A DYNAMIC LOOK-AHEAD MONTE CARLO ALGORITHM FOR PRICING AMERICAN OPTIONS

DANIEL EGLOFF, MICHAEL KOHLER, AND NEBOJSA TODOROVIC

ABSTRACT. Pricing of American options can be achieved by solving optimal stopping problems. This in turn can be done by computing so-called continuation values, which we represent as regression functions defined by the aid of a cash flow for the next few time periods. We use Monte Carlo to generate data and apply nonparametric least squares regression estimates to estimate the continuation values from these data. The parameters of the regression estimates and of the underlying regression problems are chosen data-dependent. Results concerning consistency and rate of convergence of these estimates are presented, and the resulting pricing of American options is illustrated by the aid of simulated data.

1. INTRODUCTION

Many financial contracts allow for early exercise before expiry. Most of the exchange traded option contracts are of American type which allows the holder to choose any exercise date before expiry, or Bermudan with exercise dates restricted to a predefined discrete set of dates. Mortgages have embedded prepayment options such that the mortgage can be amortized or repayed. Also life insurance contracts may allow for early surrender. In this paper we are interested in pricing options with early exercise features. For simplicity we restrict ourselves to Bermudan options, which can be considered as a discrete time approximation of American options. It is well-known that in complete and arbitrage free markets the price of a derivative security can be represented as an expected value with respect to the so called martingale measure, see for instance [17]. More generally the price of a Bermudan option can be represented as an optimal stopping problem

$$V_0 = \sup_{\tau \in \mathcal{T}(0,\dots,T)} \mathbf{E}\left\{d_{0,\tau} f_\tau(X_\tau)\right\}, \tag{1.1}$$

where f_t is the payoff function, $X_0, X_1, \dots, X_T$ is the underlying stochastic process, $\mathcal{T}(0,\dots,T)$ is the class of all $\{0,\dots,T\}$–stopping times, and $d_{s,t}$ are nonnegative $\mathcal{F}(X_s,\dots,X_t)$–measurable discount factors satisfying $d_{0,t} = d_{0,s} \cdot d_{s,t}$ for $s < t$. In practice, the stochastic process $X_0, X_1, \dots, X_T$ might be, e.g., determined by Black Scholes model or by nonparametric estimation of a time series from observed data. In the sequel we assume that $X_0, X_1, \dots, X_T$ is a $[-A, A]^d$–valued Markov process recording all necessary information about financial variables including prices of the underlying assets as well as additional risk factors driving stochastic volatility or stochastic interest rates. Neither the Markov property nor the form of the payoff as a function of the state of X_t is restrictive and can always be achieved by including supplementary variables. For instance in case of an Asian option we add the running mean as an additional variable into X_t. Usually in Black Scholes models or nonparametric estimation of time series from observed data the underlying stochastic process will be modelled by an unbounded stochastic process. If the Markov process X_t is not localized to the bounded set $[-A, A]^d$ we replace it with the process $X_t^A = X_{t\wedge\tau_A}$ killed at first exit from $[-A, A]^d$, where $\tau_A = \inf\{s \geq 0 \mid X_t \notin [-A, A]^d\}$. It can

2000 *Mathematics Subject Classification.* Primary 91B28, 60G40, 93E20; Secondary 65C05, 93E24, 62G05.
Key words and phrases. Optimal stopping, American options, nonparametric regression, Monte Carlo methods.

be shown that the Markov semigroup of X_t^A converges for $A \to \infty$ to the one of X_t in a suitable sense. We refer to [12] for details. Moreover, the corresponding solutions of the optimal stopping problems converge as well, see [19].

The boundedness assumption $X_t \in [-A, A]^d$ enables us to estimate the price of the American option from samples of polynomial size in the number of free parameters, in contrast to Monte Carlo estimation from standard Black Scholes models, where Glasserman and Yu [14] showed that samples of exponential size in the number of free parameters are needed.

The computation of (1.1) can be done via determination of an optimal stopping rule $\tau^* \in \mathcal{T}(0, \ldots, T)$ satisfying

$$V_0 = \mathbf{E}\left\{d_{0,\tau^*} f_{\tau^*}(X_{\tau^*})\right\}. \tag{1.2}$$

Let

$$q_t(x) = \sup_{\tau \in \mathcal{T}(t+1,\ldots,T)} \mathbf{E}\left\{d_{t,\tau} f_\tau(X_\tau) | X_t = x\right\} \tag{1.3}$$

be the so-called continuation value describing the value of the option at time t given $X_t = x$ and subject to the constraint of holding the option at time t rather than exercising it. The general theory of optimal stopping for Markov processes, see for instance [4, 22, 26, 11], implies that

$$\tau^* = \inf\{s \geq 1 : q_s(X_s) \leq f_s(X_s)\}$$

is an optimal stopping time, i.e., τ^* satisfies (1.2). Therefore, computing the continuation values (1.3) solves the optimal stopping problem (1.1).

Explicit solutions of (1.1) do not exist, except in very rare cases, but there are a variety of numerical procedures to solve optimal stopping problems, each with its strength and weaknesses. In this paper we study a concrete simulation algorithm. The first attempts to use simulation are [1, 28, 2]. Longstaff and Schwartz [21] introduce a new algorithm for Bermudan options in discrete time. It combines Monte Carlo simulation with multivariate function approximation. Tsitsiklis and Van Roy [29] independently propose an alternative parametric approximation algorithm using stochastic approximation to derive the weights of the approximation. Both algorithms approximate the value function or the early exercise rule and therefore provide a lower bound for the true optimal stopping value. Upper bounds based on the dual problem are derived in [24, 16]. More details and further references can be found in [3] and [13]. Also, the article [20] compares several Monte Carlo approaches empirically.

In this paper we enhance the approach of [21] and its generalization presented in [10]. We construct estimates $\hat{q}_t$ of q_t, approximate the optimal stopping rule τ^* by

$$\hat{\tau} = \inf\{s \geq 1 : \hat{q}_s(X_s) \leq f_s(X_s) \tag{1.4}$$

and estimate the price V_0 of the American option by the Monte Carlo estimate of

$$\mathbf{E}\{d_{0,\hat{\tau}} f_{\hat{\tau}}(X_{\hat{\tau}})\}. \tag{1.5}$$

To estimate q_t, we represent q_t as a regression function of a distribution (X_t, Y_t), where Y_t depends on the partial sample path $X_{t+1}, \ldots, X_{t+w+1}$ and $q_{t+1}, \ldots, q_{t+w+1}$ for some tunable parameter $w \in \{0, 1, \ldots, T-t-1\}$. This distribution will in turn be approximated by $(X_t, \hat{Y}_t)$, where $\hat{Y}_t$ depends on $X_{t+1}, \ldots, X_{t+w+1}$ and $\hat{q}_{t+1}, \ldots, \hat{q}_{t+w+1}$. We construct an estimate $\hat{q}_t$ of q_t with nonparametric regression techniques applied to a Monte Carlo sample of the distribution $(X_t, \hat{Y}_t)$ and use this estimate together with $\hat{q}_{t+1}, \ldots, \hat{q}_{t+w}$ to compute recursively estimates of $q_{t-1}, \ldots, q_0$. All parameters of the estimates and the parameter w of the distribution of (X_t, Y_t) are chosen data dependent. We present results concerning consistency and rate of convergence of the estimates, and illustrate them by the aid of simulated data.

In Section 2 we describe in detail the connection between discrete time optimal stopping problems and recursive regression. The dynamic look-ahead Monte Carlo algorithm for solving optimal stopping problems is introduced in Section 3. The main theoretical results are presented in Section 4, and the finite sample properties of the proposed algorithm are illustrated in Section 5 by the aid of simulated data. Section 6 contains the proofs.

2. Discrete Time Optimal Stopping and Recursive Regression

Let $\mathbf{X} = (X_t)_{t=0,\ldots,T}$ be a discrete time $[-A, A]^d$-valued Markov process, μ_t the law induced by X_t on $\mathbb{R}^d$, and $\mathbb{F} = (\mathcal{F}_t)$ be the induced filtration where

$$\mathcal{F}_t = \mathcal{F}(X_0, \ldots, X_t) = \bigvee_{s \leq t} \sigma(X_s), \tag{2.1}$$

is the sigma algebra generated by the random variables $\{X_s \mid s \leq t\}$. The solution of the discrete time optimal stopping problem for nonnegative reward or payoff functions f_t is given by the value function

$$v_t(x) = \sup\nolimits_{\tau \in \mathcal{T}(t,\ldots,T)} E\big[f_\tau(X_\tau) \mid X_t = x\big]. \tag{2.2}$$

The supremum runs over the class $\mathcal{T}(t, \ldots, T)$ of all $\mathbb{F}$-stopping times with values in $\{t, \ldots, T\}$. By definition, each $\tau \in \mathcal{T}(t, \ldots, T)$ satisfies $\{\tau = k\} \in \mathcal{F}(X_0, \ldots, X_k)$ for $k \in \{t, \ldots, T\}$. Here and in the sequel we assume for notational simplicity that f_t contains already the discount factor occurring in (1.1). Once the value function has been determined, the smallest optimal stopping time can be derived as

$$\tau_t^* = \inf\{s \geq t \mid v_s(X_s) \leq f_s(X_s)\}. \tag{2.3}$$

The optimal stopping problem can also be characterized in terms of the so called continuation value, which is given by

$$q_t(x) = \sup\nolimits_{\tau \in \mathcal{T}(t+1,\ldots,T)} E\big[f_\tau(X_\tau) \mid X_t = x\big] = E\big[f_{\tau_{t+1}^*}(X_{\tau_{t+1}^*}) \mid X_t = x\big]. \tag{2.4}$$

The value function and the continuation value are related by

$$v_t(X_t) = \max\big(f_t(X_t), q_t(X_t)\big), \quad q_t(X_t) = E\big[v_{t+1}(X_{t+1}) \mid X_t\big]. \tag{2.5}$$

From now on we primarily consider q_t. The continuation value satisfies the dynamic programming equations

$$\begin{aligned} q_T(x) &= 0, \\ q_t(x) &= E\big[\max(f_{t+1}(X_{t+1}), q_{t+1}(X_{t+1})) \mid X_t = x\big]. \end{aligned} \tag{2.6}$$

The recursion for the optimal stopping rules is given by

$$\begin{aligned} \tau_T^* &= T, \\ \tau_t^* &= t\, 1_{\{q_t(X_t) \leq f_t(X_t)\}} + \tau_{t+1}^* 1_{\{q_t(X_t) > f_t(X_t)\}}. \end{aligned} \tag{2.7}$$

The dynamic programming equations (2.6) show that the optimal stopping problem in discrete time is essentially equivalent to a series of regression problems. Equation (2.4) provides a different regression representation of the continuation value, once the optimal stopping rule of the next future period is known. These representations are in a sense extreme cases as we will explain in the following. Let $h_t \in L_1(\mu_t)$ with $h_T = f_T$ and introduce the indicator functions

$$\theta_{f,t}(h) = \theta(f_t - h_t) = 1_{\{f_t - h_t \geq 0\}}, \quad \theta_{f,t}^-(h) = 1 - \theta(f_t - h_t) = 1_{\{f_t - h_t < 0\}}. \tag{2.8}$$

Then define on $\mathbb{R}^{(w+1)d} = \times_{w+1}\mathbb{R}^d$ the function

$$\vartheta_{t:w}(f, h_t, \dots, h_{t+w})(x_t, \dots, x_{t+w}) = \sum_{s=t}^{t+w} f_s(x_s)\,\theta_{f,s}(h)(x_s) \prod_{r=t}^{s-1} \theta^-_{f,r}(h)(x_r) + h_{t+w}(x_{t+w}) \prod_{r=t}^{t+w} \theta^-_{f,r}(h)(x_r), \quad (2.9)$$

where we follow the convention that the product over an empty index set is equal to one. In the following, to reduce notational overhead, we simply write

$$\vartheta_{t:w}(f, h) = \vartheta_{t:w}(f, h_t, \dots, h_{t+w}), \quad (2.10)$$

thereby implicitly assuming that $\vartheta_{t:w}(f, h)$ is solely depending on $h_t, \dots, h_{t+w}$.

In a financial context the function $\vartheta_{t:w}(f, h)$ has a natural interpretation as the future payoff we would get by holding the Bermudan option for at most w periods, applying the stopping rule $\tau_t(h)$ which is defined recursively by

$$\begin{aligned} \tau_T(h) &= T, \\ \tau_t(h) &= t\,\theta_{f,t}(h)(X_t) + \tau_{t+1}(h)\,\theta^-_{f,t}(h)(X_t), \end{aligned} \quad (2.11)$$

and selling the option at time $t + w$ for the price $h_{t+w}(X_{t+w})$, if it is not exercised before.

We now come back to the generalization of the regression representations (2.4) and (2.6). First note that $\max(f_{t+1}, q_{t+1}) = \vartheta_{t+1:0}(f, q)$ and therefore

$$q_t(x) = E\big[\vartheta_{t+1:0}(f, q)(X_{t+1}) \mid X_t = x\big]. \quad (2.12)$$

On the other hand the recursive formula (2.7) for the optimal stopping rule τ_t^* shows that

$$f_{\tau^*_{t+1}}(X_{\tau^*_{t+1}}) = f_{\tau_{t+1}(q)}(X_{\tau_{t+1}(q)}) = \vartheta_{t+1:T-t-1}(f, q)(X_{t+1}, \dots, X_T),$$

such that we also have (cf., (2.4))

$$q_t(x) = E\big[\vartheta_{t+1:T-t-1}(f, q)(X_{t+1}, \dots, X_T) \mid X_t = x\big]. \quad (2.13)$$

More generally, we have for any $0 \le w \le T - t - 1$ the representation

$$q_t(x) = E\big[\vartheta_{t+1:w}(f, q)(X_{t+1}, \dots, X_{t+w+1}) \mid X_t = x\big]. \quad (2.14)$$

To prove (2.14) we start with

$$\begin{aligned} q_t(X_t) &= E\big[\max(f_{t+1}(X_{t+1}), q_{t+1}(X_{t+1})) \mid X_t\big] \\ &= E\big[f_{t+1}(X_{t+1})\theta_{f,t+1}(q)(X_{t+1}) + q_{t+1}(X_{t+1})\theta^-_{f,t+1}(q)(X_{t+1}) \mid \mathcal{F}_t\big], \end{aligned} \quad (2.15)$$

where we have used the Markov property in the second equality. Then we expand $q_{t+1}(X_{t+1})$ in (2.15) by

$$E\big[f_{t+2}(X_{t+2})\theta_{f,t+2}(q)(X_{t+2}) + q_{t+2}(X_{t+2})\theta^-_{f,t+2}(q)(X_{t+2}) \mid \mathcal{F}_{t+1}\big]$$

and proceed recursively up to $t + w + 1$. Equation (2.14) follows from the projection property $E\big[E[\cdot \mid \mathcal{F}_{t+1}] \mid \mathcal{F}_t\big] = E[\cdot \mid \mathcal{F}_t]$ of conditional expectations and by another application of the Markov property.

3. Monte Carlo algorithms for optimal stopping

Equation (2.14) shows that the continuation value q_t at time t can be obtained as the regression function of $\vartheta_{t+1:w}(f, q)$ for some $0 \le w \le T - t - 1$. Least squares Monte-Carlo methods pioneered by [21], and extended in [10] to arbitrary w, recursively estimate the regression function q_t from independent sample paths of the underlying Markov process X_t. Let

$$X_{t+1:w} = (X_{t+1}, \dots, X_{t+w+1}) \quad (3.1)$$

be the partial sample path of length w starting at $t+1$. When it comes to estimation of the continuation value q_t, these algorithms use the previously determined estimates $\hat{q}_{t+1}, \ldots, \hat{q}_{t+w+1}$ for $q_{t+1}, \ldots, q_{t+w+1}$ to construct

$$\hat{Y}_t = \vartheta_{t+1:w}(f, \hat{q})(X_{t+1:w}) = \vartheta_{t+1:w}(f, \hat{q}_{t+1}, \ldots, \hat{q}_{t+w+1})(X_{t+1:w}), \tag{3.2}$$

which takes the role of the dependent variable of the regression problem for time step t. The random variable $\hat{Y}_t$ is a estimate of the unknown optimal reward

$$Y_t = \vartheta_{t+1:w}(f, q)(X_{t+1:w}) = \vartheta_{t+1:w}(f, q_{t+1}, \ldots, q_{t+w+1})(X_{t+1:w}). \tag{3.3}$$

Given independent sample paths

$$\mathbf{X}_i = (X_{i,t})_{t=0,\ldots,T}, \quad i = 1, \ldots, n \tag{3.4}$$

of the underlying Markov process $\mathbf{X}$, the least squares estimate of q_t is obtained as

$$\hat{q}_{n,t} = \underset{h \in \mathcal{H}_{n,t}}{\arg\min} \ \frac{1}{n} \sum_{i=1}^{n} |h(X_{i,t}) - \hat{Y}_{i,t}|^2, \tag{3.5}$$

where

$$\hat{Y}_{i,t} = \vartheta_{t+1:w}(f, \hat{q})(X_{i,t+1:w}), \quad X_{i,t+1:w} = (X_{i,t+1}, \ldots, X_{i,t+w+1}) \tag{3.6}$$

and $\mathcal{H}_{n,t}$ is a set of functions $h : \mathbb{R}^d \to \mathbb{R}$.

With $w = 0$ the above algorithm corresponds to the Tsitsiklis–Van Roy algorithm [29], while the use of $w = T - t - 1$ was proposed by [21]. The idea of using an intermediate value $w \in \{0, 1, \ldots, T - t - 1\}$ in order to "interpolate" between these two algorithms was introduced in [10]. There results concerning consistency and rate of convergence of the above algorithm were derived for fixed w and fixed convex and uniformly bounded function spaces $\mathcal{H}_{n,t}$.

The boundedness assumption on $\mathcal{H}_{n,t}$ makes computation of the least squares estimate in (3.5) difficult because it leads to constrained optimization problems, see for instance [15, Section 10.1]. In addition, the convexity assumption excludes promising choices like spaces of polynomial splines with free knots or spaces of artificial neural networks, which require restrictions on the number of knots or the number of hidden neurons, respectively, to control the "complexity" of the function spaces. The resulting function spaces violate the convexity assumptions. Taking the convex hull instead is not an option because it would lead to function classes with a complexity that is much to high. Furthermore, in view of applications it is desirable to choose parameters of the functions spaces and also the parameter w of the underlying regression problems data dependent. In this paper we modify the above algorithm such that this is possible. For simplicity we restrict ourselves to function spaces, which are linear vector spaces, however, it is straightforward to derive similar results for spaces of polynomial splines with free knots or spaces of artificial neural networks.

The main problem in analyzing the estimates $\hat{q}_{n,t}$ is the control of the error propagation, i.e, to answer the question how the errors of $\hat{q}_{n,t+1}, \ldots, \hat{q}_{n,t+w+1}$ influence the error of $\hat{q}_{n,t}$. It is this point were the convexity assumption on $\mathcal{H}_{n,t}$ was used in [10] to bound the L_2-error in terms of the approximation error and a sample error derived from a suitably centered loss function. The difficulty for obtaining error estimates is the fact that $\hat{q}_{t+1}, \ldots, \hat{q}_{t+w+1}$ depend on a single set of sample paths (3.4) and are thus dependent. Clément, Lamberton, Protter [5] face the same difficulty while deriving a Central Limit Theorem for the Longstaff-Schwartz algorithm with linear approximation.

In the sequel we use a trick to simplify the analysis of the error propagation. Instead of using the partial sample path $X_{t+1:w}$ of our training data, which was used in part already in construction of the estimates $\hat{q}_{n,t+1}, \ldots, \hat{q}_{n,t+w+1}$, we generate new data $X_{t+1:w}^{t,new}$ for $X_{t+1:w}$ which are conditionally independent from all previously used data of time $s > t$ given X_t at time point t. We then construct samples of the distribution of $(X_t, \hat{Y}_t^{w,new})$ where

$$\hat{Y}_t^{w,new} = \vartheta_{t+1:w}(f, \hat{q}_{n,t+1}, \ldots, \hat{q}_{n,t+w+1})(X_{t+1:w}^{t,new}).$$

Since for given X_t the random variable $X_{t+1:w}^{t,new}$ is independent from all previously used data from time periods $s > t$, it is in particular independent from the data used in the construction of $\hat{q}_{n,t+1}, \ldots, \hat{q}_{n,t+w+1}$. Set

$$q_t^{w,new}(x) = \mathbf{E}^*\{\hat{Y}_t^{w,new} | X_t = x\},$$

where in $\mathbf{E}^*\{\cdot | X_t = x\}$ we take the conditional expectation with respect to fixed $X_t = x$ and with all the data fixed which were used in the construction of $\hat{q}_{n,t+1}, \ldots, \hat{q}_{n,t+w+1}$. Proposition 6.4 in [10] implies

$$\left\{ \int |q_t^{w,new}(x) - q_t(x)|^2 \mu_t(dx) \right\}^{1/2} \leq \sum_{s=t+1}^{t+w+1} \left\{ \int |\hat{q}_{n,s}(x) - q_s(x)|^2 \mu_s(dx) \right\}^{1/2}. \tag{3.7}$$

This allows us to control the error propagation. By induction, assume that we have

$$\mathbf{P}\left\{ \int |\hat{q}_{n,s}(x) - q_s(x)|^2 \mu_s(dx) > \sum_{r=s}^{T-1} c \cdot \left(\delta_{n,r} + \min_{h \in \mathcal{H}_{n,r}} \int |h(x) - q_r(x)|^2 \mu_r(dx) \right) \right\} \to 0 \quad (n \to \infty) \tag{3.8}$$

for $s \in \{t+1, \ldots, t+w+1\}$. Assume in addition, that we are able to show

$$\mathbf{P}\left\{ \int |\hat{q}_{n,t}(x) - q_t^{w,new}(x)|^2 \mu_t(dx) > c \cdot \left(\delta_{n,t} + \min_{h \in \mathcal{H}_{n,t}} \int |h(x) - q_t^{w,new}(x)|^2 \mu_t(dx) \right) \right\} \to 0 \quad (n \to \infty), \tag{3.9}$$

which is for suitable $\delta_{n,t}$ (depending on the "complexity" of the function spaces $\mathcal{H}_{n,t}$) a standard rate of convergence result for least squares estimates from a sample of size n where in the sample the response variables are independent given the predictor variables and where the predictor variables are independent, see [30] or [18].

It can be shown that (3.7)–(3.9) imply

$$\mathbf{P}\left\{ \int |\hat{q}_{n,t}(x) - q_t(x)|^2 \mu_t(dx) > \bar{c} \cdot \sum_{s=t}^{T-1} \left(\delta_{n,s} + \min_{h \in \mathcal{H}_{n,s}} \int |h(x) - q_s(x)|^2 \mu_s(dx) \right) \right\} \to 0 \quad (n \to \infty)$$

(details concerning related arguments can be found in the proofs of Theorems 4.1 and 4.3 below).

The main difference between our work here and the algorithms used in [21] and [10] is that we generate new data to construct samples of $\hat{Y}_t^{w,new}$. Because of this the data used for estimation of $q_t^{w,new}$ is conditionally independent given the sample of X_t, which enables us to to conclude (3.9) from standard rate-of-convergence results in nonparametric regression. The generation of the new, independent data is similar to the data generation in the random tree method (see, for instance, Section 8.3 in [13]). However, in contrast to the random tree method we use nonparametric regression techniques to estimate the regression function, while in the random tree method simple averages are used to estimate the regression function point by point. As a consequence, the number of data points needed grows exponential in T in the random tree method, while for our method it grows only linearly in T.

In the sequel we explain the definition of the estimates in detail. Let n be the size of the samples which we generate for our regression estimates, and let $w_{max} \in \{0, 1, \ldots, T-1\}$ be the maximal look-ahead which we use. We start with generating n independent sample paths

$$\mathbf{X}_i = (X_{i,t})_{t=0,\dots,T} \quad (i = 1, \dots, n)$$

of the underlying Markov process $\mathbf{X}$. Then we set

$$\hat{q}_T = \hat{q}_{n,T} = 0$$

and construct successively estimates of $q_{T-1}, \dots, q_0$ as follows: Fix $t \in \{0, 1, \dots, T-1\}$ and assume that estimates $\hat{q}_{n,t+1}, \dots, \hat{q}_{n,T-1}$ of $q_{t+1}, \dots, q_{T-1}$ are already constructed. Let

$$w_{max}(t) = \min\{w_{max}, T - t - 1\}$$

be the maximal look-ahead of time period t. Generate independent sample paths

$$X^{t,new}_{i,t:w_{max}(t)+1} = \left(X^{t,new}_{i,s}\right)_{s=t,\dots,t+w_{max}(t)+1} \quad (i = 1, \dots, n)$$

starting at $X^{t,new}_{i,t} = X_{i,t}$ for every $i \in \{1, \dots, n\}$ such that for all i, the partial sample paths

$$X^{t,new}_{i,t:w_{max}(t)+1} \tag{3.10}$$

have the same distribution as $X_{i,t:w_{max}(t)+1}$, and such that, given $X_{1,t}, \dots, X_{n,t}$, this data is independent from all previously generated data points for time periods $s > t$. Then set for every $w \in \{0, \dots, w_{max}(t)\}$

$$\hat{Y}^{w,new}_{i,t} = \vartheta_{t+1:w}(f, \hat{q}_{n,t+1}, \dots, \hat{q}_{n,t+w+1})(X^{t,new}_{i,t+1}, \dots, X^{t,new}_{i,t+w+1})$$

and apply a nonparametric least squares estimate to the data

$$\left((X_{i,t}, \hat{Y}^{w,new}_{i,t})\right)_{i=1,\dots,n} \tag{3.11}$$

to construct estimates $\hat{q}^w_{n,t}$ of q_t. Finally choose

$$\hat{w}_t \in \{0, 1, \dots, w_{max}(t)\}$$

and set

$$\hat{q}_{n,t} = \hat{q}^{\hat{w}_t}_{n,t}.$$

Next we explain how to define the nonparametric least squares estimates applied to the data (3.11), and how to choose $\hat{w}_t$ in a data dependent way. To do this we split our sample in three parts: a learning sample of size n_l, a testing sample of size n_t, and a validation sample of size n_v, where $n = n_l + n_t + n_v$. Furthermore we assume that we have given a finite set $\mathcal{P}_n$ of parameters and sets $\mathcal{H}_{n,p}$ of functions $h : \mathbb{R}^d \to \mathbb{R}$ for each $p \in \mathcal{P}_n$.

We start with explaining the definition of $\hat{q}^w_{n,t}$ for fixed w. For $p \in \mathcal{P}_n$ let

$$\tilde{q}^{w,p}_{n,t}(\cdot) = \arg\min_{h \in \mathcal{H}_{n,p}} \frac{1}{n_l} \sum_{i=1}^{n_l} |h(X_{i,t}) - \hat{Y}^{w,new}_{i,t}|^2 \tag{3.12}$$

be the least squares estimate of $q^{w,new}_t$ in $\mathcal{H}_{n,p}$, which we use as an estimate of q_t. Here we assume for notational simplicity that the minimum exists, however we do not require that it is unique. Depending on a truncation parameter $\beta_n > 0$, which we will choose later such that q_t is bounded in absolute value by β_n, we set

$$\hat{q}^{w,p}_{n,t}(x) = T_{\beta_n} \tilde{q}^{w,p}_{n,t}(x) \quad (x \in \mathbb{R}^d), \tag{3.13}$$

where $T_L z = \max\{-L, \min\{L, z\}\}$ for $z \in \mathbb{R}$ and $L > 0$. Next we use splitting of the sample, see for instance Chapter 7 in [15], to choose the parameter p. We set

$$\hat{q}^w_{n,t}(x) = \hat{q}^{w,\hat{p}^w_t}_{n,t}(x) \quad (x \in \mathbb{R}^d), \tag{3.14}$$

where $\hat{p}_t^w \in \mathcal{P}_n$ satisfies

$$\frac{1}{n_t}\sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w,\hat{p}_t^w}(X_{i,t}) - \hat{Y}_{i,t}^{w,new}|^2 = \min_{p\in\mathcal{P}_n} \frac{1}{n_t}\sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w,p}(X_{i,t}) - \hat{Y}_{i,t}^{w,new}|^2.$$

Finally, we explain our choice of w. For each $w \in \{0,1,\dots,w_{max}(t)\}$ we have already defined an estimate $\hat{q}_{n,t}^w$ of q_t. The idea is to compute from these estimates an approximately optimal stopping rule which provides a lower bound on the solution of the optimal stopping problem at time t. We then choose w such that this lower bound is maximized, i.e., we set

$$\hat{w}_t = \arg\max_{w\in\{0,1,\dots,w_{max}(t)\}} \frac{1}{n_v}\sum_{i=n_l+n_t+1}^{n} f_{\hat{\tau}_t^w(X_{i,t:T-t-1}^{t,new})}(X_{i,\hat{\tau}_t^w(X_{i,t:T-t-1}^{t,new})}^{t,new}), \tag{3.15}$$

where for $w \in \{0,1,\dots,w_{max}(t)\}$ the approximately optimal stopping rule $\hat{\tau}_t^w$ is defined via

$$\hat{\tau}_t^w = \tau_t(\hat{q}_{n,t}^w, \hat{q}_{n,t+1}, \dots, \hat{q}_{n,T-2}, \hat{q}_{T-1}) \tag{3.16}$$

where $\tau_t(h)$ is recursively defined by (2.11). With this choice of w we define our estimate of q_t by

$$\hat{q}_{n,t} = \hat{q}_{n,t}^{\hat{w}_t}. \tag{3.17}$$

4. Main theoretical results

In the sequel we derive results concerning consistency of our estimate under the assumption

$$X_t \in [-A,A]^d \quad a.s. \quad (t \in \{0,1,\dots,T\}). \tag{4.1}$$

In addition we assume that the payoff f_s is bounded on $[-A,A]^d$ by some constant $L > 0$, i.e. we assume

$$|f_s(x)| \le L \quad \text{for } x \in [-A,A]^d \text{ and } s \in \{0,1,\dots,T\}. \tag{4.2}$$

Observe that (4.2) implies $|q_t(x)| \le L$ for $x \in [-A,A]^d$ and $t \in \{0,1,\dots,T\}$, therefore we use in the sequel $\beta_n = L$ as truncation parameter of the estimate.

Because of boundedness of X_t, this assumptions is, e.g., fulfilled (for some L depending on A) if $f_t(x) = f(x)$ for some $f: \mathbb{R}^d \to \mathbb{R}$ satisfies

$$|f(u)| \le const \cdot \|u\|^r \quad (u \in \mathbb{R}^d)$$

for some $r > 0$.

In the sequel we use polynomial splines to define the function spaces $\mathcal{H}_{n,p} = \mathcal{H}_p$. Here $\mathcal{H}_{n,p}$ will depend in an application on the sample size n via the parameter $p = (M,\alpha) \in \mathbb{N}_0 \times (0,\infty)$ which we will choose in a concrete application depending on the sample.

Depending on these parameters set $u_k = k\cdot\alpha$ ($k \in \mathbb{Z}$). For $k \in \mathbb{Z}$ and $M \in \mathbb{N}_0$ let $B_{k,M}: \mathbb{R} \to \mathbb{R}$ be the univariate B-spline of degree M with knot sequence $(u_l)_{l\in\mathbb{Z}}$ and support $supp(B_{k,M}) = [u_k, u_{k+M+1}]$. In case $M = 0$ this means that $B_{k,0}$ is the indicator function of the interval $[u_k, u_{k+1})$, and for $M = 1$ we have the so-called hat-functions

$$B_{k,1}(x) = \begin{cases} \frac{x-u_k}{u_{k+1}-u_k} & , u_k \le x \le u_{k+1}, \\ \frac{u_{k+2}-x}{u_{k+2}-u_{k+1}} & , u_{k+1} < x \le u_{k+2}, \\ 0 & , \text{else}. \end{cases}$$

The general definition of $B_{k,M}$ can be found, e.g., in [7], or in Section 14.1 of [15]. These B-splines are basis functions of sets of univariate piecewise polynomials of degree M, which are globally $(M-1)$–times continuously differentiable and where the M-th derivative of the functions have jump points only at the knots u_l.

For $\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{Z}^d$ we define the tensor product B-spline $B_{\mathbf{k},M} : \mathbb{R}^d \to \mathbb{R}$ by

$$B_{\mathbf{k},M}(x^{(1)}, \dots, x^{(d)}) = B_{k_1,M}(x^{(1)}) \cdots B_{k_d,M}(x^{(d)}) \quad (x^{(1)}, \dots, x^{(d)} \in \mathbb{R}).$$

With these functions we define $\mathcal{H}_{n,p}$ as the set of all linear combinations of all those of the above tensor product B-splines, where the support has nonempty intersection with $[-A, A]^d$, i.e., we set

$$\mathcal{H}_{n,p} = \left\{ \sum_{\mathbf{k} \in \mathbb{Z}^d : supp(B_{\mathbf{k},M}) \cap [-A,A]^d \neq \emptyset} a_{\mathbf{k}} \cdot B_{\mathbf{k},M} \; : \; a_{\mathbf{k}} \in \mathbb{R} \right\}.$$

It can be shown by using standard arguments from spline theory, that the functions in $\mathcal{H}_{n,p}$ are in each component $(M-1)$-times continuously differentiable, that they are equal to a (multivariate) polynomial of degree less than or equal to M (in each component) on each rectangular

$$[u_{k_1}, u_{k_1+1}) \times \cdots \times [u_{k_d}, u_{k_d+1}) \quad (\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{Z}^d), \tag{4.3}$$

and that they vanish on all of those rectangles (4.3) where $\mathbf{k} \in \mathbb{Z}^d$ satisfies for some $j \in \{1, \dots, d\}$

$$k_j > 0 \quad \text{and} \quad u_{k_j - M} > A$$

or

$$k_j < 0 \quad \text{and} \quad u_{k_j + M + 1} < -A.$$

So $\mathcal{H}_{n,p}$ is a set of functions which are piecewise polynomials with respect to a equidistant partition of $\mathbb{R}^d$ in cubes of side length α and which vanish outside a compact set.

For the set $\mathcal{P}_n$ of parameters p of the functions spaces we use

$$\mathcal{P}_n \quad = \quad \left\{ (M, \alpha) \; : \; M \in \mathbb{N}_0, M \leq \lceil \log(n) \rceil, \alpha = 2^k \text{ for some } k \in \mathbb{Z}, |k| \leq \lceil \log(n) \rceil \right\}.$$

Here log denotes the natural logarithm, and for $z \in \mathbb{R}$ we denote by $\lceil z \rceil$ the smallest integer greater than or equal to z.

Let $\hat{q}_{n,t}$ be defined as in Section 3 with $\mathcal{P}_n$ and $\mathcal{H}_{n,p}$ as above and with $n_v = n_t = \lfloor n/3 \rfloor$ and $n_l = n - n_v - n_t$.

Our first result concerns consistency of the estimate.

Theorem 4.1. *Assume (4.1), (4.2) and let the estimate $\hat{q}_{n,t}$ be defined as above with $\beta_n = L$. Then*

$$\mathbf{E} \int |\hat{q}_{n,t}(x) - q_t(x)|^2 \mu_t(dx) \to 0 \quad (n \to \infty)$$

for all $t \in \{0, 1, \dots, T-1\}$.

Next we study the rate of convergence of our estimate. It is well known in nonparametric regression, that without smoothness assumptions on the regression function the rate of convergence can be arbitrarily slow (cf., e.g., [6], [8] or [15, Chapter 3]. In the sequel we assume that the continuation values q_t are (p, C)–smooth according to the following definition.

Definition 4.2. Let $p = k + \beta$ for some $k \in \mathbb{N}_0$, $\beta \in (0, 1]$, and let $C > 0$. A function $f : \mathbb{R}^d \to \mathbb{R}$ is called (p, C)–smooth, if all partial derivatives

$$\frac{\partial f}{\partial^{\alpha_1} x^{(1)} \dots \partial^{\alpha_d} x^{(d)}}$$

of total order $\alpha_1 + \cdots + \alpha_d = k$ exist and satisfy

$$\left| \frac{\partial f}{\partial^{\alpha_1} x^{(1)} \dots \partial^{\alpha_d} x^{(d)}}(x) - \frac{\partial f}{\partial^{\alpha_1} x^{(1)} \dots \partial^{\alpha_d} x^{(d)}}(z) \right| \leq C \cdot \|x - z\|^\beta$$

for all $x, z \in \mathbb{R}^d$.

The next theorem contains our main result concerning rate of convergence of the estimate.

Theorem 4.3. *Let $p = k + \beta$ for some $k \in \mathbb{N}_0$, $\beta \in (0, 1]$, and let $C > 0$. Assume $k \le M_{max}$, (4.1), (4.2) and*

$$q_t \quad (p, C) - smooth$$

for all $t \in \{0, 1, \dots, T-1\}$. Let the estimate $\hat{q}_{n,t}$ be defined as above with $\beta_n = L$. Then for every $t \in \{0, 1, \dots, T-1\}$

$$\mathbf{E} \int |\hat{q}_{n,t}(x) - q_t(x)|^2 \mu_t(dx) \le const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}}.$$

Remark 4.4. We would like to stress that in the theorems above there is no assumption on the distribution of X besides the assumption (4.1). In particular it is not required that X has a density with respect to the Lebesgue-Borel-measure.

Remark 4.5. It is well-known that for estimation of (p, C)–smooth functions $n^{-2p/(2p+d)}$ is the optimal rate of convergence (see, e.g., [27] or [15, Chapter 3]). So the rate of convergence in Theorem 4.3 is optimal up to a logarithmic factor.

Remark 4.6. The definition of the estimate in Theorem 4.3 above does not depend on (p, C), however the estimate achieves nevertheless the optimal rate of convergence for this particular smoothness of the continuation value. In this sense the estimate is able to adapt automatically to the smoothness of the continuation value, in contrast to the estimates in [10].

Remark 4.7. Assume $X_0 = x_0$ a.s. for some $x_0 \in \mathbb{R}$. We can estimate the price

$$V_0 = v_0(x_0) = \max\{f_0(x_0), q_0(x_0)\}$$

(cf., (1.1), (2.2), (2.5)) of the American option by

$$\hat{V}_0 = \max\{f_0(x_0), \hat{q}_{n,0}(x_0)\}.$$

Since the distribution μ_0 of X_0 is concentrated on x_0, under the assumptions of Theorem 4.3 we have the following error bound:

$$\begin{aligned} \mathbf{E}\{|\hat{V}_0 - V_0|^2\} &= \mathbf{E}\{|\max\{f_0(x_0), q_{n,0}(x_0)\} - \max\{f_0(x_0), q_0(x_0)\}|^2\} \\ &\le \mathbf{E}\{|q_{n,0}(x_0) - q_0(x_0)|^2\} \\ &\le const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}}. \end{aligned}$$

5. Application to simulated data

In this section, we illustrate the finite sample behaviour of our algorithm by comparing it with the Longstaff–Schwartz [21] and Tsitsiklis–Van Roy [29] algorithm.

For a comparison we apply all three algorithms to 100 independently generated sets of paths. We compute for each algorithm 100 Monte Carlo estimates (MCE) of the price (1.5) of the approximate optimal stopping rule (1.4) with a sample size of 4000. Figure 2 and 4 show boxplots of the MCE of the price (1.5) for an ordinary put and for a more complicated strangle spread payoff. Because all three algorithms provide lower bounds to the optimal stopping value, and since we evaluate the approximative optimal stopping rule on newly generated data, a higher MCE indicates a better performance of the algorithm.

We simulate the paths of the underlying stocks with a simple Black-Scholes-model. The time to maturity is assumed to be 1 year. We discretize the time interval $[0, 1]$ with m time steps.

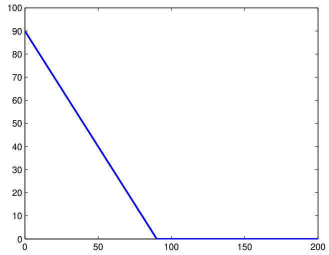

FIGURE 1. Put-payoff with exercise price 90

The prices of the underlying stocks at the time points $0, \frac{1}{m}, \ldots, \frac{m-1}{m}, 1$ are then given by

$$X_{i,j} = X_i\Big(\frac{j}{m}\Big)$$

where

$$X_{i,j} = X_0 \cdot \exp\Big((r - \frac{1}{2}\sigma^2) \cdot \frac{j}{m} + \frac{\sigma}{\sqrt{m}} \cdot W_{i,j}\Big) \quad (i = 1, \ldots, n\ , j = 1, \ldots, m)\ . \tag{5.1}$$

Here, X_0 is the initial stock price at time 0, r is the risk-free interest rate, σ the instantaneous volatility, and $W_{i,j}$ is discretized Brownian motion

$$W_{i,j} = \sum_{l=1}^{j} Z_{i,l}$$

where $Z_{i,l}$ $(i = 1, \ldots, n\ , l = 1, \ldots, m)$ are independent standard normally distributed random variables. For all of the subsequent simulations we choose $X_0 = 100$, $r = 0.05$, $m = 12$, a sample size of $n = 10000$, and discount factors given by $d_{0,t} = e^{-rt}$.

We set for our algorithm the number of learning, training, and validation samples to $n_l = 6000$, $n_t = 2000$ and $n_v = 2000$, respectively. To simplify the implementation we select the degree M, the knot distance α, and the look-ahead parameter w in a data-dependent manner as described in Section 3 from the sets $M \in \{0, 1, 2\}$, $\alpha \in \{\frac{100}{2}, \frac{100}{2^2}, \frac{100}{2^3}, \frac{100}{2^4}\}$, and $w(t) \in \{0, 4, T - t - 1\}$. For the Longstaff–Schwartz and Tsitsiklis–Van Roy algorithm we use polynomials of degree 3.

We first analyze a standard put-payoff with exercise price 90, illustrated in Figure 1, and simulate the paths of the underlying stock with an instantaneous volatility of $\sigma = 0.25$. As we can see from Figure 2, our algorithm is slightly better than the Longstaff–Schwartz algorithm and comparable to the algorithm of Tsitsiklis–Van Roy. This is not surprising, since it is well known that for simple payoff functions the Longstaff–Schwartz as well as the Tsitsiklis–Van Roy algorithm perform very well.

Next, we make the pricing problem more difficult. We consider $m = 48$ time steps, a strangle spread payoff with strikes 50, 90, 110 and 150 as illustrated in Figure 3, and a large volatility of $\sigma = 0.5$. This time our algorithm is clearly superior to Longstaff–Schwartz and Tsitsiklis–Van Roy algorithm. Figure 4 show that our dynamic look-ahead algorithm provides a higher MCE of the option price.

6. PROOFS

In the proofs we will need an auxiliar result concerning properties of the splitting of the sample method, which we formulate and prove for the sake of generality in a fixed design regression model.

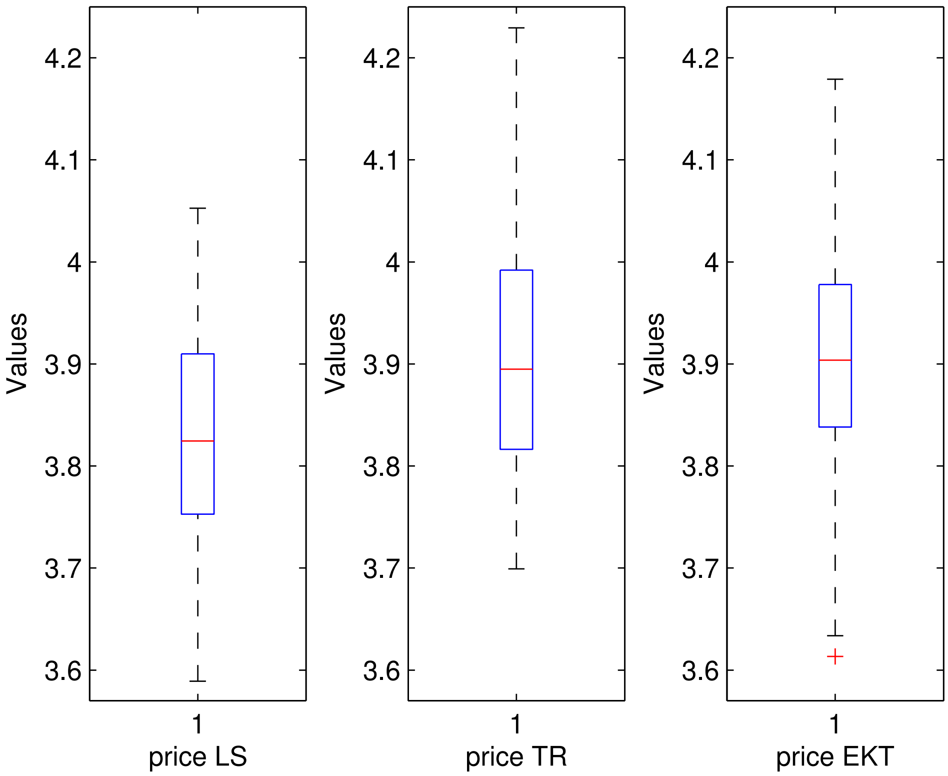

FIGURE 2. Boxplots for realized option prices for the put-payoff of the Longstaff–Schwartz (price LS), Tsitsiklis–Van Roy (price TR), and our algorithm (price EKT). In the boxplots the box stretches from the 25th percentile to the 75th percentile and the median is shown as a line across the box.

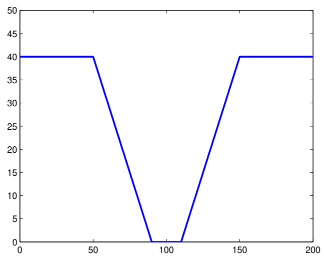

FIGURE 3. Strangle spread payoff with strike prices 50, 90, 110 and 150

Let $x_1, \dots, x_n \in \mathbb{R}^d$ and let $Y_1, \dots, Y_n$ be independent square integrable random variables which satisfy

$$\mathbf{E}Y_i = m(x_i) \quad (i = 1, \dots, n)$$

for some function $m : \mathbb{R}^d \to \mathbb{R}$. Let $\mathcal{P}_n$ be a finite set of parameters and assume that for each $p \in \mathcal{P}_n$ an estimate $m_p : \mathbb{R}^d \to \mathbb{R}$ is given. Choose $p^* \in \mathcal{P}_n$ by minimizing the empirical L_2 risk

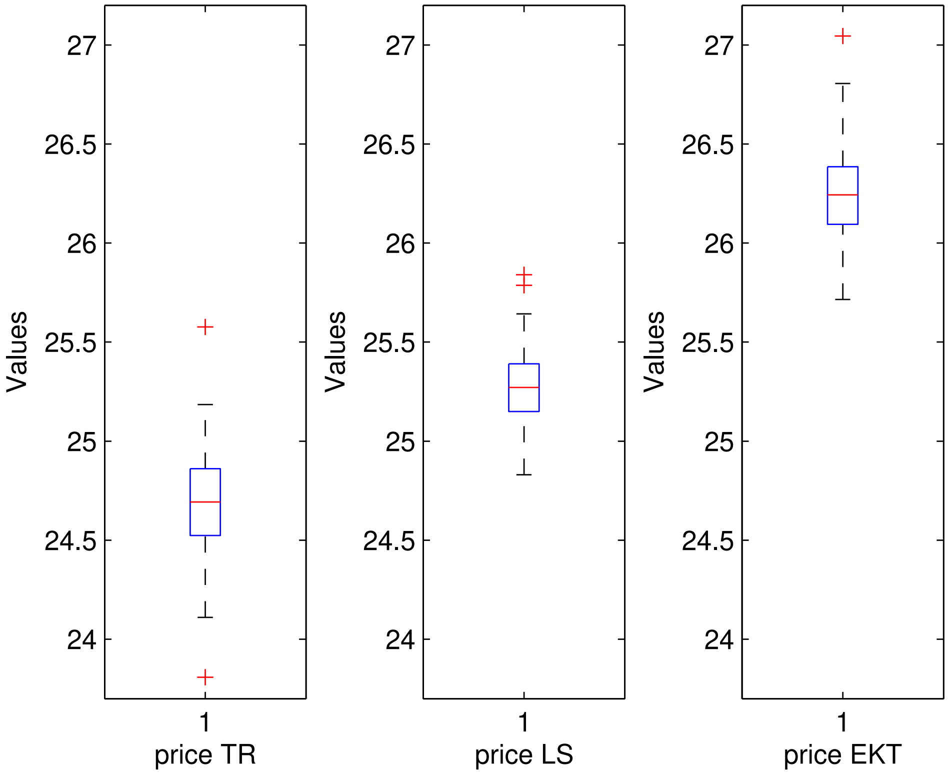

FIGURE 4. Realized option prices for the strangle spread-payoff of the Longstaff–Schwartz (price LS), Tsitsiklis–Van Roy (price TR) and our algorithm (price EKT)

on the sample $(x_1, Y_1), \ldots, (x_n, Y_n)$, i.e., assume

$$\frac{1}{n}\sum_{i=1}^{n}|m_{p^*}(x_i) - Y_i|^2 = \min_{p\in\mathcal{P}_n}\frac{1}{n}\sum_{i=1}^{n}|m_p(x_i) - Y_i|^2.$$

Then the following bound on the error

$$\frac{1}{n}\sum_{i=1}^{n}|m_{p^*}(x_i) - m(x_i)|^2$$

of m_{p^*} holds:

Lemma 6.1. *Under the above assumptions we have for each $\epsilon > 0$*

$$\mathbf{P}\left\{\frac{1}{n}\sum_{i=1}^{n}|m_{p^*}(x_i) - m(x_i)|^2 > \epsilon + 18\cdot\min_{p\in\mathcal{P}_n}\frac{1}{n}\sum_{i=1}^{n}|m_p(x_i) - m(x_i)|^2\right\}$$
$$\leq c_1\cdot\max_{i=1,\ldots,n}\mathbf{E}Y_i^2\cdot\frac{|\mathcal{P}_n|}{\epsilon\cdot n}$$

for some constant c_1 which does not depend on n or ϵ.

Proof. Set
$$m^* = \arg \min_{f\in\{m_p : p\in\mathcal{P}_n\}} \frac{1}{n}\sum_{i=1}^n |f(x_i)-m(x_i)|^2.$$
By Lemma 1 in [18] (or standard results from the book [30], see proof of Theorem 10.11 in [30]) we have
$$\begin{aligned}
&\mathbf{P}\left\{\frac{1}{n}\sum_{i=1}^n |m_{p^*}(x_i)-m(x_i)|^2 > \epsilon + 18\cdot \min_{p\in\mathcal{P}_n}\frac{1}{n}\sum_{i=1}^n |m_p(x_i)-m(x_i)|^2\right\}\\
&\leq \mathbf{P}\left\{\frac{\epsilon}{2} < \frac{1}{n}\sum_{i=1}^n |m_{p^*}(x_i)-m^*(x_i)|^2 \leq \frac{16}{n}\sum_{i=1}^n (m_{p^*}(x_i)-m^*(x_i))\cdot(Y_i-m(x_i))\right\}\\
&\leq |\mathcal{P}_n|\cdot\max_{p\in\mathcal{P}_n}\mathbf{P}\left\{\frac{\epsilon}{2} < \frac{1}{n}\sum_{i=1}^n |m_{p}(x_i)-m^*(x_i)|^2 \leq \frac{16}{n}\sum_{i=1}^n (m_{p}(x_i)-m^*(x_i))\cdot(Y_i-m(x_i))\right\}\\
&\leq |\mathcal{P}_n|\cdot\max_{p\in\mathcal{P}_n}\sum_{s=0}^\infty \mathbf{P}\Bigg\{2^{s-1}\epsilon < \frac{1}{n}\sum_{i=1}^n |m_{p}(x_i)-m^*(x_i)|^2 \leq 2^s\epsilon,\\
&\qquad\qquad \frac{1}{n}\sum_{i=1}^n |m_{p}(x_i)-m^*(x_i)|^2 \leq \frac{16}{n}\sum_{i=1}^n (m_{p}(x_i)-m^*(x_i))\cdot(Y_i-m(x_i))\Bigg\}\\
&\leq |\mathcal{P}_n|\cdot\sum_{s=0}^\infty \max_{\substack{p\in\mathcal{P}_n,\\ \frac{1}{n}\sum_{i=1}^n |m_p(x_i)-m^*(x_i)|^2\leq 2^s\epsilon}} \mathbf{P}\left\{\frac{1}{n}\sum_{i=1}^n (m_{p}(x_i)-m^*(x_i))\cdot(Y_i-m(x_i)) > \frac{2^s\epsilon}{32}\right\}.
\end{aligned}$$
Using
$$\mathbf{V}\left(\frac{1}{n}\sum_{i=1}^n (m_{p}(x_i)-m^*(x_i))\cdot(Y_i-m(x_i))\right) \leq \frac{1}{n^2}\sum_{i=1}^n (m_p(x_i)-m^*(x_i))^2\cdot\max_{i=1,\dots,n}\mathbf{E}Y_i^2$$
we can bound the right-hand side above via Chebyshev inequality by
$$|\mathcal{P}_n|\cdot\sum_{s=0}^\infty \frac{\frac{1}{n}\cdot 2^s\cdot\epsilon\cdot\max_{i=1,\dots,n}\mathbf{E}Y_i^2}{(2^s\epsilon/32)^2} = \frac{|\mathcal{P}_n|}{n}\cdot\frac{\max_{i=1,\dots,n}\mathbf{E}Y_i^2}{\epsilon}\cdot\sum_{s=0}^\infty\frac{32^2}{2^s}$$
□

Proof of Theorem 4.1. Because of
$$\mathbf{E}\int |\hat{q}_{n,t}(x)-q_t(x)|^2\mu_t(dx) \leq \sum_{w=0}^{w_{max}(t)}\mathbf{E}\int |\hat{q}^w_{n,t}(x)-q_t(x)|^2\mu_t(dx)$$
it suffices to show
$$\mathbf{E}\int |\hat{q}^w_{n,t}(x)-q_t(x)|^2\mu_t(dx)\to 0 \quad (n\to\infty) \tag{6.1}$$
for every $t\in\{0,1,\dots,T-1\}$ and every $w\in\{0,1,\dots,w_{max}(t)\}$.

Fix $t\in\{0,1,\dots,T-1\}$ and assume (by induction) that we have for every $s\in\{t+1,\dots,T-1\}$ and every $v\in\{0,1,\dots,w_{max}(s)\}$
$$\mathbf{E}\int |\hat{q}^v_{n,s}(x)-q_s(x)|^2\mu_t(dx)\to 0 \quad (n\to\infty). \tag{6.2}$$
Fix $w\in\{0,1,\dots,w_{max}(t)\}$. In the sequel we show
$$\mathbf{E}\int |\hat{q}^w_{n,t}(x)-q_t(x)|^2\mu_t(dx)\to 0 \quad (n\to\infty). \tag{6.3}$$

To do this, we use for fixed $p_n \in \mathcal{P}_n$ the error decomposition

$$\begin{aligned}
& \int |\hat{q}^w_{n,t}(x) - q_t(x)|^2 \mu_t(dx) \\
= & \int |\hat{q}^w_{n,t}(x) - q_t(x)|^2 \mu_t(dx) - \frac{1}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^w_{n,t}(X_{i,t}) - q_t(X_{i,t})|^2 \\
& + \frac{1}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^w_{n,t}(X_{i,t}) - q_t(X_{i,t})|^2 - \frac{2}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^w_{n,t}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \\
& + \frac{2}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^w_{n,t}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 - \frac{36}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^{w,p_n}_{n,t}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \\
& + \frac{36}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^{w,p_n}_{n,t}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 - \frac{72}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^{w,p_n}_{n,t}(X_{i,t}) - q_t(X_{i,t})|^2 \\
& + \frac{72}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^{w,p_n}_{n,t}(X_{i,t}) - q_t(X_{i,t})|^2 \\
= & \sum_{j=1}^{5} T_{j,n}
\end{aligned}$$

and observe that it suffices to show

$$\limsup_{n\to\infty} \mathbf{E} T_{j,n} \le 0 \tag{6.4}$$

for $j \in \{1, 2, \ldots, 5\}$.

In the sequel we denote by $\mathcal{D}^T_{n,t+1}$ the set of all data used in the construction of the estimates $\hat{q}^{w,p}_{n,s}$ for $s > t$, $w \in \{0, 1, \ldots, w_{max}(s)\}$ and $p \in \mathcal{P}_n$.

Using boundedness of $\hat{q}^w_{n,t}$ and q_t by L we can conclude from Hoeffding inequality (see, for instance, Lemma A.3 in [15])

$$\begin{aligned}
& \mathbf{P}\left\{ T_{1,n} > \epsilon | X^{t,new}_{i,t:w_{max}(t)+1} \quad (i = 1, \ldots, n_l), \mathcal{D}^T_{n,t+1} \right\} \\
& \le |\mathcal{P}_n| \cdot \max_{p \in \mathcal{P}_n} \mathbf{P}\Bigg\{ \int |\hat{q}^{w,p}_{n,t}(x) - q_t(x)|^2 \mu_t(dx) - \frac{1}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^{w,p}_{n,t}(X_{i,t}) - q_t(X_{i,t})|^2 > \epsilon \\
& \qquad\qquad \Bigg| X^{t,new}_{i,t+w_{max}(t)+1} \, (i = 1, \ldots, n_l), \mathcal{D}^T_{n,t+1} \Bigg\} \\
& \le |\mathcal{P}_n| \cdot \exp\left(-\frac{2 n_t \epsilon^2}{(4L^2)^2} \right) = \exp\left(\log(|\mathcal{P}_n|) - \frac{2 n_t \epsilon^2}{16 L^4} \right),
\end{aligned}$$

thus

$$\begin{aligned}
\mathbf{E}T_{1,n} &\le \int_0^\infty \mathbf{P}\{T_{1,n} > s\}ds \\
&= \int_0^\infty \mathbf{E}\left\{\mathbf{P}\left\{T_{1,n} > s | X^{t,new}_{i,t+w_{max}(t)+1} \, (i = 1,\ldots,n_l), \mathcal{D}^T_{n,t+1}\right\}\right\} ds \\
&\le 4L^2\sqrt{\log(|\mathcal{P}_n|)/n_t} + \int_{4L^2\sqrt{\log(|\mathcal{P}_n|)/n_t}}^\infty \exp\left(-\frac{n_t s^2}{16L^4}\right) ds \\
&\le 4L^2\sqrt{\log(|\mathcal{P}_n|)/n_t} + \int_{4L^2\sqrt{\log(|\mathcal{P}_n|)/n_t}}^\infty \exp\left(-\frac{n_t \cdot 4L^2\sqrt{\log(|\mathcal{P}_n|)/n_t}}{16L^2} \cdot s\right) ds \\
&\le 4L^2\sqrt{\log(|\mathcal{P}_n|)/n_t} + \frac{4L^2}{n_t\sqrt{\log(|\mathcal{P}_n|)/n_t}} \cdot \exp(-\log(|\mathcal{P}_n|)) \to 0 \quad (n \to \infty).
\end{aligned}$$

Furthermore, by $a^2 = (a-b+b)^2 \le 2(a-b)^2 + 2b^2$ we get

$$T_{2,n} \le \frac{2}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |q^{w,new}_t(X_{i,t}) - q_t(X_{i,t})|^2,$$

from which we conclude by (3.7) and (6.2)

$$\begin{aligned}
\mathbf{E}T_{2,n} &= \mathbf{E}\left\{\mathbf{E}\left\{T_{2,n} | X^{t,new}_{i,t:w_{max}(t)+1} \quad (i = 1,\ldots,n_l), \mathcal{D}^T_{n,t+1}\right\}\right\} \\
&\le 2\mathbf{E}\int |q^{w,new}_t(x) - q_t(x)|^2 \mu_t(dx) \to 0 \quad (n \to \infty).
\end{aligned}$$

Similarly we get

$$\mathbf{E}T_{4,n} \le 72\mathbf{E}\int |q^{w,new}_t(x) - q_t(x)|^2 \mu_t(dx) \to 0 \quad (n \to \infty).$$

To bound $T_{3,n}$ we use Lemma 6.1, which shows

$$\begin{aligned}
&\mathbf{P}\left\{T_{3,n} > \epsilon | X^{t,new}_{i,t:w_{max}(t)+1} \quad (i = 1,\ldots,n_l), \mathcal{D}^T_{n,t+1}\right\} \\
&\le \mathbf{P}\Bigg\{\frac{1}{n_t}\sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^w_{n,t}(X_{i,t}) - q^{w,new}_t(X_{i,t})|^2 > \frac{\epsilon}{2} + 18 \cdot \min_{p\in\mathcal{P}_n} \frac{1}{n_t}\sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^{w,p}_{n,t}(X_{i,t}) - q^{w,new}_t(X_{i,t})|^2 \\
&\qquad\qquad \Bigg| X^{t,new}_{i,t:w_{max}(t)+1} \quad (i = 1,\ldots,n_l), \mathcal{D}^T_{n,t+1}\Bigg\} \\
&\le c_2 \cdot \frac{|\mathcal{P}_n|}{\epsilon \cdot n_t}.
\end{aligned}$$

From this we get for $u > 0$

$$\begin{aligned}
\mathbf{E}T_{3,n} &\le \int_0^\infty \mathbf{P}\{T_{3,n} > \epsilon\}d\epsilon \\
&\le \int_0^\infty \mathbf{E}\left\{\mathbf{P}\left\{T_{3,n} > \epsilon | X^{t,new}_{i,t:w_{max}(t)+1} \quad (i = 1,\ldots,n_l), \mathcal{D}^T_{n,t+1}\right\}\right\} d\epsilon \\
&\le u + \int_u^{const} c_2 \cdot \frac{|\mathcal{P}_n|}{\epsilon \cdot n_t} d\epsilon \\
&= u + c_2 \cdot \frac{|\mathcal{P}_n|}{n_t} \cdot (\log(const) - \log u),
\end{aligned}$$

where we have used that (3.13) and boundedness of f (which implies boundedness of $q_t^{w,new}$) yield

$$T_{3,n} \leq \frac{2}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \leq const.$$

For $u = |\mathcal{P}_n|/n_t$ we get

$$\limsup_{n\to\infty} \mathbf{E} T_{3,n} \leq 0.$$

Furthermore

$$\mathbf{E}T_{5,n} = \mathbf{E}\left\{\mathbf{E}\left\{T_{5,n}|X_{i,t:w_{max}(t)+1}^{t,new} \quad (i=1,\ldots,n_l), \mathcal{D}_{n,t+1}^T\right\}\right\} = 72\cdot\mathbf{E}\int |\hat{q}_{n,t}^{w,p_n}(x) - q_t(x)|^2 \mu_t(dx).$$

So it remains to show

$$\mathbf{E}\int |\hat{q}_{n,t}^{w,p_n}(x) - q_t(x)|^2 \mu_t(dx) \to 0 \quad (n\to\infty) \tag{6.5}$$

for some $p_n \in \mathcal{P}_n$.

To show this we set $p_n = (0, 2^{-\lceil \log_2(n)/(2+d)\rceil})$ (where $\log_2$ denotes the logarithm with base 2) and consider the error decomposition

$$\begin{aligned}
& \int |\hat{q}_{n,t}^{w,p_n}(x) - q_t(x)|^2 \mu_t(dx) \\
= & \int |\hat{q}_{n,t}^{w,p_n}(x) - q_t(x)|^2 \mu_t(dx) - \frac{2}{n_l}\sum_{i=1}^{n_l} |\hat{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 \\
& + \frac{2}{n_l}\sum_{i=1}^{n_l} |\hat{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 - \frac{2}{n_l}\sum_{i=1}^{n_l} |\tilde{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 \\
& + \frac{2}{n_l}\sum_{i=1}^{n_l} |\tilde{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 - \frac{4}{n_l}\sum_{i=1}^{n_l} |\tilde{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \\
& + \frac{4}{n_l}\sum_{i=1}^{n_l} |\tilde{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \\
= & \sum_{j=6}^{9} T_{j,n}.
\end{aligned}$$

Because of boundedness of q_t by L we have

$$T_{7,n} \leq 0 \quad \text{and} \quad \mathbf{E}T_{7,n} \leq 0.$$

Furthermore, as for $T_{2,n}$ we get by (3.7) and (6.2)

$$\begin{aligned}
\mathbf{E}T_{8,n} & \leq 4\cdot\mathbf{E}\left\{\mathbf{E}\left\{\frac{1}{n_l}\sum_{i=1}^{n_l} |q_t(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \middle| \mathcal{D}_{n,t+1}^T\right\}\right\} \\
& = 4\cdot\mathbf{E}\int |q_t(x) - q_t^{w,new}(x)|^2 \mu_t(dx) \to 0 \quad (n\to\infty),
\end{aligned}$$

where the last equality follows from the fact that the conditional expectation $q_t^{w,new}(x)$ does not depend on data from time t.

Next we bound $T_{6,n}$. The functions $\hat{q}_{n,t}^{w,p_n}$ and q_t are bounded in absolute value by L, and $\tilde{q}_{n,t}^{w,p_n}$ belongs to the linear vector space $\mathcal{H}_{n,p_n}$, whose dimension D_n is bounded by some constant (depending on A) times $n^{d/(2+d)}$. As in the proof of Theorem 11.3 in [15] (see there proof of inequality (11.6)) this implies

$$\mathbf{E}T_{6,n} = \mathbf{E}\left\{\mathbf{E}\left\{T_{6,n}|\mathcal{D}_{n,t+1}^T\right\}\right\} \leq c_3 L^2 \frac{(\log n_l + 1)\cdot n^{d/(2+d)}}{n_l} \to 0 \quad (n \to \infty).$$

Finally we bound $T_{9,n}$. With

$$\sigma^2 = \sup_{x\in\mathbb{R}^d} \mathbf{E}^*\left\{|\hat{Y}_{1,t}^{w,new}|^2 | X_{1,t} = x\right\} \leq 4L^2 < \infty$$

we can conclude from Theorem 11.1 in [15]

$$\begin{aligned}&\mathbf{E}\left\{T_{9,n}|X_{i,t} \quad (i=1,\dots,n_l), \mathcal{D}_{n,t+1}^T\right\}\\ &\leq 4\sigma^2 \frac{c_4 n^{d/(2+d)}}{n_l} + 4 \min_{h\in\mathcal{H}_{n,p_n}} \frac{1}{n_l}\sum_{i=1}^{n_l} |h(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2,\end{aligned}$$

so

$$\begin{aligned}&\mathbf{E}T_{9,n}\\ &= \mathbf{E}\left\{\mathbf{E}\left\{T_{9,n}|X_{i,t} \quad (i=1,\dots,n_l), \mathcal{D}_{n,t+1}^T\right\}\right\}\\ &\leq 4\sigma^2 \frac{c_4 n^{d/(2+d)}}{n_l} + 4 \min_{h\in\mathcal{H}_{n,p_n}} \mathbf{E}\int |h(x) - q_t^{w,new}(x)|^2 \mu_t(dx)\\ &\leq 4\sigma^2 \frac{c_4 n^{d/(2+d)}}{n_l} + 8\mathbf{E}\int |q_t^{w,new}(x) - q_t(x)|^2\mu_t(dx) + 8 \min_{h\in\mathcal{H}_{n,p_n}} \int |h(x) - q_t(x)|^2\mu_t(dx).\end{aligned}$$

Because of (3.7), (6.2) and

$$\int |q_t(x)|^2 \mu_t(dx) \leq L^2 < \infty,$$

which implies that q_t can be approximated arbitrarily closely by functions from $\mathcal{H}_{n,p_n}$ (which follows from Theorem A.1 in [15] and the fact that any continuous function can be approximated in supremum norm on the compact set $[-A,A]^d$ arbitrarily closely by the piecewise constant functions in $\mathcal{H}_{n,p_n}$ for $n \to \infty$), the right hand side of the above inequality tends to zero for $n \to \infty$.

The proof is complete. □

Proof of Theorem 4.3. Because of

$$\mathbf{E}\int |\hat{q}_{n,t}(x) - q_t(x)|^2 \mu_t(dx) \leq \sum_{w=0}^{w_{max}(t)} \mathbf{E}\int |\hat{q}_{n,t}^w(x) - q_t(x)|^2\mu_t(dx)$$

it suffices to show

$$\mathbf{E}\int |\hat{q}_{n,t}^w(x) - q_t(x)|^2\mu_t(dx) \leq const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}} \tag{6.6}$$

for every $t \in \{0,1,\dots,T-1\}$ and every $w \in \{0,1,\dots,w_{max}(t)\}$.

Fix $t \in \{0,1,\dots,T-1\}$ and assume (by induction) that we have for every $s \in \{t+1,\dots,T-1\}$ and every $v \in \{0,1,\dots,w_{max}(s)\}$

$$\mathbf{E}\int |\hat{q}_{n,s}^v(x) - q_s(x)|^2\mu_t(dx) \leq const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}}. \tag{6.7}$$

Fix $w \in \{0, 1, \ldots, w_{max}(t)\}$. In the sequel we show

$$\mathbf{E} \int |\hat{q}_{n,t}^{w}(x) - q_t(x)|^2 \mu_t(dx) \leq const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n} \right)^{\frac{2p}{2p+d}}. \tag{6.8}$$

To do this, we use for fixed $p_n \in \mathcal{P}_n$ the error decomposition

$$\begin{aligned}
& \int |\hat{q}_{n,t}^{w}(x) - q_t(x)|^2 \mu_t(dx) \\
= & \int |\hat{q}_{n,t}^{w}(x) - q_t(x)|^2 \mu_t(dx) - \frac{2}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w}(X_{i,t}) - q_t(X_{i,t})|^2 \\
& + \frac{2}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w}(X_{i,t}) - q_t(X_{i,t})|^2 - \frac{4}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \\
& + \frac{4}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 - \frac{72}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \\
& + \frac{72}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 - \frac{144}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 \\
& + \frac{144}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 \\
= & \sum_{j=1}^{5} T_{j,n}
\end{aligned}$$

and observe that it suffices to show

$$\mathbf{E} T_{j,n} \leq const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n} \right)^{\frac{2p}{2p+d}} \tag{6.9}$$

for $j \in \{1, 2, \ldots, 5\}$.

We can conclude from Bernstein inequality (see, for instance, Lemma A.2 in [15]) by using boundedness of $\hat{q}_{n,t}^{w}$ and q_t by L together with

$$\begin{aligned}
\sigma^2 & = \mathbf{V}\Big(|\hat{q}_{n,t}^{w,p}(X_{n_l+1,t}) - q_t(X_{n_l+1,t})|^2 \,\Big|\, X_{i,t:w_{max}(t)+1}^{t,new} \, (i = 1, \ldots, n_l), \mathcal{D}_{n,t+1}^{T} \Big) \\
& \leq \mathbf{E}\Big(|\hat{q}_{n,t}^{w,p}(X_{n_l+1,t}) - q_t(X_{n_l+1,t})|^4 \,\Big|\, X_{i,t:w_{max}(t)+1}^{t,new} \, (i = 1, \ldots, n_l), \mathcal{D}_{n,t+1}^{T} \Big) \\
& \leq 4L^2 \mathbf{E}\Big(|\hat{q}_{n,t}^{w,p}(X_{n_l+1,t}) - q_t(X_{n_l+1,t})|^2 \,\Big|\, X_{i,t:w_{max}(t)+1}^{t,new} \, (i = 1, \ldots, n_l), \mathcal{D}_{n,t+1}^{T} \Big) \\
& = 4L^2 \int |\hat{q}_{n,t}^{w,p}(x) - q_t(x)|^2 \mu_t(dx)
\end{aligned}$$

$$
\begin{aligned}
&\mathbf{P}\left\{T_{1,n} > \epsilon | X^{t,new}_{i,t:w_{max}(t)+1} \quad (i=1,\ldots,n_l), \mathcal{D}^T_{n,t+1}\right\} \\
&\leq |\mathcal{P}_n| \cdot \max_{p\in\mathcal{P}_n} \mathbf{P}\Bigg\{ \int |\hat{q}^{w,p}_{n,t}(x) - q_t(x)|^2 \mu_t(dx) - \frac{2}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^{w,p}_{n,t}(X_{i,t}) - q_t(X_{i,t})|^2 > \epsilon \\
&\qquad\qquad \Bigg| X^{t,new}_{i,t:w_{max}(t)+1} \, (i=1,\ldots,n_l), \mathcal{D}^T_{n,t+1} \Bigg\} \\
&= |\mathcal{P}_n| \cdot \max_{p\in\mathcal{P}_n} \mathbf{P}\Bigg\{ \int |\hat{q}^{w,p}_{n,t}(x) - q_t(x)|^2 \mu_t(dx) - \frac{1}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^{w,p}_{n,t}(X_{i,t}) - q_t(X_{i,t})|^2 \\
&\qquad > \frac{\epsilon}{2} + \frac{1}{2} \int |\hat{q}^{w,p}_{n,t}(x) - q_t(x)|^2 \mu_t(dx) \, \Bigg| X^{t,new}_{i,t:w_{max}(t)+1} \, (i=1,\ldots,n_l), \mathcal{D}^T_{n,t+1} \Bigg\} \\
&\leq |\mathcal{P}_n| \cdot \max_{p\in\mathcal{P}_n} \mathbf{P}\Bigg\{ \int |\hat{q}^{w,p}_{n,t}(x) - q_t(x)|^2 \mu_t(dx) - \frac{1}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |\hat{q}^{w,p}_{n,t}(X_{i,t}) - q_t(X_{i,t})|^2 \\
&\qquad > \frac{\epsilon}{2} + \frac{1}{2} \cdot \frac{\sigma^2}{4L^2} \, \Bigg| X^{t,new}_{i,t:w_{max}(t)+1} \, (i=1,\ldots,n_l), \mathcal{D}^T_{n,t+1} \Bigg\} \\
&\leq |\mathcal{P}_n| \cdot \exp\left(- \frac{n_t(\frac{\epsilon}{2} + \frac{\sigma^2}{8L^2})^2}{2\sigma^2 + 2(\frac{\epsilon}{2} + \frac{\sigma^2}{8L^2}) \cdot \frac{4L^2}{3}} \right) \\
&\leq |\mathcal{P}_n| \cdot \exp\left(- \frac{n_t(\frac{\epsilon}{2} + \frac{\sigma^2}{8L^2})^2}{(16L^2 + \frac{8L^2}{3})(\frac{\epsilon}{2} + \frac{\sigma^2}{8L^2})} \right) \\
&\leq |\mathcal{P}_n| \cdot \exp\left(- \frac{1}{32 + \frac{16}{3}} \cdot \frac{n_t \epsilon}{L^2} \right) = |\mathcal{P}_n| \cdot \exp\left(- \frac{3}{112} \cdot \frac{n_t \epsilon}{L^2} \right),
\end{aligned}
$$

thus

$$
\begin{aligned}
\mathbf{E} T_{1,n} &\leq \int_0^\infty \mathbf{P}\{T_{1,n} > s\} ds \\
&= \int_0^\infty \mathbf{E}\left\{ \mathbf{P}\left\{ T_{1,n} > s | X^{t,new}_{i,t:w_{max}(t)+1} \, (i=1,\ldots,n_l), \mathcal{D}^T_{n,t+1} \right\} \right\} ds \\
&\leq |\mathcal{P}_n| \cdot \int_0^\infty \exp\left(- \frac{3n_t}{112L^2} \cdot s \right) ds \\
&\leq \frac{112L^2}{3} \cdot \frac{|\mathcal{P}_n|}{n_t} \quad \leq \; const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n} \right)^{\frac{2p}{2p+d}}.
\end{aligned}
$$

Furthermore, by $a^2 = (a-b+b)^2 \leq 2(a-b)^2 + 2b^2$ we get

$$
T_{2,n} \leq \frac{4}{n_t} \sum_{i=n_l+1}^{n_l+n_t} |q^{w,new}_t(X_{i,t}) - q_t(X_{i,t})|^2,
$$

from which we conclude by (3.7) and (6.7)

$$\begin{aligned}\mathbf{E}T_{2,n} &= \mathbf{E}\left\{\mathbf{E}\left\{T_{2,n}|X^{t,new}_{i,t:w_{max}(t)+1} \quad (i=1,\dots,n_l),\mathcal{D}^T_{n,t+1}\right\}\right\}\\ &\le 4\mathbf{E}\int |q^{w,new}_t(x)-q_t(x)|^2\mu_t(dx)\\ &\le const\cdot C^{\frac{2d}{2p+d}}\cdot\left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}}.\end{aligned}$$

Similarly we get

$$\mathbf{E}T_{4,n}\le const\cdot C^{\frac{2d}{2p+d}}\cdot\left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}}.$$

To bound $T_{3,n}$ we use Lemma 6.1, which shows

$$\begin{aligned}&\mathbf{P}\left\{T_{3,n}>\epsilon|X^{t,new}_{i,t:w_{max}(t)+1} \quad (i=1,\dots,n_l),\mathcal{D}^T_{n,t+1}\right\}\\ &\le \mathbf{P}\Bigg\{\frac{1}{n_t}\sum_{i=n_l+1}^{n_l+n_t}|\hat q^w_{n,t}(X_{i,t})-q^{w,new}_t(X_{i,t})|^2>\frac{\epsilon}{4}+18\cdot\min_{p\in\mathcal{P}_n}\frac{1}{n_t}\sum_{i=n_l+1}^{n_l+n_t}|\hat q^{w,p}_{n,t}(X_{i,t})-q^{w,new}_t(X_{i,t})|^2\\ &\qquad\qquad\Bigg|X^{t,new}_{i,t:w_{max}(t)+1} \quad (i=1,\dots,n_l),\mathcal{D}^T_{n,t+1}\Bigg\}\\ &\le c_5\cdot\frac{|\mathcal{P}_n|}{\epsilon\cdot n_t}.\end{aligned}$$

From this we get for $u>0$

$$\begin{aligned}\mathbf{E}T_{3,n} &\le \int_0^\infty \mathbf{P}\{T_{3,n}>\epsilon\}d\epsilon\\ &\le \int_0^\infty \mathbf{E}\left\{\mathbf{P}\left\{T_{3,n}>\epsilon|X^{t,new}_{i,t:w_{max}(t)+1} \quad (i=1,\dots,n_l),\mathcal{D}^T_{n,t+1}\right\}\right\}d\epsilon\\ &\le u+\int_u^{const} c_5\cdot\frac{|\mathcal{P}_n|}{\epsilon\cdot n_t}d\epsilon\\ &= u+c_5\cdot\frac{|\mathcal{P}_n|}{n_t}\cdot(\log(const)-\log u),\end{aligned}$$

where we have used that (3.13) and boundedness of f (which implies boundedness of $q^{w,new}_t$) yield

$$T_{3,n}\le\frac{4}{n_t}\sum_{i=n_l+1}^{n_l+n_t}|\hat q^w_{n,t}(X_{i,t})-q^{w,new}_t(X_{i,t})|^2\le const.$$

With $u=\log(n)/n$ we get

$$\mathbf{E}T_{3,n}\le\frac{\log n}{n}\left(1+c_6\left(\log(const)-\log\left(\frac{\log n}{n}\right)\right)\right)\le const\cdot C^{\frac{2d}{2p+d}}\cdot\left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}}.$$

Furthermore

$$\mathbf{E}T_{5,n}=\mathbf{E}\left\{\mathbf{E}\left\{T_{5,n}|X^{t,new}_{i,t:w_{max}(t)+1} \quad (i=1,\dots,n_l),\mathcal{D}^T_{n,t+1}\right\}\right\}=144\cdot\mathbf{E}\int|\hat q^{w,p_n}_{n,t}(x)-q_t(x)|^2\mu_t(dx).$$

So it remains to show

$$\mathbf{E}\int|\hat q^{w,p_n}_{n,t}(x)-q_t(x)|^2\mu_t(dx)\le const\cdot C^{\frac{2d}{2p+d}}\cdot\left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}} \tag{6.10}$$

for some $p_n\in\mathcal{P}_n$.

To bound $\mathbf{E}T_{5,n}$ we use the error decomposition

$$\begin{aligned}
& \int |\hat{q}_{n,t}^{w,p_n}(x) - q_t(x)|^2 \mu_t(dx) \\
= & \int |\hat{q}_{n,t}^{w,p_n}(x) - q_t(x)|^2 \mu_t(dx) - \frac{2}{n_l}\sum_{i=1}^{n_l} |\hat{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 \\
& + \frac{2}{n_l}\sum_{i=1}^{n_l} |\hat{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 - \frac{2}{n_l}\sum_{i=1}^{n_l} |\tilde{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 \\
& + \frac{2}{n_l}\sum_{i=1}^{n_l} |\tilde{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t(X_{i,t})|^2 - \frac{4}{n_l}\sum_{i=1}^{n_l} |\tilde{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \\
& + \frac{4}{n_l}\sum_{i=1}^{n_l} |\tilde{q}_{n,t}^{w,p_n}(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \\
= & \sum_{j=6}^{9} T_{j,n}
\end{aligned}$$

with

$$p_n = (k, 2^l) \quad \text{where} \quad l = \left\lceil \log_2\left(C^{-2/(2p+d)} (n/\log(n))^{-1/(2p+d)}\right)\right\rceil.$$

Because of boundedness of q_t by L we have

$$T_{7,n} \le 0 \quad \text{and} \quad \mathbf{E}T_{7,n} \le 0.$$

Furthermore, as for $T_{2,n}$ we get by (3.7) and (6.7)

$$\begin{aligned}
\mathbf{E}T_{8,n} & \le 4\mathbf{E}\left\{\mathbf{E}\left\{\frac{1}{n_l}\sum_{i=1}^{n_l} |q_t(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \middle| \mathcal{D}_{n,t+1}^T\right\}\right\} \\
& = 4\mathbf{E}\int |q_t(x) - q_t^{w,new}(x)|^2 \mu_t(dx) \\
& \le const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}},
\end{aligned}$$

where the last equality follows from the fact that the conditional expectation $q_t^{w,new}(x)$ does not depend on data from time t.

Next we bound $T_{6,n}$. The functions $\hat{q}_{n,t}^{w,p_n}$ and q_t are bounded in absolute value by L, and $\tilde{q}_{n,t}^{w,p_n}$ belongs to the linear vector space $\mathcal{H}_{n,p_n}$, whose dimension D_n is bounded by some constant (depending on A and k) times $C^{2d/(2p+d)} \cdot (n/\log(n))^{d/(2p+d)}$. As in the proof of Theorem 11.3 in [15] (see there proof of inequality (11.6)) this implies

$$\begin{aligned}
\mathbf{E}T_{6,n} = \mathbf{E}\left\{\mathbf{E}\left\{T_{6,n} \middle| \mathcal{D}_{n,t+1}^T\right\}\right\} & \le c_7 L^2 \frac{(\log n_l + 1) \cdot C^{2d/(2p+d)} \cdot (n/\log(n))^{d/(2p+d)}}{n_l} \\
& \le const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}}.
\end{aligned}$$

Finally we bound $T_{9,n}$. With

$$\sigma^2 = \sup_{x \in \mathbb{R}^d} \mathbf{E}^* \left\{ |\hat{Y}_{1,t}^{w,new}|^2 | X_{1,t} = x \right\} \le 4L^2 < \infty$$

we can conclude from Theorem 11.1 in [15]

$$\begin{aligned}
&\mathbf{E}\left\{T_{9,n}|X_{i,t} \quad (i=1,\ldots,n_l),\mathcal{D}_{n,t+1}^T\right\} \\
&\leq 4\sigma^2 \cdot \frac{D_n}{n_l} + 4 \min_{h\in\mathcal{H}_{n,p_n}} \frac{1}{n_l}\sum_{i=1}^{n_l} |h(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2 \\
&\leq 4\sigma^2 \cdot C^{2d/(2p+d)} \cdot \frac{c_8}{n^{2p/(2p+d)} \cdot \log(n)^{d/(2p+d)}} + 4 \min_{h\in\mathcal{H}_{n,p_n}} \frac{1}{n_l}\sum_{i=1}^{n_l} |h(X_{i,t}) - q_t^{w,new}(X_{i,t})|^2,
\end{aligned}$$

so

$$\begin{aligned}
&\mathbf{E}T_{9,n} \\
&= \mathbf{E}\left\{\mathbf{E}\left\{T_{9,n}|X_{i,t} \quad (i=1,\ldots,n_l),\mathcal{D}_{n,t+1}^T\right\}\right\} \\
&\leq 12\sigma^2 \cdot C^{2d/(2p+d)} \cdot \left(\frac{\log n}{n}\right)^{2p/(2p+d)} + 4 \min_{h\in\mathcal{H}_{n,p_n}} \mathbf{E}\int |h(x) - q_t^{w,new}(x)|^2 \mu_t(dx) \\
&\leq 12\sigma^2 \cdot C^{2d/(2p+d)} \cdot \left(\frac{\log n}{n}\right)^{2p/(2p+d)} + 8\mathbf{E}\int |q_t^{w,new}(x) - q_t(x)|^2 \mu_t(dx) \\
&\quad +8 \min_{h\in\mathcal{H}_{n,p_n}} \int |h(x) - q_t(x)|^2 \mu_t(dx).
\end{aligned}$$

Notice, that for the last term in the last inequality (without the factor 8) we get

$$\min_{h\in\mathcal{H}_{n,p}} \int |h(x) - q_t(x)|^2 \mu_t(dx) \leq \min_{h\in\mathcal{H}_{n,p}} \sup_{x\in[-A,A]^d} |h(x) - q_t(x)|^2.$$

Because we have assumed the (p,C)-smoothness of q_t, there exist a $h \in \mathcal{H}_{n,p}$ with

$$\sup_{x\in[-A,A]^d} |h(x) - q_t(x)| \leq c_9 \cdot C \cdot \delta_n^p$$

where $\delta_n = C^{-2/(2p+d)} \cdot (n/\log(n))^{-1/(2p+d)}$ is the side-length in the cubic partition used in the definition of the spline space, see Theorem 12.8 in [25]. From this we can conclude

$$\begin{aligned}
\min_{h\in\mathcal{H}_{n,p}} \int |h(x) - q_t(x)|^2 \mu_t(dx) &\leq c_9^2 \cdot C^2 \cdot \delta_n^{2p} \\
&= c_9^2 \cdot C^2 \cdot C^{-\frac{4p}{2p+d}} \cdot (n/\log(n))^{\frac{-2p}{2p+d}} \\
&\leq const \cdot C^{\frac{2d}{2p+d}} \cdot \left(\frac{\log n}{n}\right)^{\frac{2p}{2p+d}}.
\end{aligned}$$

With (3.7), (6.7) and the above inequality we get that the right hand side of $\mathbf{E}T_{9,n}$ has the required property.

The proof is complete. □

References

1. P. Boessarts, *Simulation estimators of optimal early exercise*, Working paper, Carnegie-Mellon University, 1989.
2. M. Broadie and P. Glasserman, *Pricing American-style securities using simulation*, J. of Economic Dynamics and Control **21** (1997), 1323–1352.
3. ———, *Monte Carlo methods for pricing high-dimensional American options: An overview*, Monte Carlo Methodologies and Applications for Pricing and Risk Management, Risk Books, 1998, pp. 149–161.
4. Y. S. Chow, H. Robbins, and D. Siegmund, *Great expectations: The theory of optimal stopping*, Houghton Mifflin, Boston, 1971.

5. E. Clément, D. Lamberton, and P. Protter, *An analysis of the Longstaff-Schwartz algorithm for American option pricing*, Finance and Stochastics **6** (2002), no. 4, 449–471.
6. T. M. Cover, *Rates of convergence of nearest neighbor procedures*, Proceedings of the Hawaii International Conference on Systems Sciences (1968), 413–415.
7. C. de Boor, *A practical guid to splines*, Springer-Verlag, 1978.
8. L. Devroye, *Necessary and sufficient conditions for the almost everywhere convergence of nearest neighbor regression function estimates*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete (1982), no. 61, 467–481.
9. R. M. Dudley, *Uniform central limit theorems*, Cambridge Studies in advanced mathematics, vol. 63, Cambridge Univ. Press, 1999.
10. D. Egloff, *Monte Carlo algorithms for optimal stopping and statistical learning*, Annals of Applied Probability **15** (2005), no. 2, 1–37.
11. N. El Karoui, *Les aspects probabilistes du contrôle stochastique*, Lecture Notes in Mathematics, vol. 876, Springer, 1981, pp. 74–239.
12. S. N. Ethier and T. G. Kurtz, *Markov processes: Characterization and convergence,*, John Wiley, New York, 1986.
13. P. Glasserman, *Monte Carlo methods in financial engineering*, Applications of Mathematics, vol. 53, Springer, 2004.
14. P. Glasserman and B. Yu, *Number of paths versus number of basis functions in American option pricing*, Annals of Applied Probability **14** (2004), no. 4, 1–30.
15. L. Györfi, M. Kohler, A. Krzyżak, and H. Walk, *A distribution-free theory of nonparametric regression*, Springer Series in Statistics, Springer, 2002.
16. M. Haugh and L. Kogan, *Pricing american options: A duality approach*, Operations Research **52** (2004), no. 2, 258–270.
17. I. Karatzas and S. E. Shreve, *Methods of mathematical finance*, Applications of Math., vol. 39, Springer-Verlag, 1998.
18. M. Kohler, *Nonparametric regression with additional measurement errors in the dependent variable*, Journal of Statistical Planning and Inference (2005).
19. D. Lamberton and G. Pagès, *Sur l'approximation des réduites*, Ann. Inst. Henri Poincarés, Probab. Statist. **26** (1990), no. 2, 331–355.
20. S. B. Laprise, Y. Su, R. Wu, M. C. Fu, and D. B. Madan, *Pricing American options: A comparision of Monte Carlo simulation approaches*, Journal Comput. Finance **4** (2001), no. 3, 39–88.
21. F. A. Longstaff and E. S. Schwartz, *Valuing American options by simulation: A simple least-square approach*, Review of Financial Studies **14** (2001), no. 1, 113–147.
22. J. Neveu, *Discrete-parameter martingales*, second ed., Wiley Series in Probability and Math. Statistics, North-Holland Publ. Comp., 1975.
23. D. Pollard, *Convergence of stochastic processes*, Springer Series in Statistics, Springer-Verlag, 1984.
24. L. C. G. Rogers, *Monte Carlo valuing of American options*, Mathematical Finance **12** (2002), 271–286.
25. L. Schumaker, *Spline functions: Basic theory*, Wiley, New-York, 1981.
26. A. N. Shiryayev, *Optimal stopping rules*, Applications of Mathematics, vol. 8, Springer Verlag, 1978.
27. C. J. Stone, *Optimal rates of convergence for nonparametric regression*, Annals of Statist. **10** (1982), no. 4, 1040–1053.
28. J. A. Tilley, *Valuing American options in a path simulation model*, Transactions of the Society of Actuaries **45** (1993), 83–104.
29. 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 Autom. Control **44** (1999), no. 10, 1840–1851.
30. S. van de Geer, *Empirical process in M-estimation*, Cambridge University Press, New-York, 2000.
31. A. W. Van der Vaart and J. A. Wellner, *Weak convergence and empirical processes with applications to statistics*, Springer Series in Statistics, Springer, 1996.

Wasserfuristrasse 42, 8542 Wiesendangen, Switzerland

Current address: Zurich Cantonal Bank, P.O. Box, CH-8010 Zurich, Switzerland. Phone: +41 1 292 45 33

E-mail address: daniel.egloff@zkb.ch

Department of Mathematics, University of Saarbrücken, Postfach 151150, D-66041 Saarbrücken, Germany

E-mail address: kohler@math.uni-sb.de

Department of Mathematics, University of Saarbrücken, Postfach 151150, D-66041 Saarbrücken, Germany

E-mail address: todorovic@math.uni-sb.de