth” optimal stopping boundaries. Third, our method allows the user to utilize her *domain knowledge* during the optimization step. In most practical applications, the user already has a guess regarding an optimal stopping rule and the numerical computations are used as a refinement and precision tool. However, most optimal stopping algorithms rely on a “brute force” scheme to obtain an optimal stopping rule. By permitting custom input for the optimization step, our scheme should heuristically lead to reduced computational efforts and increased accuracy.

Finally, maintaining the simulation paradigm throughout the solution allows us to integrate

the filtering and Snell envelope computations. In particular, by carrying along a high-dimensional approximation of  $\pi_t$ , the initial filtering errors can be minimized in a flexible and anticipative way with respect to the subsequent optimization step. Thus, the introduction of filtering errors is delayed for as long as possible. This is important for optimal stopping where the forward-propagated errors (such as the filtering error) strongly affect the subsequent backward recursion solution for  $\tau^*$ . To summarize, our scheme should be viewed as an even more flexible alternative for the optimal quantization method of [36].

*Remark 1.* To our knowledge the idea of integrated stochastic filtering and optimization was conceived in [34], in the context of utility maximization with partially observed state variables. Muller et al. [34] proposed to use the Markov Chain Monte Carlo (MCMC) methods and an auxiliary randomized pseudo-control variable to do both steps at once. These ideas were then further analyzed in [3, 41] for a portfolio optimization problem with unobserved drift parameter and unobserved stochastic volatility, respectively. In fact, Viens et al. [41] utilized a particle filter but then relied on discretizing the control and observation processes to obtain a finite-dimensional problem with discrete scenarios. While of the same flavor, this approach must be modified for optimal stopping problems like (3), as the control variable  $\tau$  is infinite-dimensional. Indeed, stopping rules  $\tau$  are in one-to-one correspondence with stopping regions, i.e. subsets of the space-time state space. Such objects do not admit easy discretization. Moreover, the explicit presence of time-dimension as part of our control makes MCMC simulation difficult. Thus, we maintain the probabilistic backward recursion solution method instead.

The rest of the paper is organized as follows. In Section 2 we recall the general filtering paradigm for our model and the Snell envelope formulation of the optimal stopping problem (3). Section 3 describes in detail the new algorithm, including the variance-minimizing branching particle filter in Section 3.1, and the regression Monte Carlo approach to compute the Snell envelope in Section 3.2. We devote Section 4 to the error analysis of our scheme and to the proof of the overall convergence of the algorithm. Section 5 then illustrates our scheme on a numerical example; a further computational example is provided in Section 6 which discusses the extension of our method to discretely observed stochastic volatility models. Finally, Section 7 concludes.

Before proceeding, we now give a small list of applications of the model (1)-(3).

### 1.1. Applications

#### *Optimal Investment under Partial Information*

The following investment timing problem arises in the theory of *real options*. A manager is planning to launch a new project, whose value ( $Y_t$ ) evolves according to

$$dY_t = X_t dt + \sigma_Y dU_t,$$

where the drift parameter ( $X_t$ ) is unobserved and ( $U_t$ ) is an  $\mathbb{R}$ -valued Wiener process. The environment variable  $X_t$  represents the current economic conditions; thus when the economy is booming, potential project value grows quickly, whereas it may be declining during a recession. At launch time  $\tau$  the received profit  $g(\cdot)$  is a function of current project value  $Y_\tau$ , as well as extra uncertainty that depends on the environment state. For instance, consider  $g(\cdot) = Y_\tau \cdot (a_0 + a_1 X_\tau + b_0 \epsilon)$ ,  $\epsilon \sim \mathcal{N}(0, 1)$  independent, where the second term models the profit multiplier based on economy state. Conditioning on the realization of  $\epsilon$ , expected profit is  $g(\tau, X_\tau, Y_\tau) = Y_\tau (a_0 + a_1 X_\tau)$ . Such a model with continuous-time observations was considered by [11] in the *static* case where  $X_0 \in \{0, 1\}$  and  $dX_t = 0$ . A similar problem was studied in [31] with an additional consumption control.

Using the methods below, we can treat this problem for general  $X$ -dynamics of the type (1), under both continuous and discrete observations.

### *Stochastic Convenience Yield Models*

Compared to holding of financial futures, physical ownership of commodities entails additional benefits and costs. Accordingly, the rate of return on the commodity spot contract will be different from the risk-free rate. The stochastic convenience yield models [6, 37] postulate that the drift of the asset price ( $Y_t$ ) under the pricing measure  $\mathbb{P}$  is itself a stochastic process,

$$\begin{cases} dY_t = Y_t(X_t dt + \sigma_Y dU_t), \\ dX_t = b(X_t) dt + \alpha(X_t) dU_t + \sigma_X(X_t) dW_t. \end{cases}$$

One may now consider the pricing of American Put options on asset  $Y$  with maturity  $T$  and strike  $K$ ,

$$\sup_{\tau \leq T} \mathbb{E} [e^{-r\tau} (K - Y_\tau)_+],$$

where the convenience yield  $X$  is unobserved and must be dynamically inferred. To learn about  $X$  it is also possible to filter other observables beyond  $Y$ , e.g. futures contracts, see [6].

### *Reliability Models with Continuous Review*

Quality control models in industrial engineering [19] can also be viewed as examples of (3). Let  $X_t$  represent the current quality of the manufacturing process. This quality fluctuates due to machinery state and also external disturbances, such as current workforce effort, random shocks, etc. When quality is high, the revenue stream  $Y$  is increasing; conversely poor quality may decrease revenues. Because revenues are also subject to random disturbances, current quality is never observed directly. In this context, it is asked to find an optimal time  $\tau$  to replace the machinery (at cost  $g(X_t)$ ) and reset the quality process  $X$ . Assuming “white noise” shocks to the system and continuous monitoring of revenue stream this leads again to (1)-(2)-(3). The case where  $Y$  is discretely observed and  $X$  is a finite-state Markov chain was treated by Jensen and Hsu [19].

## **2. Optimization Problem**

### *2.1. Notation*

We will use the following notation throughout the paper:

- For  $x \in \mathbb{R}$ , we write  $x = \lfloor x \rfloor + \{x\}$  to denote the largest integer smaller than  $x$  and the fractional part of  $x$ , respectively.
- $\delta_x$  denotes the Dirac measure at point  $x$ .
- $C_b^k$  denotes the space of all real-valued, bounded, continuous functions with bounded continuous derivatives up to order  $k$  on  $\mathbb{R}^d$ . We endow  $C_b^k(\mathbb{R}^d)$  with the following norm

$$\|f\|_{m,\infty} = \sum_{|\alpha| \leq m} \sup_{x \in \mathbb{R}^d} |D_\alpha f(x)|, \quad f \in C_b^k(\mathbb{R}^d), \quad m \leq k,$$

where  $\alpha = (\alpha_1, \dots, \alpha_d)$  is a multi-index and derivatives are written as  $D_\alpha f \triangleq \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d} f$ .

- $W_p^k = \{f : D_\alpha f \in L^p(\mathbb{R}^d), |\alpha| \leq k\}$  denotes the Sobolev space of functions with  $p$ -integrable derivatives up to order  $k$ .
- $\mathcal{P}(\mathbb{R}^d)$  is the space of all probability measures over the Borel  $\sigma$ -algebra  $\mathcal{B}(\mathbb{R}^d)$ . For  $\mu \in \mathcal{P}$ ,  $\mu(f) \triangleq \int_{\mathbb{R}^d} f(x)\mu(dx)$ . We endow  $\mathcal{P}$  with the weak topology;  $\mu_n \rightarrow \mu$  weakly if  $\forall f \in C_b^0$ ,  $\mu_n(f) \rightarrow \mu(f)$ .

## 2.2. Filtering Model

In this section we briefly review the theory of nonlinear filtering as applied to problem (3). We follow [7] in our presentation.

Before we begin, we make the following technical assumption regarding the coefficients in (1) and (2).

**Assumption 1.** *The coefficients of (1) satisfy:  $b(x) \in C_b^3(\mathbb{R}^d)$ ,  $\alpha(x) \in C_b^3(\mathbb{R}^{d \times d_Y})$ ,  $\sigma(x) \in C_b^3(\mathbb{R}^{d \times d_W})$  and moreover,  $\alpha$  and  $\sigma$  are strictly positive-definite matrices of size  $d \times d_Y$  and  $d \times d_W$  respectively. Similarly, in (2),  $h(x) \in C_b^4(\mathbb{R}^d)$ .*

This assumption in particular guarantees the existence of a unique strong solution to (1), (2). We also assume that

**Assumption 2.** *The payoff function  $g$  is bounded and twice jointly continuously differentiable  $g \in C_b^2([0, T] \times \mathbb{R}^d \times \mathbb{R}^{d_Y})$ .*

The latter condition is often violated in practice where payoffs can be unbounded. However, one may always truncate  $g$  at some high level  $\bar{G}$  without violating the applicability of the model.

We begin by considering the conditional distribution of  $X$  given  $\mathcal{F}_t^Y$ . Namely, for  $f \in C_b^2(\mathbb{R}^d)$  define

$$\pi_t f \triangleq \mathbb{E}[f(X_t) | \mathcal{F}_t^Y]. \quad (4)$$

It is well-known [2] that  $\pi_t f$  is a Markov,  $\mathcal{F}^Y$ -adapted process that solves the Kushner-Stratonovich equation

$$d(\pi_t f) = \pi_t(Af) dt + \sum_{k=1}^{d_Y} \left[ \pi_t(h_k \cdot f) - \pi_t(h_k) \cdot \pi_t(f) + \pi_t(B^k f) \right] [dY_t^k - \pi_t(h_k) dt], \quad (5)$$

where the action of the differential operators  $A$  and  $B^k$  on a test function  $f \in C_b^2(\mathbb{R}^d)$  is defined by

$$\begin{cases} Af(x) \triangleq \frac{1}{2} \sum_{i,j}^d \left( \sum_{k=1}^{d_U} \alpha_{ik}(x) \alpha_{jk}(x) + \sum_{k=1}^{d_W} \sigma_{ik}(x) \sigma_{jk}(x) \right) \partial_i \partial_j f(x) + \sum_{i=1}^d b_i(x) \partial_i f(x), \\ B^k f(x) \triangleq \sum_{i=1}^d \alpha_{ik}(x) \partial_i f(x). \end{cases} \quad (6)$$

In other words,  $\pi_t$  is a probability measure-valued process solving the stochastic partial differential equation (spde) corresponding to the adjoint of (5). To avoid the nonlinearities in (5), a

simpler linear version is obtained by utilizing the reference probability measure device. Define a  $\mathbb{P}$ -equivalent probability measure  $\tilde{\mathbb{P}}$  by

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\Big|_{\mathcal{F}_\tau} = \zeta_\tau \triangleq \exp\left(-\sum_{k=1}^{d_U} \int_0^\tau h_k(X_s) dU_s^k - \frac{1}{2} \sum_{k=1}^{d_U} \int_0^\tau h_k^2(X_s) ds\right). \quad (7)$$

From the Girsanov change of measure theorem (recall that  $h$  is bounded so that  $\mathbb{E}[\zeta_t] = 1$ ), it follows that under  $\tilde{\mathbb{P}}$  the observation  $Y$  is a Brownian motion and the signal  $X$  satisfies

$$dX_t = (b(X_t) - ah(X_t)) dt + \alpha(X_t) dY_t + \sigma(X_t) dW_t. \quad (8)$$

We now set

$$\rho_t f \triangleq \tilde{\mathbb{E}}\left[f(X_t)\zeta_t^{-1}\Big|\mathcal{F}_t^Y\right], \quad (9)$$

with  $\zeta_t$  defined in (7). Then by Bayes formula,  $\pi_t f = \frac{\rho_t f}{\rho_t}$  and moreover,  $\rho_t f$  solves the *linear* stochastic differential equation

$$d(\rho_t f) = \rho_t (A f) dt + \sum_{k=1}^{d_Y} [\rho_t (h_k f) + \rho_t (B^k f)] dY_t^k, \quad (10)$$

with  $A, B^k$  from (6). The measure-valued Markov process  $\rho_t$  is called the *unnormalized* conditional distribution of  $X$  and will play a major role in the subsequent analysis. Under the given smoothness assumptions, it is known [2] that  $\pi_t$  (and  $\rho_t$ ) will possess a smooth density in  $W_p^1$  for all  $p > 1$  and  $t > 0$ .

Returning to our optimal stopping problem (3), let us define the value function  $V$  by

$$V(t, \xi, y; T) \triangleq \sup_{\tau \leq T, \mathcal{F}^Y\text{-adapted}} \mathbb{E}\left[g(\tau, X_\tau, Y_\tau) \Big| X_t \sim \xi, Y_t = y\right].$$

Economically,  $V$  denotes the optimal reward that can be obtained on the horizon  $[t, T]$  starting with initial condition  $Y_t = y$  and  $X_t \sim \xi$ . Using conditional expectations we may write,

$$\begin{aligned} V(t, \xi, y) &= \sup_{t \leq \tau \leq T} \mathbb{E}^{t, \xi, y} [\pi_\tau g(\tau, \cdot, Y_\tau)] \\ &= \sup_{t \leq \tau \leq T} \tilde{\mathbb{E}}^{t, \xi, y} [\rho_\tau g(\tau, \cdot, Y_\tau)] \\ &= \sup_{t \leq \tau \leq T} \tilde{\mathbb{E}}^{t, \xi, y} [G(\tau, \rho_\tau, Y_\tau)], \quad \text{where } G(t, \xi, y) \triangleq \int_{\mathbb{R}^d} g(t, x, y) \xi(dx), \end{aligned} \quad (11)$$

and where  $\tilde{\mathbb{E}}^{t, \xi, y}$  denotes  $\tilde{\mathbb{P}}$ -expectation conditional on  $\rho_t = \xi, Y_t = y$ .

Equation (11) achieved two key transformations. First, its right-hand-side is now a standard optimal stopping problem featuring the Markov hyperstate  $(\rho_t, Y_t)$ . Secondly, (11) has separated the filtering and optimization steps by introducing the fully observed problem through the new state variable  $\rho_t$ . However, this new formulation remains complex as  $\rho_t$  is an infinite-dimensional object. With a slight abuse of notation, we will write  $V(t, \rho_t, Y_t)$  to denote the value function in terms of the current unnormalized distribution  $\rho_t$ .

As can be seen from the last two lines of (11), one may solve (3) either under the original *physical* measure  $\mathbb{P}$  using  $\pi_t$ , or equivalently under the reference measure  $\tilde{\mathbb{P}}$  using  $\rho_t$ . In our

approach we will work with the latter formulation due to the simpler dynamics of  $\rho_t$  and more importantly due to the fact that under  $\tilde{\mathbb{P}}$  one can separate the evolution of  $Y$  and  $X$ . In particular, under  $\tilde{\mathbb{P}}$ ,  $Y$  is a Brownian motion and can be simulated entirely on its own. In contrast, under  $\mathbb{P}$ , the evolutions of  $\pi_t$  and  $Y$  are intrinsically tied together due to the *joint* (and unobserved) noise source ( $U_t$ ).

### 2.3. Snell Envelope

Let us briefly summarize the Snell envelope theory of optimal stopping in our setting. All our results are stated under the  $\tilde{\mathbb{P}}$  reference measure, following the formulation in (11). For any  $\mathcal{F}^Y$ -stopping time  $\sigma$ , define

$$\hat{\mathcal{Z}}(\sigma) = \sup_{\sigma \leq \tau \leq T} \tilde{\mathbb{E}} \left[ G(\tau, \rho_\tau, Y_\tau) | \mathcal{F}_\sigma^Y \right]. \quad (12)$$

**Proposition 1** ([29]). *The set  $(\hat{\mathcal{Z}}_\sigma)$  form a supermartingale family, i.e. there exists a continuous process  $\mathcal{Z}$ , such that  $\hat{\mathcal{Z}}(\sigma) = \mathcal{Z}_\sigma$ ,  $\mathcal{Z}$  stopped at time  $\sigma$ . Moreover, an optimal time  $\tau$  for (11) exists and is given by  $\tau = \inf\{t: \mathcal{Z}_t = G(t, \rho_t, Y_t)\}$ .*

The above process  $\mathcal{Z}$  is called the Snell envelope of the optimal stopping problem (11). The proposition implies that to solve (11) it suffices to compute the Snell envelope  $\mathcal{Z}$ . We denote by  $t \leq \tau_t^* \leq T$  an optimal stopping time achieving the supremum in  $\mathcal{Z}_t = \tilde{\mathbb{E}}[G(\tau_t^*, \rho_{\tau_t^*}, Y_{\tau_t^*}) | \mathcal{F}_t^Y]$ . By virtue of the (strong) Markov property of  $(\rho_t, Y_t)$  and the fact that  $\rho_t$  is a sufficient statistic for the distribution of  $X_t | \mathcal{F}_t^Y$  it follows that  $V(t, \rho_t, Y_t) = \sup_{t \leq \tau \leq T} \tilde{\mathbb{E}}[G(\tau, \rho_\tau, Y_\tau) | \mathcal{F}_t^Y] = \mathcal{Z}_t$  and (3) is equivalent to finding  $\tau_0^*$  above. Mazziotto [29] also gave a formal proof of the equivalence of the Snell envelopes under  $\mathbb{P}$  and  $\tilde{\mathbb{P}}$  that we discussed in the end of the previous section.

To make computational progress in computing  $\tau_0^*$ , it will be eventually necessary to discretize time. Thus, we restrict possible stopping times to lie in the set  $\mathcal{S}^\Delta = \{0, \Delta t, 2\Delta t, \dots, T\}$ , and label the corresponding value function (of the so-called Bermudan problem) as

$$V^\Delta(t, \xi, y) = \sup\{\tilde{\mathbb{E}}^{t, \xi, y}[G(\tau, \rho_\tau, Y_\tau)] : t \leq \tau \text{ is } \mathcal{S}^\Delta\text{-valued, } \mathcal{F}^Y\text{-adapted}\}.$$

In this discrete version, since one either stops at  $t$  or waits till  $t + \Delta t$ , the dynamic programming principle implies that the Snell envelope satisfies

$$V^\Delta(t, \rho_t, Y_t) = \max\left(G(t, \rho_t, Y_t), \tilde{\mathbb{E}}\left[V^\Delta(t + \Delta t, \rho_{t+\Delta t}, Y_{t+\Delta t}) | \mathcal{F}_t^Y\right]\right). \quad (13)$$

### 2.4. Continuation Values and Cashflow Functions

For notational convenience we now write  $Z_t \equiv (t, \rho_t, Y_t)$  and  $G_t = G(Z_t)$ . Let

$$q_t = q_t(Z_t) \triangleq \tilde{\mathbb{E}}[V^\Delta(Z_{t+\Delta t}) | \mathcal{F}_t^Y],$$

denote the *continuation* value. Then the Snell envelope property (13) implies that  $q_t$  satisfies the recursive equation

$$q_t = \tilde{\mathbb{E}}\left[\max(G_{t+\Delta t}, q_{t+\Delta t}) | \mathcal{F}_t^Y\right]. \quad (14)$$

The optimal stopping time  $\tau_t^*$  also satisfies a recursion, namely

$$\tau_t^* = \tau_{t+\Delta t}^* \mathbf{1}_{\{q_t > G_t\}} + t \mathbf{1}_{\{q_t \leq G_t\}}. \quad (15)$$

In other words, when the continuation value is bigger than the immediate expected reward, it is optimal to wait; otherwise it is optimal to stop. Equation (15) also highlights the fact that the continuation value  $q_t$  serves as a *threshold* in making the stopping decision. Associated with a stopping rule  $\tau^*$  defined above is the future cashflow function. Denote  $B_t(q) \triangleq 1_{\{q_t \leq G_t\}}$  and its complement by  $B_t^c(q) \equiv 1 - B_t(q)$ , and starting from the timepoint  $t$ , define the expected future cashflow as

$$\vartheta_t(q)(\mathbf{Z}) \triangleq \sum_{s=t}^T G(Z_s) \cdot 1_{B_s(q)} \cdot 1_{B_t^c(q) \cdot B_{t+\Delta t}^c(q) \cdots B_{s-\Delta t}^c(q)}. \quad (16)$$

$\vartheta_t(q)$  is a path function whose value depends on the realization of  $(Z_t)$  between  $t$  and  $T$ , as well as the threshold function  $q$ . Note that (16) can be defined for *any* threshold rule  $q'$  by simply using  $B_t(q')$ , etc. instead. In discrete time using the fact that  $\tau_t^*$  is an  $\mathcal{F}^Y$ -stopping time and (15) we get

$$\begin{aligned} q_t(Z_t) &= \tilde{\mathbb{E}}[V^\Delta(Z_{t+\Delta t}) | \mathcal{F}_t^Y] = \tilde{\mathbb{E}}[G(\tau_{t+\Delta t}^*, \rho_{\tau_{t+\Delta t}^*}, Y_{\tau_{t+\Delta t}^*}) | \mathcal{F}_t^Y] \\ &= \tilde{\mathbb{E}} \left[ \sum_{s=t+\Delta t}^T G(s, \rho_s, Y_s) 1_{\{\tau_{t+\Delta t}^* = s\}} \middle| \mathcal{F}_t^Y \right] = \tilde{\mathbb{E}} \left[ \vartheta_{t+\Delta t}(q)(\mathbf{Z}) \middle| \mathcal{F}_t^Y \right]. \end{aligned} \quad (17)$$

It follows that knowing  $\vartheta(q)$ , one can back-out the continuation values  $q$  and then recover the value function itself from  $V^\Delta(Z_t) = \max(G(Z_t), q(Z_t))$ . In particular, for  $t = 0$ , we obtain  $V^\Delta(0, \xi_0, y_0) = \max(G(Z_0), q_0(Z_0))$ . The approximation algorithm will compute  $q$  and the associated  $\vartheta$  by repeatedly evaluating the conditional expectation in (17) and updating (16). The advantage in using cashflows  $\vartheta(q)$  rather than  $q$  itself is that an error in computing  $q$  is not propagated backwards *unless* it leads to a wrong stopping decision for (15). As a result, the numerical scheme is more stable.

*Remark 2.* Egloff [13] discusses a slightly more general situation, where the look-ahead cashflows  $\vartheta$  are taken not on the full horizon  $[t, T]$  but only some number  $w$  of steps ahead. This then produces

$$\vartheta_{t,w}(q)(\mathbf{Z}) = \sum_{s=t}^{t+w\Delta t} G(Z_s) \cdot 1_{B_t^c B_{t+\Delta t}^c \cdots B_s} + q_{t+w\Delta t}(Z_{t+w\Delta t}) \cdot 1_{B_t^c B_{t+\Delta t}^c \cdots B_{t+w\Delta t}^c}, \quad (18)$$

and one still has  $q_t(Z_t) = \tilde{\mathbb{E}}[\vartheta_{t+1,w}(q)(\mathbf{Z}) | \mathcal{F}_t^Y]$  for any  $w = 0, \dots, T - t - 1$ . In particular, the case  $w = 0$  is the Tsitsiklis-van Roy [40] algorithm,

$$\vartheta_{t,0}(q) = G(Z_t) 1_{B_t} + q_t(Z_t) 1_{B_t^c} = \max(G_t, q_t). \quad (19)$$

To compute (17), the corresponding conditional expectation will be approximated by a finite-dimensional projection  $\mathcal{H}$ . Indeed, by definition of conditional expectation with respect to the Markov state  $(\rho_t, Y_t)$ , we have  $q_t(Z_t) = \tilde{\mathbb{E}}[\vartheta_{t+\Delta t}(q)(\mathbf{Z}) | \mathcal{F}_t^Y] = F(\rho_t, Y_t)$  for some function  $F$ . Let  $(B_j)_{j=1}^\infty$  be a (Schauder) basis for the Banach space  $\mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d)$ . Then as  $r \rightarrow \infty$ ,  $F$  (and  $q_t$ ) can be approximated arbitrarily well by the truncated sum

$$q_t(Z_t) \simeq \hat{q}_t(Z_t) \triangleq \sum_{j=1}^r \alpha_j B_j(\rho_t, Y_t) = \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}} \left[ \vartheta_{t+\Delta t}(q)(\mathbf{Z}) \middle| \mathcal{F}_t^Y \right], \quad (20)$$

where the projection manifold (or architecture) is  $\mathcal{H} = \text{span}(B_j(\xi, y), j = 1, \dots, r)$ . As long as (20) does not modify much the resulting stopping sets  $B_t(\hat{q})$ , one expects that the resulting cashflow function  $\vartheta(\hat{q})$  will be close to the true one  $\vartheta(q)$ . In our filtering context, the extra modification is that  $Z_t$  must itself be approximated by a finite-dimensional filter  $Z_t^n$ . However, if the approximation is high-dimensional, then it should have very little effect on the projection step of the Snell envelope in (20).

### 2.5. Analytic Approach

We briefly recall the analytic approach to optimal stopping theory which characterizes the value function  $V(t, \xi, y)$  in terms of a parabolic-type free boundary problem. This is in direct counterpart to standard optimal stopping problems for diffusion models.

The major difficulty is the infinite-dimensional nature of the state variable  $\pi$ . Limited results exist for the corresponding optimal stopping problems on Polish spaces, see e.g. [30, 29]. In particular, [30] characterize  $V$  as the minimal excessive function dominating  $G$  in terms of the (Feller) transition semigroups of  $(\pi_t, Y_t)$ . A more direct theory is available when  $\pi_t \in H$  belongs to a Hilbert space; this will be the case if  $\xi_0$  (and therefore  $\pi_t$  for all  $t$ ) admits a smooth  $L^2$ -density. Even then, since the smoothness properties of  $V$  with respect to  $\xi$  are unknown, one must work with viscosity solutions to second-order pdes as is common in general stochastic control theory. The following proposition is analogous to Theorem 2.2 in [16]. Denote by  $D$  the Fréchet derivative operator and for a twice Fréchet differentiable test function  $\phi(t, \xi, y)$  let

$$\mathcal{L}\phi = \frac{1}{2} \text{tr}((\sigma\sigma^T + \alpha\alpha^T)D_{\xi\xi}^2\phi) + \langle b, D_\xi\phi \rangle + h\partial_y\phi + \alpha D_\xi\phi \cdot \partial_y\phi, \quad (21)$$

(with  $\langle \cdot, \cdot \rangle$  denoting the inner product in  $H$ ) be the infinitesimal generator of the Markov process  $(\pi_t, Y_t)$ .

**Proposition 2.** *The value function  $V(t, \pi, y)$  is the unique viscosity solution of*

$$\begin{cases} V_t + \mathcal{L}V \leq 0, \\ V(t, \pi, y) \geq G(t, \pi, y). \end{cases} \quad (22)$$

*Moreover,  $V$  is bounded and locally Lipschitz (with respect to the Hilbert norm).*

In principle the infinite-dimensional free boundary problem (22) can be tackled by a variety of numerical methods including the projection approach that passes to a finite-dimensional subset of  $L^2(\mathbb{R}^d)$ . We will return to (22) in Section 5.1.

### 3. New Algorithm

In this section we describe a new numerical simulation algorithm to solve (11). This algorithm will be a combination of the minimal-variance branching particle filter algorithm for approximating  $\pi_t$  and  $\rho_t$ , described in Section 3.1, and the regression Monte Carlo algorithm described in Section 3.2.

### 3.1. Particle Filtering

The main idea of particle filters is to approximate the measure-valued conditional distribution  $\pi_t$  by a discrete system of point masses that follows a mutation-selection algorithm to reproduce the dynamics of (10). In what follows we summarize the particular algorithm proposed in [7, 9, 8]. We assume that we are given (1)-(2) with continuous observation of  $(Y_t)$ . Fix  $n > 0$ ; we shall approximate  $\pi_t$  by a particle system  $\pi_t^n$  of  $n$  particles. The interacting particle system consists of a collection of  $n$  weights  $a_j^n(t)$  and corresponding locations  $v_j^n(t) \in \mathbb{R}^d$ ,  $j = 1, \dots, n$ . We think of  $v_j^n$  as describing the evolution of the  $n$ -th particle and of  $a_j^n(t) \in \mathbb{R}_+$  as its importance in the overall system. Begin by initializing the system by independently drawing  $v_j^n(0)$  from the initial distribution  $X_0 \sim \xi_0$  and taking  $a_j^n(0) = 1 \forall j$ . Let  $\delta$  be a parameter indicating the frequency of mutations; the description below is for a generic time step  $t \in [m\delta, (m+1)\delta)$ , assuming that we already have  $v_j^n(m\delta)$  and  $a_j^n(m\delta) \equiv 1$ .

First, for  $m\delta \leq t < (m+1)\delta$  we have

$$v_j^n(t) = v_j^n(m\delta) + \int_{m\delta}^t (b - \alpha h)(v_j^n(s)) ds + \int_{m\delta}^t \alpha(v_j^n(s)) dY_s + \int_{m\delta}^t \sigma(v_j^n(s)) dW_s^{(j)}, \quad (23)$$

where  $W^{(j)}$  are  $n$  independent  $\tilde{\mathbb{P}}$ -Wiener processes. Thus, each particle location evolves independently according to the law of  $X$  under  $\tilde{\mathbb{P}}$ . The unnormalized weights  $a_j^n(\cdot)$  are given by the stochastic exponentials

$$a_j^n(t) = 1 + \sum_{k=1}^{d_Y} \int_{m\delta}^t a_j^n(s) h_k(v_j^n(s)) dY_s^k = \exp\left(\sum_{k=1}^{d_Y} \int_{m\delta}^t h_k(v_j^n(s)) dY_s^k - \frac{1}{2} \sum_{k=1}^{d_Y} \int_{m\delta}^t h_k(v_j^n(s))^2 ds\right). \quad (24)$$

Let

$$\bar{a}_j^n((m+1)\delta-) \triangleq \frac{a_j^n((m+1)\delta-)}{\sum_j a_j^n((m+1)\delta-)} \in (0, 1),$$

denote the normalized weights just before the next mutation time. Then at  $t = (m+1)\delta$  each particle produces  $o_j^n((m+1)\delta)$  offspring inheriting the parent's location, with the branching carried out such that

$$\begin{cases} o_j^n((m+1)\delta) = \begin{cases} \lfloor n\bar{a}_j^n((m+1)\delta-) \rfloor & \text{with prob. } 1 - \{n\bar{a}_j^n((m+1)\delta-)\}, \\ 1 + \lfloor n\bar{a}_j^n((m+1)\delta-) \rfloor & \text{with prob. } \{n\bar{a}_j^n((m+1)\delta-)\}, \end{cases} \\ \sum_{j=1}^n o_j^n((m+1)\delta) = n, \end{cases} \quad (25)$$

where  $\{x\}$  denotes the fractional part of  $x \in \mathbb{R}$ . Note that the different  $o_j^n$ 's are correlated so that the total number of particles always stays constant at  $n$ . One way to generate such  $o_j^n$ 's is given in the Appendix of [7]. Following the mutation, particle weights are reset to  $a_j^n((m+1)\delta) = 1$  and one proceeds with the next propagation step.

With this construction we now set for  $m\delta \leq t < (m+1)\delta$ ,

$$\begin{cases} \pi_t^n \triangleq \sum_{j=1}^n \frac{na_j^n(t)}{\sum_{\ell=1}^n a_\ell^n(t)} \delta_{v_j^n(t)}(\cdot); \\ \rho_t^n \triangleq \left[ \prod_{\ell=1}^m \left( \frac{1}{n} \sum_{j=1}^n a_j^n(\ell\delta-) \right) \right] \cdot \left( \frac{1}{n} \sum_{j=1}^n a_j^n(t) \delta_{v_j^n(t)}(\cdot) \right). \end{cases} \quad (26)$$

Interpreted as a probability measure on  $\mathbb{R}^d$ ,  $\pi_t^n$  ( $\rho_t^n$ ) is an approximation to the true  $\pi_t$  (resp.  $\rho_t$ ) as indicated by the following

**Proposition 3** ([7], Theorem 5). *There exist constants  $C_1(t), C_2(t)$  such that for any  $f \in C_b^1(\mathbb{R}^d)$ ,*

$$\mathbb{E}\left[\left(\frac{\rho_t^n f - \rho_t f}{\rho_t 1}\right)^2\right] \leq \frac{C_1(t)}{n} \|f\|_{1,\infty}^2, \quad (27)$$

which in turn implies that (since  $\mathbb{E}[\zeta_t^2]$  is bounded)

$$\mathbb{E}\left[(\pi_t^n f - \pi_t f)^2\right] \leq \frac{C_2(t)}{n} \|f\|_{1,\infty}^2, \quad (28)$$

with  $C_i(t) = O(e^t \cdot t)$ .

Similar results can be obtained under the assumption that  $Y$  is observed *discretely* every  $\delta$  time units. In that case one simply takes,

$$a_j^n((m+1)\delta-) = \exp\left(\sum_{k=1}^{d_Y} h_k(v_j^n(m\delta)) \cdot (Y_{(m+1)\delta}^k - Y_{m\delta}^k) - \frac{1}{2} \sum_{k=1}^{d_Y} h_k(v_j^n(m\delta))^2 \cdot \delta\right),$$

with the rest of the algorithm remaining unchanged.

The use of discrete point masses in the interacting particle filter renders the analytical results based on Hilbert-space theory (e.g. (22)) inapplicable. This can be overcome by considering regularized particle filters [24], where point masses are replaced by smooth continuous distributions and the particle branching procedure switches back to a true re-sampling step.

### 3.2. Regression Monte Carlo

The main idea of our algorithm is to simulate  $N$  paths of the  $Z$  process (or rather the particle approximation ( $Z^n$ )), yielding a sample  $(z_t^k)$ ,  $k = 1, 2, \dots, N$ ,  $t = 0, \Delta t, \dots, T$ . To simulate  $(z_t^i)$ , we first simulate the Brownian motion  $(Y_t)$  under  $\mathbb{P}$ , and then re-compute  $\rho_t^n$  along the simulated paths as described in the previous subsection. Using this sample and approximation architectures  $\mathcal{H}_t$  of (20), we approximate the projection  $\text{pr}_{\mathcal{H}_t}$  through an empirical least-squares regression. Namely, an empirical continuation value is computed according to

$$\hat{q}_t = \arg \min_{f \in \mathcal{H}_t} \frac{1}{N} \sum_{i=1}^N |f(z_t^i) - \vartheta^N(\hat{q})(z_{t+\Delta t}^i)|^2 \simeq \text{pr}_{\mathcal{H}_t} \circ \mathbb{E}[\vartheta_{t+\Delta t}^N(\hat{q})(\mathbf{Z}^n) | \mathcal{F}_t^Y], \quad (29)$$

where  $\vartheta^N$  is the empirical cashflow function along simulated paths obtained using the future  $\hat{q}$ 's. One then updates pathwise  $\vartheta^N$  and  $\tau$  using (16) and (15) respectively and proceeds recursively backwards in time. This is the same idea as the celebrated regression Monte Carlo algorithm of Longstaff and Schwartz [26]. The resulting error between  $\hat{q}$  and the true  $q$  will be studied in Section 4 below.

Many choices exist regarding the selection of basis functions  $B_j(\rho_t, Y_t)$  for the regression step. As a function of  $y$ , one may pick any basis for  $L^2(\mathbb{R}^{d_Y}, \mathbb{P})$ , e.g. the Laguerre polynomials. As a function of  $\rho$ , a natural probabilistic choice involves the moments of  $X_t | \mathcal{F}_t^Y$ , i.e.  $\sum_i \alpha_i(\rho_t x^i)$ . It is also known that using a basis function of the form  $EUR(z) \triangleq \mathbb{E}_t[G(Z_T)]$  (the conditional expectation of the terminal reward or the ‘‘European’’ counterpart,) is a good empirical choice.

*Remark 3.* If one only uses the first two conditional moments of  $X$ ,  $\rho_t x$  and  $\rho_t x^2$  inside the basis functions, then our algorithm can be seen as the non-Markovian analogue of applying the extended Kalman filter for the partial observations of  $X$  and then computing the (pseudo)-Snell envelope of (3). In that sense, our approach generalizes previous filtering projection methods [5, 23] for (3).

### 3.3. Overall Algorithm

For the reader's convenience, we now summarize the overall algorithm for solving (3).

- Select model parameters  $N$  (number of paths);  $n$  (number of particles per path);  $\Delta t$  (time step for Snell envelope);  $\delta$  (time step for observations and particle mutation);  $B_i$  (regression basis functions);  $r$  (number of basis functions).
- Simulate  $N$  paths of  $(y_t^k)$  under  $\tilde{\mathbb{P}}$  (which is a Brownian motion) with fixed initial condition  $y_0^k = y_0$ .
- Given the path  $(y_t^k)$ , use the particle filter algorithm (23)-(24)-(26) to compute  $\rho_t^{n,k}$  along that path, starting with  $\rho_0^{n,k} \sim \xi_0$ .
- Initialize  $\hat{q}^k(T) = \vartheta_T^{N,k}(\hat{q}) = G(z_T^k)$ ,  $\tau^k(T) = T$ ,  $k = 1, \dots, N$ .
- Repeat for  $t = (M - 1)\Delta t, \dots, \Delta t, 0$ :

- Evaluate the basis functions  $B_i(z_t^k)$ , for  $i = 1, \dots, r$  and  $k = 1, \dots, N$ .
- Regress

$$\alpha_t^N \triangleq \arg \min_{\alpha \in \mathbb{R}^r} \sum_{k=1}^N \left| \vartheta_{t+\Delta t}^{N,k}(\hat{q}) - \sum_{i=1}^r \alpha^i B_i(z_t^k) \right|^2.$$

- For each  $k = 1, \dots, N$  do the following steps: Set  $\hat{q}^k(t) = \sum_{i=1}^r \alpha_t^{N,i} B_i(z_t^k)$ .
- Compute  $G(z_t^k) = \rho_t^{n,k} g(t, \cdot, y_t^k)$ .
- Update  $\vartheta_t^{N,k}(\hat{q}) = \begin{cases} G(z_t^k) & \text{if } \hat{q}_t^k < G(z_t^k); \\ \vartheta_{t+\Delta t}^{N,k}(\hat{q}) & \text{otherwise.} \end{cases}$
- Update  $\tau^k(t) = \begin{cases} t & \text{if } \hat{q}_t^k < G(z_t^k); \\ \tau^k(t + \Delta t) & \text{otherwise.} \end{cases}$

- End loop;
- Return  $V^\Delta(0, \xi_0, y_0) \simeq \frac{1}{N} \sum_{k=1}^N \vartheta_0^{N,k}(\hat{q})$ .

Note that it is not necessary to save the entire particle systems  $(v_j^{n,k}(m\Delta t))_{j=1}^n$  after the simulation step; rather one needs to keep around just the evaluated basis functions  $(B_i(z_t^k))_{i=1}^r$ , so that the total memory requirements are  $O(N \cdot M \cdot r)$ . In terms of number of operations the overall algorithm complexity is  $O(M \cdot N \cdot (n^2 + r^3))$ , with the most intensive steps being the resampling of the filter particles and the regression step against the  $r$  basis functions.

#### 4. Error Analysis

This section is devoted to the error analysis of the algorithm proposed in Section 3.3. Looking back, our numerical scheme involves three main errors. These are:

- Error in computing  $\rho_t$  which arises from using a finite number of particles and the resampling error of the particle filter  $\rho_t^n$ ;
- Error in projecting the cashflow function  $\vartheta$  onto the span of basis functions  $\mathcal{H}$  and the subsequent wrong stopping decisions;
- Error in computing projection coefficients  $\alpha^i$  due to the use of finite-sample least-squares regression.

We note that the filtering error is propagated forward, while the projection and empirical errors are propagated backwards. In that sense, the filtering error is more severe and should be controlled tightly. The projection error is the most difficult to deal with since we only have crude estimates on the dependence of the value function on  $\rho_t$ . Consequently, the provable error estimates are very pessimistic. Heuristic considerations would imply that this error is in fact likely to be small. Indeed, the approximate decision rule will be excellent as long as  $\tilde{\mathbb{P}}(\{q_t > G_t\} \cap \{\hat{q}_t \leq G_t\})$  is small, since the given event is the only way that the optimal cashflows are computed incorrectly. By applying domain knowledge the above probability can be controlled through customizing the projection architecture  $\mathcal{H}_t$ . For instance, as mentioned above, using  $EUR(z)$  as one of the basis functions is often useful.

The sample regression error is compounded due to the fact that we do not use the true basis functions but rather approximations based on  $Z^n$ . This implies the presence of *error-in-variable* during the regression step from the pathwise filtering errors. It is well-known (see e.g. [15]) that this leads to attenuation in the computed regression result, i.e.  $|\alpha^{N,i}| \leq |\alpha^i|$ . An extensive statistical literature treats error reduction methods to counteract this effect, a topic that we leave to future research.

As a notational shorthand, in the remainder of this section we write  $\tilde{\mathbb{E}}_t$  to denote expectations (as a function on  $\mathbb{R}^d \times \mathbb{R}^{d_r}$ ) conditional on  $Y_t = y$  and  $\rho_t = \xi$ . We recall that the optimal cashflows satisfy

$$q_t = \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q)(\mathbf{Z})],$$

while the approximate cashflows are

$$\hat{q}_t = \text{pr}_{\mathcal{H}}^N \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}^N(\hat{q})(\mathbf{Z}^n)].$$

Note that inside the algorithm,  $\hat{q}_t$  is evaluated not at the true value  $Z_t = (\rho_t, Y_t)$ , but at the approximate point  $Z_t^n$ . To emphasize the process under consideration we denote by  $q_t^n \equiv q_t^n(Z_t^n)$  the continuation function resulting from working with the  $Z^n$ -process. Observe that the difference between  $q_t^n$  and the true  $q$  is solely due to the inaccurate recursive evaluation of the reward  $G$  (since  $Y$  is simulated exactly); thus if the original reward  $g$  in (3) is independent of  $X$  then  $q_t^n \equiv q$ .

The error analysis will be undertaken in two steps. In the first step, we consider the mean-squared error between the continuation value  $q_t$  based on the true filter  $\rho_t$  and the continuation value  $q_t^n$  based on the approximate filter  $\rho_t^n$ . In the second step, we will study the difference between  $q_t^n$  and the approximate  $\hat{q}_t$  above. Throughout this section,  $\|\cdot\|_2 \equiv \tilde{\mathbb{E}}[\|\cdot\|^2]^{1/2}$ .

**Lemma 1.** *There exists a constant  $C(T)$ , such that for all  $t \leq T$*

$$\left\| \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n) - \vartheta_{t+\Delta t}(q)] \right\|_2 \leq \frac{(T-t) \cdot C(T)}{\Delta t \cdot \sqrt{n}} \cdot \|g\|_{1,\infty}. \quad (30)$$

*Proof.* Suppose without loss of generality that  $q^n(Z_t^n) > q(Z_t)$ . Let  $\tau$  be an optimal stopping time for the problem represented by  $q^n$ . Clearly such  $\tau$  is sub-optimal for  $q$ ; moreover since both  $Z$  and  $Z^n$  are  $\mathcal{F}^Y$ -adapted,  $\tau$  is admissible for  $q$ . Therefore,

$$\begin{aligned} (q^n(Z_t^n) - q(Z_t))^2 &\leq \tilde{\mathbb{E}}_t[G(Z_\tau^n) - G(Z_\tau)]^2 \\ &= \left\{ \sum_{s=t+\Delta t}^T \tilde{\mathbb{E}}_t[(G(Z_s^n) - G(Z_s)) \cdot 1_{\{\tau=s\}}] \right\}^2 \\ &\leq \sum_{s=t+\Delta t}^T \frac{T-t}{\Delta t} \cdot \tilde{\mathbb{E}}_t[|G(Z_s^n) - G(Z_s)|^2], \end{aligned}$$

where the last line is due to Jensen's inequality. Averaging over the realizations of  $(Z_t^n, Z_t)$  we then obtain

$$\begin{aligned} \tilde{\mathbb{E}}[|q^n(Z_t^n) - q(Z_t)|^2] &\leq \sum_{s=t+\Delta t}^T \frac{T-t}{\Delta t} \cdot \tilde{\mathbb{E}}[|G(Z_s^n) - G(Z_s)|^2] \\ &\leq \sum_{s=t+\Delta t}^T \frac{(T-t)C(T)}{\Delta t \cdot n} \|g\|_{1,\infty}^2 = \frac{(T-t)^2 \cdot C(T)}{\Delta t^2 \cdot n} \|g\|_{1,\infty}^2, \end{aligned}$$

using Proposition 3.

Note that this error explodes as  $\Delta t \rightarrow 0$  due to the fact that we do not have tight bounds for  $\tilde{\mathbb{E}}_t[|G(Z_s^n) - G(Z_s)|^2 1_{\{\tau=s\}}]$ . In general, one expects that  $\tilde{\mathbb{E}}_t[|G(Z_s^n) - G(Z_s)|^2 1_{\{\tau=s\}}] \simeq \tilde{\mathbb{E}}_t[|G(Z_s^n) - G(Z_s)|^2] \cdot \mathbb{P}(\tau = s)$  which would eliminate the  $\Delta t^{-2}$  term on the last line above.  $\square$

In the second step we study the  $L^2$ -difference of the unnormalized continuation values,  $\|q_t^n - \hat{q}_t\|_2 \equiv \tilde{\mathbb{E}}[(q_t^n(Z_t^n) - \hat{q}_t(Z_t))^2]^{1/2}$ . This total error can be decomposed as

$$\begin{aligned} \|\hat{q}_t - q_t\|_2 &\leq \underbrace{\left\| \text{pr}_{\mathcal{H}}^N \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}^N(\hat{q})(\mathbf{Z}^n)] - \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})(\mathbf{Z}^n)] \right\|_2}_{\mathcal{E}_1} \\ &\quad + \underbrace{\left\| \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})(\mathbf{Z}^n)] - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})(\mathbf{Z}^n)] \right\|_2}_{\mathcal{E}_2} + \underbrace{\left\| \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})(\mathbf{Z}^n)] - \vartheta_{t+\Delta t}(q)(\mathbf{Z}^n) \right\|_2}_{\mathcal{E}_3}. \end{aligned} \quad (31)$$

The three error terms  $\mathcal{E}_i$  on the right-hand-side of (31) are respectively the empirical error  $\mathcal{E}_1$ , the projection error  $\mathcal{E}_2$ , and the recursive error from the next time step  $\mathcal{E}_3$ . Each of these terms is considered in turn in the next several lemmas with the final result summarized in Theorem 1. The first two lemmas have essentially appeared in [13] and the proofs below are provided for completeness.

**Lemma 2** ([13, Lemma 6.3]). *Define the centered loss random variable*

$$\ell_t(\hat{q})(\mathbf{Z}^n) = |\hat{q}_t - \vartheta_{t+\Delta t}(\hat{q})|^2 - |\text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q})|^2. \quad (32)$$

Then

$$\mathcal{E}_1^2 = \|\hat{q}_t - \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]\|_2^2 \leq \tilde{\mathbb{E}}[\ell_t(\hat{q})(\mathbf{Z}^n)]. \quad (33)$$

*Proof.* First note that

$$\|\hat{q}_t - \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]\|_2^2 + \|\text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]\|_2^2 \leq \|\hat{q}_t - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]\|_2^2, \quad (34)$$

because  $\hat{q}_t \in \mathcal{H}_t$  belongs to the convex space  $\mathcal{H}_t$ , while  $\text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] \in \mathcal{H}_t$  is the projection of  $\vartheta_{t+\Delta t}(\hat{q})$ . Therefore the three respective vectors form an obtuse triangle in  $L^2$ :

$$\tilde{\mathbb{E}} \left[ (\hat{q}_t - \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]) \cdot (\text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]) \right] \leq 0.$$

Direct expansion using the tower property of conditional expectations and the fact that  $\hat{q}_t \in \mathcal{F}_t^Y$  shows that  $\tilde{\mathbb{E}}[(\hat{q}_t - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]) \cdot (\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q}))] = 0$ , so that

$$\tilde{\mathbb{E}} \left[ |\hat{q}_t - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]|^2 \right] + \tilde{\mathbb{E}} \left[ |\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q})|^2 \right] = \tilde{\mathbb{E}} \left[ |\hat{q}_t - \vartheta_{t+\Delta t}(\hat{q})|^2 \right]. \quad (35)$$

Similarly,

$$\tilde{\mathbb{E}} \left[ (\text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]) \cdot (\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q})) \right] = 0,$$

and so

$$\begin{aligned} \tilde{\mathbb{E}} \left[ |\text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]|^2 \right] &= \tilde{\mathbb{E}} \left[ |\text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q})|^2 \right] \\ &\quad - \tilde{\mathbb{E}} \left[ |\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q})|^2 \right]. \end{aligned} \quad (36)$$

Combining (34)-(35)-(36) we find

$$\begin{aligned} \left\| \hat{q}_t - \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] \right\|_2^2 &\leq \tilde{\mathbb{E}} \left[ |\hat{q}_t - \vartheta_{t+\Delta t}(\hat{q})|^2 - |\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q})|^2 \right] \\ &\quad - \left\{ |\text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q})|^2 - |\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q})|^2 \right\} = \tilde{\mathbb{E}}[\ell_t(\hat{q})(\mathbf{Z}^n)]. \end{aligned}$$

□

The above lemma shows that the squared error  $\mathcal{E}_1^2$  resulting from the empirical regression used to obtain  $\hat{q}_t$  (which recall is a proxy for  $\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})]$ ) can be expressed as the difference between the expected actual difference  $|\hat{q}_t - \vartheta_{t+\Delta t}(\hat{q})|^2$  versus the theoretical best average error after the projection  $|\text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \vartheta_{t+\Delta t}(\hat{q})|^2$ .

**Lemma 3** (cf. [13, Proposition 6.1]). *We have  $\mathcal{E}_2 \leq 2\|\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n) - \vartheta_{t+\Delta t}(\hat{q})]\|_2 + \inf_{f \in \mathcal{H}_t} \|f - q_t^n\|_2$ .*

*Proof.* We re-write,

$$\begin{aligned} \mathcal{E}_2 &= \left\| \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] \right\|_2 \leq \left\| \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(\hat{q})] - \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n)] \right\|_2 \\ &\quad + \left\| \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n)] - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n)] \right\|_2 + \left\| \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n)] - \vartheta_{t+\Delta t}(\hat{q}) \right\|_2 \\ &\leq 2\|\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n) - \vartheta_{t+\Delta t}(\hat{q})]\|_2 + \inf_{f \in \mathcal{H}_t} \|f - \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n)]\|_2 \\ &= 2\|\tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n) - \vartheta_{t+\Delta t}(\hat{q})]\|_2 + \inf_{f \in \mathcal{H}_t} \|f - q_t^n\|_2, \end{aligned}$$

where the second inequality uses the contraction property of the projection map  $\text{pr}_{\mathcal{H}}$  and the definition of projection onto the manifold  $\mathcal{H}_t$ . □

**Lemma 4.** We have for any  $p > 1$

$$\left\| \tilde{\mathbb{E}}_t[\vartheta_{t+\Delta t}(q^n) - \vartheta_{t+\Delta t}(\hat{q})] \right\|_p \leq \sum_{s=t+\Delta t}^T \|\hat{q}_s - q_s^n\|_p. \quad (37)$$

*Proof.* To simplify notation we drop the function arguments and also write  $q_{t+1}, G_{t+1}$ , etc., to mean  $q_{t+\Delta t}$ , etc. in the proof below. By definition of the cashflow function,  $\mathcal{E}_3 := \left\| \tilde{\mathbb{E}}_t[\vartheta_{t+1}(q^n) - \vartheta_{t+1}(\hat{q})] \right\|_p =$

$$\begin{aligned} & \left\| \tilde{\mathbb{E}}_t[G_{t+1}1_{\{q_{t+1}^n \leq G_{t+1}\}} + \vartheta_{t+2}(q^n)1_{\{q_{t+1}^n > G_{t+1}\}} - G_{t+1}1_{\{\hat{q}_{t+1} \leq G_{t+1}\}} - \vartheta_{t+2}(\hat{q})1_{\{\hat{q}_{t+1} > G_{t+1}\}}] \right\|_p \\ &= \left\| \tilde{\mathbb{E}}_t[G_{t+1}(1_{\{G_{t+1} \geq q_{t+1}^n\}} - 1_{\{G_{t+1} \geq \hat{q}_{t+1}\}}) + \vartheta_{t+2}(q^n)1_{\{q_{t+1}^n > G_{t+1}\}} - \vartheta_{t+2}(\hat{q})1_{\{\hat{q}_{t+1} > G_{t+1}\}}] \right\|_p \\ &\leq \left\| \tilde{\mathbb{E}}_t[A_1] \right\|_p + \left\| \tilde{\mathbb{E}}_t[q_{t+1}^n(1_{\{G_{t+1} \geq q_{t+1}^n\}} - 1_{\{G_{t+1} \geq \hat{q}_{t+1}\}}) + \vartheta_{t+2}(q^n)1_{\{G_{t+1} < q_{t+1}^n\}} - \vartheta_{t+2}(\hat{q})1_{\{G_{t+1} < \hat{q}_{t+1}\}}] \right\|_p, \end{aligned}$$

where

$$\begin{aligned} A_1 &= (G_{t+1} - q_{t+1}^n) \cdot (1_{\{G_{t+1} \geq q_{t+1}^n\}} - 1_{\{G_{t+1} \geq \hat{q}_{t+1}\}}) \\ &= (G_{t+1} - q_{t+1}^n) (1_{\{\hat{q}_{t+1} > G_{t+1} \geq q_{t+1}^n\}} - 1_{\{q_{t+1}^n > G_{t+1} \geq \hat{q}_{t+1}\}}) \\ &\leq (\hat{q}_{t+1} - q_{t+1}^n) 1_{\{\hat{q}_{t+1} > G_{t+1} \geq q_{t+1}^n\}} - (\hat{q}_{t+1} - q_{t+1}^n) 1_{\{q_{t+1}^n > G_{t+1} \geq \hat{q}_{t+1}\}} \\ &\leq |\hat{q}_{t+1} - q_{t+1}^n|. \end{aligned}$$

For the remaining terms, using the fact that  $q_{t+1}^n = \tilde{\mathbb{E}}[\vartheta_{t+2}(q^n) | \mathcal{F}_{t+1}^Y]$  we obtain

$$\tilde{\mathbb{E}}_t[q_{t+1}^n(1_{\{G_{t+1} \geq q_{t+1}^n\}} - 1_{\{G_{t+1} \geq \hat{q}_{t+1}\}})] = \tilde{\mathbb{E}}_t[\vartheta_{t+2}(q^n)(1_{\{G_{t+1} \geq q_{t+1}^n\}} - 1_{\{G_{t+1} \geq \hat{q}_{t+1}\}})],$$

and therefore

$$\begin{aligned} \left\| \tilde{\mathbb{E}}_t[\vartheta_{t+1}(q^n) - \vartheta_{t+1}(\hat{q})] \right\|_p &\leq \left\| \tilde{\mathbb{E}}_t[\vartheta_{t+2}(q^n)(1_{\{G_{t+1} < q_{t+1}^n\}} + 1_{\{G_{t+1} \geq q_{t+1}^n\}} - 1_{\{G_{t+1} \geq \hat{q}_{t+1}\}}) \right. \\ &\quad \left. - \vartheta_{t+2}(\hat{q})1_{\{G_{t+1} < \hat{q}_{t+1}\}}] \right\|_p + \left\| \tilde{\mathbb{E}}_t[|\hat{q}_{t+1} - q_{t+1}^n|] \right\|_p \\ &\leq \|\hat{q}_{t+1} - q_{t+1}^n\|_p + \left\| \tilde{\mathbb{E}}_t[(\vartheta_{t+2}(q^n) - \vartheta_{t+2}(\hat{q}))1_{\{G_{t+1} < \hat{q}_{t+1}\}}] \right\|_p \\ &\leq \|\hat{q}_{t+1} - q_{t+1}^n\|_p + \|\vartheta_{t+2}(q^n) - \vartheta_{t+2}(\hat{q})\|_p. \end{aligned}$$

By induction,  $\mathcal{E}_3 \leq \sum_{s=t+1}^T \|\hat{q}_s - q_s^n\|_p$  follows.  $\square$

Based on Lemmas 1-2-3-4, we obtain the main

**Theorem 1.** We have

$$\|\hat{q}_t(Z_t^n) - q_t(Z_t)\|_2 \leq 4^{(T-t)/\Delta t} \max_{t \leq s \leq T} \left\{ \inf_{f \in \mathcal{H}_s} \|f - q_s^n\|_2 + \sqrt{\tilde{\mathbb{E}}[l_s(\hat{q})]} \right\} + \frac{C(T)(T-t)}{\Delta t \cdot \sqrt{n}} \|g\|_{1,\infty}. \quad (38)$$

*Proof.* Combining Lemmas 2-3-4 we find that

$$\|\hat{q}_t - q_t^n\|_2 \leq \sqrt{\tilde{\mathbb{E}}[l_t(\hat{q})]} + \inf_{f \in \mathcal{H}_t} \|f - q_t^n\|_2 + 3 \sum_{s=t+\Delta t}^T \|\hat{q}_s - q_s^n\|_2.$$

Therefore, iterating

$$\begin{aligned} \|\hat{q}_t - q_t^n\|_2 &\leq \sqrt{\tilde{\mathbb{E}}[l_t(\hat{q})]} + \inf_{f \in \mathcal{H}_t} \|f - q_t^n\|_2 + 3 \cdot \left( \sqrt{\tilde{\mathbb{E}}[l_{t+\Delta t}(\hat{q})]} + \inf_{f \in \mathcal{H}_{t+\Delta t}} \|f - q_{t+\Delta t}^n\|_2 \right) \\ &\quad + 9 \cdot \left( \sqrt{\tilde{\mathbb{E}}[l_{t+2\Delta t}(\hat{q})]} + \inf_{f \in \mathcal{H}_{t+2\Delta t}} \|f - q_{t+2\Delta t}^n\|_2 \right) + \dots \\ &\leq 4^{(T-t)/\Delta t} \max_{t \leq s \leq T} \left\{ \sqrt{\tilde{\mathbb{E}}[l_s(\hat{q})]} + \inf_{f \in \mathcal{H}_s} \|f - q_s^n\|_2 \right\}. \end{aligned}$$

Finally, we have  $\|\hat{q}_t - q_t\|_2 \leq \|\hat{q}_t - q_t^n\|_2 + \|q_t^n - q_t\|_2$ , and applying Lemma 1 the result (38) follows.  $\square$

#### 4.1. Convergence

To obtain convergence, one proceeds as follows. First, taking  $n \rightarrow \infty$  eliminates the filtering error so that  $\mathbf{Z}^n \rightarrow \mathbf{Z}$  and the corresponding errors in evaluating  $G$  vanish. Next, one takes  $N \rightarrow \infty$ , reducing the empirical error and the respective centered loss term  $\tilde{\mathbb{E}}[l_t(\hat{q})]$ . Thirdly, one increases the number of basis functions  $r \rightarrow \infty$  in order to eliminate the projection error  $\inf_{f \in \mathcal{H}_t} \|f - q_t^n\|$ . Finally, taking  $\Delta t \rightarrow 0$  we remove the Snell envelope discretization error.

The performed error analysis shows the major trade-off regarding the approximation architectures  $\mathcal{H}_t$ . On the one hand,  $\mathcal{H}_t$  should be large in order to minimize the projection errors  $\min_{f \in \mathcal{H}_t} \|f - q_t^n\|$ . On the other hand,  $\mathcal{H}_t$  should be small to control the empirical variance of the regression coefficients. With many basis functions, one requires a very large number of paths to ensure that  $\hat{q}$  is close to  $q$ . Finally,  $\mathcal{H}_t$  should be smooth in order to further bound the empirical regression errors and the filtering error-in-variable accumulated when computing the regression coefficients.

In the original finite-dimensional study of [13], the size of  $\mathcal{H}_t$  was described in terms of the Vapnik-Cervonenkis (VC) dimensions  $n_{VC}$  and the corresponding covering numbers. Using this theory, [13] showed that overall convergence can be obtained for example by using the polynomial basis for  $\mathcal{H}_t$  and taking the number of paths as  $N = r^{d+2k}$  where  $r$  is the number of basis functions,  $d$  is the dimension of the state variable and  $k$  is the smoothness of the payoff function  $g \in W_p^k$ . In the infinite-dimensional setting of our model, the VC-dimension is meaningless and therefore such estimates do not apply. One could trivially treat  $\rho_t^n$  as an  $n$ -dimensional object, but then the resulting bounds are absurdly poor. It appears difficult to state a useful result on the required relationship between the number of basis functions and the number of paths needed for convergence.

*Remark 4.* A possible alternative is to apply the Tsitsiklis-van Roy algorithm [40], which directly approximates  $q_t$  (rather than  $\vartheta$ ) using the recursion formula (19):  $q_t = \tilde{\mathbb{E}}_t[\max(G(Z_{t+\Delta t}), q(Z_{t+\Delta t}))]$ . Like in Section 3, the approximate algorithm would consist in computing via regression Monte Carlo the empirical continuation value

$$\hat{q}_t = \text{pr}_{\mathcal{H}_t}^N \circ \tilde{\mathbb{E}}_t[\max(G(Z_{t+\Delta t}^n), \hat{q}(Z_{t+\Delta t}^n))].$$

In such a case, the error between  $\hat{q}_t$  and  $q_t$  admits the simpler decomposition (using  $\max(a, b) \leq$

$a + b)$

$$\begin{aligned}
\|\hat{q}_t(Z_t^n) - q_t(Z_t)\|_2 &\leq \left\| \text{pr}_{\mathcal{H}}^N \circ \tilde{\mathbb{E}}_t[\max(G(Z_{t+\Delta t}^n), \hat{q}(Z_{t+\Delta t}^n))] - \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\max(G(Z_{t+\Delta t}^n), \hat{q}(Z_{t+\Delta t}^n))] \right\|_2 \\
&\quad + \left\| \text{pr}_{\mathcal{H}} \circ \tilde{\mathbb{E}}_t[\max(G(Z_{t+\Delta t}^n), \hat{q}(Z_{t+\Delta t}^n))] - \tilde{\mathbb{E}}_t[\max(G(Z_{t+\Delta t}^n), \hat{q}(Z_{t+\Delta t}^n))] \right\|_2 \\
&\quad + \left\| \tilde{\mathbb{E}}_t[G(Z_{t+\Delta t}^n) - G(Z_{t+\Delta t})] \right\|_2 + \left\| \tilde{\mathbb{E}}_t[\hat{q}(Z_{t+\Delta t}^n) - q(Z_{t+\Delta t})] \right\|_2 \\
&\quad + \left\| \tilde{\mathbb{E}}_t[q(Z_{t+\Delta t}^n) - q(Z_{t+\Delta t})] \right\|_2.
\end{aligned} \tag{39}$$

We identify the first two terms as the empirical  $\mathcal{E}_1$  and projection  $\mathcal{E}_2$  errors (as in Lemmas 2 and 3), the third term as the  $G$ -evaluation error, the fourth term as the next-step recursive error, and finally the last term as the sensitivity error of  $q$  with respect to  $Z$ . Controlling the latter error requires understanding the properties of the continuation (or value) function in terms of current state. This seems difficult in our infinite-dimensional setting and is left to future work. Nevertheless, proceeding as in the previous subsection and iterating (39), we obtain for some constants  $C_3, C_4$

$$\|\hat{q}_t - q_t\|_2 \leq \frac{C_3 \cdot (T - t)}{\Delta t} \cdot \max_{t \leq s \leq T} \left\{ \inf_{f \in \mathcal{H}_s} \|f - \hat{q}_s\|_2 + \sqrt{\tilde{\mathbb{E}}[I_s(\hat{q})]} + \frac{C_4}{\sqrt{n}} \|g\|_{1,\infty} + \|q(Z_s^n) - q(Z_s)\|_2 \right\},$$

so that the total error is *linear* rather than exponential in number of steps  $T/\Delta t$  as in Theorem 1. Even though this theoretical result appears to be better, empirical evidence shows that the original algorithm is more stable thanks to its use of  $\vartheta$ .

## 5. Examples

To illustrate the ideas of Section 3 and to benchmark the described algorithm, we consider a model where an explicit finite-dimensional solution is possible. Let

$$\begin{cases} dX_t = -\kappa X_t dt + \sigma_X(\rho dW_t + \sqrt{1 - \rho^2} dU_t); \\ dY_t = (X_t - a) dt + \sigma_Y dW_t, \end{cases} \tag{40}$$

with  $(U, W)$  being two standard independent one-dimensional Brownian motions. Thus,  $Y$  is a linear diffusion with a stochastic, zero-mean-reverting Gaussian drift  $X$ . We study the finite horizon optimal stopping problem of the form

$$V(t, \xi, y) = \sup_{\tau \leq T} \mathbb{E}^{t, \xi, y} [e^{-r\tau} g(X_\tau, Y_\tau)] \triangleq \sup_{\tau \leq T} \mathbb{E}^{t, \xi, y} [e^{-r\tau} (Y_\tau(c_1 + X_\tau) - c_2)_+], \quad c_i \in \mathbb{R}, \tag{41}$$

which can be viewed as an exotic Call option on  $Y$ , see the first example in Section 1.1. Note that the payoff is guaranteed to be non-negative even if the controller stops when  $Y_\tau(c_1 + X_\tau) < c_2$ . In this example, under the reference measure  $\tilde{\mathbb{P}}$ , we have

$$\begin{cases} dY_t = \sigma_Y d\bar{W}_t; \\ dX_t = [-\kappa X_t - \rho(\sigma_X/\sigma_Y) \cdot (X_t - a)] dt + \rho\sigma_X d\bar{W}_t + \sqrt{1 - \rho^2} \sigma_X dW_t^\perp; \\ d\rho_t(x) = \frac{1}{2} \sigma_X^2 \rho_t''(x) + \kappa x \rho_t'(x) + \kappa \rho_t(x) + \left[ \frac{x - a}{\sigma_Y^2} - \frac{\rho\sigma_X}{\sigma_Y} \right] dY_t, \end{cases}$$

where  $W^\perp$  is a  $\tilde{\mathbb{P}}$ -Wiener process independent of  $\bar{W}$ .

Below we carry out a numerical study with parameter values taken as

Parameter	$\kappa$	$a$	$\sigma_Y$	$\sigma_X$	$T$	$r$	$\rho$	$c_1$	$c_2$
Value	2	0.05	0.1	0.3	1	0.1	0.6	1	2

Since on average  $X_t$  is around  $0 < a$ ,  $Y$  tends to decrease, so that in (41) it is optimal to stop early. However, the drift process  $X$  is highly volatile and quite often  $X_t > a$  produces positive drift for  $Y$ , in which case one should wait. Consequently, the stopping region will be highly sensitive to the conditional distribution  $\pi_t$ .

### 5.1. Kalman Filter Formulation

The model (40) also fits into the Kalman-Bucy [21] filter framework. Thus, if the initial distribution  $X_0 \sim \mathcal{N}(m_0, P_0)$  is a Gaussian density, then  $X_t | \mathcal{F}_t^Y \sim \mathcal{N}(m_t, P_t)$  is conditionally Gaussian, where

$$\begin{cases} dm_t = -\kappa m_t dt + (\rho\sigma_X + P_t/\sigma_Y) d\bar{W}_t, & d\bar{W}_t = \frac{dY_t - (m_t - a) dt}{\sigma_Y}, \\ dP_t = (-2\kappa P_t + \sigma_X^2 - (\rho\sigma_X + P_t/\sigma_Y)^2) dt. \end{cases} \quad (42)$$

Note that the conditional variance  $P_t$  is deterministic and solves the Riccati ode on the second line of (42). In (42),  $\bar{W}$  is a  $\mathbb{P}$ -Brownian motion, the so-called innovation process. Moreover, as shown by [25, Section 12.1],  $\mathcal{F}_t^Y = \mathcal{F}_t^{\bar{W}, \xi_0}$ , so that we may equivalently write

$$dY_t = (m_t - a) dt + \sigma_Y d\bar{W}_t.$$

The pair  $(m_t, P_t)$  are sufficient statistics for the conditional distribution of  $X_t | \mathcal{F}_t^Y$  and the corresponding payoff can be computed as

$$\begin{aligned} \mathbb{E}[g(X_t, Y_t) | \mathcal{F}_t^Y] &= \mathbb{E}\left[(y(c_1 + m_t + \sqrt{P_t}X) - c_2)_+\right], & \text{where } X &\sim \mathcal{N}(0, 1) \\ &= \int_{x^*}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \left\{((c_1 + m_t)y - c_2) + y\sqrt{P_t}x\right\} dx \\ &= \frac{y\sqrt{P_t}}{\sqrt{2\pi}} \cdot e^{-(x^*)^2/2} + ((c_1 + m_t)y - c_2) \cdot (1 - \Phi(x^*)) =: G(m_t, P_t, Y_t), \end{aligned}$$

where  $x^* = \frac{c_2 - (c_1 + m_t)y}{y\sqrt{P_t}}$ , and  $\Phi(x)$  is the standard normal cumulative distribution function. Thus, the original problem is reduced to

$$V(t, m, p, y) = \sup_{\tau \leq T} \mathbb{E}\left[e^{-r\tau} G(m_\tau, P_\tau, Y_\tau) \mid m_0 = m, P_0 = p, Y_0 = y\right]. \quad (43)$$

This two-dimensional problem (recall that  $(P_t)$  is deterministic) can be solved numerically using a pde solver applied to the corresponding version of the free boundary problem (22). Namely  $V$  of (43) is characterized by the quasi-variational inequality

$$\begin{cases} \max\left\{V_t + (m - a)V_y + \frac{1}{2}\sigma_Y^2 V_{yy} - \kappa m V_m + \frac{1}{2}(\rho\sigma_X + P_t/\sigma_Y)^2 V_{mm} \right. \\ \quad \left. + (\rho\sigma_X\sigma_Y + P_t)V_{my} - rV, G(m, p, y) - V(t, m, p, y)\right\} = 0, \\ V(T, m, p, y) = G(m, p, y). \end{cases} \quad (44)$$

## 5.2. Numerical Results

To benchmark the proposed algorithm we proceed to compare two solutions of (41), namely (i) a simulation algorithm of Section 3.3 and (ii) a finite-differences pde solver of (44). The Monte Carlo implementation used  $N = 30\,000$  paths with  $n = 500$ ,  $\delta = 0.01$ ,  $\Delta t = 0.05$  or twenty time-steps. For basis functions we used the set  $\{1, y, y^2, \rho_t x, \rho_t g, \rho_t EUR\}$ , where  $EUR(t, \xi, y) = \tilde{\mathbb{E}}^{t, \xi, y}[e^{-r(T-t)} g(X_T, Y_T)]$  is the conditional expectation of terminal payoff. A straightforward code written in Matlab with minimal optimization took about three minutes to run on a desktop PC. The pde solver utilized a basic explicit scheme and used a  $400 \times 400$  grid with 8000 timesteps. In order to allow a fair comparison, the pde solver also allowed only  $T/\Delta t = 20$  exercise opportunities by enforcing the barrier condition  $V(t, m, p, y) \geq G(m, p, y)$  only for  $t = m\Delta t$ ,  $m = 0, 1, \dots, 20$ . In financial lingo, we thus studied the Bermudan variant of (41) with  $\Delta t = 0.05$ .

The obtained results are summarized in Table 1 for a variety of initial conditions  $(\xi_0, Y_0)$ . Using the pde solver as a proxy for the true answer, we find that our algorithm was generally within 2% of the correct value which is acceptable performance. Interestingly, our algorithm performed worst for “in-the-money” options (such as when  $Y_0 = 1.8$ ,  $X_0 \sim \mathcal{N}(0.2, 0.05^2)$ ), i.e. when it is optimal to stop early. As expected, our method produced an underestimate of true  $V$  since the computed stopping rule is necessarily sub-optimal. We found that the distribution of the computed  $\tau^*$  was quite uniform on  $\{\Delta t, 2\Delta t, \dots, (M-1)\Delta t\}$  showing that this was a nontrivial stopping problem. For comparison, Table 1 also lists the European option price assuming that early exercise is no longer possible. This column shows that our algorithm captured about 85-90% of the *time value of money*, i.e. the extra benefit due to early stopping.

To further illustrate the structure of the solution, Figure 1 compares the optimal stopping regions computed by each algorithm at a fixed time point  $t = 0.5$ . Note that since the value function  $V$  is typically not very sensitive to the choice of a stopping rule, direct comparison of optimal stopping regions is more relevant (and more important for a practicing controller). As we can see, an excellent fit is obtained through our non-parametric method. Figure 1 also reveals that both  $\{q_t > G_t\} \cap \{\hat{q}_t \leq G_t\}$  and  $\{q_t \leq G_t\} \cap \{\hat{q}_t > G_t\}$  are non-empty (in other words, sometimes our algorithm stops too early; sometimes it stops too late). Recall that the simulation solver works under  $\tilde{\mathbb{P}}$  and therefore the empirical distribution of  $Y_t$  in Figure 1 is different from the actual realizations under  $\mathbb{P}$  that would be observed by the controller.

While the pde formulation (42)-(44) is certainly better for the basic example above, it is crucially limited in its applicability. For instance, (42) assumes Gaussian initial condition; any other  $\xi_0$  renders it invalid. Similarly, perturbations to the dynamics (40) will at the very least require re-derivation of (42)-(44), or more typically lead to the case where no finite-dimensional sufficient statistics of  $X_t | \mathcal{F}_t^Y$  exist. In stark contrast to such difficulties, the particle filter algorithm can be used without any modifications for any  $\xi_0$ , and would need only minor adjustments to accommodate a different version of (40). A simple illustration is shown in the last two rows of Table 1 where we consider a uniform and a discrete initial distribution, respectively. Heuristically, for a given mean and variance,  $V$  should be increasing with respect to the kurtosis of  $\xi_0$ , as a more spread-out initial distribution of  $X_t$  leads to more optionality. Indeed, as confirmed by Table 1,  $V(0, \xi^1, y_0) < V(0, \xi^2, y_0) < V(0, \xi^3, y_0)$ , where  $\xi^1 = \mathcal{N}(0, 0.05^2)$ ,  $\xi^2 = \text{Unif}[-0.05\sqrt{3}, 0.05\sqrt{3}]$ ,  $\xi^3 = 0.5(\delta_{-0.05} + \delta_{0.05})$  are three initial distributions of  $X$  normalized to  $\int_{\mathbb{R}} x \xi^i(dx) = 0$ ,  $\int_{\mathbb{R}} x^2 \xi^i(dx) = 0.05^2$ .

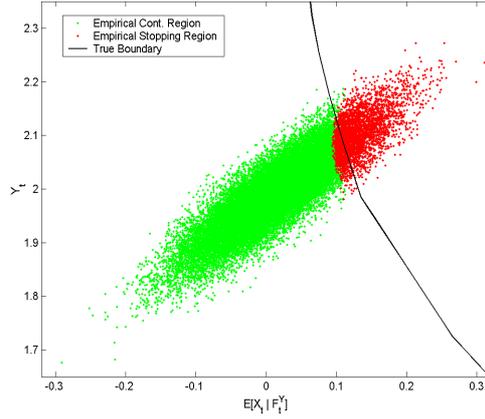


Figure 1: Comparison of optimal stopping regions for the pde and Monte Carlo solvers. The solid line shows the optimal stopping boundary as a function of  $m_t$ ; the color-coded points show the values  $z_j^t(t)$ ,  $j = 1, \dots, N$  projected onto  $\mathbb{E}[X_t | \mathcal{F}_t^Y] = (\rho_t x) \cdot (\rho_t 1)^{-1}$ . Here  $X_0 \sim \mathcal{N}(0, 0.05^2)$ ,  $Y_0 = 2$ ,  $N = 30\,000$  and  $t = 0.5$ .

$\xi_0$	$y_0$	Simulation solver	pde solver	European option
$\mathcal{N}(0, 0.05^2)$	2	0.1810	0.1853	0.1331
$\mathcal{N}(-0.12, 0.05^2)$	2.24	0.2566	0.2661	0.2136
$\mathcal{N}(0.2, 0.05^2)$	1.8	0.1862	0.1904	0.1052
$\mathcal{N}(0, 0.1^2)$	2	0.1852	0.1919	0.1349
$\delta_0$	2	0.1723	0.1832	0.1325
$Unif_{[-0.05\sqrt{3}, 0.05\sqrt{3}]}$	2	0.1827	--	0.1347
$0.5(\delta_{-0.05} + \delta_{0.05})$	2	0.1853	--	0.1332

Table 1: Comparison of the Monte Carlo scheme of Section 3.3 versus the Bermudan pde solver for the stochastic drift example of Section 5. Standard error of the Monte Carlo solver was about 0.001.

## 6. American Option Pricing under Stochastic Volatility

Our method can also be applied to *stochastic volatility* models. Such asset pricing models are widely used in financial mathematics to represent stock dynamics and assume that the local volatility of the underlying stock is itself stochastic. While under continuous observations the local volatility is perfectly known through the quadratic variation process, under discrete observations this leads to a partially observed model similar to (3).

To be concrete, let  $Y_t$  represent the log-price of a stock at time  $t$  under the given (pricing) measure  $\mathbb{P}$ , and let  $X_t$  be the instantaneous volatility of  $Y$  at time  $t$ . We postulate that  $(X, Y)$  satisfy the following system of sde's (known as the Stein-Stein model),

$$\begin{cases} dY_t = (r - \frac{1}{2}X_t^2) dt + X_t dU_t, \\ dX_t = \kappa(\bar{\sigma} - X_t) dt + \rho\alpha dU_t + \sqrt{1 - \rho^2}\alpha dW_t. \end{cases} \quad (45)$$

The stock price  $Y$  is only observed at the discrete time instances  $\tilde{T} = \{\Delta t, 2\Delta t, \dots\}$  with  $\tilde{\mathcal{F}}_t^Y = \sigma(Y_0, Y_{\Delta t}, \dots, Y_{\lfloor t/\Delta t \rfloor \Delta t})$ . The American (Put) option pricing problem consists in finding the optimal  $\tilde{\mathcal{F}}^Y$ -adapted and  $\tilde{T}$ -valued stopping time  $\tau$  for

$$\sup_{\tau \in \tilde{T}} \mathbb{E}[e^{-r\tau}(K - e^{Y_\tau})_+]. \quad (46)$$

A variant of (45)-(46) was recently studied by Sellami et al. [36]. More precisely, [36] considered the American option pricing model in a simplified discrete setting where  $(X_t)$  of the Stein-Stein model (45) was replaced with a corresponding 3-state Markov chain approximation. In a related vein, Viens et al. [41] considered the filtering and portfolio optimization problem where the second line of (45) was replaced with the Heston model

$$d(X_t^2) = \kappa(\bar{\sigma} - X_t^2) dt + \rho\alpha X_t dU_t + \sqrt{1 - \rho^2}\alpha X_t dW_t. \quad (47)$$

In general, the problem of estimation of  $X_t$  is well-known, see e.g. [10, 14, 39]. Observe that while (45) is linear, the square-root dynamics in (47) are highly non-linear and no finite-dimensional sufficient statistics exist for  $\pi_t$  in the latter case.

In the presence of stochastic volatility, one may no longer use the reference probability measure  $\tilde{\mathbb{P}}$ . Indeed, there is no way to obtain a Brownian motion from the observation process  $Y$  whose increments are now tied with the values of the unobserved  $X$ . Accordingly,  $\zeta$  is no longer defined and consequently we cannot use it as an importance weight during the particle branching step in (24).

A way out of this difficulty is provided by Del Moral et al. [12]. The idea is to propagate particles independently of observations and to compute a *candidate observation* for each propagated particle. The weights are then assigned by comparing the candidates with the actual observation. Let  $\phi$  be a smooth bounded function with  $\int_{\mathbb{R}} \phi(x) dx = 1$  and  $\int_{\mathbb{R}} |x|\phi(x) dx < \infty$  (e.g.  $\phi(x) = \exp(-x^2/2) \cdot (2\pi)^{-1/2}$ ). The propagated particles and candidates are obtained by

$$\begin{cases} v_j^n(t) = v_j^n(m\Delta t) + \int_{m\Delta t}^t \kappa(\bar{\sigma} - v_j^n(s)) ds + \int_{m\Delta t}^t \rho\alpha dU_s^{(j)} + \int_{m\Delta t}^t \sqrt{1 - \rho^2}\alpha dW_s^{(j)}, \\ Y^{(j)}(t) = \int_{m\Delta t}^t (r - \frac{1}{2}(v_j^n(s))^2) ds + \int_{m\Delta t}^t v_j^n(s) dU_s^{(j)}, \quad m\Delta t \leq t \leq (m+1)\Delta t, \end{cases}$$

where  $(U^{(j)}, W^{(j)})_{j=1}^n$  are  $n$  independent copies of bivariate Wiener processes. The branching weights are then given by

$$\bar{a}_j^n((m+1)\Delta t) = n^{1/3} \phi \left( n^{1/3} (Y_{(m+1)\Delta t}^{(j)} - Y_{(m+1)\Delta t}) \right). \quad (48)$$

Hence, particles whose candidates  $Y_{(m+1)\Delta t}^{(j)}$  are close to the true observed  $Y_{(m+1)\Delta t}$  get high weights, while those particles that produced poor candidates are likely to be killed off. The rest of the algorithm remains the same as in Section 3.1. As shown in [12, Theorem 5.1], the resulting filter satisfies for any bounded payoff function  $f \in C_b^0(\mathbb{R}^d)$

$$\mathbb{E}[|\pi_t^n f - \pi_t f|] \leq \frac{C(t)}{n^{1/3}} \|f\|_{0,\infty}, \quad \text{with } C(t) = \mathcal{O}(e^t). \quad (49)$$

Note that compared to (28), the error in (49) as a function of number of particles  $n$  is worse. This is due to the higher re-sampling variance produced by the additional randomness in  $Y^{(j)}$ 's.

### 6.1. Numerical Example

Recently [36] considered the above model (45) with the parameter values

Parameter	$Y_0$	$X_0$	$K$	$\kappa$	$\bar{\sigma}$	$\alpha$	$T$	$r$	$\rho$
Value	110	0.15	100	1	0.15	0.1	1	0.05	0

Plugging-in the above parameters and using the modification (48), we implemented our algorithm with  $N = 30,000$ ,  $n = 1000$ . Since no other solver of (46) is available, as in [36] we compare the Monte Carlo solver of the partially-observed problem to a pde solver for the fully observed case (in which case the Bermudan option price is easily computed using the quasi-variational formulation based directly on (45)). Table 2 shows the results as we vary the observation frequency  $\Delta t$ . Since  $\Delta t$  is also the frequency of the stopping decisions, smaller  $\Delta t$  increases both the partially and fully observed value functions. Moreover, as  $\Delta t$  gets smaller, the information set becomes richer and the handicap of partial information vanishes.

In this example where the payoff  $K - \exp(Y_t)$  is a function of the observable  $Y$  only, our algorithm obtains excellent performance. Also, we see that partial information has apparently only a mild effect on potential earnings (difference of less than 1.5% even if  $Y$  is observed just five times). To give an idea of the corresponding time value of money, the European option price in this example was 1.570. Comparison with the results obtained in [36] (first two columns of Table 2) is complicated because the latter paper immediately discretizes  $X$  and constructs a three-state Markov chain  $(\tilde{X}_t)$ . This discrete version takes on the values  $\tilde{X}_{m\Delta t} \in \{0.1, 0.15, 0.2\}$  and therefore does not exhibit the asymmetric behavior of very small  $X_t$ -realizations that dampen the volatility of  $Y$  and drastically reduce Put profits. In contrast, our algorithm operates on the original continuous-state formulation in (45). Consequently, as can be seen in Table 2, the full observation prices of the two models are quite different.

## 7. Conclusion

In this paper we have presented a new numerical scheme to solve partially observable optimal stopping problems. Our method is entirely simulation-based and only requires the ability to simulate the state processes. Consequently, we believe it is more robust than other proposals in the existing literature.

$\Delta t$	Discrete Model		Continuous Model	
	Full Obs.	Partial Obs.	Full Obs.	Partial Obs.
0.2	1.575	0.988	1.665	1.646
0.1	1.726	1.306	1.686	1.673
0.05	1.912	1.596	1.696	1.685

Table 2: Comparison of discrete and continuous models for (45) under full and partial observations. The first two columns are reproduced from [36, Table 3].

While our analysis was stated in the most simple setting of multi-dimensional diffusions, it can be considerably extended. First, as explained in Section 6, our algorithm can be easily adjusted to take into account *discrete* observations which is often the more realistic setup. Second, the assumption of diffusion state processes is not necessary from a numerical point of view; one may consider other cases such as models with jumps, or even discrete-time formulations given in terms of general transition semigroups. For an example using a particle filter to filter a stable Lévy process  $X$ , see [22]. Third, one may straightforwardly incorporate *state* constraints on the unobserved factor  $X$ . For instance, some applications imply that  $X_t \geq 0$  is an extra constraint on top of (1) (in other words the observable filtration is generated by  $Y$  and  $1_{\{X_t \geq 0\}}$ ). Such a restriction can be added by assigning zero weights to particles that violate state constraints so that they are not propagated during the next branching step. Finally, if one uses the modification (48) from [12] then many other noise formulations can be chosen beyond (2).

## Acknowledgment

I thank René Carmona for early advice and encouragement of this paper.

## References

- [1] V. E. Beneš. Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics*, 5(1-2):65–92, 1981.
- [2] A. Bensoussan. *Stochastic control of partially observable systems*. Cambridge University Press, Cambridge, 1992.
- [3] M. W. Brandt, A. Goyal, P. Santa-Clara, and J. R. Stroud. A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *Review of Financial Studies*, 18:831–873, 2005.
- [4] D. Brigo, B. Hanzon, and F. Le Gland. Approximate nonlinear filtering by projection on exponential manifolds of densities. *Bernoulli*, 5(3):495–534, 1999.
- [5] D. Brigo, B. Hanzon, and F. LeGland. A differential geometric approach to nonlinear filtering: the projection filter. *IEEE Trans. Automat. Control*, 43(2):247–252, 1998.
- [6] R. Carmona and M. Ludkovski. Spot convenience yield models for the energy markets. In *Mathematics of finance*, volume 351 of *Contemp. Math.*, pages 65–79. Amer. Math. Soc., Providence, RI, 2004.
- [7] D. Crisan. Particle approximations for a class of stochastic partial differential equations. *Appl. Math. Optim.*, 54(3):293–314, 2006.
- [8] D. Crisan, J. Gaines, and T. Lyons. Convergence of a branching particle method to the solution of the Zakai equation. *SIAM J. Appl. Math.*, 58(5):1568–1590 (electronic), 1998.
- [9] D. Crisan and T. Lyons. A particle approximation of the solution of the Kushner-Stratonovitch equation. *Probab. Theory Related Fields*, 115(4):549–578, 1999.
- [10] J. Cvitanić, R. Liptser, and B. Rozovskii. A filtering approach to tracking volatility from prices observed at random times. *Ann. Appl. Probab.*, 16(3):1633–1652, 2006.
- [11] J.-P. Décamps, T. Mariotti, and S. Villeneuve. Investment timing under incomplete information. *Math. Oper. Res.*, 30(2):472–500, 2005.
- [12] P. Del Moral, J. Jacod, and P. Protter. The Monte-Carlo method for filtering with discrete-time observations. *Probab. Theory Related Fields*, 120(3):346–368, 2001.

- [13] D. Egloff. Monte Carlo algorithms for optimal stopping and statistical learning. *Ann. Appl. Probab.*, 15(2):1396–1432, 2005.
- [14] R. Frey and W. J. Runggaldier. A nonlinear filtering approach to volatility estimation with a view towards high frequency data.
- [15] W. A. Fuller. *Measurement error models*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1987.
- [16] D. Gątarek and A. Świąch. Optimal stopping in Hilbert spaces and pricing of American options. *Math. Methods Oper. Res.*, 50(1):135–147, 1999.
- [17] F. Gozzi, E. Rouy, and A. Świąch. Second order Hamilton-Jacobi equations in Hilbert spaces and stochastic boundary control. *SIAM J. Control Optim.*, 38(2):400–430 (electronic), 2000.
- [18] K. Ito and K. Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Trans. Automat. Control*, 45(5):910–927, 2000.
- [19] U. Jensen and G.-H. Hsu. Optimal stopping by means of point process observations with applications in reliability. *Mathematics of Operations Research*, 18(3):645–657, 1993.
- [20] G. Kallianpur. *Stochastic filtering theory*, volume 13 of *Applications of Mathematics*. Springer-Verlag, New York, 1980.
- [21] R. Kalman and R. Bucy. New results in linear filtering and prediction theory. *Journal of Basic Engineering*, 83D:95–108, 1961.
- [22] M. A. Kouritzin and W. Sun. Rates for branching particle approximations of continuous-discrete filters. *Ann. Appl. Probab.*, 15(4):2739–2772, 2005.
- [23] H. J. Kushner and A. S. Budhiraja. A nonlinear filtering algorithm based on an approximation of the conditional distribution. *IEEE Trans. Automat. Control*, 45(3):580–585, 2000.
- [24] F. Le Gland and N. Oudjane. Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters. *Ann. Appl. Probab.*, 14(1):144–187, 2004.
- [25] R. S. Liptser and A. N. Shiryaev. *Statistics of random processes. II*, volume 6 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, expanded edition, 2001.
- [26] F. Longstaff and E. Schwartz. Valuing American options by simulations: a simple least squares approach. *Rev. Finan. Studies*, 14:113–148, 2001.
- [27] S. Lototsky, R. Mikulevicius, and B. L. Rozovskii. Nonlinear filtering revisited: a spectral approach. *SIAM J. Control Optim.*, 35(2):435–461, 1997.
- [28] S. V. Lototsky. Wiener chaos and nonlinear filtering. *Appl. Math. Optim.*, 54(3):265–291, 2006.
- [29] G. Mazziotto. Approximations of the optimal stopping problem in partial observation. *J. Appl. Probab.*, 23(2):341–354, 1986.
- [30] G. Mazziotto, L. Stettner, J. Szpirglas, and J. Zabczyk. On impulse control with partial observation. *SIAM J. Control Optim.*, 26(4):964–984, 1988.
- [31] J. Miao and N. Wang. Experimentation under uninsurable idiosyncratic risk: An application to entrepreneurial survival,. Technical report.
- [32] R. Mikulevicius and B. L. Rozovskii. Fourier-Hermite expansions for nonlinear filtering. *Teor. Veroyatnost. i Primenen.*, 44(3):675–680, 1999.
- [33] G. Moscarini and L. Smith. The optimal level of experimentation. *Econometrica*, 69(6):1629–1644, 2001.
- [34] P. Muller, B. Sanso, and M. De Iorio. Optimal Bayesian design by inhomogeneous Markov chain simulation. *Journal of the American Statistical Association*, 99:788–798, 2004.
- [35] G. Pagès and H. Pham. Optimal quantization methods for nonlinear filtering with discrete-time observations. *Bernoulli*, 11(5):893–932, 2005.
- [36] H. Pham, W. Runggaldier, and A. Sellami. Approximation by quantization of the filter process and applications to optimal stopping problems under partial observation. *Monte Carlo Methods Appl.*, 11(1):57–81, 2005.
- [37] E. Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance*, LII(3):922–973, 1997.
- [38] A. N. Shiryaev. *Optimal stopping rules*. Springer-Verlag, Berlin, 1978.
- [39] J. R. Stroud, N. G. Polson, and P. Müller. Practical filtering for stochastic volatility models. In *State space and unobserved component models*, pages 236–247. Cambridge Univ. Press, Cambridge, 2004.
- [40] J. N. Tsitsiklis and B. Van Roy. Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives. *IEEE Trans. Automat. Control*, 44(10):1840–1851, 1999.
- [41] F. Viens, R. Desai, and T. Lele. A Monte-Carlo method for portfolio optimization under partially observed stochastic volatility. In *IEEE International Conference on Computational Intelligence for Financial Engineering, 2003*, pages 257–263, 2003.