Pathwise Dynamic Programming

Christian Bender[1,2], Christian Gärtner[2], and Nikolaus Schweizer[3]

December 20, 2016

Abstract

We present a novel method for deriving tight Monte Carlo confidence intervals for solutions of stochastic dynamic programming equations. Taking some approximate solution to the equation as an input, we construct pathwise recursions with a known bias. Suitably coupling the recursions for lower and upper bounds ensures that the method is applicable even when the dynamic program does not satisfy a comparison principle. We apply our method to three nonlinear option pricing problems, pricing under bilateral counterparty risk, under uncertain volatility, and under negotiated collateralization.

Keywords: stochastic dynamic programming; Monte Carlo; confidence bounds; option pricing

MSC 2000: 90C39, 65C05, 90C46, 91B28.

1 Introduction

We study the problem of computing Monte Carlo confidence intervals for solutions of discrete-time, finite horizon stochastic dynamic programming equations. There is a random terminal value, and a nonlinear recursion which allows to compute the value at a given time from expectations about the value one step ahead. Equations of this type are highly prominent in the analysis of multistage sequential decision problems under uncertainty (Bertsekas, 2005; Powell, 2011). Yet they also arise in a variety of further applications such as financial option pricing (e.g. Guyon and Henry-Labordère, 2013), the evaluation of recursive utility functionals (e.g. Kraft and Seifried, 2014) or the numerical solution of partial differential equations (e.g. Fahim et al., 2011).

The key challenge when computing solutions to stochastic dynamic programs numerically stems from a high order nesting of conditional expectations operators in the backward recursion. The solution at each time step depends on an expectation of what happens one step ahead, which in turn depends on expectations of what happens at later dates. If the system is driven by a

[1]Corresponding author.

[2]Saarland University, Department of Mathematics, Postfach 151150, D-66041 Saarbrücken, Germany, bender@math.uni-sb.de; gaertner@math.uni-sb.de.

[3]Tilburg University, Department of Econometrics and OR, NL-5000 LE Tilburg, PO box 90153, n.f.f.schweizer@uvt.nl.

Markov process with a high-dimensional state space – as is the case in typical applications – a naive numerical approach quickly runs into the curse of dimensionality. In practice, conditional expectations thus need to be replaced by an approximate conditional expectations operator which can be nested several times at moderate computational costs, e.g., by a least-squares Monte Carlo approximation. For an introduction and overview of this “approximate dynamic programming” approach, see, e.g., Powell (2011). The error induced by these approximations is typically hard to quantify and control. This motivates us to develop a posteriori criteria which use an approximate or heuristic solution to the dynamic program as an input in order to compute tight upper and lower bounds on the true solution. In this context, “tight” means that the bounds match when the input approximation coincides with the true solution.

Specifically, we study stochastic dynamic programming equations of the form

$$Y_j = G_j(E_j[\beta_{j+1}Y_{j+1}], F_j(E_j[\beta_{j+1}Y_{j+1}])) \tag{1}$$

with random terminal condition $Y_J = \xi$. Here, $E_j[\cdot]$ denotes the conditional expectation given the information at time step j. F_j and G_j are real-valued, adapted, random functions. F_j is convex while G_j is concave and increasing in its second argument. In particular, a possible dependence on state and decision variables is modeled implicitly through the functions F_j and G_j. The stochastic weight vectors β_j form an adapted process, i.e., unlike F_j and G_j, β_{j+1} is known at time $j+1$ but not yet at time j. Such weight processes frequently arise in dynamic programs for financial option valuation problems where, roughly speaking, they are associated with the first two derivatives, Delta and Gamma, of the value process Y. Similarly, Fahim et al. (2011) derive a probabilistic scheme for fully nonlinear second-order parabolic partial differential equations which involves stochastic weights in the approximation of the space derivatives. This scheme is of the form (1), when the nonlinearity has a concave-convex structure. This is the case for Hamilton-Jacobi-Bellman equations arising from continuous-time stochastic control problems and for some Hamilton-Jacobi-Isaacs equations arising from stochastic differential games.

There are two main motivations for assuming the concave-convex structure of the right hand side of (1). First, since convexity and concavity may stem, respectively, from maximization or minimization, the equation is general enough to cover not only dynamic programs associated with classical optimization problems but also dynamic programs arising from certain stochastic two-player games or robust optimization problems under model uncertainty. Second, many functions which are themselves neither concave nor convex may be written in this concave-convex form. This can be seen, for instance, in the application to bilateral counterparty credit risk in Section 7.1 below.

Traditionally, a posteriori criteria for this type of recursion were not derived directly from equation (1) but rather from a primal-dual pair of optimization problems associated with it. For instance, with the choices $G_j(z, y) = y$, $F_j(z) = \max\{S_j, z\}$, $\beta_j = 1$ and $\xi = S_J$, (1) becomes the dynamic programming equation associated with the optimal stopping problem for a process $(S_j)_j$.

From an approximate solution of (1) one can conclude an approximation of the optimal stopping strategy. Testing this strategy by simulation yields lower bounds on the value of the optimal stopping problem. The derivation of dual upper bounds starts by considering strategies which allow to look into the future, i.e., by considering strategies which allow to optimize pathwise rather than in conditional expectation. The best strategy of this type is easy to simulate and implies an upper bound on the value of the usual optimal stopping problem. This bound can be made tight by penalizing the use of future information in an appropriate way. Approximately optimal information penalties can be derived from approximate solutions to (1). Combining the resulting low-biased and high-biased Monte Carlo estimators for Y_0 yields Monte Carlo confidence intervals for the solution of the dynamic program. This information relaxation approach was developed in the context of optimal stopping independently by Rogers (2002) and Haugh and Kogan (2004). The approach was extended to general discrete-time stochastic optimization problems by Brown et al. (2010) and Rogers (2007).

Recently, Bender et al. (2015) have proposed a posteriori criteria which are derived directly from a dynamic programming recursion like (1). The obvious advantage is that, in principle, this approach requires neither knowledge nor existence of associated primal and dual optimization problems. This paper generalizes the approach of Bender et al. (2015) in various directions. For example, we merely require that the functions F_j and G_j are of polynomial growth while the corresponding condition in the latter paper is Lipschitz continuity. Moreover, we assume that the weights β_j are sufficiently integrable rather than bounded, and introduce the concave-convex functional form of the right hand side of (1) which is much more flexible than the particular convex functional form considered there.

Conceptually, our main departure from the approach of Bender et al. (2015) is that we do not require that the recursion (1) satisfies a comparison principle. Suppose that two adapted processes Y^{low} and Y^{up} fulfill the analogs of (1) with "$=$" replaced, respectively, by "$\leq$" and "$\geq$". We call such processes subsolutions and supersolutions to (1). The comparison principle postulates that subsolutions are always smaller than supersolutions. Relying on such a comparison principle, the approach of Bender et al. (2015) mimics the classical situation in the sense that their lower and upper bounds can be interpreted as stemming, respectively, from a primal and a dual optimization problem constructed from (1).

The bounds of the present paper apply regardless of whether a comparison principle holds or not. This increased applicability is achieved at negligible additional numerical costs. The key idea is to construct a pathwise recursion associated with particular pairs of super- and subsolutions. These pairs remain ordered even when such an ordering is not a generic property of super- and subsolutions, i.e., when comparison is violated in general. Roughly speaking, whenever a violation of comparison threatens to reverse the order of the upper and lower bound, the upper and lower bound are simply exchanged on the respective path. Consequently, lower and upper bounds must always be computed simultaneously. In particular, they can no longer be viewed as the respective solutions of distinct primal and dual optimization problems.

For some applications like optimal stopping, the comparison principle is not an issue. For others like the nonlinear pricing applications studied in Bender et al. (2015), the principle can be set in force by a relatively mild truncation of the stochastic weights β. Dispensing with the comparison principle allows us to avoid any truncation and the corresponding truncation error. This matters because there is also a class of applications where the required truncation levels are so low that truncation would fundamentally alter the problem. This concerns, in particular, backward recursions associated with fully nonlinear second-order partial differential equations, when the space derivatives are approximated in a probabilistic way by so-called Malliavin Monte-Carlo weights, see Fahim et al. (2011). Here, it may easily happen that, due to the application of the gamma weight for the approximation of the second space derivative, the discrete time version fails a comparison principle, although the comparison principle is satisfied (in the sense of viscosity solutions) by the continuous time partial differential equation. We explain this phenomenon in more detail for the problem of pricing under volatility uncertainty which serves as a running example throughout the paper. Solving equation (1) for the uncertain volatility model is well-known as a challenging numerical problem (Guyon and Henry-Labordère, 2011; Alanko and Avellaneda, 2013), making an a posteriori evaluation of solutions particularly desirable. Summarizing, by dispensing with the comparison principle, the setting of the present paper covers certain probabilistic discretizations for Hamilton-Jacobi-Bellman equations arising in stochastic control, where the controls may influence the drift and the diffusion part, while the setting in Bender et al. (2015) is tailored to construct confidence bounds for the corresponding discrete time equations in the special case of stochastic drift control.

We finally extend the results to systems of dynamic programming equations with convex nonlinearities. As a special case we recover the martingale dual for multiple stopping due to Schoenmakers (2012), but, more importantly, our results also allow to compute tight bounds for probabilistic discretization schemes of systems of parabolic equations. These discretization schemes can be seen to violate a componentwise comparison principle, unless there is no coupling through the space derivatives and the coupling of one component of the solution into the other component equations is nondecreasing. Thus, by avoiding the comparison principle in our approach, we can remove these restrictive assumptions on the coupling of one equation into the others. As an example, we can compute tight bounds for the pricing problem of collateralized contracts under funding costs in the theoretical framework of Nie and Rutkowski (2016).

The paper is organized as follows: Section 2 provides further motivation for the stochastic dynamic programming equation (1) by discussing how it arises in a series of examples. Section 3 states our setting and key assumptions in detail and explores a partial connection between equation (1) and a class of two-player games which makes some of the constructions of the later sections easier to interpret. Section 4 presents our main results. We first discuss the restrictiveness of assuming the comparison principle in the applications we are interested in. Afterwards, we proceed to Theorem 4.5, the key result of the paper, providing a new pair of upper and lower bounds which is applicable in the absence of a comparison principle. In Section 5, we relate our results to the information relaxation duals of Brown et al. (2010). In particular, we show that our bounds in the *presence* of the comparison principle can be reinterpreted in terms of information relaxation bounds for stochastic two-player zero-sum games, complementing recent work by Haugh and Wang (2015). In Section 6, we discuss how to extend Theorem 4.5 to situations where the dynamic programming equation is itself multidimensional. Finally, in Section 7, we apply our results in the context of three nonlinear valuation problems, option pricing in a four-dimensional interest rate model with bilateral counterparty risk, in the uncertain volatility model, and in a five-dimensional Black-Scholes model with negotiated collateralization. While such nonlinearities in pricing have received increased attention since the financial crisis, tight Monte Carlo confidence bounds for these three problems were previously unavailable. All proofs are postponed to the Appendix.

2 Examples

In this section, we briefly discuss several examples of backward dynamic programming equations arising in option pricing problems, which are covered by the framework of our paper. These include nonlinearities due to early exercise features, counterparty risk, and model uncertainty.

Example 2.1 (Bermudan options). Our first example is the optimal stopping problem, or, in financial terms, the pricing problem of a Bermudan option. Given a stochastic process S_j, $j = 0, \ldots, J$, which is adapted to the information given in terms of a filtration on an underlying filtered probability space, one wishes to maximize the expected reward from stopping S, i.e.,

$$Y_0 := \sup_{\tau} E[S_\tau], \tag{2}$$

where τ runs through the set of $\{0, \ldots, J\}$-valued stopping times. If the expectation is taken with respect to a pricing measure, under which all tradable and storable securities in an underlying financial market are martingales, then Y_0 is a fair price of a Bermudan option with discounted payoff process given by S. It is well-known and easy to check that Y_0 is the initial value of the dynamic program

$$Y_j = \max\{E_j[Y_{j+1}], S_j\}, \quad Y_J = S_J, \tag{3}$$

which is indeed of the form (1). In the traditional primal-dual approach, the primal maximization problem (2) is complemented by a dual minimization problem. This is the information relaxation dual which was initiated for this problem independently by works of Rogers (2002) and Haugh and Kogan (2004). It states that

$$Y_0 = \inf_M E\left[\sup_{j=0,\ldots,J}(S_j - M_j)\right], \tag{4}$$

where M runs over the set of martingales which start in zero at time zero. In order to illustrate our pathwise dynamic programming approach, we shall present an alternative proof of this dual representation for optimal stopping in Example 4.2 below, which, in contrast to the arguments in Rogers (2002) and Haugh and Kogan (2004), does not make use of specific properties of the optimal stopping problem (2), but rather relies on the convexity of the max-operator in the associated dynamic programming equation (3).

Example 2.2 (Credit value adjustment). In order to explain the derivation of a discrete-time dynamic programming equation of the form (1) for pricing under counterparty risk, we restrict ourselves to a simplified setting, but refer, e.g., to the monographs by Brigo et al. (2013) or Crépey et al. (2014) for more realistic situations: A party A and a counterparty B trade several derivatives, all of which mature at the same time T. The random variable ξ denotes the (possibly negative) cumulative payoff of the basket of derivatives which A receives at time T. We assume that ξ is measurable with respect to the market's reference information which is given by a filtration $(\mathcal{F}_t)_{0\le t\le T}$, and that only the counterparty B can default. Default occurs at an exponential waiting time τ with parameter γ which is independent of the reference filtration. The full filtration $(\mathcal{G}_t)_{0\le t\le T}$ is the smallest one which additionally to the reference information contains observation of the default event, i.e. it satisfies $\{\tau \le t\} \in \mathcal{G}_t$ for every $t \in [0, T]$. A recovery scheme $X = (X_t)_{0\le t\le T}$ describes the (possibly negative) amount of money which A receives, if B defaults at time t. The process X is assumed to be adapted to the reference filtration. We denote by $D(t, s; \kappa) = e^{-\kappa(s-t)}$, $t \le s$, the discount factor for a rate κ and stipulate that the constant $r \ge 0$ is a good proxy to a risk-free rate. Then, standard calculations in this reduced form approach (e.g. Duffie et al., 1996) show that the fair price for these derivatives under default risk at time t is given by

$$\begin{aligned} V_t &= \mathbf{1}_{\{\tau>t\}} E\left[\mathbf{1}_{\{\tau>T\}} D(t,T,r)\xi + \mathbf{1}_{\{\tau\le T\}} D(t,\tau,r) X_\tau \,\middle|\, \mathcal{G}_t\right] \\ &= \mathbf{1}_{\{\tau>t\}} E\left[D(t,T,r+\gamma)\left(\xi + \int_t^T D(s,T,-(r+\gamma))\gamma X_s ds\right)\middle|\, \mathcal{F}_t\right] =: \mathbf{1}_{\{\tau>t\}}\mathcal{Y}_t, \end{aligned} \tag{5}$$

where expectation is again taken with respect to a pricing measure. We assume that, upon default, the contract (consisting of the basket of derivatives) is replaced by the same one with a new counterparty which has not yet defaulted but is otherwise identical to the old counterparty B. This suggests the choice $X_t = -(\mathcal{Y}_t)_- + \mathfrak{r}(\mathcal{Y}_t)_+ = \mathcal{Y}_t - (1-\mathfrak{r})(\mathcal{Y}_t)_+$ for some recovery rate $\mathfrak{r} \in [0, 1]$. Here $(\cdot)_\pm$ denote the positive part and the negative part, respectively. Substituting this definition of X into (5) and integrating by parts, leads to the following nonlinear backward stochastic differential equation:

$$\mathcal{Y}_t = E\left[\xi - \int_t^T r\mathcal{Y}_s + \gamma(1-\mathfrak{r})(\mathcal{Y}_s)_+ ds \middle| \mathcal{F}_t\right]. \tag{6}$$

Note that, by Proposition 3.4 in El Karoui et al. (1997), $\mathcal{Y}_0 = \inf_\rho E\left[e^{-\int_0^T \rho_s ds}\xi\right]$, where ρ runs over the set of adapted processes with values in $[r, r+\gamma(1-\mathfrak{r})]$. Thus, the party A can minimize the price by choosing the interest rate which is applied for the discounting within a range determined by the riskless rate r, the credit spread γ of the counterparty, and the recovery rate $\mathfrak{r}$.

Discretizing (6) in time, we end up with the backward dynamic programming equation

$$Y_j = (1-r\Delta)E[Y_{j+1}|\mathcal{F}_{t_j}] - \gamma(1-\mathfrak{r})\left(E[Y_{j+1}|\mathcal{F}_{t_j}]\right)_+\Delta, \quad Y_J = \xi,$$

where $t_j = jh$, $j = 0, \ldots, J$, and $\Delta = T/J$. This equation is of the form (1) with $\beta_j \equiv 1$, $G_j(z,y) = (1-r\Delta)z - \gamma(1-\mathfrak{r})\Delta\,(z)_+$, and $F_j(z) \equiv 1$. We note that, starting with the work by Zhang (2004), time discretization schemes for backward stochastic differential equations have been intensively studied, see, e.g., the literature overview in Bender and Steiner (2012).

If one additionally takes the default risk and the funding cost of the party A into account, the corresponding dynamic programming equation has both concave and convex nonlinearities, and Y_0 corresponds to the equilibrium value of a two-player game, see the discussion at the end of Section 3. This is one reason to consider the more flexible concave-convex form (1). A numerical example in the theoretical framework of Crépey et al. (2013) with random interest and default rates is presented in Section 7 below.

Example 2.3 (Uncertain volatility). Dynamic programming equations of the form (1) also appear in the discretization of Hamilton-Jacobi-Bellman equations for stochastic control problems, see, e.g., Fahim et al. (2011). We illustrate such an approach by the pricing problem of a European option on a single asset under uncertain volatility, but the underlying discretization methodology of Fahim et al. (2011) can, of course, be applied to a wide range of (multidimensional) stochastic control problems in continuous time.

We suppose that the price of the asset is given under risk-neutral dynamics and in discounted units by

$$X_t^\sigma = x_0 \exp\left\{\int_0^t \sigma_u dW_u - \frac{1}{2}\int_0^t \sigma_u^2 du\right\},$$

where W denotes a Brownian motion, and the volatility σ is a stochastic process which is adapted to the Brownian filtration $(\mathcal{F}_t)_{0\leq t\leq T}$. The pricing problem of a European option with payoff function g under uncertain volatility then becomes the optimization problem

$$\mathcal{Y}_0 = \sup_\sigma E[g(X_T^\sigma)], \tag{7}$$

where σ runs over the nonanticipating processes with values in $[\sigma_{low}, \sigma_{up}]$ for some given constants

$\sigma_{low} < \sigma_{up}$. This is the worst case price over all stochastic volatility processes ranging within the interval $[\sigma_{low}, \sigma_{up}]$, and thus reflects the volatility uncertainty. The study of this so-called uncertain volatility model was initiated by Avellaneda et al. (1995) and Lyons (1995).

The corresponding Hamilton-Jacobi-Bellman equation can easily be transformed in such a way that $\mathcal{Y}_0 = v(0,0)$ where v solves the fully nonlinear Cauchy problem

$$\begin{aligned} v_t(t,x) &= -\frac{1}{2}v_{xx}(t,x) - \max_{\sigma\in\{\sigma_{low},\sigma_{up}\}} \frac{1}{2}\left(\frac{\sigma^2}{\hat\rho^2}-1\right)(v_{xx}(t,x)-\hat\rho v_x(t,x)), \quad (t,x)\in[0,T)\times\mathbb{R} \\ v(T,x) &= g(x_0 e^{\hat\rho x - \frac{1}{2}\hat\rho^2 T}), \quad x\in\mathbb{R}, \end{aligned} \tag{8}$$

for any (constant) reference volatility $\hat\rho > 0$ of one's choice. Under suitable assumptions on g, there exists a unique classical solution to (8) (satisfying appropriate growth conditions), see Pham (2009). Let $0 = t_0 < t_1 < \ldots < t_J = T$ again be an equidistant partition of the interval $[0,T]$, where $J \in \mathbb{N}$, and set $\Delta = T/J$. We approximate v by an operator splitting scheme, which on each small time interval first solves the linear subproblem, i.e., a Cauchy problem for the heat equation, and then corrects for the nonlinearity by plugging in the solution of the linear subproblem. Precisely, for fixed J, let

$$\begin{aligned} y^J(x) &= g(x_0 e^{\hat\rho x - \frac{1}{2}\hat\rho^2 T}), \quad x\in\mathbb{R} \\ \bar y^j_t(t,x) &= -\frac{1}{2}\bar y^j_{xx}(t,x), \quad \bar y^j(t_{j+1},x) = y^{j+1}(x), \quad t\in[t_j,t_{j+1}),\ x\in\mathbb{R},\ j=J-1,\ldots,0 \\ y^j(x) &= \bar y^j(t_j,x) + \Delta \max_{\sigma\in\{\sigma_{low},\sigma_{up}\}} \frac{1}{2}\left(\frac{\sigma^2}{\hat\rho^2}-1\right)\left(\bar y^j_{xx}(t_j,x)-\hat\rho \bar y^j_x(t_j,x)\right), \quad x\in\mathbb{R}. \end{aligned}$$

Evaluating $y^j(x)$ along the Brownian paths leads to $Y_j := y^j(W_{t_j})$. Applying the Feynman-Kac representation for the heat equation, see e.g. Karatzas and Shreve (1991), repeatedly, one observes that $\bar y^j(t_j, W_{t_j}) = E_j[Y_{j+1}]$, where E_j denotes the expectation conditional on $\mathcal{F}_{t_j}$. It is well-known from Fournié et al. (1999) that the space derivatives of $\bar y^j(t_j,\cdot)$ can be expressed via the so-called Malliavin Monte-Carlo weights as

$$\bar y^j_x(t_j, W_{t_j}) = E_j\left[\frac{\Delta W_{j+1}}{\Delta}Y_{j+1}\right], \quad \bar y^j_{xx}(t_j, W_{t_j}) = E_j\left[\left(\frac{\Delta W^2_{j+1}}{\Delta^2}-\frac{1}{\Delta}\right)Y_{j+1}\right],$$

where $\Delta W_j = W_{t_j} - W_{t_{j-1}}$. (In our particular situation, one can verify this simply by re-writing the conditional expectations as integrals on $\mathbb{R}$ with respect to the Gaussian density and integrating by parts.) Hence, one arrives at the following discrete-time dynamic programming equation:

$$\begin{aligned} Y_J &= g(X^{\hat\rho}_T), \\ Y_j &= E_j[Y_{j+1}] + \Delta \max_{\sigma\in\{\sigma_{low},\sigma_{up}\}}\left(\frac{1}{2}\left(\frac{\sigma^2}{\hat\rho^2}-1\right)E_j\left[\left(\frac{\Delta W^2_{j+1}}{\Delta^2}-\hat\rho\frac{\Delta W_{j+1}}{\Delta}-\frac{1}{\Delta}\right)Y_{j+1}\right]\right), \end{aligned} \tag{9}$$

where $X_T^{\hat{\rho}}$ is the price of the asset at time T under the constant reference volatility $\hat{\rho}$. This type of time-discretization scheme was analyzed for a general class of fully nonlinear parabolic PDEs by Fahim et al. (2011). In the particular case of the uncertain volatility model, the scheme was suggested by Guyon and Henry-Labordère (2011) by a slightly different derivation.

Let $G_j(z,y) = y$, $s_\iota = \frac{1}{2}(\frac{\sigma_\iota^2}{\hat{\rho}^2} - 1)$ for $\iota \in \{up, low\}$, and $F_j(z) = z^{(1)} + \Delta \max_{s \in \{s_{low}, s_{up}\}} s z^{(2)}$, where $z^{(i)}$ denotes the i-th component of the two-dimensional vector z, and define the $\mathbb{R}^2$-valued process β by

$$\beta_j = \left(1, \frac{\Delta W_j^2}{\Delta^2} - \hat{\rho}\frac{\Delta W_j}{\Delta} - \frac{1}{\Delta}\right)^\top, \quad j = 1, \ldots, J,$$

where $(\cdot)^\top$ denotes matrix transposition. With these choices, the dynamic programming equation for Y in (9) is of the form (1). We emphasize that, by Example 4.4, the weight

$$1 + \Delta \frac{1}{2}\left(\frac{\sigma^2}{\hat{\rho}^2} - 1\right)\left(\frac{\Delta W_{j+1}^2}{\Delta^2} - \hat{\rho}\frac{\Delta W_{j+1}}{\Delta} - \frac{1}{\Delta}\right) \tag{10}$$

becomes negative with positive probability for any choice of the reference volatility $\hat{\rho}$, when $\sigma_{low} > \sigma_{up}/\sqrt{3}$. Then, by Theorem 4.3 below, the discrete time dynamic programming equation (9) fails a comparison principle for super- and subsolutions, although the limiting partial differential equation satisfies the comparison principle in the theory of viscosity solutions. As a second consequence of negative weights in (10), Y may fail to be the value process of a discrete time control problem. This is because the Hamilton-Jacobi-Bellman equation is discretized and not the continuous time control problem. Hence, in such situations, the general theory of information relaxation duals due to Brown et al. (2010) cannot be applied to construct upper bounds on Y (though, of course, it can be applied to direct discretizations of the control problem), but the pathwise dynamic programming approach, which we present in Section 4, still applies.

This discussion of the uncertain volatility model can be seen as prototypical for fully nonlinear parabolic PDEs of the form

$$-v_t(t,x) - H(t,x,v,v_x,v_{xx}) = 0, \quad (t,x) \in [0,T) \times \mathbb{R}^d, \quad v(T,\cdot) = h, \tag{11}$$

where $H : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d\times d} \to \mathbb{R}$. After applying the discretization scheme of Fahim et al. (2011), we end up with a dynamic programming equation of the form (1), if H is e.g. convex in v and in the space derivatives of v, which is the case for Hamilton-Jacobi-Bellman equations. Here the β-weight takes values in $\mathbb{R}^{1+d+d^2}$ and consists of the 1, the d delta weights for the first space derivatives, and the d^2 gamma weights for the second space derivatives. We emphasize that the comparison principle for the discretized equation does typically only hold, when the nonlinearity in the second derivative is mild, and it is particularly restrictive in high-dimensional problems. If, e.g., H depends only on the Laplacian of v and is increasing (to make the PDE parabolic) and convex, then a similar argument as in Example 4.4 below shows that either H is linear or the resulting discrete-time equation fails the comparison principle for every choice of the reference volatility, when the space dimension d is sufficiently large. We note that the consistency proof of the discretization scheme of Fahim et al. (2011), which basically only builds on Itô's formula, works under very mild assumptions. While assuming the discrete time comparison principle (resp. the equivalent monotonicity condition in the convex case, cp. Theorem 4.3) facilitates the proof of the numerical stability in Fahim et al. (2011), the numerical experiments in Fahim et al. (2011) and ours in Section 7 below indicate that the scheme is stable, if the reference volatility is sufficiently strong.

3 Setting

In this section, we introduce the general setting of the paper, into which the examples of the previous section are accommodated. We then explain, how the dynamic programming equation (1) relates to a class of stochastic two-player games, which have a sound interpretation from the point of view of option pricing. While establishing these connections has some intrinsic interest, their main motivation is to make the constructions and results of Section 4 more tangible.

Throughout the paper we study the following type of concave-convex dynamic programming equation on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_j)_{j=0,\dots J}, P)$ in discrete time:

$$\begin{aligned} Y_J &= \xi, \\ Y_j &= G_j(E_j[\beta_{j+1}Y_{j+1}], F_j(E_j[\beta_{j+1}Y_{j+1}])), \quad j = J-1, \dots, 0 \end{aligned} \tag{12}$$

where $E_j[\cdot]$ denotes the conditional expectation with respect to $\mathcal{F}_j$. We assume

(C): For every $j = 0, \dots, J-1$, $G_j : \Omega \times \mathbb{R}^D \times \mathbb{R} \to \mathbb{R}$ and $F_j : \Omega \times \mathbb{R}^D \to \mathbb{R}$ are measurable and, for every $(z, y) \in \mathbb{R}^D \times \mathbb{R}$, the processes $(j, \omega) \mapsto G_j(\omega, z, y)$ and $(j, \omega) \mapsto F_j(\omega, z)$ are adapted. Moreover, for every $j = 0, \dots, J-1$ and $\omega \in \Omega$, the map $(z, y) \mapsto G_j(\omega, z, y)$ is concave in (z, y) and non-decreasing in y, and the map $z \mapsto F_j(\omega, z)$ is convex in z.

(R):
- G and F are of polynomial growth in (z, y) in the following sense: There exist a constant $q \geq 0$ and a nonnegative adapted process (α_j) such that for all $(z, y) \in \mathbb{R}^{D+1}$ and $j = 0, \dots, J-1$

$$|G_j(z, y)| + |F_j(z)| \leq \alpha_j(1 + |z|^q + |y|^q), \quad P\text{-a.s.},$$

  and $\alpha_j \in L^p(\Omega, P)$ for every $p \geq 1$.
- $\beta = (\beta_j)_{j=1,\dots,J}$ is an adapted process such that $\beta_j \in L^p(\Omega, P)$ for every $p \geq 1$ and $j = 1, \dots, J$.
- The terminal condition ξ is an $\mathcal{F}_J$-measurable random variable such that $\xi \in L^p(\Omega, P)$ for every $p \geq 1$.

In the following, we abbreviate by

- $L^{\infty-}(\mathbb{R}^N)$ the set of $\mathbb{R}^N$-valued random variables that are in $L^p(\Omega, P)$ for all $p \geq 1$.
- $L_j^{\infty-}(\mathbb{R}^N)$ the set of $\mathcal{F}_j$-measurable random variables that are in $L^{\infty-}(\mathbb{R}^N)$.
- $L_{ad}^{\infty-}(\mathbb{R}^N)$ the set of adapted processes Z such that $Z_j \in L_j^{\infty-}(\mathbb{R}^N)$ for every $j = 0, \ldots, J$.

Thanks to the integrability assumptions on the terminal condition ξ and the weight process β and thanks to the polynomial growth condition on F and G, it is straightforward to check recursively that the (P-a.s. unique) solution Y to (12) belongs to $L_{ad}^{\infty-}(\mathbb{R})$ under assumptions (C) and (R). In order to simplify the exposition, we shall also use the following convention: Unless otherwise noted, all equations and inequalities are supposed to hold P-almost surely.

For our main results in Section 4, we rewrite equation (12) relying on convex duality techniques. In the remainder of this section, we thus recall some concepts from convex analysis and show that – in some cases – these techniques allow us to interpret the dynamic programming equation (12) in terms of a class of two-player games. We recall that the convex conjugate of F_j is given by

$$F_j^{\#}(u) := \sup_{z \in \mathbb{R}^D} (u^\top z - F_j(z)), \quad u \in \mathbb{R}^D.$$

Note that $F_j^{\#}$ can take the value $+\infty$ and that the maximization takes place pathwise, i.e., ω by ω. Analogously, for G_j the concave conjugate at $\left(v^{(1)}, v^{(0)}\right) \in \mathbb{R}^{D+1}$ can be defined in terms of the convex conjugate of $-G_j$ as

$$G_j^{\#}\left(v^{(1)}, v^{(0)}\right) := -(-G_j)^{\#}\left(-v^{(1)}, -v^{(0)}\right) = \inf_{(z,y) \in \mathbb{R}^{D+1}} \left(\left(v^{(1)}\right)^\top z + v^{(0)} y - G_j(z,y)\right),$$

which can take the value $-\infty$. By the Fenchel-Moreau theorem, e.g. in the form of Theorem 12.2 in Rockafellar (1970), one has, for every $j = 0, \ldots, J-1$, $z \in \mathbb{R}^D$, $y \in \mathbb{R}$, and every $\omega \in \Omega$

$$G_j(\omega, z, y) = \inf_{(v^{(1)}, v^{(0)}) \in \mathbb{R}^{D+1}} \left(\left(v^{(1)}\right)^\top z + v^{(0)} y - G_j^{\#}(\omega, v^{(1)}, v^{(0)})\right), \tag{13}$$

$$F_j(\omega, z) = \sup_{u \in \mathbb{R}^D} \left(u^\top z - F_j^{\#}(\omega, u)\right). \tag{14}$$

Here, the minimization and maximization can, of course, be restricted to the effective domains

$$D_{G_j^{\#}(\omega,\cdot)} = \{(v^{(1)}, v^{(0)}) \in \mathbb{R}^{D+1},\ G_j^{\#}(\omega, v^{(1)}, v^{(0)}) > -\infty\},\ D_{F_j^{\#}(\omega,\cdot)} = \{u \in \mathbb{R}^D,\ F_j^{\#}(\omega, u) < +\infty\}$$

of $G_j^{\#}$ and $F_j^{\#}$. Noting that $v^{(0)} \geq 0$ for $(v^{(1)}, v^{(0)}) \in D_{G_j^{\#}(\omega,\cdot)}$ by the monotonicity assumption on G in the y-variable, the dynamic programming equation (12) can be rewritten in the form

$$Y_j = \inf_{(v^{(1)}, v^{(0)}) \in D_{G_j^{\#}}} \sup_{u \in D_{F_j^{\#}}} \left(\left(v^{(1)} + v^{(0)} u\right)^\top E_j[\beta_{j+1} Y_{j+1}] - v^{(0)} F_j^{\#}(u) - G_j^{\#}(v^{(1)}, v^{(0)})\right),$$

which formally looks like the dynamic programming equation for a two-player game with a random ($\mathcal{F}_{j+1}$-measurable!) and controlled 'discount factor' $\left(v^{(1)}+v^{(0)}u\right)^\top \beta_{j+1}$ between time j and $j+1$.

To make this intuition precise, we next introduce sets of adapted 'controls' (or strategies) in terms of F and G by

$$\begin{aligned}
\mathcal{A}_j^F &= \left\{ (r_i)_{i=j,\ldots,J-1} \middle| r_i \in L_i^{\infty-}(\mathbb{R}^D),\ F_i^{\#}(r_i) \in L^{\infty-}(\mathbb{R}) \text{ for } i=j,\ldots,J-1 \right\}, \\
\mathcal{A}_j^G &= \left\{ \left(\rho_i^{(1)},\rho_i^{(0)}\right)_{i=j,\ldots,J-1} \middle| \left(\rho_i^{(1)},\rho_i^{(0)}\right) \in L_i^{\infty-}(\mathbb{R}^{D+1}), \right. \\
&\qquad \left. G_i^{\#}\left(\rho_i^{(1)},\rho_i^{(0)}\right) \in L^{\infty-}(\mathbb{R}) \text{ for } i=j,\ldots,J-1 \right\}. \qquad (15)
\end{aligned}$$

In Appendix A, we show that these sets of controls $\mathcal{A}_j^G$ and $\mathcal{A}_j^F$ are nonempty under the standing assumptions (C) and (R). For $j=0,\ldots,J$, we consider the following multiplicative weights (corresponding to the 'discounting' between time 0 and j)

$$w_j(\omega, v^{(1)}, v^{(0)}, u) = \prod_{i=0}^{j-1} (v_i^{(1)} + v_i^{(0)} u_i)^\top \beta_{i+1}(\omega) \qquad (16)$$

for $\omega \in \Omega$, $v^{(1)} = (v_0^{(1)}, \ldots, v_{J-1}^{(1)}) \in (\mathbb{R}^D)^J$, $v^{(0)} = (v_0^{(0)}, \ldots, v_{J-1}^{(0)}) \in \mathbb{R}^J$ and $u = (u_0, \ldots, u_{J-1}) \in (\mathbb{R}^D)^J$. Assuming (for simplicity) that $\mathcal{F}_0$ is trivial, and (crucially) that the positivity condition

$$(v^{(1)} + v^{(0)} u)^\top \beta_{j+1}(\omega) \geq 0 \qquad (17)$$

on the 'discount factors' holds for every $\omega \in \Omega$, $(v^{(1)}, v^{(0)}) \in D_{G_j^{\#}(\omega,\cdot)}$ and $u \in D_{F_j^{\#}(\omega,\cdot)}$, Y_0 is indeed the equilibrium value of a (in general, non-Markovian) stochastic two-player zero-sum game, namely,

$$\begin{aligned}
Y_0 &= \inf_{(\rho^{(1)},\rho^{(0)}) \in \mathcal{A}_0^G} \sup_{r \in \mathcal{A}_0^F} E\left[w_J(\rho^{(1)},\rho^{(0)},r)\xi - \sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},r) \left(\rho_j^{(0)} F_j^{\#}(r_j) + G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)}) \right) \right] \\
&= \sup_{r \in \mathcal{A}_0^F} \inf_{(\rho^{(1)},\rho^{(0)}) \in \mathcal{A}_0^G} E\left[w_J(\rho^{(1)},\rho^{(0)},r)\xi - \sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},r) \left(\rho_j^{(0)} F_j^{\#}(r_j) + G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)}) \right) \right]
\end{aligned}$$

as is shown in Appendix C.2. In financial terms, we may think of $w_j(\rho^{(1)},\rho^{(0)},r)$ as a discrete-time price deflator (or as an approximation of a continuous-time price deflator given in terms of a stochastic exponential which can incorporate both discounting in the real-world sense and a change of measure). Then, the first term in

$$E\left[w_J(\rho^{(1)},\rho^{(0)},r)\xi - \sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},r) \left(\rho_j^{(0)} F_j^{\#}(r_j) + G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)}) \right) \right]$$

corresponds to the fair price of an option with payoff ξ in the price system determined by the deflator $w_J(\rho^{(1)}, \rho^{(0)}, r)$, which is to be chosen by the two players. The choice may come with an additional running reward or cost which is formulated via the convex conjugates of F and G in the second term of the above expression. With this interpretation, Y_0 is the equilibrium price for an option with payoff ξ, on which the two players agree. Of course, the above stochastic two-player game degenerates to a stochastic maximization (resp. minimization) problem, if G (resp. F) is linear, and so one of the control sets becomes a singleton.

As multiplication with a negative weight changes minimization into maximization and vice versa, one cannot expect, in general, that the equilibrium value of the above two-player zero-sum game can be described by the dynamic programming equation (12) in absence of the positivity condition (17). Hence the validity of this game-theoretic interpretation of the dynamic programming equation (12) crucially depends on this positivity assumption. For most of the paper, we shall not assume the positivity condition (17), and hence go beyond the above game-theoretic setting. Nonetheless, by a slight abuse of terminology, we will also refer to members of the sets $\mathcal{A}_j^G$ and $\mathcal{A}_j^F$ as controls, when the positivity condition is violated.

4 Main results

In this section, we state and discuss the main contribution of the paper, the pathwise dynamic programming approach of Theorem 4.5, which avoids the use of nested conditional expectations, in order to construct tight upper and lower bounds on the solution Y_j to the dynamic program (12). Extending the ideas in Bender et al. (2015), this construction will be based on the concept of supersolutions and subsolutions to the dynamic program (12).

Definition 4.1. A process Y^{up} (resp. Y^{low}) $\in L^{\infty-}_{ad}(\mathbb{R})$ is called *supersolution* (resp. *subsolution*) to the dynamic program (12) if $Y_J^{up} \geq Y_J$ (resp. $Y_J^{low} \leq Y_J$) and for every $j = 0, \ldots, J-1$ it holds

$$Y_j^{up} \geq G_j([E_j[\beta_{j+1} Y_{j+1}^{up}], F_j(E_j[\beta_{j+1} Y_{j+1}^{up}])),$$

(and with ' $\geq$' replaced by ' $\leq$' for a subsolution).

Before we explain a general construction method for super- and subsolutions, we first illustrate how the dual representation for optimal stopping in Example 2.1 can be derived this way.

Example 4.2. Fix a martingale M. The dynamic version of (4) suggests to consider

$$\Theta_j^{up} := \max_{i=j,\ldots,J} S_i - (M_i - M_j), \quad Y_j^{up} = E_j[\Theta_j^{up}].$$

The anticipative discrete-time process Θ^{up} clearly satisfies the recursive equation (pathwise dynamic program)

$$\Theta_j^{up} = \max\{S_j, \Theta_{j+1}^{up} - (M_{j+1} - M_j)\}, \quad \Theta_J^{up} = S_J.$$

Hence, by convexity of the max-operator and Jensen's inequality, we obtain, thanks to the martingale property of M and the tower property of conditional expectation,

$$Y_j^{up} \geq \max\{S_j, E_j[\Theta_{j+1}^{up} - (M_{j+1} - M_j)]\} = \max\{S_j, E_j[Y_{j+1}^{up}]\}, \quad Y_J^{up} = S_J.$$

Thus, Y^{up} is a supersolution to the dynamic program (3). By the monotonicity of the max-operator, we observe by backward induction that $Y_j^{up} \geq Y_j$, because (assuming that the claim is already proved at time $j+1$)

$$Y_j^{up} \geq \max\{S_j, E_j[Y_{j+1}^{up}]\} \geq \max\{S_j, E_j[Y_{j+1}]\} = Y_j.$$

In particular, writing $Y^{up}(M)$ instead of Y^{up} in order to emphasize the dependence on the fixed martingale M, we get

$$Y_0 \leq \inf_M Y_0^{up}(M) = \inf_M E\left[\max_{i=0,\ldots,J} S_i - (M_i - M_0)\right]. \tag{18}$$

Defining M^* to be the Doob martingale of Y, i.e. $M_{i+1}^* - M_i^* = Y_{i+1} - E_i[Y_{i+1}]$, and $\Theta^{up,*} = \Theta^{up}(M^*)$, it is straightforward to check by backward induction, that, for every $j = J, \ldots, 0$, $\Theta_j^{up,*} = Y_j$, as (assuming again that the claim is already proved for time $j+1$)

$$\Theta_j^{up,*} = \max\{S_j, \Theta_{j+1}^{up,*} - (Y_{j+1} - E_j[Y_{j+1}])\} = \max\{S_j, E_j[Y_{j+1}]\} = Y_j.$$

This turns the inequality in (18) into an equality.

This alternative proof of the dual representation for optimal stopping is by no means shorter or simpler than the original one by Rogers (2002) and Haugh and Kogan (2004) which relies on the optional sampling theorem and the supermartingale property of Y. It comes, however, with the advantage that it generalizes immediately to dynamic programming equations which share the same convexity and monotonicity properties. Indeed, if $G_j(z, y) = y$ in our general setting and we, thus, consider the convex dynamic program

$$Y_j = F_j(E_j[\beta_{j+1} Y_{j+1}]), \quad j = J-1, \ldots, 0, \quad Y_J = \xi, \tag{19}$$

we can write the corresponding pathwise dynamic program

$$\Theta_j^{up} = F_j(\beta_{j+1}\Theta_{j+1}^{up} - (M_{j+1} - M_j)), \quad \Theta_J^{up} = \xi \tag{20}$$

for $M \in \mathcal{M}_D$, i.e., for D-dimensional martingales which are members of $L_{ad}^{\infty-}(\mathbb{R}^D)$. Then, the same argument as above, based on convexity and Jensen's inequality implies that $Y_j^{up} = E_j[\Theta_j^{up}]$ is a supersolution to the dynamic program (19). A similar construction of a subsolution can be built on classical Fenchel duality. Indeed, consider, for $r \in \mathcal{A}_0^F$, the linear pathwise dynamic program

$$\Theta_j^{low} = r_j^\top \beta_{j+1} \Theta_{j+1}^{up} - F_j^{\#}(r_j), \quad \Theta_J^{low} = \xi. \tag{21}$$

Then, by (14), it is easy to check that $Y_j^{low} = E_j[\Theta_j^{low}]$ defines a subsolution.

These observations already led to the construction of tight upper and lower bounds for discrete-time backward stochastic differential equations (BSDEs) with convex generators in Bender et al. (2015). In the latter paper, the authors exploit that in their setting the special form of F, stemming from a Lipschitz BSDE, implies – after a truncation of Brownian increments – a similar monotonicity property than in the Bermudan option case.

In the general setting of the present paper, we would have to assume the following comparison principle to ensure that the supersolution Y_j^{up} and the subsolution Y_j^{low} indeed constitute an upper bound and a lower bound for Y_j.

(Comp): For every supersolution Y^{up} and every subsolution Y^{low} to the dynamic program (12) it holds that

$$Y_j^{up} \geq Y_j^{low}, \quad P\text{-a.s.}, \quad \text{for every } j = 0, \ldots, J.$$

The following characterization shows that the comparison principle can be rather restrictive and is our motivation to remove it below through a more careful construction of suitable pathwise dynamic programs.

Theorem 4.3. *Suppose (R) and that the dynamic program is convex, i.e., (C) is in force with $G_j(z, y) = y$ for $j = 0, \ldots, J-1$. Then, the following assertions are equivalent:*
(a) The comparison principle (Comp) holds.
(b) For every $r \in \mathcal{A}_0^F$ the following positivity condition is fulfilled: For every $j = 0, \ldots, J-1$

$$r_j^\top \beta_{j+1} \geq 0, \quad P\text{-a.s.}$$

(c) For every $j = 0, \ldots, J-1$ and any two random variables $Y^{(1)}, Y^{(2)} \in L^{\infty-}(\mathbb{R})$ with $Y^{(1)} \geq Y^{(2)}$ P-a.s., the following monotonicity condition is satisfied:

$$F_j(E_j[\beta_{j+1} Y^{(1)}]) \geq F_j(E_j[\beta_{j+1} Y^{(2)}]), \quad P\text{-a.s.}$$

Theorem 4.3 provides three equivalent formulations for the comparison principle. Formulation (b) illustrates the restrictiveness of the principle most clearly. For instance, when β contains unbounded entries with random sign, comparison can only hold in degenerate cases. In some applications, problems of this type can be resolved by truncation of the weights at the cost of a small additional error. Yet in applications such as pricing under volatility uncertainty, the comparison principle may fail (even under truncation), as shown in the following example.

Example 4.4. Recall the setting and the dynamic programming equation (9) from Example 2.3. In this setting, thanks to condition (c) in Theorem 4.3, comparison boils down to the requirement that the prefactor

$$1 + s_j \left(\frac{\Delta W_{j+1}^2}{\Delta} - \hat{\rho}\Delta W_{j+1} - 1 \right)$$

of Y_{j+1} in equation (9) for Y_j is P-almost surely nonnegative for both of the feasible values of s_j,

$$s_j \in \left\{ \frac{1}{2}\left(\frac{\sigma_{low}^2}{\hat{\rho}^2} - 1 \right), \frac{1}{2}\left(\frac{\sigma_{up}^2}{\hat{\rho}^2} - 1 \right) \right\}.$$

For $s_j > 1$, this requirement is violated for realizations of ΔW_{j+1} sufficiently close to zero, while for $s_j < 0$ violations occur for sufficiently negative realizations of the Brownian increment – and this violation also takes place if one truncates the Brownian increments at $\pm const.\sqrt{\Delta}$ with an arbitrarily large constant. Consequently, we arrive at the necessary conditions $\hat{\rho} \leq \sigma_{low}$ and $\hat{\rho} \geq \sigma_{up}/\sqrt{3}$ for comparison to hold. For $\sigma_{low} = 0.1$ and $\sigma_{up} = 0.2$, the numerical test case in Guyon and Henry-Labordère (2011) and Alanko and Avellaneda (2013), these two conditions cannot hold simultaneously, ruling out the possibility of a comparison principle.

Our aim is, thus, to modify the above construction of supersolutions and subsolutions in such a way that comparison still holds for these particular pairs of supersolutions and subsolutions, although the general comparison principle (Comp) may be violated. At the same time, we generalize from the convex structure to the concave-convex structure.

To this end, we consider the following 'coupled' pathwise recursion. Given $(\rho^{(1)}, \rho^{(0)}) \in \mathcal{A}_j^G$, $r \in \mathcal{A}_j^F$ and $M \in \mathcal{M}_D$, define the (in general) nonadapted processes $\theta_i^{up} = \theta_i^{up}(\rho^{(1)}, \rho^{(0)}, r, M)$ and $\theta_i^{low} = \theta_i^{low}(\rho^{(1)}, \rho^{(0)}, r, M)$, $i = j, \ldots, J$, via the 'pathwise dynamic program'

$$\begin{aligned}
\theta_J^{up} &= \theta_J^{low} = \xi, \\
\theta_i^{up} &= \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1} \right)_+ \theta_{i+1}^{up} - \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1} \right)_- \theta_{i+1}^{low} - \left(\rho_i^{(1)}\right)^\top \Delta M_{i+1} \\
&\quad + \rho_i^{(0)} \max_{\iota \in \{up, low\}} F_i(\beta_{i+1}\theta_{i+1}^{\iota} - \Delta M_{i+1}) - G_i^{\#}\left(\rho_i^{(1)}, \rho_i^{(0)}\right) \\
\theta_i^{low} &= \min_{\iota \in \{up, low\}} G_i \Big(\beta_{i+1}\theta_{i+1}^{\iota} - \Delta M_{i+1}, \left(r_i^\top \beta_{i+1}\right)_+ \theta_{i+1}^{low} - \left(r_i^\top \beta_{i+1}\right)_- \theta_{i+1}^{up} \\
&\quad - r_i^\top \Delta M_{i+1} - F_i^{\#}(r_i) \Big), \qquad (22)
\end{aligned}$$

where $\Delta M_i = M_i - M_{i-1}$. We emphasize that these two recursion formulas are coupled, as θ_{i+1}^{low} enters the defining equation for θ_i^{up} and θ_{i+1}^{up} enters the defining equation for θ_i^{low}. Roughly speaking, the rationale of this coupled recursion is to replace θ_{j+1}^{up} by θ_{j+1}^{low} in the upper bound recursion at time j, whenever the violation of the comparison principle threatens to reverse the order between upper bound recursion and lower bound recursion. Due to this coupling the elementary argument based on convexity and Fenchel duality outlined at the beginning of this section does not apply anymore, but a careful analysis is required to disentangle the influences of the upper and lower bound recursion (see the proofs in the appendix).

Our main result on the construction of tight upper and lower bounds for the concave-convex dynamic program (12) in absence of the comparison principle now reads as follows:

Theorem 4.5. *Suppose (C) and (R). Then, for every $j = 0, \dots, J$,*

$$\begin{aligned} Y_j &= \operatorname*{essinf}_{(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_j^G,\, r\in\mathcal{A}_j^F,\, M\in\mathcal{M}_D} E_j[\theta_j^{up}(\rho^{(1)},\rho^{(0)},r,M)] \\ &= \operatorname*{esssup}_{(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_j^G,\, r\in\mathcal{A}_j^F,\, M\in\mathcal{M}_D} E_j[\theta_j^{low}(\rho^{(1)},\rho^{(0)},r,M)], \quad P\text{-a.s.} \end{aligned}$$

For any $(\rho^{(1)},\rho^{(0)}) \in \mathcal{A}_j^G$, $r \in \mathcal{A}_j^F$, and $M \in \mathcal{M}_D$, we have the P-almost sure relation

$$\theta_i^{low}(\rho^{(1)},\rho^{(0)},r,M) \le \theta_i^{up}(\rho^{(1)},\rho^{(0)},r,M) \tag{23}$$

for every $i = j, \dots, J$. Moreover,

$$Y_j = \theta_j^{up}(\rho^{(1,*)},\rho^{(0,*)},r^*,M^*) = \theta_j^{low}(\rho^{(1,*)},\rho^{(0,*)},r^*,M^*) \tag{24}$$

P-almost surely, for every $(\rho^{(1,*)},\rho^{(0,*)}) \in \mathcal{A}_j^G$ and $r^* \in \mathcal{A}_j^F$ satisfying the duality relations

$$\begin{aligned} \left(\rho_i^{(1,*)}\right)^\top E_i\left[\beta_{i+1}Y_{i+1}\right] + \rho_i^{(0,*)} F_i\left(E_i[\beta_{i+1}Y_{i+1}]\right) - G_i^{\#}\left(\rho_i^{(1,*)},\rho_i^{(0,*)}\right) \\ = G_i\left(E_i\left[\beta_{i+1}Y_{i+1}\right], F_i(E_i[\beta_{i+1}Y_{i+1}])\right) \end{aligned} \tag{25}$$

and

$$\left(r_i^*\right)^\top E_i\left[\beta_{i+1}Y_{i+1}\right] - F_i^{\#}\left(r_i^*\right) = F_i\left(E_i\left[\beta_{i+1}Y_{i+1}\right]\right) \tag{26}$$

P-almost surely for every $i = j, \dots, J-1$, and with M^* being the Doob martingale of βY, i.e.,

$$M_k^* = \sum_{i=0}^{k-1} \beta_{i+1}Y_{i+1} - E_i[\beta_{i+1}Y_{i+1}], \quad P\text{-a.s., for every } k = 0, \dots J.$$

We note that, by Lemma A.1, optimizers $(\rho^{(1,*)},\rho^{(0,*)}) \in \mathcal{A}_j^G$ and $r^* \in \mathcal{A}_j^F$ which solve the duality relations (25) and (26) do exist. These duality relations also provide some guidance on how to choose nearly optimal controls in numerical implementations of the method: When an approximate solution $\tilde{Y}$ of the dynamic program is available, controls can be chosen such that they (approximately) fulfill the analogs of (25) and (26) with $\tilde{Y}$ in place of the unknown true solution. Likewise, M can be chosen as (an approximate) Doob martingale of $\beta\tilde{Y}$. Moreover, the almost sure optimality property in (24) suggests that (as is the case for the dual upper bounds for optimal stopping) a Monte Carlo implementation benefits from a low variance, when the approximate solution $\tilde{Y}$ to the dynamic program (12) is sufficiently close to the true solution Y.

As shown in the next proposition, the processes $E_j[\theta_j^{up}]$ and $E_j[\theta_j^{low}]$ in Theorem 4.5 indeed define super- and subsolutions to the dynamic program (12) which are ordered even though the comparison principle need not hold:

Proposition 4.6. *Under the assumptions of Theorem 4.5, the processes Y^{up} and Y^{low} given by $Y_j^{up} = E_j[\theta_j^{up}(\rho^{(1)}, \rho^{(0)}, r, M)]$ and $Y_j^{low} = E_j[\theta_j^{low}(\rho^{(1)}, \rho^{(0)}, r, M)]$, $j = 0, \ldots, J$ are, respectively, super- and subsolutions to (12) for every $(\rho^{(1)}, \rho^{(0)}) \in \mathcal{A}_0^G$, $r \in \mathcal{A}_0^F$, and $M \in \mathcal{M}_D$.*

We close this section by stating a simplified and decoupled pathwise dynamic programming approach, which can be applied if the comparison principle is known to be in force.

Theorem 4.7. *Suppose (C), (R), and (Comp) and fix $j = 0, \ldots, J$. For every $(\rho^{(1)}, \rho^{(0)}) \in \mathcal{A}_j^G$ and $M \in \mathcal{M}_D$ define $\Theta_j^{up} = \Theta_j^{up}(\rho^{(1)}, \rho^{(0)}, M)$ by the pathwise dynamic program*

$$\Theta_i^{up} = \left(\rho_i^{(1)}\right)^\top \left(\beta_{i+1}\Theta_{i+1}^{up} - (M_{i+1} - M_i)\right) + \rho_i^{(0)} F_i(\beta_{i+1}\Theta_{i+1}^{up} - (M_{i+1} - M_i)) - G_i^{\#}(\rho_i^{(1)}, \rho_i^{(0)}),$$

for $i = J-1, \ldots, j$, initiated at $\Theta_J^{up} = \xi$. Moreover, for $r \in \mathcal{A}_j^F$ and $M \in \mathcal{M}_D$ define $\Theta_j^{low} = \Theta_j^{low}(r, M)$ for $i = J-1, \ldots, j$ by the pathwise dynamic program

$$\Theta_i^{low} = G_i\left(\beta_{i+1}\Theta_{i+1}^{low} - (M_{i+1} - M_i), r_i^\top \left(\beta_{i+1}\Theta_{i+1}^{low} - (M_{i+1} - M_i)\right) - F_i^{\#}(r_i)\right)$$

initiated at $\Theta_J^{low} = \xi$. Then,

$$Y_j = \operatorname*{essinf}_{(\rho^{(1)}, \rho^{(0)}) \in \mathcal{A}_j^G,\, M \in \mathcal{M}_D} E_j[\Theta_j^{up}(\rho^{(1)}, \rho^{(0)}, M)] = \operatorname*{esssup}_{r \in \mathcal{A}_j^F,\, M \in \mathcal{M}_D} E_j[\Theta_j^{low}(r, M)], \; P\text{-a.s.,}$$

Moreover,

$$Y_j = \Theta_j^{up}(\rho^{(1,*)}, \rho^{(0,*)}, M^*) = \Theta_j^{low}(r^*, M^*) \tag{27}$$

P-almost surely, for every $(\rho^{(1,*)}, \rho^{(0,*)}) \in \mathcal{A}_j^G$ and $r^* \in \mathcal{A}_j^F$ satisfying the duality relations (25)–(26) and with M^* being the Doob martingale of βY.

Remark 4.8. Assume that the dynamic programming equation (12) is convex, i.e. $G_j(z, y) = y$ for every $j = 0, \ldots, J-1$. Then, it is straightforward to check that $E_j[\Theta_j^{low}(\rho^{(1)}, \rho^{(0)}, M)]$ does not depend on the choice of the martingale M in Theorem 4.7, i.e. M merely acts as a control variate for the lower bound. In the language of information relaxation duals, this can be rephrased as using dual feasible penalties (which satisfy a suitable martingale property) as control variates in the primal problem and has been suggested in different contexts, see e.g. Remark 3.6 in Bender et al. (2015) and Section 2.3.1 in Brown and Haugh (2016). In contrast, the proof of Theorem 4.5 crucially depends on the pathwise comparison property (23), which requires that the same martingale is applied in the recursion for θ^{up} and θ^{low}.

5 Relation to information relaxation duals

In this section, we relate a special case of our pathwise dynamic programming approach to the information relaxation dual for the class of stochastic two-player games discussed in Section 3. Throughout the section, we assume that the positivity condition (17) holds and that no information is available at time zero, i.e., $\mathcal{F}_0$ is trivial. Then, by virtue of Proposition B.1, the comparison principle (Comp) is in force and we can apply the simplified pathwise dynamic programming approach of Theorem 4.7.

Under assumption (17), Y_0 is the equilibrium value of a stochastic two-player zero-sum game, namely,

$$\begin{aligned} Y_0 &= \inf_{(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_0^G} \sup_{r\in\mathcal{A}_0^F} E\left[w_J(\rho^{(1)},\rho^{(0)},r)\xi - \sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},r)\left(\rho_j^{(0)}F_j^{\#}(r_j) + G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)})\right)\right] \\ &= \sup_{r\in\mathcal{A}_0^F} \inf_{(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_0^G} E\left[w_J(\rho^{(1)},\rho^{(0)},r)\xi - \sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},r)\left(\rho_j^{(0)}F_j^{\#}(r_j) + G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)})\right)\right], \end{aligned}$$

as seen in Section 3 where the multiplicative weights w_j are defined in (16). Now, assume that player 1 fixes her control $(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_0^G$. Then,

$$Y_0 \le \sup_{r\in\mathcal{A}_0^F} E\left[w_J(\rho^{(1)},\rho^{(0)},r)\xi - \sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},r)\left(\rho_j^{(0)}F_j^{\#}(r_j) + G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)})\right)\right].$$

One can next apply the information relaxation dual of Brown et al. (2010) (with strong duality), which states that the maximization problem on the right-hand side equals

$$\inf_{\mathfrak{p}} E\left[\sup_{u\in(\mathbb{R}^D)^J} \left(w_J(\rho^{(1)},\rho^{(0)},u)\xi - \sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},u)\left(\rho_j^{(0)}F_j^{\#}(u_j) + G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)})\right) - \mathfrak{p}(u)\right)\right],$$

where the infimum runs over the set of all dual-feasible penalties, i.e., all mappings $\mathfrak{p}:\Omega\times(\mathbb{R}^D)^J\to\mathbb{R}\cup\{+\infty\}$ which satisfy $E[\mathfrak{p}(r)]\le 0$ for every $r\in\mathcal{A}_0^F$, i.e., for every adapted and admissible control r of player 2. Notice that, for every martingale $M\in\mathcal{M}_D$, the map

$$\mathfrak{p}_M : u \mapsto \sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},u)(\rho_j^{(1)} + \rho_j^{(0)}u_j)^\top \Delta M_{j+1}$$

is a dual-feasible penalty, since, for adapted controls $r\in\mathcal{A}_0^F$,

$$E[\mathfrak{p}_M(r)] = \sum_{j=0}^{J-1} E\left[w_j(\rho^{(1)},\rho^{(0)},r)(\rho_j^{(1)} + \rho_j^{(0)}r_j)^\top E_j[\Delta M_{j+1}]\right] = 0$$

by the martingale property of M and the tower property of the conditional expectation. We shall refer to penalties of that particular form as martingale penalties. It turns out (see Appendix C.1 for the detailed argument) that, given the fixed controls $(\rho^{(1)}, \rho^{(0)}) \in \mathcal{A}_0^G$ of player 1, one has, for every martingale $M \in \mathcal{M}_D$,

$$\begin{aligned} &\Theta_0^{up}(\rho^{(1)}, \rho^{(0)}, M) \qquad (28) \\ = &\sup_{u \in (\mathbb{R}^D)^J} \left(w_J(\rho^{(1)}, \rho^{(0)}, u)\xi - \sum_{j=0}^{J-1} w_j(\rho^{(1)}, \rho^{(0)}, u) \left(\rho_j^{(0)} F_j^{\#}(u_j) + G_j^{\#}(\rho_j^{(1)}, \rho_j^{(0)}) \right) - \mathfrak{p}_M(u) \right). \end{aligned}$$

Hence, Theorem 4.7 implies

$$\begin{aligned} Y_0 = &\inf_{(\rho^{(1)}, \rho^{(0)}) \in \mathcal{A}_j^G,\, M \in \mathcal{M}_D} E\Bigg[\sup_{u \in (\mathbb{R}^D)^J} \Bigg(w_J(\rho^{(1)}, \rho^{(0)}, u)\xi \\ &- \sum_{j=0}^{J-1} w_j(\rho^{(1)}, \rho^{(0)}, u) \left(\rho_j^{(0)} F_j^{\#}(u_j) + G_j^{\#}(\rho_j^{(1)}, \rho_j^{(0)}) \right) - \mathfrak{p}_M(u) \Bigg) \Bigg] \end{aligned}$$

i.e., one still obtains strong duality, when the minimization is restricted from the set of all dual-feasible penalties to the subset of martingale penalties $\{\mathfrak{p}_M;\ M \in \mathcal{M}_D\}$. Summarizing, when assuming the positivity condition (17), we can interpret the upper bound $E[\Theta_0^{up}(\rho^{(1)}, \rho^{(0)}, M)]$ in such a way that, first, player 1 fixes her strategy $(\rho^{(1)}, \rho^{(0)})$ and the penalty by the choice of the martingale M, while, then, the player 2 is allowed to maximize the penalized problem pathwise. For the analogous representation of the lower bound $E[\Theta_0^{low}(r, M)]$, we refer to Appendix C.1.

In this way, we end up with the information relaxation dual of Brown et al. (2010) for each player given that the other player has fixed a control, but with the additional twist that, for our particular class of (possibly non-Markovian) two-player games, strong duality still holds for the subclass of martingale penalties. The general procedure explained above is analogous to the recent information relaxation approach by Haugh and Wang (2015) for two-player games in a classical Markovian framework which dates back to Shapley (1953).

We stress that, in the general theory of information relaxation duals, solving the pathwise optimization problem often turns out to be the computational bottleneck, compare, e.g., the discussion in Section 4.2 of Brown and Smith (2011) and in Section 2.3 of Haugh and Lim (2012). In contrast, in the framework of the current paper, the implementation is as straightforward as in the optimal stopping case: One first chooses a (D-dimensional) martingale M and then solves the pathwise maximization problem on the right-hand side of (28) for the martingale penalty $\mathfrak{p}_M$ simply by evaluating the explicit pathwise recursion for $\Theta_0^{up}(\rho^{(1)}, \rho^{(0)}, M)$ in Theorem 4.7.

As an important example, the setting of this section covers pricing of convertible bonds. In its simplest form it is a stopping game of two adapted processes $L_j \leq U_j$ governed by the equation

$$Y_j = \min\{U_j, \max\{L_j, E_j[Y_{j+1}]\}\}, \quad Y_J = L_J.$$

For this problem, representations in the sense of pathwise optimal control were previously studied by Kühn et al. (2007) in continuous time and by Beveridge and Joshi (2011) in discrete time.

6 Systems of dynamic programming equations

In this section, we generalize Theorem 4.5 to the case of systems of dynamic programming equations. Our main motivation is to construct confidence bounds for the components of suitable discretization schemes for systems of semilinear parabolic partial differential equations, although our approach also covers the classical multiple stopping case. In order to simplify the notation, we focus on the convex case, but note that the concave-convex case can be handled analogously.

More precisely, we consider systems of the form

$$\begin{aligned} Y_J^{(\nu)} &= \xi^{(\nu)} \\ Y_j^{(\nu)} &= F_j^{(\nu)}\left(E_j\left[\beta_{j+1}Y_{j+1}^{(1)}\right], ..., E_j\left[\beta_{j+1}Y_{j+1}^{(N)}\right]\right), \quad \nu = 1, ..., N,\ j = J-1, ..., 0, \end{aligned} \tag{29}$$

where β and each terminal condition $\xi^{(\nu)}$ satisfy **(R)** and each $F_j^{(\nu)} : \Omega \times \mathbb{R}^{ND} \to \mathbb{R}$ satisfies the conditions on F_j imposed in **(C)**, **(R)** (with D replaced by ND). In particular, each $F_j^{(\nu)}$ is convex in its ND space variables. For a vector $x \in \mathbb{R}^{ND}$, we denote by $x^{[n]}$ the vector in $\mathbb{R}^D$ consisting of the $((n-1)D+1)$-th up to the (nD)-th entry of x, i.e. $x = (x^{[1]}, \ldots, x^{[N]})$.

Now, we fix $j = 0, \ldots, J-1$, a martingale $M \in \mathcal{M}_{ND}$ and controls $r^{(\nu)} \in \mathcal{A}_j^{F^{(\nu)}}$, $\nu = 1, \ldots, N$. Note that each $r_i^{(\nu)}$ takes values in $\mathbb{R}^{ND}$ and we denote by $r_i^{(\nu),[n]}$ the N blocks of $\mathbb{R}^D$-valued random variables. The pathwise recursions for the system (29) now read as follows:

$$\begin{aligned} \theta_J^{(up,\nu)} &= \theta_J^{(low,\nu)} = \xi^{(\nu)} \\ \theta_i^{(up,\nu)} &= \max_{\iota \in \{up,low\}^N} F_i^{(\nu)}\left(\beta_{i+1}\theta_{i+1}^{(\iota_1,1)} - \Delta M_{i+1}^{[1]}, ..., \beta_{i+1}\theta_{i+1}^{(\iota_N,N)} - \Delta M_{i+1}^{[N]}\right) \\ \theta_i^{(low,\nu)} &= \sum_{n=1}^N \left(\left(r_i^{(\nu),[n]}\right)^\top \beta_{i+1}\right)_+ \theta_{i+1}^{(low,n)} - \sum_{n=1}^N \left(\left(r_i^{(\nu),[n]}\right)^\top \beta_{i+1}\right)_- \theta_{i+1}^{(up,n)} \\ &\quad - \sum_{n=1}^N \left(r_i^{(\nu),[n]}\right)^\top \Delta M_{i+1}^{[n]} - F_i^{(\nu,\#)}\left(r_i^{(\nu)}\right), \end{aligned} \tag{30}$$

$i = j, \ldots, J-1$, $\nu = 1, \ldots, N$. As before, we sometimes write e.g. $\theta_i^{(up,\nu)}(r^{(1)}, ..., r^{(N)}, M)$ to stress the dependence on the martingale and the controls. In the setting described above, the following variant of Theorem 4.5 for systems of dynamic programming equations holds true.

Theorem 6.1. *For every $j = 0, \ldots, J$ and $\nu = 1, ..., N$,*

$$Y_j^{(\nu)} = \operatorname*{essinf}_{\substack{r^{(1)} \in \mathcal{A}_j^{F^{(1)}}, ..., r^{(N)} \in \mathcal{A}_j^{F^{(N)}}, \\ M \in \mathcal{M}_{ND}}} E_j[\theta_j^{(up,\nu)}(r^{(1)}, ..., r^{(N)}, M)]$$

$$= \operatorname*{esssup}_{\substack{r^{(1)}\in\mathcal{A}_j^{F^{(1)}},\ldots,r^{(N)}\in\mathcal{A}_j^{F^{(N)}},\\ M\in\mathcal{M}_{ND}}} E_j[\theta_j^{(low,\nu)}(r^{(1)},...,r^{(N)},M)], \quad P\text{-a.s.}$$

Moreover, we have the P-almost sure relation

$$\theta_i^{(low,\nu)}(r^{(1)},...,r^{(N)},M) \le \theta_i^{(up,\nu)}(r^{(1)},...,r^{(N)},M) \tag{31}$$

for every $i = j, \ldots, J$, Finally,

$$Y_j^{(\nu)} = \theta_j^{(up,\nu)}(r^{(1,*)},...,r^{(N,*)},M^*) = \theta_j^{low}(r^{(1,*)},...,r^{(N,*)},M^*) \tag{32}$$

P-almost surely, whenever each $r^{(\nu,*)}$ satisfies the duality relation

$$\sum_{n=1}^{N}\left(r_i^{(\nu,*),[n]}\right)^\top E_i\left[\beta_{i+1}Y_{i+1}^{(n)}\right] - F_i^{(\nu,\#)}\left(r_i^{(\nu,*)}\right) = F_i^{(\nu)}\left(E_i\left[\beta_{i+1}Y_{i+1}^{(1)}\right],...,E_i\left[\beta_{i+1}Y_{i+1}^{(N)}\right]\right) \tag{33}$$

P-almost surely for every $i = j, \ldots, J-1$ and each $M^{*,[\nu]}$ is the Doob martingale of $\beta Y^{(\nu)}$.

Remark 6.2. The proof of Theorem 6.1 follows componentwise the same line of arguments as the one of Theorem 4.5 applying the following reformulation of the upper bound recursion

$$\begin{aligned}\theta_i^{(up,\nu)} \quad = \quad & \sup_{u\in\mathbb{R}^{ND}}\Bigg(\sum_{n=1}^{N}\left(\left(u^{[n]}\right)^\top\beta_{i+1}\right)_+\theta_{i+1}^{(up,n)} - \sum_{n=1}^{N}\left(\left(u^{[n]}\right)^\top\beta_{i+1}\right)_-\theta_{i+1}^{(low,n)} \\ & -\sum_{n=1}^{N}\left(u^{[n]}\right)^\top\Delta M_{i+1}^{[n]} - F_i^{(\nu,\#)}(u)\Bigg)\end{aligned}$$

(which is the componentwise version of Proposition B.3 in the convex case). As the sup can be restricted to the effective domain of $F_i^{(\nu,\#)}$ this alternative representation can be beneficial in the numerical implementation in situations, when N is very large (and, hence, maximizing over the set $\{up, low\}^N$ is prohibitive), but the effective domain has a low-dimensional parametrization.

We illustrate the scope of Theorem 6.1 by two examples.

Example 6.3. The standard multiple stopping problem (or, pricing problem of a swing option) of a stochastic process $S \in L_{ad}^{\infty-}(\mathbb{R})$ is governed by the system of dynamic programming equations

$$Y_j^{(\nu)} = \max\{E_j[Y_{j+1}^{(\nu)}],\ S_j + E_j[Y_{j+1}^{(\nu-1)}]\}, \quad Y_J^{(\nu)} = \nu S_J,$$

$\nu = 1, \ldots, N$ with the convention $Y^{(0)} \equiv 0$. $Y_0^{(\nu)}$ describes the value of the problem of maximizing the expected reward from stopping the process S at ν times, where all stops must be at different times (except that all remaining rights are automatically executed at time J). One can e.g. think of J as an artificial time point and define $S_J = 0$ in order to model the convention that remaining rights at the end of the time horizon are worthless. In this example, $D = 1$ and $\beta \equiv 1$. Due to the monotonicity of the max-operator and the pathwise comparison (31), the upper bound recursion decouples from the lower bound recursion and we get for any $\mathbb{R}^N$-valued martingale M,

$$\theta_j^{(up,\nu)} = \max\left\{\theta_{j+1}^{(up,\nu)} - \Delta M_{j+1}^{[\nu]}, S_j + \theta_{j+1}^{(up,\nu-1)} - \Delta M_{j+1}^{[\nu-1]}\right\}, \quad \theta_J^{(up,\nu)} = \nu S_J, \quad \theta^{(up,0)} \equiv 0,$$

$(\nu = 1, \dots, N)$. Solving this system of pathwise recursive equations explicitly yields

$$\theta_j^{(up,\nu)} = \max_{\substack{j \le i_1 \le \dots \le i_\nu, \\ i_k = i_{k+1} \Rightarrow i_k = J}} \sum_{k=1}^{\nu} \left(S_{i_k} - M_{i_k}^{[\nu-k+1]} - M_{i_{k-1}}^{[\nu-k+1]}\right), \quad i_0 := j.$$

Hence, Theorem 6.1 recovers the pure martingale dual for multiple stopping of Schoenmakers (2012) as special case, which can also be derived from the general theory of information relaxation duals, see Chandramouli and Haugh (2012).

Example 6.4. We now consider systems of semilinear parabolic PDEs of the form

$$\begin{aligned} & v_t^{(\nu)}(t,x) + \frac{1}{2}\sum_{k,l=1}^{d} (\sigma\sigma^\top)_{k,l}(t,x) v_{x_k,x_l}^{(\nu)}(t,x) + \sum_{k=1}^{d} b_k(t,x) v_{x_k}^{(\nu)}(t,x) \\ = \; & -H^{(\nu)}(t,x,v^{(1)}(t,x),\sigma(t,x)\nabla_x v^{(1)}(t,x),\dots,v^{(N)}(t,x),\sigma(t,x)\nabla_x v^{(N)}(t,x)), \end{aligned}$$

$(t,x) \in [0,T) \times \mathbb{R}^d$, $\nu = 1, \dots, N$ with terminal conditions $v^{(\nu)}(T,x) = h^{(\nu)}(x)$. A two-dimensional system of this type arises, e.g., for pricing under negotiated collateralization as shown in Nie and Rutkowski (2016) and will be discussed in more detail in Section 7.

Under suitable conditions on the coefficients $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $H^{(\nu)} : [0,T] \times \mathbb{R}^d \times \mathbb{R}^{N(1+d)} \to \mathbb{R}$, and $h^{(\nu)} : \mathbb{R}^d \to \mathbb{R}$, this system has a unique classical solution satisfying certain growth conditions, see e.g. Chapter 9 of Friedman (1964). Exploiting the link between semilinear parabolic PDEs and backward stochastic differential equations (see e.g. Pardoux, 1998), the parabolic system above can be discretized as follows: Given a partition $\pi = (t_0, \dots, t_J)$ of $[0,T]$ and a d-dimensional Brownian motion W with increments $\Delta W_{i+1} := W_{t_{i+1}} - W_{t_i}$ over time lags of size $\Delta_{i+1} = t_{i+1} - t_i$, consider the Euler type scheme $(j = 0, \dots, J-1,\ \nu = 1, \dots, N)$

$$\begin{aligned} X_{j+1} &= X_j + b(t_j, X_j)\Delta_{j+1} + \sigma(t_j, X_j)\Delta W_{j+1}, \quad X_0 = x, \\ Y_j^{(\nu)} &= E_j[Y_{j+1}^{(\nu)}] + H^{(\nu)}(t_j, X_j, E_j[\beta_{j+1}Y_{j+1}^{(1)}], \dots, E_j[\beta_{j+1}Y_{j+1}^{(N)}])\Delta_{j+1}, \quad Y_J^{(\nu)} = h(X_J), \\ \beta_{j+1} &= \left(1, \frac{\Delta W_{j+1}^{(1)}}{\Delta_{j+1}}, \dots, \frac{\Delta W_{j+1}^{(d)}}{\Delta_{j+1}}\right)^\top, \end{aligned}$$

where E_j denotes the conditional expectation with respect to the information generated by the Brownian motion up to time t_j. Under standard Lipschitz conditions, the approximation error $\sup_{\nu=1,\dots,N} |v^{(\nu)}(0,x) - Y_0^{(\nu)}|$ is known to converge at order $1/2$ in the mesh size of the partition, see e.g. the arguments in Zhang (2004). If H is convex in the last $N(d+1)$ variables and the coefficients satisfy suitable growth conditions, then this discretization scheme is of the form (29).

In order to obtain a componentwise comparison principle for this discretization scheme at fine partitions, it is necessary that, for every ν, $H^{(\nu)}$ does not depend on the gradient of $v^{(n)}$ for $n \neq \nu$ (even if the Brownian increments are truncated in a standard way) and that it depends on $v^{(n)}$, $n \neq \nu$, in a monotonically increasing way. This can be shown analogously to the proof of Theorem 4.3. Here, we mean by a componentwise comparison principle that the inequalities in the definition of super- and subsolutions and in the definition of the comparison principle are supposed to hold for each of the N components. Removing these restrictive assumptions on the coupling in a system of equations is another important reason not to assume a discrete time comparison principle, and instead to work with the coupled pathwise dynamic programming equations for upper bounds and lower bounds in Theorem 6.1.

7 Applications

In this section, we apply the pathwise dynamic programming methodology to calculate upper and lower bounds for three nonlinear pricing problems, pricing of a long-term interest rate product under bilateral counterparty risk, option pricing under volatility uncertainty, and under negotiated collateralization. Traditionally, the valuation of American options was by far the most prominent nonlinear pricing problem both in practice and in academia. In the wake of the financial crisis, various other sources of nonlinearity such as model uncertainty, default risk, liquidity problems or transaction costs have received considerable attention, see the recent monographs Brigo et al. (2013), Crépey (2013), and Guyon and Henry-Labordère (2013).

7.1 Bilateral counterparty risk

Suppose that two counterparties have agreed to exchange a stream of payments C_{t_j} over the sequence of time points $t_0, \ldots, t_J$. For many common interest rate products such as swaps, the sign of C is random – so that the agreed payments may flow in either direction. Therefore, a consistent pricing approach must take into account bilateral default risk, thus introducing nonlinearities into the recursive pricing equations which are in general neither convex nor concave. In this respect, the present setting is conceptually similar but more complex than the model of unilateral counterparty risk in Example 2.2. We refer to Crépey et al. (2013) for technical background and an in-depth discussion of the intricacies of pricing under bilateral counterparty risk and funding. We notice that by discretizing their equations (2.14) and (3.8), we arrive at the following nonlinear backward dynamic program for the value of the product Y_j at time t_j (given that there was no default prior to t_j) associated with the payment stream C, namely, $Y_J = C_{t_J}$ and

$$Y_j \quad = \quad (1 - \Delta(r_{t_j} + \gamma_{t_j}(1-\mathfrak{r})(1-2p_{t_j}) + \overline{\lambda}))E_j[Y_{j+1}]$$

$$+\Delta(\gamma_{t_j}(1-\mathfrak{r})(1-3p_{t_j})+\overline{\lambda}-\lambda)E_j[Y_{j+1}]_+ + C_{t_j}.$$

Here E_j denotes the expectation conditional on the market's reference filtration up to time t_j (i.e., the full market information is obtained by enlarging this filtration with the information whether default has yet occurred or not). In focusing on equation (3.8) in Crépey et al. (2013), we consider a pre-default CSA recovery scheme without collateralization, see their paper for background. In the pricing equation, r_t denotes (a proxy to) the risk-less short rate at time t. The rate at which a default of either side occurs at time t is denoted by γ_t. Moreover, p_t is the associated conditional probability that it is the counterparty who defaults, if default occurs at time t. We rule out simultaneous default so that own default happens with conditional probability $1-p_t$, and assume that the three parameters ρ, $\overline{\rho}$ and $\mathfrak{r}$ associated with partial recovery from Crépey et al. (2013) are identical and denoted by $\mathfrak{r}$. Finally, λ and $\overline{\lambda}$ are two constants associated with the costs of external lending and borrowing, and $\Delta = t_{j+1}-t_j$ is the stepsize of the equidistant time partition.

Defining $g_j = 1-\Delta(r_{t_j}+\gamma_{t_j}(1-\mathfrak{r})(1-2p_{t_j})+\overline{\lambda})$ and $h_j = \Delta(\gamma_{t_j}(1-\mathfrak{r})(1-3p_{t_j})+\overline{\lambda}-\lambda)$, we can express the recursion for Y in terms of a concave function G_j and a convex function F_j by setting $F_j(z) = z_+$ and

$$G_j(z,y) = g_j\, z + (h_j)_+\, y - (h_j)_-\, z_+ + C_{t_j},$$

so that $Y_j = G_j(E_j[Y_{j+1}], F_j(E_j[Y_{j+1}]))$. Hence, $D=1$ and $\beta \equiv 1$ in this example. Denote by $(\tilde{Y}_j)$ a numerical approximation of the process (Y_j), by $(\tilde{Q}_j)$ a numerical approximation of $(E_j[Y_{j+1}])$ and by $(\tilde{M}_j)$ a martingale which we think of as an approximation to the Doob martingale of Y. In terms of these inputs, the pathwise recursions (22) for upper and lower bound are given by

$$\begin{aligned}
\theta_j^{up} &= \left(\tilde{\rho}_j^{(1)}\right)_+\left(\theta_{j+1}^{up}-\Delta\tilde{M}_{j+1}\right) - \left(\tilde{\rho}_j^{(1)}\right)_-\left(\theta_{j+1}^{low}-\Delta\tilde{M}_{j+1}\right) + \tilde{\rho}_j^{(0)}\left(\theta_{j+1}^{up}-\Delta\tilde{M}_{j+1}\right)_+ + C_{t_j} \\
\theta_j^{low} &= \min_{\iota\in\{up,low\}} g_j\left(\theta_{j+1}^{\iota}-\Delta\tilde{M}_{j+1}\right) - (h_j)_-\left(\theta_{j+1}^{\iota}-\Delta\tilde{M}_{j+1}\right)_+ + (h_j)_+\,\tilde{s}_j\left(\theta_{j+1}^{low}-\Delta\tilde{M}_{j+1}\right) + C_{t_j}
\end{aligned}$$

where

$$(\tilde{\rho}_j^{(1)},\tilde{\rho}_j^{(0)},\tilde{s}_j) = \begin{cases} (g_j,(h_j)_+,0)\,, & \tilde{Q}_j<0, \\ (g_j-(h_j)_-,(h_j)_+,1)\,, & \tilde{Q}_j\geq 0. \end{cases}$$

For the payment stream C_{t_j}, we consider a swap with notional N, fixed rate R and an equidistant sequence of tenor dates $\mathcal{T} = \{T_0,\dots,T_K\} \subseteq \{t_0,\dots t_J\}$. Denote by δ the length of the time interval between T_i and T_{i+1} and by $P(T_{i-1},T_i)$ the T_{i-1}-price of a zero-bond with maturity T_i. Then, the payment process C_{t_j} is given by

$$C_{T_i} = N\cdot\left(\frac{1}{P(T_{i-1},T_i)} - (1+R\delta)\right)$$

for $T_i \in \mathfrak{T} \setminus \{T_0\}$ and $C_{t_j} = 0$ otherwise, see Brigo and Mercurio (2006), Chapter 1.

For r and γ, we implement the model of Brigo and Pallavicini (2007), assuming that the risk-neutral dynamics of r is given by a two-factor Gaussian short rate model, a reparametrization of the two-factor Hull-White model, while γ is a CIR process. For the conditional default probabilities p_t we assume $p_t = 0 \wedge \tilde{p}_t \vee 1$ where $\tilde{p}$ is an Ornstein-Uhlenbeck process. In continuous time, this corresponds to the system of stochastic differential equations

$$dx_t = -\kappa_x x_t dt + \sigma_x dW_t^x, \quad dy_t = -\kappa_y y_t dt + \sigma_y dW_t^y$$
$$d\gamma_t = \kappa_\gamma(\mu_\gamma - \gamma_t)dt + \sigma_\gamma \sqrt{\gamma_t} dW_t^\gamma, \quad d\tilde{p}_t = \kappa_p(\mu_p - \tilde{p}_t)dt + \sigma_p dW_t^p$$

with $r_t = r_0 + x_t + y_t$, $x_0 = y_0 = 0$. Here, W^x, W^y and W^γ are Brownian motions with instantaneous correlations ρ_{xy}, $\rho_{x\gamma}$ and $\rho_{y\gamma}$. In addition, we assume that $W_t^p = \rho_{\gamma p} W_t^\gamma + \sqrt{1 - \rho_{\gamma p}^2} W_t$ where the Brownian motion W is independent of (W^x, W^y, W^γ). We choose the filtration generated by the four Brownian motions as the reference filtration.

For the dynamics of x, y and $\tilde{p}$, exact time discretizations are available in closed form. We discretize γ by $(\tilde{\gamma}_{t_j})_+$, where $\tilde{\gamma}$ denotes the fully truncated scheme of Lord et al. (2010). The bond prices $P(t, s)$ are given as an explicit function of x_t and y_t in this model. This implies that the swap's "clean price", i.e., the price in the absence of counterparty risk is given in closed form as well, see Sections 1.5 and 4.2 of Brigo and Mercurio (2006).

We consider 60 half-yearly payments over a horizon of $T = 30$ years, i.e., $\delta = 0.5$. J is always chosen as an integer multiple of 60 so that δ is an integer multiple of $\Delta = T/J$. For the model parameters, we choose

$$(r_0, \kappa_x, \sigma_x, \kappa_y, \sigma_y) = (0.03, 0.0558, 0.0093, 0.5493, 0.0138),$$
$$(\gamma_0, \mu_\gamma, \kappa_\gamma, \sigma_\gamma, p_0, \mu_p, \kappa_p, \sigma_p) = (0.0165,\ 0.026,\ 0.4, 0.14,\ 0.5,\ 0.5,\ 0.8,\ 0.2),$$
$$(\rho_{xy}, \rho_{x\gamma}, \rho_{y\gamma}, \mathfrak{r}, \lambda, \overline{\lambda}, N) = (-0.7,\ 0.05,\ -0.7,\ 0.4,\ 0.015,\ 0.045,\ 1).$$

We thus largely follow Brigo and Pallavicini (2007) for the parametrization of r and γ but leave out their calibration to initial market data and choose slightly different correlations to avoid the extreme cases of a perfect correlation or independence of r and γ. The remaining parameters J, R and $\rho_{\gamma p}$ are varied in the numerical experiments below.

In order to pre-compute the input approximation $(\tilde{Y}, \tilde{Q})$ we apply here (and in the other numerical examples below) a variant of the regression-later algorithm of Glasserman and Yu (2004), which was developed there for the optimal stopping case. This algorithm performs, in each step backwards in time, a simulation-based empirical least-squares regression on a set of basis functions. Our variant is described in detail in Section 5.1 of Bender et al. (2016). It requires that, for each basis function at time $(i+1)$, say η_{i+1}, the one-step conditional expectation $E_i[\beta_{i+1}\eta_{i+1}]$ can be computed in closed form, as well as the one-step conditional expectation $E_{J-1}[\beta_J \tilde{Y}_J]$ of the terminal condition $\tilde{Y}_J$. As a main benefit of regression-later approaches, the Doob martingale

J	Clean Price	$\rho_{\gamma p} = 0.8$		$\rho_{\gamma p} = 0$		$\rho_{\gamma p} = -0.8$		Run Time ($\rho_{\gamma p} = 0$)
120 ($N_r = 10^5$)	0	21.34 (0.02)	21.39 (0.02)	24.88 (0.02)	24.94 (0.02)	28.33 (0.02)	28.41 (0.02)	33 sec.
120 ($N_r = 10^6$)	0	21.32 (0.02)	21.37 (0.02)	24.90 (0.02)	24.96 (0.02)	28.29 (0.02)	28.38 (0.02)	244 sec.
360 ($N_r = 10^5$)	0	21.27 (0.02)	21.33 (0.02)	24.85 (0.02)	24.90 (0.02)	28.26 (0.02)	28.35 (0.02)	90 sec.
360 ($N_r = 10^6$)	0	21.27 (0.02)	21.32 (0.02)	24.84 (0.02)	24.91 (0.02)	28.26 (0.02)	28.35 (0.02)	709 sec.
720 ($N_r = 10^5$)	0	21.26 (0.02)	21.31 (0.02)	24.84 (0.02)	24.90 (0.02)	28.26 (0.02)	28.35 (0.02)	182 sec.
720 ($N_r = 10^6$)	0	21.26 (0.02)	21.31 (0.02)	24.86 (0.02)	24.92 (0.02)	28.25 (0.02)	28.34 (0.02)	1429 sec.
1440 ($N_r = 10^5$)	0	21.24 (0.02)	21.29 (0.02)	24.84 (0.02)	24.90 (0.02)	28.23 (0.02)	28.32 (0.02)	366 sec.
1440 ($N_r = 10^6$)	0	21.27 (0.02)	21.32 (0.02)	24.83 (0.02)	24.89 (0.02)	28.25 (0.02)	28.34 (0.02)	2693 sec.

Table 1: Lower and upper bound estimators for varying values of $\rho_{\gamma p}$, J and N_r with $R = 275.12$ basis points (b.p.), $N_e = 5 \cdot 10^5$. Prices and standard deviations (in brackets) are given in b.p.

$\tilde{M}$ of $\beta\tilde{Y}$, which we apply in the pathwise recursions, is available in closed form and need not be approximated by a nested simulation as in Andersen and Broadie (2004) type algorithms.

We initialize the regression at $\tilde{Y}_J = C_{t_J}$ and choose, at each time $1 \leq j \leq J-1$, the four basis functions 1, $\tilde{\gamma}_{t_j}$, $\tilde{\gamma}_{t_j} \cdot \tilde{p}_{t_j}$ and the closed-form expression for the clean price at time t_j of the swap's remaining payment stream. The regression is run with N_r independent trajectories of the underlying discrete-time Markov process $(x_{t_j}, y_{t_j}, \tilde{\gamma}_{t_j}, \tilde{p}_{t_j}, x_{T(j)}, y_{T(j)})$, the so-called regression paths, where $T(j)$ denotes the largest tenor date which is strictly smaller than t_j. Storing the coefficients from the regression, we next simulate $N_e = 5 \cdot 10^5$ new independent trajectories of the underlying Markov process, which we call evaluation paths (and which are independent of the regression paths). We can then go through the recursion for θ^{up} and θ^{low} along each evaluation path. Denote by $\hat{Y}_0^{up}$ and $\hat{Y}_0^{low}$ the resulting empirical means as Monte Carlo estimators of $E[\theta_0^{up}]$ and $E[\theta_0^{low}]$ and their associated empirical standard deviations by $\hat{\sigma}^{up}$ and $\hat{\sigma}^{low}$. Then, an asymptotic 95%-confidence interval for Y_0 is given by

$$[\hat{Y}_0^{low} - 1.96\hat{\sigma}^{low}, \hat{Y}_0^{up} + 1.96\hat{\sigma}^{up}].$$

Table 1 displays upper and lower bound estimators with their standard deviations for different step sizes of the time discretization, for two choices of the number of regression paths, $N_r \in \{10^5, 10^6\}$, and for different correlations between γ and p. Here, R is chosen as the fair swap rate in the absence of default risk, i.e., it is chosen such that the swap's clean price at $j = 0$ is zero. The four choices of J correspond to a quarterly, monthly, bi-weekly, and weekly time discretization, respectively. In all cases, the width of the resulting confidence interval is about 0.6% of the value. We note that the regression estimates $\tilde{Y}_0$ (which we do not report here) are

$\rho_{\gamma p}$	Adjusted Fair Swap Rate	Clean Price	Bounds	
0.8	290.82	-31.53	-0.02 (0.02)	0.05 (0.02)
0	293.65	-37.22	-0.05 (0.02)	0.03 (0.02)
-0.8	296.39	-42.71	-0.06 (0.02)	0.04 (0.02)

Table 2: Adjusted fair swap rates and lower and upper bound estimators for varying values of $\rho_{\gamma p}$ with $N_r = 10^5$, $N_e = 5 \cdot 10^5$ and $J = 360$. Rates, prices and standard deviations (in brackets) are given in b.p.

more stable for 10^6 paths in the case of weekly and bi-weekly time discretizations. Nonetheless, the resulting upper and lower confidence bounds do not vary significantly for the two choices of regression paths. Moreover, the differences in the bounds for the monthly, the bi-weekly, and the weekly time discretization can all be explained by the standard deviations, while the quarterly time discretization yields significantly larger price bounds. These results indicate that a monthly time discretization (i.e., 360 discretization steps) and 10^5 regression are sufficient to accurately price this long-dated swap under bilateral default risk. The total run time is 90 seconds for this parameter choice (Matlab implementation of the method on a 24 core 2.5 GHz Intel Xeon processor). Run times are reported in the last column of Table 1 for the case of zero correlation between γ and p. We note that the choice of this correlation parameter has little to no influence on the run time and that the regression requires, in all cases, about 1/4 of the total run time for $N_r = 10^5$ regression paths and 90% of the total run time for $N_r = 10^6$ regression paths. The effect of varying the correlation parameter of γ and p also has the expected direction. Roughly, if $\rho_{\gamma p}$ is positive then larger values of the overall default rate go together with larger conditional default risk of the counterparty and smaller conditional default risk of the party, making the product less valuable to the party. While this effect is not as pronounced as the overall deviation from the clean price, the bounds are easily tight enough to differentiate between the three cases.

We next compare our numerical results with the ‘generic method’ of Section 5 in Bender et al. (2015). While the latter paper focuses on convex nonlinearities, it also suggests a generic local approximation of Lipschitz nonlinearities by convex nonlinearities, which can be applied for the problem of bilateral default risk (after suitable truncations). Based on the same input approximations as above (computed by the regression-later approach with $N_r = 10^5$ regression paths), this algorithm produced a 95%-confidence interval of $[-0.0675, 0.6825]$ for the case $J = 360$ and $\rho_{\gamma p} = 0$ with about the same total run time of 90 seconds as in our new algorithm. The length of this confidence interval is several magnitudes wider than the one computed from Table 1, and it cannot even significantly distinguish between the clean price and the price under default risk. These results demonstrate the importance of exploiting the concave-convex structure for pricing under bilateral default risk.

Finally, Table 2 displays the adjusted fair swap rates accounting for counterparty risk and funding for the three values of $\rho_{\gamma p}$, i.e., the values of R which set the adjusted price to zero in the three different correlation scenarios. To identify these rates, we fix a set of evaluation and regression paths and define $\mu(R)$ as the midpoint of the confidence interval we obtain when running the algorithm with these paths and rate R for the fixed leg of the swap. We apply a standard bisection method to find the zero of $\mu(R)$. The confidence intervals for the prices in Table 2 are then obtained by validating these swap rates with a new set of evaluation paths. We observe that switching from a clean valuation to the adjusted valuation with $\rho = 0.8$ increases the fair swap rate by 16 basis points (from 275 to 291). Changing ρ from 0.8 to -0.8 leads to a further increase by 5 basis points.

7.2 Uncertain Volatility Model

In this section, we apply our numerical approach to the uncertain volatility model of Example 2.3. Let $0 = t_0 < t_1 < ... < t_J = T$ be an equidistant partition of the interval $[0, T]$, where $T \in \mathbb{R}_+$. Recall that for an adapted process $(\sigma_t)_t$, the price of the risky asset X^σ at time t is given by

$$X_t^\sigma = x_0 \exp\left\{ \int_0^t \sigma_u dW_u - \frac{1}{2} \int_0^t \sigma_u^2 du \right\},$$

where W is a Brownian motion. Further, let g be the payoff a European option written on the risky asset. Then, by Example 2.3, the discretized value process of this European option under uncertain volatility is given by $Y_J = g(X_T^{\hat{\rho}})$,

$$\Gamma_j = E_j\left[B_{j+1}^{\hat{\rho}} Y_{j+1}\right] \quad \text{and} \quad Y_j = E_j[Y_{j+1}] + \Delta \max_{s \in \{s_{low}, s_{up}\}} s\,\Gamma_j, \tag{34}$$

where

$$B_{j+1}^{\hat{\rho}} = \frac{\Delta W_{j+1}^2}{\Delta^2} - \hat{\rho}\frac{\Delta W_{j+1}}{\Delta} - \frac{1}{\Delta}, \quad \text{and } s_\iota = \frac{1}{2}\left(\frac{\sigma_\iota^2}{\hat{\rho}^2} - 1\right)$$

for $\iota \in \{low, up\}$. Moreover, $X_T^{\hat{\rho}}$ is the price of the asset at time T under the constant reference volatility $\hat{\rho} > 0$. Notice that the reference volatility $\hat{\rho}$ is a choice parameter in the discretization. The basic idea is to view the uncertain volatility model as a suitable correction of a Black-Scholes model with volatility $\hat{\rho}$.

As in Section 7.1, we denote in the following by $(\tilde{Y}_j)$ a numerical approximation of the process (Y_j), by $(\tilde{Q}_j)$ a numerical approximation of $(E_j[Y_{j+1}])$ and by $(\tilde{\Gamma}_j)$ a numerical approximation of (Γ_j). Furthermore, $(\tilde{M}_j) = (\tilde{M}_j^{(1)}, \tilde{M}_j^{(2)})^\top$ denotes an approximation of the Doob martingale M of $\beta_j Y_j$, which is given by

$$M_{j+1} - M_j = \begin{pmatrix} Y_{j+1} - E_j[Y_{j+1}] \\ B_{j+1}^{\hat{\rho}} Y_{j+1} - E_j\left[B_{j+1}^{\hat{\rho}} Y_{j+1}\right] \end{pmatrix}.$$

The recursion (22) for θ_j^{up} and θ_j^{low}, based on these input approximations, can be written as

$$\begin{aligned}\theta_j^{up} &= \max_{\iota\in\{up,low\}} \max_{s\in\{s_{low},s_{up}\}} \left\{\theta_{j+1}^{\iota} - \Delta\tilde{M}_{j+1}^{(1)} + sB_{j+1}^{\hat{\rho}}\theta_{j+1}^{\iota}\Delta - s\Delta\tilde{M}_{j+1}^{(2)}\Delta\right\} \\ \theta_j^{low} &= \left(\tilde{r}_j^{(1)} + \tilde{r}_j^{(2)}B_{j+1}^{\hat{\rho}}\Delta\right)_+ \theta_{j+1}^{low} - \left(\tilde{r}_j^{(1)} + \tilde{r}_j^{(2)}B_{j+1}^{\hat{\rho}}\Delta\right)_- \theta_{j+1}^{up} - \tilde{r}_j^{(1)}\Delta\tilde{M}_{j+1}^{(1)} - \tilde{r}_j^{(2)}\Delta\tilde{M}_{j+1}^{(2)}\Delta,\end{aligned}$$

where $\tilde{r}_j = (\tilde{r}_j^{(1)}, \tilde{r}_j^{(2)})$ is given by

$$\tilde{r}_j = \begin{cases}(1, s_{low}), & \tilde{\Gamma}_j < 0 \\ (1, s_{up}), & \tilde{\Gamma}_j \geq 0.\end{cases}$$

For the payoff, we consider a European call-spread option with strikes K_1 and K_2, i.e.,

$$g(x) = (x - K_1)_+ - (x - K_2)_+,$$

which is also studied in Guyon and Henry-Labordère (2011), Alanko and Avellaneda (2013), and Kharroubi et al. (2014). Following their setting, we choose the maturity $T = 1$, the volatility bounds $\sigma_{low} = 0.1$ and $\sigma_{up} = 0.2$ as well as $K_1 = 90$ and $K_2 = 110$ and $x_0 = 100$. The reference volatility $\hat{\rho}$ is varied in our numerical experiments. This example is by now a standard test case for Monte Carlo implementations of Hamilton-Jacobi-Bellman equations. The option price in the continuous time limit can be calculated in closed form and equals 11.2046, see Vanden (2006).

The input approximation $(\tilde{Y}, \tilde{Q}, \tilde{\Gamma})$ is again computed by the regression-later approach. We first simulate $N_r = 10^5$ regression paths of the process $(X_j^{\hat{\rho}})$ under the constant volatility $\hat{\rho}$. For the regression, we do not start all paths at x_0 but rather start $N_r/200$ trajectories at each of the points $31, \ldots, 230$. Since X is a geometric Brownian motion under $\hat{\rho}$, it can be simulated exactly. Starting the regression paths at multiple points allows to reduce the instability of regression coefficients arising at early time points. See Rasmussen (2005) for a discussion of this stability problem and of the method of multiple starting points. For the empirical regression we choose 162 basis functions. The first three (at time $j+1$) are 1, $X_{j+1}^{\hat{\rho}}$ and $E[g(X_J^{\hat{\rho}})|X_{j+1}^{\hat{\rho}}]$. Note that the third one is simply the Black-Scholes price (under $\hat{\rho}$) of the spread option g. For the remaining 159 basis functions, we also choose Black-Scholes prices of spread options with respective strikes $K^{(k)}$ and $K^{(k+1)}$ for $k = 1, \ldots, 159$, where the numbers $K^{(1)}, \ldots, K^{(160)}$ increase from 20.5 to 230.5. The one-step conditional expectations of these basis functions after multiplication with the second derivative weight $B^{\hat{\rho}}$ are just (differences of) Black-Scholes Gammas at time j. As already noticed in Bender and Steiner (2012), the regression-later approach benefits from huge variance reduction effects in the presence of Malliavin Monte Carlo weights for the space derivatives. Indeed, regression-now estimates for conditional expectations involving the weight $B^{\hat{\rho}}$ suffer from the large variance of this weight, while in the regression-later approach the corresponding conditional expectations are computed in closed form. As in Section 7.1, we then simulate $(X_j^{\hat{\rho}}, \tilde{Y}_j, \tilde{Q}_j, \tilde{\Gamma}_j)$

along $N_e = 10^5$ evaluation paths started in x_0. As pointed out above, the Doob martingale $\tilde{M}$ of $\beta_j \tilde{Y}_j$ is available in closed form in the regression-later approach. Finally, the recursions for θ^{up} and θ^{low} are calculated backwards in time on each evaluation path.

Table 3 shows the approximated prices $\tilde{Y}_0$ as well as upper and lower bounds for $\hat{\rho} = 0.2/\sqrt{3} \approx 0.115$ depending on the time discretization. This is the smallest choice of $\hat{\rho}$, for which the monotonicity condition in Theorem 4.3 can only be violated when the absolute values of the Brownian increments are large, cp. Example 4.4. As before, $\hat{Y}_0^{up}$ and $\hat{Y}_0^{low}$ denote Monte Carlo estimates of $E[\theta_0^{up}]$ respectively $E[\theta_0^{low}]$. The numerical results suggest convergence from below towards the continuous-time limit for finer time discretizations. This is intuitive in this example, since finer time discretizations allow for richer choices of the process (σ_t) in the maximization problem (7). The run time for the results in Table 3 is linear in J and takes about 85 seconds for $J = 24$. It is about equally split between the regression and the computation of the bounds. We notice that the bounds are fairly tight (with, e.g., a relative width of 1.3% for the 95% confidence interval with $J = 21$ time discretization points), although the upper bound begins to deteriorate as $\tilde{Y}_0$ approaches its limiting value. The impact of increasing $\hat{\rho}$ to 0.15 (as proposed in Guyon and Henry-Labordère, 2011; Alanko and Avellaneda, 2013) is shown in Table 4. The relative width of the 95%-confidence interval is now about 0.6% for up to $J = 35$ time steps, but also the convergence to the continuous-time limit appears to be slower with this choice of $\hat{\rho}$.

J	3	6	9	12	15	18	21	24
$\tilde{Y}_0$	10.8553	11.0500	11.1054	11.1340	11.1484	11.1585	11.1666	11.1710
$\hat{Y}_0^{up}$	10.8591 (0.0001)	11.0536 (0.0002)	11.1116 (0.0005)	11.1438 (0.0007)	11.1728 (0.0056)	11.2097 (0.0088)	11.2764 (0.0173)	11.5593 (0.0984)
$\hat{Y}_0^{low}$	10.8550 (0.0001)	11.0502 (0.0002)	11.1060 (0.0005)	11.1344 (0.0002)	11.1486 (0.0002)	11.1585 (0.0006)	11.1666 (0.0003)	11.1683 (0.0032)

Table 3: Approximated price as well as lower and upper bounds for $\hat{\rho} = 0.2/\sqrt{3}$ for different time discretizations. Standard deviations are given in brackets

J	5	10	15	20	25	30	35	40
$\tilde{Y}_0$	10.8153	10.9982	11.0684	11.1027	11.1241	11.1386	11.1479	11.1554
$\hat{Y}_0^{up}$	10.8167 (0.0001)	11.0028 (0.0001)	11.0728 (0.0001)	11.1092 (0.0002)	11.1379 (0.0008)	11.1687 (0.0030)	11.2058 (0.0047)	12.0483 (0.6464)
$\hat{Y}_0^{low}$	10.8153 (0.0001)	10.9983 (0.0001)	11.0683 (0.0001)	11.1023 (0.0001)	11.1237 (0.0001)	11.1376 (0.0002)	11.1462 (0.0002)	11.0101 (0.1391)

Table 4: Approximated price as well as lower and upper bounds for $\hat{\rho} = 0.15$ for different time discretizations. Standard deviations are given in brackets

Comparing Table 4 with the results in Alanko and Avellaneda (2013), we observe that their point estimates for Y_0 at time discretization levels $J = 10$ and $J = 20$ do not lie in our confidence intervals which are given by $[10.9981, 11.0030]$ and $[11.1021, 11.1096]$, indicating that their (regression-now) least-squares Monte Carlo estimator may still suffer from large variances (although they apply control variates). The dependence of the time discretization error on the choice of the reference volatility $\hat{\rho}$ is further illustrated in Table 5, which displays the mean and the standard deviation of 30 runs of the regression-later algorithm for different choices of $\hat{\rho}$ and up to 640 time steps. By and large, convergence is faster for smaller choices of $\hat{\rho}$, but the algorithm becomes unstable when the reference volatility is too small.

J	10	20	40	80	160	320	640
$\hat{\rho} = 0.06$	79.7561 (5.1739)	$1.6421 \cdot 10^5$ $(8.4594 \cdot 10^5)$	$1.7010 \cdot 10^{16}$ $(6.3043 \cdot 10^{16})$	$3.2151 \cdot 10^{24}$ $(1.7603 \cdot 10^{25})$	$1.8613 \cdot 10^{24}$ $(7.0234 \cdot 10^{24})$	$4.5672 \cdot 10^{39}$ $(2.5016 \cdot 10^{40})$	$7.0277 \cdot 10^{39}$ $(3.7590 \cdot 10^{40})$
$\hat{\rho} = 0.08$	11.6463 (0.2634)	12.3183 (1.5447)	130.2723 (372.2625)	11.8494 (1.1846)	11.6951 (1.5772)	$5.4389 \cdot 10^3$ $(1.8766 \cdot 10^4)$	$1.0153 \cdot 10^8$ $(3.5571 \cdot 10^8)$
$\hat{\rho} = 0.1$	11.1552 (0.0031)	11.1823 (0.0040)	11.1942 (0.0005)	11.1999 (0.0002)	11.2026 (0.0001)	11.2047 (0.0001)	11.2057 (0.0001)
$\hat{\rho} = 0.15$	10.9985 (0.0006)	11.1027 (0.0006)	11.1556 (0.0004)	11.1821 (0.0003)	11.1952 (0.0003)	11.2017 (0.0003)	11.2047 (0.0002)
$\hat{\rho} = 0.2$	10.7999 (0.0007)	10.9746 (0.0005)	11.0819 (0.0005)	11.1455 (0.0006)	11.1802 (0.0005)	11.1980 (0.0006)	11.2073 (0.0006)
$\hat{\rho} = 0.5$	9.7088 (0.0004)	9.9652 (0.0005)	10.2306 (0.0009)	10.4945 (0.0015)	10.7453 (0.0047)	10.9635 (0.0076)	11.1248 (0.0103)

Table 5: Mean of $L = 30$ simulations of $\tilde{Y}_0$ for different $\hat{\rho}$ and discretizations. Standard deviations are given in brackets.

7.3 Negotiated collateral

We now consider the problem of pricing under negotiated collateralization in the presence of funding costs as discussed in Nie and Rutkowski (2016). Collateralized contracts differ from 'standard' contracts in the way that the involved parties not only agree on a payment stream until maturity but also on the collateral posted by both parties. By providing collateral, both parties can reduce the possible loss resulting from a default of the respective counterparty prior to maturity. In the following, we apply our approach to the valuation of a contract under negotiated collateral, i.e. the imposed collateral depends on the valuations of the contract made by the two parties. More precisely, the party ('hedger') wishes to perfectly hedge the stream of payments consisting of the option payoff and the posted collateral under funding costs, while the counterparty hedges the negative payment stream under funding costs. As hedging under funding costs is known to be nonlinear, both hedges do not cancel each other. Hence, one ends up with a coupled system of two equations where the coupling is due to the fact that the counterparty's hedging strategy influences the hedger's payment stream due to the negotiated collateral and vice versa.

We next translate the original backward SDE formulation of the problem in Nie and Rutkowski (2016) into the parabolic PDE setting of Example 6.4. To this end let $h : \mathbb{R}^d \to \mathbb{R}$ be a function of polynomial growth which represents the payoff of a European-style option written on d risky assets with maturity T. The dynamics of the risky assets $X = (X^{(1)}, ..., X^{(d)})$ under the pricing measure are given by independent identically distributed Black-Scholes models

$$X_t^{(k)} = x_0 \exp\left\{ \left(R^L - \frac{1}{2}\sigma^2 \right) t + \sigma W_t^{(k)} \right\}, \quad k = 1, \dots, d,$$

where $R^L \geq 0$ is the lending rate of the bond, $\sigma > 0$ is the assets volatility, and $W = (W^{(1)}, ..., W^{(d)})$ is a d-dimensional Brownian motion. We, moreover, denote by R^B the borrowing rate of the bond and by R^C the collateralization rate. Hence, $R^B \geq R^L$. As in Example 3.2 in Nie and Rutkowski (2016) we consider the case that the collateral is a convex combination $\bar{q}(v^{(1)}, -v^{(2)}) = \alpha v^{(1)} + (1-\alpha)(-v^{(2)})$ of the hedger's price $v^{(1)}$ (i.e., the party's hedging cost) and the counterparty's price $-v^{(2)}$ (i.e, the negative of the counterparty's hedging cost) for some $\alpha \in [0,1]$. Following Proposition 3.3 in Nie and Rutkowski (2016) with zero initial endowment the system of PDEs then reads as follows:

$$v_t^{(\nu)}(t,x) + \frac{1}{2}\sum_{k,l=1}^{d} v_{x_k,x_l}^{(\nu)}(t,x) = -H^{(\nu)}(v^{(1)}(t,x), \nabla_x v^{(1)}(t,x), v^{(2)}(t,x), \nabla_x v^{(2)}(t,x)),$$

$(t,x) \in [0,T) \times \mathbb{R}^d$, with terminal conditions

$$v^{(\nu)}(T,x) = (-1)^{\nu-1} h\left(\left(x_0 \exp\left\{\left(R^L - \frac{1}{2}\sigma^2\right)t + \sigma x^{(k)}\right\}\right)_{k=1,\ldots,d}\right), \; x = (x^{(1)}, \ldots, x^{(d)}) \in \mathbb{R}^d$$

and nonlinearities given by

$$\begin{aligned} & H^{(\nu)}(v^{(1)}(t,x), \nabla_x v^{(1)}(t,x), v^{(2)}(t,x), \nabla_x v^{(2)}(t,x)) \\ = \; & -R^L a_\nu (v^{(1)}(t,x) + v^{(2)}(t,x)) + (-1)^\nu R^C(\alpha v^{(1)}(t,x) - (1-\alpha) v^{(2)}(t,x)) \\ & +(R^B - R^L)\left(a_\nu(v^{(1)}(t,x) + v^{(2)}(t,x)) - \frac{1}{\sigma}(\nabla_x v^{(\nu)}(t,x))^\top \mathbf{1}\right)_-, \quad \nu = 1,2. \end{aligned}$$

Here, $(a_1, a_2) = (1-\alpha, \alpha)$ and $\mathbf{1}$ is the vector in $\mathbb{R}^d$ consisting of ones. With this notation, $v^{(1)}(t, W_t)$ and $-v^{(2)}(t, W_t)$ denote the hedger's price and counterparty's price of the collateralized contract at time t. Applying the probabilistic discretization scheme presented in Example 6.4 with an equidistant time grid $\{t_0, ..., t_J\}$, we end up with the following dynamic program:

$$\begin{aligned} X_{j+1}^{(k)} &= X_j^{(k)} \exp\left\{\left(R^L - \frac{1}{2}\sigma^2\right)\Delta + \sigma \Delta W_{j+1}^{(k)}\right\}, \quad X_0^{(k)} = x_0, \quad k = 1, ..., d \\ Y_J^{(1)} &= -Y_J^{(2)} = h(X_J) \\ Z_j^{[\nu]} &= E_j\left[\Delta^{-1}\left(\Delta W_{j+1}^{(1)}, ..., \Delta W_{j+1}^{(d)}\right)^\top Y_{j+1}^{(\nu)}\right], \quad \nu = 1,2 \\ Y_j^{(1)} &= E_j[Y_{j+1}^{(1)}] - R^L(1-\alpha)(E_j[Y_{j+1}^{(1)}] + E_j[Y_{j+1}^{(2)}])\Delta - R^C(\alpha E_j[Y_{j+1}^{(1)}] - (1-\alpha)E_j[Y_{j+1}^{(2)}])\Delta \\ & \quad +(R^B - R^L)\left((1-\alpha)(E_j[Y_{j+1}^{(1)}] + E_j[Y_{j+1}^{(2)}]) - \frac{1}{\sigma}\left(Z_j^{[1]}\right)^\top \mathbf{1}\right)_- \Delta \\ Y_j^{(2)} &= E_j[Y_{j+1}^{(2)}] - R^L\alpha(E_j[Y_{j+1}^{(1)}] + E_j[Y_{j+1}^{(2)}])\Delta + R^C(\alpha E_j[Y_{j+1}^{(1)}] - (1-\alpha)E_j[Y_{j+1}^{(2)}])\Delta \\ & \quad +(R^B - R^L)\left(\alpha(E_j[Y_{j+1}^{(1)}] + E_j[Y_{j+1}^{(2)}]) - \frac{1}{\sigma}\left(Z_j^{[2]}\right)^\top \mathbf{1}\right)_- \Delta, \end{aligned} \tag{35}$$

which obviously fits into the framework of Section 6. Note that, in a slight abuse of notation, we here changed from time t_i to the time index i in the notation of the stock price models $X^{(k)}$. In view of the discussion in Example 6.4, we observe that this system fails the componentwise comparison principle, if $R^B > R^C$, which is the practically relevant case.

We next run the system of pathwise dynamic programs (30) with $N = 2$, $D = (1 + d)$ and $F_j^{(1)}, F_j^{(2)} : \mathbb{R}^{2(1+d)} \to \mathbb{R}$ given by $F_j^{(\nu)}(z_1, z_2) = z_\nu^{(0)} + H^{(\nu)}(z_1, z_2)\Delta$, $z_\iota = (z_\iota^{(0)}, ..., z_\iota^{(d)}) \in \mathbb{R}^{d+1}$. To this end, we need to construct input martingales and input controls. Again, we shall first pre-compute numerical approximations $\tilde{Y}$, $\tilde{Q}$ and $\tilde{Z}$ of the processes Y, $(E_j[Y_{j+1}])$ and Z by the regression-later approach (with basis functions to be specified later on). Given these approximations we define the controls $r^{(\nu)}$ by

$$r_j^{(\nu)} = \begin{cases} u^{(\nu)}(R^L), & a_\nu(\tilde{Q}_j^{(1)} + \tilde{Q}_j^{(2)}) - \frac{1}{\sigma}\left(\tilde{Z}_j^{[\nu]}\right)^\top \mathbf{1} \geq 0 \\ u^{(\nu)}(R^B), & a_\nu(\tilde{Q}_j^{(1)} + \tilde{Q}_j^{(2)}) - \frac{1}{\sigma}\left(\tilde{Z}_j^{[\nu]}\right)^\top \mathbf{1} < 0 \end{cases}$$

with

$$u^{(1)}(r) = \begin{pmatrix} 1 - r(1-\alpha)\Delta - R^C\alpha\Delta \\ \frac{(r-R^L)\Delta}{\sigma} \cdot \mathbf{1} \\ (R^C - r)(1-\alpha)\Delta \\ 0 \cdot \mathbf{1} \end{pmatrix} \quad \text{and} \quad u^{(2)}(r) = \begin{pmatrix} (R^C - r)\alpha\Delta \\ 0 \cdot \mathbf{1} \\ 1 - r\alpha\Delta - R^C(1-\alpha)\Delta \\ \frac{(r-R^L)\Delta}{\sigma} \cdot \mathbf{1} \end{pmatrix}.$$

We emphasize that these controls satisfy the duality relation (33), if the approximations are replaced by the solution Y, $(E_j[Y_{j+1}])$ and Z. As input martingales we use the Doob martingale of $(\beta\tilde{Y}^{(1)}, \beta\tilde{Y}^{(2)})$, where the conditional expectations in the increments of the Doob martingale are again available in closed form in our approach.

As a numerical example, we consider the valuation of a European call-spread option on the maximum of d assets with maturity T and payoff

$$h(x) = \left(\max_{k=1,\dots,d} x^{(k)} - K_1\right)_+ - 2\left(\max_{k=1,\dots,d} x^{(k)} - K_2\right)_+.$$

Except for adding the collateralization scheme (and, hence, the coupling between the hedger's and counterparty's valuation), this is the same numerical example as in Bender et al. (2015) and we follow their parameter choices

$$(x_0, d, T, K_1, K_2, \sigma, R^L, R^B, R^C, \alpha) = (100, 5, 0.25, 95, 115, 0.2, 0.01, 0.06, 0.02, 0.5)$$

adding only the values of α and R^C. The choice $\alpha = 0.5$ implies that the posted collateral is given by the average of the two parties' value processes $Y^{(1)}$ and $-Y^{(2)}$.

The regression-later algorithm for the computation of the input approximations is run with

J	5	10	15	20	25
$\hat{Y}_0^{(up,1)}$	13.8404 (0.0018)	13.8601 (0.0020)	13.8731 (0.0020)	13.8748 (0.0020)	13.8747 (0.0022)
$\hat{Y}_0^{(low,1)}$	13.8390 (0.0018)	13.8554 (0.0020)	13.8657 (0.0019)	13.8651 (0.0019)	13.8618 (0.0020)
$-\hat{Y}_0^{(low,2)}$	13.2788 (0.0015)	13.2601 (0.0016)	13.2549 (0.0016)	13.2475 (0.0016)	13.2439 (0.0017)
$-\hat{Y}_0^{(lup,2)}$	13.2779 (0.0014)	13.2569 (0.0016)	13.2499 (0.0016)	13.2411 (0.0016)	13.2351 (0.0017)

Table 6: Upper and lower bounds with $N_r = 10^3$ and $N_e = 10^4$ for different time discretizations. Standard deviations are given in brackets.

$N_r = 1,000$ regression paths. At time $i+1$ (where $0 \le i \le J-1$) we apply 7 basis functions, namely 1, $X_{i+1}^{(1)}, \dots X_{i+1}^{(5)}$, and as the seventh basis function an approximation to $E_{i+1}[h(X_J)]$. Precisely, this basis function is defined in terms of an optimal Λ-point quantization $\sum_{\kappa=1}^{\Lambda} p_\kappa \, \delta_{z_\eta}$ of a standard normal distribution $\mathcal{N}$ by

$$\begin{aligned}\eta_{i+1} \;&=\; \sum_{l=1}^{5}\sum_{\kappa=1}^{\Lambda} p_\kappa \sqrt{\frac{T-t_i}{T-t_{i+1}}}\, h\left(X_i^{(l)} e^{\left(R^L - \frac{1}{2}\sigma^2\right)(T-t_i)+\sigma z_\kappa \sqrt{T-t_i}}\right) e^{\frac{z_\kappa^2}{2} - \frac{(\sqrt{T-t_i}z_\kappa - \Delta W_{i+1}^{(l)})^2}{2(T-t_{i+1})}} \\ &\quad\times \prod_{l' \in \{1,\dots,5\}\setminus\{l\}} \mathcal{N}\left(\frac{1}{\sqrt{T-t_{i+1}}}\left(\sqrt{T-t_i}z_\kappa + \frac{\ln(X_i^{(l)}) - \ln(X_i^{(l')})}{\sigma} - \Delta W_{i+1}^{(l')}\right)\right).\end{aligned}$$

As a trade-off between accuracy and computational time, we choose $\Lambda = 25$, but note that this basis function converges to $E_{i+1}[h(X_J)]$, as Λ tends to infinity. The one-step conditional expectations can be expressed as $E_i[\eta_{i+1}] = \bar{q}(i, X_i)$ and $E_i[(\Delta W_{i+1}/\Delta)\eta_{i+1}] = \bar{z}(i, X_i)$ for deterministic functions $\bar{q}$ and $\bar{z}$ which can easily be calculated in closed form. We also apply these functions in order to initialize the regression algorithm at $\tilde{Y}_J = \bar{q}(t_{J-1}, X_{J-1}) + \bar{z}(t_{J-1}, X_{J-1}))^\top \Delta W_J$, where the first term approximates the clean price (with zero interest rate) of the payoff at time t_{J-1}, while the second one approximates the corresponding Delta hedge on the interval $[t_{J-1}, t_J]$.

In order to compute the upper and lower bounds stated in Table 6, we simulate $N_e = 10^4$ evaluation paths of (X_j) and compute the approximations $(\tilde{Y}, \tilde{Q}, \tilde{Z})$ along these paths. We denote by $\hat{Y}_0^{(up,\nu)}$ and $\hat{Y}_0^{(low,\nu)}$ the Monte Carlo estimators for $E[\theta_0^{(up,\nu)}]$ and $E[\theta_0^{(low,\nu)}]$. The run time for this algorithm is again linear in the number J of time steps and is about 10 seconds for $J = 25$.

Table 6 indicates that the quality of the upper and lower bounds is similar for $Y^{(1)}$ and $Y^{(2)}$. This is as expected since the recursions for $Y^{(2)}$ and $Y^{(1)}$ are rather symmetric. With regard to the asymptotic 95%-confidence intervals for $Y_0^{(1)}$ and $Y_0^{(2)}$, we observe two things: First, the relative length of these intervals is about 0.15% for all considered time discretizations, and 25 time steps are quite sufficient in this numerical example. Second, we see that the two parties' valuations differ by about 60 cent, corresponding to about 5 percent of the overall value. So our price bounds are clearly tight enough to distinguish between the two parties' pricing rules.

8 Conclusion

This paper proposes a new method for constructing high-biased and low-biased Monte Carlo estimators for solutions of stochastic dynamic programming equations. When applied to the optimal stopping problem, the method simplifies to the classical primal-dual approach of Rogers (2002) and Haugh and Kogan (2004) (except that the martingale penalty appears as a control variate in the lower bound). Our approach is complementary to earlier generalizations of this methodology by Rogers (2007) and Brown et al. (2010) whose approaches start out from a primal optimization problem rather than from a dynamic programming equation. The resulting representation of high-biased and low-biased estimators in terms of a pathwise recursion makes the method very tractable from a computational point of view. Suitably coupling upper and lower bounds in the recursion enables us to handle situations which are outside the scope of classical primal-dual approaches, because the dynamic programming equation fails to satisfy a comparison principle.

Acknowledgements

Financial support by the Deutsche Forschungsgemeinschaft under grant BE3933/5-1 is gratefully acknowledged.

A Preparations

In this section, we verify in our setting, i.e., under the standing assumptions (C) and (R), a number of technical properties of the control processes which are needed in the later proofs. Note first that, by continuity of F_i, $F_i^{\#}(r_i) = \sup_{z\in\mathbb{Q}^D}(r_i^\top z - F_i(z))$ is $\mathcal{F}_i$-measurable for every $r \in \mathcal{A}_j^F$ and $i = j, \ldots, J-1$ — and so is $G_i^{\#}(\rho_i^{(1)}, \rho_i^{(0)})$ for $(\rho^{(1)}, \rho^{(0)}) \in \mathcal{A}_j^G$. Moreover, the integrability condition on the controls requires that $F_i^{\#}(r_i) < \infty$ and $G_i^{\#}(\rho_i^{(1)}, \rho_i^{(0)}) > -\infty$ P-almost surely, i.e., controls take values in the effective domain of the convex conjugate of F and in the effective domain of the concave conjugate of G, respectively. In particular, by the monotonicity assumption on G in the y-variable, we observe that $\rho_i^{(0)} \geq 0$ P-almost surely.

Finally, we show existence of adapted controls, so that the sets $\mathcal{A}_j^G$ and $\mathcal{A}_j^G$ are always nonempty.

Lemma A.1. *Fix $j \in \{0, \ldots, J-1\}$ and let $f_j : \Omega \times \mathbb{R}^N \to \mathbb{R}$ be a mapping such that, for every $\omega \in \Omega$, the map $x \mapsto f_j(\omega, x)$ is convex, and for every $x \in \mathbb{R}^N$, the map $\omega \mapsto f_j(\omega, x)$ is $\mathcal{F}_j$-measurable. Moreover, suppose that f_j satisfies the following polynomial growth condition: There are a constant $q \geq 0$ and a nonnegative random variable $\alpha_j \in L_j^{\infty-}(\mathbb{R})$ such that*

$$|f_j(x)| \leq \alpha_j(1 + |x|^q), \quad P\text{-a.s.},$$

for every $x \in \mathbb{R}^N$. Then, for every $\overline{Z} \in L_j^{\infty-}(\mathbb{R}^N)$ there exists a random variable $\overline{\rho}_j \in L_j^{\infty-}(\mathbb{R}^N)$

such that $f_j^{\#}\left(\overline{\rho}_j\right) \in L_j^{\infty-}(\mathbb{R})$ *and*

$$f_j(\overline{Z}) = \overline{\rho}_j^\top \overline{Z} - f_j^{\#}(\overline{\rho}_j), \quad P\text{-a.s.} \tag{36}$$

Proof. Let $\overline{Z} \in L_j^{\infty-}(\mathbb{R}^N)$. Notice first that, since f_j is convex and closed, we have $f_j^{\#\#} = f_j$ by Theorem 12.2 in Rockafellar (1970) and thus

$$f_j(\overline{Z}) \geq \rho^\top \overline{Z} - f_j^{\#}(\rho) \tag{37}$$

holds ω-wise for any random variable ρ. We next show that there exists an $\mathcal{F}_j$-measurable random variable $\overline{\rho}_j$ for which (37) holds with P-almost sure equality. To this end, we apply Theorem 7.4 in Cheridito et al. (2015) which yields the existence of an $\mathcal{F}_j$-measurable subgradient to f_j, i.e., existence of an $\mathcal{F}_j$-measurable random variable $\overline{\rho}_j$ such that for all $\mathcal{F}_j$-measurable $\mathbb{R}^N$-valued random variables Z

$$f_j\left(\overline{Z} + Z\right) - f_j\left(\overline{Z}\right) \geq \overline{\rho}_j^\top Z, \quad P\text{-a.s.} \tag{38}$$

From (38) (with the choice $Z = z - \bar{Z}$ for $z \in \mathbb{Q}^N$), we conclude that

$$\overline{\rho}_j^\top \overline{Z} - f_j\left(\overline{Z}\right) \geq \sup_{z \in \mathbb{Q}^N} \overline{\rho}_j^\top (z) - f_j(z) = f_j^{\#}(\overline{\rho}_j), \quad P\text{-a.s.}, \tag{39}$$

by continuity of f_j, which is the converse of (37), proving P-almost sure equality for $\rho = \overline{\rho}_j$ and thus (36). We next show that $\overline{\rho}_j$ satisfies the required integrability conditions, i.e., $\overline{\rho}_j \in L^{\infty-}(\mathbb{R}^N)$ and $f_j^{\#}(\overline{\rho}_j) \in L^{\infty-}(\mathbb{R})$. To this end, we first prove that $\overline{\rho}_j^\top Z \in L^{\infty-}(\mathbb{R})$ for any $Z \in L_j^{\infty-}(\mathbb{R}^N)$. Due to (38) and the Minkowski inequality and since $a \leq b$ implies $a_+ \leq |b|$, it follows for $Z \in L_j^{\infty-}(\mathbb{R}^N)$ that, for every $p \geq 1$,

$$\left(E\left[\left|\left(\overline{\rho}_j^\top Z\right)_+\right|^p\right]\right)^{\frac{1}{p}} \leq \left(E\left[\left|f_j\left(\overline{Z} + Z\right)\right|^p\right]\right)^{\frac{1}{p}} + \left(E\left[\left|f_j\left(\overline{Z}\right)\right|^p\right]\right)^{\frac{1}{p}} < \infty,$$

since f_j is of polynomial growth with 'random constant' $\alpha_j \in L_j^{\infty-}(\mathbb{R})$ and $\overline{Z}, Z$ are members of $L_j^{\infty-}(\mathbb{R}^N)$ by assumption. Applying the same argument to $\tilde{Z} = -Z$ yields

$$E\left[\left|\left(\overline{\rho}_j^\top Z\right)_-\right|^p\right] = E\left[\left|\left(\overline{\rho}_j^\top \tilde{Z}\right)_+\right|^p\right] < \infty,$$

since (38) holds for all $\mathcal{F}_j$-measurable random variables Z and $\tilde{Z}$ inherits the integrability of Z. We thus conclude that

$$E\left[\left|\overline{\rho}_j^\top Z\right|^p\right] < \infty \quad \text{and} \quad E\left[\left|\overline{\rho}_j\right|^p\right] < \infty,$$

where the second claim follows from the first by taking $Z = \operatorname{sgn}(\overline{\rho}_j)$ with the sign function applied componentwise. In order to show that $f_j^{\#}(\overline{\rho}_j) \in L^{\infty-}(\mathbb{R})$, we start with (36) and apply the Minkowski inequality to conclude that

$$\left(E\left[\left|f_j^{\#}\left(\overline{\rho}_j\right)\right|^p\right]\right)^{\frac{1}{p}} \leq \left(E\left[\left|\overline{\rho}_j^\top \overline{Z}\right|^p\right]\right)^{\frac{1}{p}} + \left(E\left[\left|f_j\left(\overline{Z}\right)\right|^p\right]\right)^{\frac{1}{p}} < \infty.$$

□

B Proofs of Section 4

B.1 Proof of Theorem 4.3

We first show that the implications (*b*) ⇒ (*c*) ⇒ (*a*) hold in an analogous formulation without the additional assumption that $G_i(z, y) = y$.

Proposition B.1. *Suppose (R) and (C), and consider the following assertions:*

(a) The comparison principle (Comp) holds.

(b') For every $(\rho^{(1)}, \rho^{(0)}) \in \mathcal{A}_0^G$ *and* $r \in \mathcal{A}_0^F$ *the following positivity condition is fulfilled: For every* $i = 0, \dots, J-1$

$$(\rho_i^{(1)} + \rho_i^{(0)} r_i)^\top \beta_{i+1} \geq 0, \quad P\text{-a.s.}$$

(c') For every $j = 0, \dots, J-1$ *and any two random variables* $Y^{(1)}, Y^{(2)} \in L^{\infty-}(\mathbb{R})$ *with* $Y^{(1)} \geq Y^{(2)}$ P-a.s.*, the following monotonicity condition is satisfied:*

$$G_j(E_j[\beta_{j+1}Y^{(1)}], F_j(E_j[\beta_{j+1}Y^{(1)}])) \geq G_j(E_j[\beta_{j+1}Y^{(2)}], F_j(E_j[\beta_{j+1}Y^{(2)}])), \quad P\text{-a.s.}$$

Then, $(b') \Rightarrow (c') \Rightarrow (a)$.

Proof. $(b') \Rightarrow (c')$: Fix $j \in \{0, \dots, J-1\}$ and let $Y^{(1)}$ and $Y^{(2)}$ be random variables which are in $L^{\infty-}(\mathbb{R})$ and satisfy $Y^{(1)} \geq Y^{(2)}$. By Lemma A.1, there are $r \in \mathcal{A}_0^F$ and $(\rho^{(1)}, \rho^{(0)}) \in \mathcal{A}_0^G$ such that

$$\begin{aligned} F_j\left(E_j\left[\beta_{j+1}Y^{(2)}\right]\right) &= r_j^\top E_j\left[\beta_{j+1}Y^{(2)}\right] - F_j^{\#}\left(r_j\right) \\ G_j\left(E_j\left[\beta_{j+1}Y^{(1)}\right], F_j(E_j[\beta_{j+1}Y^{(1)}])\right) &= \left(\rho_j^{(1)}\right)^\top E_j\left[\beta_{j+1}Y^{(1)}\right] + \rho_j^{(0)} F_j\left(E_j[\beta_{j+1}Y^{(1)}]\right) \\ &\quad -G_j^{\#}\left(\rho_j^{(1)}, \rho_j^{(0)}\right), \end{aligned}$$

P-almost surely. Hence, by (13), (b') and (14) we obtain

$$\begin{aligned} & G_j\left(E_j\left[\beta_{j+1}Y^{(2)}\right], F_j\left(E_j\left[\beta_{j+1}Y^{(2)}\right]\right)\right) \\ \leq\ & \left(\rho_j^{(1)}\right)^\top E_j\left[\beta_{j+1}Y^{(2)}\right] + \rho_j^{(0)} F_j\left(E_j[\beta_{j+1}Y^{(2)}]\right) - G_j^{\#}\left(\rho_j^{(1)}, \rho_j^{(0)}\right) \\ =\ & E_j\left[\left(\rho_j^{(1)} + \rho_j^{(0)} r_j\right)^\top \beta_{j+1}Y^{(2)} - \rho_j^{(0)} F_j^{\#}\left(r_j\right) - G_j^{\#}\left(\rho_j^{(1)}, \rho_j^{(0)}\right)\right] \end{aligned}$$

$$
\begin{aligned}
&\leq E_j\left[\left(\rho_j^{(1)}+\rho_j^{(0)}r_j\right)^\top \beta_{j+1}Y^{(1)}-\rho_j^{(0)}F_j^{\#}(r_j)-G_j^{\#}\left(\rho_j^{(1)},\rho_j^{(0)}\right)\right] \\
&\leq \left(\rho_j^{(1)}\right)^\top E_j\left[\beta_{j+1}Y^{(1)}\right]+\rho_j^{(0)}F_j\left(E_j[\beta_{j+1}Y^{(1)}]\right)-G_j^{\#}\left(\rho_j^{(1)},\rho_j^{(0)}\right) \\
&= G_j\left(E_j\left[\beta_{j+1}Y^{(1)}\right],F_j\left(E_j\left[\beta_{j+1}Y^{(1)}\right]\right)\right).
\end{aligned}
$$

$(c') \Rightarrow (a)$: We prove this implication by backward induction. Let Y^{up} and Y^{low} respectively be super- and subsolutions of (12). Then the assertion is trivially true for $j = J$. Now assume, that the assertion is true for $j+1 \in \{1,\dots,J\}$. It follows by (c') and the definition of a sub- and supersolution that

$$
Y_j^{up} \geq G_j(E_j[\beta_{j+1}Y_{j+1}^{up}],F_j(E_j[\beta_{j+1}Y_{j+1}^{up}])) \geq G_j(E_j[\beta_{j+1}Y_{j+1}^{low}],F_j(E_j[\beta_{j+1}Y_{j+1}^{low}])) \geq Y_j^{low}.
$$

□

Proof of Theorem 4.3. Notice first, that under the additional assumption $G_i(z,y)=y$, assertions (b'), (c') in Proposition B.1 coincide with assertions (b), (c) in Theorem 4.3, because the vector $(0,...,0,1)^\top \in \mathbb{R}^{D+1}$ is the only control in $\mathcal{A}_0^G$ by linearity of G. It, hence, remains to show:

$(a) \Rightarrow (b)$: We prove the contraposition. Hence, we assume that there exists a $\bar{r} \in \mathcal{A}_0^F$ and a $j_0 \in \{0,\dots,J-1\}$ such that $P(\{\bar{r}_{j_0}^\top \beta_{j_0+1} < 0\}) > 0$. Then we define the process $\overline{Y}$ by

$$
\overline{Y}_i = \begin{cases} Y_i, & i > j_0+1 \\ Y_{j_0+1} - n1_{\{\bar{r}_{j_0}^\top \beta_{j_0+1}<0\}}, & i = j_0+1 \\ \bar{r}_i^\top E_i[\beta_{i+1}\overline{Y}_{i+1}] - F_i^{\#}(\bar{r}_i), & i \leq j_0, \end{cases}
$$

for $n \in \mathbb{N}$ which we fix later on. In view of (14), it follows easily that $\overline{Y}$ is a subsolution to (12). Now we observe that

$$
\overline{Y}_{j_0} - Y_{j_0} = E_{j_0}\left[(\bar{r}_{j_0}-r_{j_0}^*)^\top \beta_{j_0+1}Y_{j_0+1}\right] + nE_{j_0}\left[(\bar{r}_{j_0}^\top\beta_{j_0+1})_-\right] - F_{j_0}^{\#}(\bar{r}_{j_0}) + F_{j_0}^{\#}(r_{j_0}^*),
$$

where $r^* \in \mathcal{A}_{j_0}^F$ is such that for all $j = j_0,\dots,J-1$

$$
(r_j^*)^\top E_j[\beta_{j+1}Y_{j+1}] - F_j^{\#}(r_j^*) = F_j(E_j[\beta_{j+1}Y_{j+1}]),
$$

see Lemma A.1. In a next step we define the set $A_{j_0,N}$ by

$$
\begin{aligned}
A_{j_0,N} &= \left\{E_{j_0}\left[(\bar{r}_{j_0}^\top\beta_{j_0+1})_-\right] \geq \frac{1}{N}\right\} \\
&\cap \left\{E_{j_0}\left[(\bar{r}_{j_0}-r_{j_0}^*)^\top \beta_{j_0+1}Y_{j_0+1}\right] - F_{j_0}^{\#}(\bar{r}_{j_0}) + F_{j_0}^{\#}(r_{j_0}^*) > -N\right\}.
\end{aligned}
$$

For $N \in \mathbb{N}$ sufficiently large (which is fixed from now on), we get that $P(A_{j_0,N}) > 0$ and therefore

$$(\overline{Y}_{j_0} - Y_{j_0})1_{A_{j_0,N}} > -N + \frac{n}{N} = 0$$

for $n = N^2$, which means that the comparison principle is violated for the subsolution $\overline{Y}$ (with this choice of n) and the (super-)solution Y.

□

B.2 Proof of Theorem 4.5

The proof of Theorem 4.5 is prepared by two propositions. The first of them shows almost sure optimality of optimal controls and martingales. For convenience, we show this simultaneously for the processes Θ^{up} and Θ^{low} from Theorem 4.7.

Proposition B.2. *With the notation in Theorems 4.5 and 4.7, we have for every $i = j, \dots J$,*

$$Y_i = \theta_i^{up}(\rho^{(1,*)}, \rho^{(0,*)}, r^*, M^*) = \Theta_i^{up}(\rho^{(1,*)}, \rho^{(0,*)}, M^*) = \theta_i^{low}(\rho^{(1,*)}, \rho^{(0,*)}, r^*, M^*) = \Theta_i^{low}(r^*, M^*).$$

Proof. The proof is by backward induction on $i = J, \dots, j$. The case $i = J$ is obvious as all five processes have the same terminal condition ξ by construction. Now suppose the claim is already shown for $i + 1 \in \{j + 1, \dots, J\}$. Then, applying the induction hypothesis to the right hand side of the recursion formulas, we observe

$$\begin{aligned}
\Theta_i^{up,*} &:= \Theta_i^{up}(\rho^{(1,*)}, \rho^{(0,*)}, M^*) = \theta_i^{up}(\rho^{(1,*)}, \rho^{(0,*)}, r^*, M^*), \\
\Theta_i^{low,*} &:= \Theta_i^{low}(\rho^{(1,*)}, \rho^{(0,*)}, M^*) = \theta_i^{low}(\rho^{(1,*)}, \rho^{(0,*)}, r^*, M^*).
\end{aligned}$$

Exploiting again the induction hypothesis as well as the definition of the Doob martingale and the duality relation (25) we obtain

$$\begin{aligned}
\Theta_i^{up,*} &= \left(\rho_i^{(1,*)}\right)^\top \left(\beta_{i+1}Y_{i+1} - \Delta M_{i+1}^*\right) + \rho_i^{(0,*)} F_i(\beta_{i+1}Y_{i+1} - \Delta M_{i+1}^*) - G_i^\#(\rho_i^{(1,*)}, \rho_i^{(0,*)}) \\
&= \left(\rho_i^{(1,*)}\right)^\top E_i[\beta_{i+1}Y_{i+1}] + \rho_i^{(0,*)} F_i(E_i[\beta_{i+1}Y_{i+1}]) - G_i^\#(\rho_i^{(1,*)}, \rho_i^{(0,*)}) \\
&= G_i\left(E_i[\beta_{i+1}Y_{i+1}], F_i(E_i[\beta_{i+1}Y_{i+1}])\right) = Y_i.
\end{aligned}$$

An analogous argument, making use of (26), shows $\Theta_i^{low,*} = Y_i$. □

The key step in the proof of Theorem 4.5 is the following alternative recursion formula for θ_j^{up} and θ_j^{low}. This result enables us to establish inequalities between Y_j, $E_j[\theta_j^{up}]$ and $E_j[\theta_j^{low}]$, replacing, in a sense, the comparison principle (Comp).

Proposition B.3. *Suppose (R) and (C) and let $M \in \mathcal{M}_D$. Then, for every $j = 0, \dots, J$ and*

$\left(\rho^{(1)}, \rho^{(0)}\right) \in \mathcal{A}_j^G$ *and* $r \in \mathcal{A}_j^F$, *we have for all* $i = j, \ldots, J$ *the* P*-almost sure identities*

$$\theta_i^{up}(\rho^{(1)}, \rho^{(0)}, r, M) = \sup_{u \in \mathbb{R}^D} \Phi_{i+1}\left(\rho_i^{(1)}, \rho_i^{(0)}, u, \theta_{i+1}^{up}(\rho^{(1)}, \rho^{(0)}, r, M), \theta_{i+1}^{low}(\rho^{(1)}, \rho^{(0)}, r, M), \Delta M_{i+1}\right)$$

$$\theta_i^{low}(\rho^{(1)}, \rho^{(0)}, r, M) = \inf_{\left(v^{(1)}, v^{(0)}\right) \in \mathbb{R}^{D+1}} \Phi_{i+1}\left(v^{(1)}, v^{(0)}, r_i, \theta_{i+1}^{low}(\rho^{(1)}, \rho^{(0)}, r, M), \theta_{i+1}^{up}(\rho^{(1)}, \rho^{(0)}, r, M), \Delta M_{i+1}\right),$$

where $\Phi_{J+1}(v^{(1)}, v^{(0)}, u, \vartheta_1, \vartheta_2, m) = \xi$ *and*

$$\begin{aligned}
&\Phi_{i+1}\left(v^{(1)}, v^{(0)}, u, \vartheta_1, \vartheta_2, m\right) \\
&= \left(\left(v^{(1)}\right)^\top \beta_{i+1}\right)_+ \vartheta_1 - \left(\left(v^{(1)}\right)^\top \beta_{i+1}\right)_- \vartheta_2 - \left(v^{(1)}\right)^\top m \\
&\quad + v^{(0)}\left(\left(u^\top \beta_{i+1}\right)_+ \vartheta_1 - \left(u^\top \beta_{i+1}\right)_- \vartheta_2 - u^\top m - F_i^{\#}(u)\right) - G_i^{\#}\left(v^{(1)}, v^{(0)}\right)
\end{aligned}$$

for $i = j, ..., J-1$*. In particular,* $\theta_i^{low}(\rho^{(1)}, \rho^{(0)}, r, M) \leq \theta_i^{up}(\rho^{(1)}, \rho^{(0)}, r, M)$ *for every* $i = j, \ldots, J$*.*

Proof. First we fix $j \in \{0, \ldots, J-1\}$, $M \in \mathcal{M}_D$ and controls $\left(\rho^{(1)}, \rho^{(0)}\right)$ and r in $\mathcal{A}_j^G$ respectively $\mathcal{A}_j^F$ and define θ^{up} and θ^{low} by (22). To lighten the notation, we set

$$\Phi_{i+1}^{low}\left(v^{(1)}, v^{(0)}, r_i\right) = \Phi_{i+1}\left(v^{(1)}, v^{(0)}, r_i, \theta_{i+1}^{low}(\rho^{(1)}, \rho^{(0)}, r, M), \theta_{i+1}^{up}(\rho^{(1)}, \rho^{(0)}, r, M), \Delta M_{i+1}\right)$$

and define Φ_{i+1}^{up} accordingly (interchanging the roles of θ^{up} and θ^{low}). We show the assertion by backward induction on $i = J, \ldots, j$ with the case $i = J$ being trivial since $\theta_J^{up} = \theta_J^{low} = \Phi_{J+1} = \xi$ by definition. Now suppose that the assertion is true for $i+1$. For any $(v^{(1)}, v^{(0)}) \in \mathbb{R}^{D+1}$ we obtain, by (13), the following upper bound for θ_i^{low}:

$$\begin{aligned}
&\Phi_{i+1}^{low}\left(v^{(1)}, v^{(0)}, r_i\right) \\
&= \left(v^{(1)}\right)^\top \left(\beta_{i+1}\left(\theta_{i+1}^{low} 1_{\{(v^{(1)})^\top \beta_{i+1} \geq 0\}} + \theta_{i+1}^{up} 1_{\{(v^{(1)})^\top \beta_{i+1} < 0\}}\right) - \Delta M_{i+1}\right) \\
&\quad + v^{(0)}\left(\left(r_i^\top \beta_{i+1}\right)_+ \theta_{i+1}^{low} - \left(r_i^\top \beta_{i+1}\right)_- \theta_{i+1}^{up} - r_i^\top \Delta M_{i+1} - F_i^{\#}(r_i)\right) - G_i^{\#}\left(v^{(1)}, v^{(0)}\right) \\
&\geq G_i\Bigg(\beta_{i+1}\left(\theta_{i+1}^{low} 1_{\{(v^{(1)})^\top \beta_{i+1} \geq 0\}} + \theta_{i+1}^{up} 1_{\{(v^{(1)})^\top \beta_{i+1} < 0\}}\right) - \Delta M_{i+1}, \\
&\qquad \left(r_i^\top \beta_{i+1}\right)_+ \theta_{i+1}^{low} - \left(r_i^\top \beta_{i+1}\right)_- \theta_{i+1}^{up} - r_i^\top \Delta M_{i+1} - F_i^{\#}(r_i)\Bigg)
\end{aligned}$$

$$\begin{aligned}
&\geq \min_{\iota\in\{up,low\}} G_i\Big(\beta_{i+1}\theta_{i+1}^{\iota}-\Delta M_{i+1}, \left(r_i^{\top}\beta_{i+1}\right)_+\theta_{i+1}^{low}-\left(r_i^{\top}\beta_{i+1}\right)_-\theta_{i+1}^{up} \\
&\qquad -r_i^{\top}\Delta M_{i+1}-F_i^{\#}(r_i)\Big) \\
&= \theta_i^{low}.
\end{aligned}$$

We emphasize that this chain of inequalities holds for every $\omega\in\Omega$. Hence,

$$\inf_{(v^{(1)},v^{(0)})\in\mathbb{R}^{D+1}} \Phi_{i+1}^{low}\left(v^{(1)},v^{(0)},r_i\right)\geq\theta_i^{low}$$

for every $\omega\in\Omega$. To conclude the argument for θ_i^{low}, it remains to show that the converse inequality holds P-almost surely. Thanks to (13) we get

$$\begin{aligned}
&G_i\left(\beta_{i+1}\theta_{i+1}^{\iota}-\Delta M_{i+1}, \left(r_i^{\top}\beta_{i+1}\right)_+\theta_{i+1}^{low}-\left(r_i^{\top}\beta_{i+1}\right)_-\theta_{i+1}^{up}-r_i^{\top}\Delta M_{i+1}-F_i^{\#}(r_i)\right) \\
&= \inf_{(v^{(1)},v^{(0)})\in\mathbb{R}^{D+1}} \left(v^{(1)}\right)^{\top}\left(\beta_{i+1}\theta_{i+1}^{\iota}-\Delta M_{i+1}\right)+v^{(0)}\Big(\left(r_i^{\top}\beta_{i+1}\right)_+\theta_{i+1}^{low}-\left(r_i^{\top}\beta_{i+1}\right)_-\theta_{i+1}^{up} \\
&\qquad -r_i^{\top}\Delta M_{i+1}-F_i^{\#}(r_i)\Big)-G_i^{\#}\left(v^{(1)},v^{(0)}\right).
\end{aligned}$$

Together with $\theta_{i+1}^{up}\geq\theta_{i+1}^{low}$ P-a.s. (by the induction hypothesis) we obtain

$$\begin{aligned}
\theta_i^{low} &= \min_{\iota\in\{up,low\}} G_i\Big(\beta_{i+1}\theta_{i+1}^{\iota}-\Delta M_{i+1}, \left(r_i^{\top}\beta_{i+1}\right)_+\theta_{i+1}^{low}-\left(r_i^{\top}\beta_{i+1}\right)_-\theta_{i+1}^{up} \\
&\qquad -r_i^{\top}\Delta M_{i+1}-F_i^{\#}(r_i)\Big) \\
&= \min_{\iota\in\{up,low\}} \inf_{(v^{(1)},v^{(0)})\in\mathbb{R}^{D+1}} \left(v^{(1)}\right)^{\top}\beta_{i+1}\theta_{i+1}^{\iota}-\left(v^{(1)}\right)^{\top}\Delta M_{i+1} \\
&\qquad +v^{(0)}\left(\left(r_i^{\top}\beta_{i+1}\right)_+\theta_{i+1}^{low}-\left(r_i^{\top}\beta_{i+1}\right)_-\theta_{i+1}^{up}-r_i^{\top}\Delta M_{i+1}-F_i^{\#}(r_i)\right)-G_i^{\#}\left(v^{(1)},v^{(0)}\right) \\
&\geq \inf_{(v^{(1)},v^{(0)})\in\mathbb{R}^{D+1}} \left(\left(v^{(1)}\right)^{\top}\beta_{i+1}\right)_+\theta_{i+1}^{low}-\left(\left(v^{(1)}\right)^{\top}\beta_{i+1}\right)_-\theta_{i+1}^{up}-\left(v^{(1)}\right)^{\top}\Delta M_{i+1} \\
&\qquad +v^{(0)}\left(\left(r_i^{\top}\beta_{i+1}\right)_+\theta_{i+1}^{low}-\left(r_i^{\top}\beta_{i+1}\right)_-\theta_{i+1}^{up}-r_i^{\top}\Delta M_{i+1}-F_i^{\#}(r_i)\right)-G_i^{\#}\left(v^{(1)},v^{(0)}\right) \\
&= \inf_{(v^{(1)},v^{(0)})\in\mathbb{R}^{D+1}} \Phi_{i+1}^{low}\left(v^{(1)},v^{(0)},r_i\right), \quad P\text{-a.s.}
\end{aligned}$$

We next turn to θ_i^{up} where the overall strategy of proof is similar. Recall first that the monotonicity of G in the y-component implies existence of a set $\bar{\Omega}_\rho$ (depending on $\rho^{(0)}$) of full P-measure such that $\rho_k^{(0)}(\omega)\geq 0$ for every $\omega\in\bar{\Omega}_\rho$ and $k=j,\ldots,J-1$. By (14) we find that, for any $u\in\mathbb{R}^D$,

$\Phi_{i+1}^{up}(\rho_i^{(0)},\rho_i^{(1)},u)$ is a lower bound for θ_i^{up} on $\bar{\Omega}_\rho$:

$$\begin{aligned}
&\Phi_{i+1}^{up}\left(\rho_i^{(1)},\rho_i^{(0)},u\right)\\
&= \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_+ \theta_{i+1}^{up} - \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_- \theta_{i+1}^{low} - \left(\rho_i^{(1)}\right)^\top \Delta M_{i+1}\\
&\quad +\rho_i^{(0)}\left(\left(u^\top \beta_{i+1}\right)_+ \theta_{i+1}^{up} - \left(u^\top \beta_{i+1}\right)_- \theta_{i+1}^{low} - u^\top \Delta M_{i+1} - F_i^{\#}(u)\right) - G_i^{\#}\left(\rho_i^{(1)},\rho_i^{(0)}\right)\\
&\le \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_+ \theta_{i+1}^{up} - \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_- \theta_{i+1}^{low} - \left(\rho_i^{(1)}\right)^\top \Delta M_{i+1}\\
&\quad +\rho_i^{(0)} F_i\left(\beta_{i+1}\left(\theta_{i+1}^{up}1_{\{u^\top\beta_{i+1}\ge 0\}} + \theta_{i+1}^{low}1_{\{u^\top\beta_{i+1}<0\}}\right) - \Delta M_{i+1}\right) - G_i^{\#}\left(\rho_i^{(1)},\rho_i^{(0)}\right)\\
&\le \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_+ \theta_{i+1}^{up} - \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_- \theta_{i+1}^{low} - \left(\rho_i^{(1)}\right)^\top \Delta M_{i+1}\\
&\quad +\rho_i^{(0)} \max_{\iota\in\{up,low\}} F_i(\beta_{i+1}\theta_{i+1}^{\iota} - \Delta M_{i+1}) - G_i^{\#}\left(\rho_i^{(1)},\rho_i^{(0)}\right)\\
&= \theta_i^{up}.
\end{aligned}$$

Hence,

$$\sup_{u\in\mathbb{R}^D} \Phi_{i+1}^{up}\left(\rho_i^{(1)},\rho_i^{(0)},u\right) \le \theta_i^{up}$$

on $\bar{\Omega}_\rho$, and, thus, P-almost surely. To complete the proof of the proposition, we show the converse inequality. As $\theta_{i+1}^{up} \ge \theta_{i+1}^{low}$ P-a.s., we conclude, by (14),

$$\begin{aligned}
\theta_i^{up} &= \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_+ \theta_{i+1}^{up} - \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_- \theta_{i+1}^{low} - \left(\rho_i^{(1)}\right)^\top \Delta M_{i+1}\\
&\quad +\rho_i^{(0)} \max_{\iota\in\{up,low\}} F_i(\beta_{i+1}\theta_{i+1}^{\iota} - \Delta M_{i+1}) - G_i^{\#}\left(\rho_i^{(1)},\rho_i^{(0)}\right)\\
&= \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_+ \theta_{i+1}^{up} - \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_- \theta_{i+1}^{low} - \left(\rho_i^{(1)}\right)^\top \Delta M_{i+1}\\
&\quad +\rho_i^{(0)} \max_{\iota\in\{up,low\}} \sup_{u\in\mathbb{R}^D}\left(u^\top\beta_{i+1}\theta_{i+1}^{\iota} - u^\top \Delta M_{i+1} - F_i^{\#}(u)\right) - G_i^{\#}\left(\rho_i^{(1)},\rho_i^{(0)}\right)\\
&\le \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_+ \theta_{i+1}^{up} - \left(\left(\rho_i^{(1)}\right)^\top \beta_{i+1}\right)_- \theta_{i+1}^{low} - \left(\rho_i^{(1)}\right)^\top \Delta M_{i+1}\\
&\quad +\rho_i^{(0)} \sup_{u\in\mathbb{R}^D}\left(\left(u^\top \beta_{i+1}\right)_+ \theta_{i+1}^{up} - \left(u^\top \beta_{i+1}\right)_- \theta_{i+1}^{low} - u^\top \Delta M_{i+1} - F_i^{\#}(u)\right) - G_i^{\#}\left(\rho_i^{(1)},\rho_i^{(0)}\right)\\
&= \sup_{u\in\mathbb{R}^D} \Phi_{i+1}^{up}\left(\rho_i^{(1)},\rho_i^{(0)},u\right), \quad P\text{-a.s.}
\end{aligned}$$

As $\Phi_{i+1}\left(v^{(1)},v^{(0)},u,\vartheta_1,\vartheta_2,m\right)$ is increasing in ϑ_1 and decreasing in ϑ_2, we finally get

$$\theta_i^{up} = \sup_{u\in\mathbb{R}^D} \Phi_{i+1}\left(\rho_i^{(1)},\rho_i^{(0)},u,\theta_{i+1}^{up},\theta_{i+1}^{low},\Delta M_{i+1}\right) \ge \sup_{u\in\mathbb{R}^D} \Phi_{i+1}\left(\rho_i^{(1)},\rho_i^{(0)},u,\theta_{i+1}^{low},\theta_{i+1}^{up},\Delta M_{i+1}\right)$$

$$\geq \inf_{\left(v^{(1)},v^{(0)}\right)\in\mathbb{R}^{D+1}} \Phi_{i+1}\left(v^{(1)},v^{(0)},r_i,\theta_{i+1}^{low},\theta_{i+1}^{up},\Delta M_{i+1}\right)=\theta_i^{low}, \quad P\text{-a.s.},$$

as $\theta_{i+1}^{up} \geq \theta_{i+1}^{low}$ P-a.s. by the induction hypothesis. □

We are now in the position to complete the proof of Theorem 4.5.

Proof of Theorem 4.5. Let $j \in \{0,\dots,J-1\}$ be fixed from now on. Due to Propositions B.2 and B.3, it only remains to show that $E_i[\theta_i^{low}] \leq Y_i \leq E_i[\theta_i^{up}]$ for $i = j,...,J$. We prove this by backward induction on i. To this end, we fix $M \in \mathcal{M}_D$ and controls $\left(\rho^{(1)},\rho^{(0)}\right)$ and r in $\mathcal{A}_j^G$ respectively $\mathcal{A}_j^F$, as well as 'optimizers' $\left(\rho^{(1,*)},\rho^{(0,*)}\right)$ and r^* in $\mathcal{A}_j^G$ respectively $\mathcal{A}_j^F$ which satisfy the duality relations (25) and (26). By definition of θ^{up} and θ^{low} the assertion is trivially true for $i = J$. Suppose that the assertion is true for $i+1$. Recalling Proposition B.3 and applying the tower property of the conditional expectation as well as the induction hypothesis, we get

$$\begin{aligned}
E_i[\theta_i^{low}] &= E_i\Bigg[\inf_{\left(v^{(1)},v^{(0)}\right)\in\mathbb{R}^{D+1}}\Bigg(\left(\left(v^{(1)}\right)^\top\beta_{i+1}\right)_+\theta_{i+1}^{low}-\left(\left(v^{(1)}\right)^\top\beta_{i+1}\right)_-\theta_{i+1}^{up} \\
&\quad -\left(v^{(1)}\right)^\top\Delta M_{i+1}+v^{(0)}\left(\left(r_i^\top\beta_{i+1}\right)_+\theta_{i+1}^{low}-\left(r_i^\top\beta_{i+1}\right)_-\theta_{i+1}^{up}-r_i^\top\Delta M_{i+1}\right. \\
&\quad \left.-F_i^{\#}(r_i)\right)-G_i^{\#}\left(v^{(1)},v^{(0)}\right)\Bigg)\Bigg] \\
&\leq E_i\Bigg[\left(\left(\rho_i^{(1,*)}\right)^\top\beta_{i+1}\right)_+E_{i+1}\left[\theta_{i+1}^{low}\right]-\left(\left(\rho_i^{(1,*)}\right)^\top\beta_{i+1}\right)_-E_{i+1}\left[\theta_{i+1}^{up}\right] \\
&\quad -\left(\rho_i^{(1,*)}\right)^\top\Delta M_{i+1}+\rho_i^{(0,*)}\left(\left(r_i^\top\beta_{i+1}\right)_+E_{i+1}\left[\theta_{i+1}^{low}\right]-\left(r_i^\top\beta_{i+1}\right)_-E_{i+1}\left[\theta_{i+1}^{up}\right]\right. \\
&\quad \left.-r_i^\top\Delta M_{i+1}-F_i^{\#}(r_i)\right)-G_i^{\#}\left(\rho_i^{(1,*)},\rho_i^{(0,*)}\right)\Bigg] \\
&\leq E_i\left[\left(\rho_i^{(1,*)}\right)^\top\beta_{i+1}Y_{i+1}+\rho_i^{(0,*)}\left(r_i^\top\beta_{i+1}Y_{i+1}-F_i^{\#}(r_i)\right)-G_i^{\#}\left(\rho_i^{(1,*)},\rho_i^{(0,*)}\right)\right] \\
&\leq G_i\left(E_i\left[\beta_{i+1}Y_{i+1}\right],F_i\left(E_i\left[\beta_{i+1}Y_{i+1}\right]\right)\right)=Y_i.
\end{aligned}$$

Here, the last inequality is an immediate consequence of (14), the nonnegativity of $\rho_i^{(0,*)}$ and the duality relation (25). Applying an analogous argument, we obtain that $E_i[\theta_i^{up}] \geq Y_i$. Indeed,

$$\begin{aligned}
E_i[\theta_i^{up}] &= E_i\Bigg[\left(\left(\rho_i^{(1)}\right)^\top\beta_{i+1}\right)_+\theta_{i+1}^{up}-\left(\left(\rho_i^{(1)}\right)^\top\beta_{i+1}\right)_-\theta_{i+1}^{low}-\left(\rho_i^{(1)}\right)^\top\Delta M_{i+1} \\
&\quad +\rho_i^{(0)}\sup_{u\in\mathbb{R}^D}\left(\left(u^\top\beta_{i+1}\right)_+\theta_{i+1}^{up}-\left(u^\top\beta_{i+1}\right)_-\theta_{i+1}^{low}-u^\top\Delta M_{i+1}-F_i^{\#}(u)\right) \\
&\quad -G_i^{\#}\left(\rho_i^{(1)},\rho_i^{(0)}\right)\Bigg] \\
&\geq E_i\Bigg[\left(\left(\rho_i^{(1)}\right)^\top\beta_{i+1}\right)_+E_{i+1}\left[\theta_{i+1}^{up}\right]-\left(\left(\rho_i^{(1)}\right)^\top\beta_{i+1}\right)_-E_{i+1}\left[\theta_{i+1}^{low}\right]-\left(\rho_i^{(1)}\right)^\top\Delta M_{i+1}
\end{aligned}$$

$$
\begin{aligned}
& +\rho_i^{(0)}\left(\left(\left(r_i^*\right)^\top \beta_{i+1}\right)_+ E_{i+1}\left[\theta_{i+1}^{up}\right]-\left(\left(r_i^*\right)^\top \beta_{i+1}\right)_- E_{i+1}\left[\theta_{i+1}^{low}\right]-\left(r_i^*\right)^\top \Delta M_{i+1}\right. \\
& \left.\left.-F_i^{\#}\left(r_i^*\right)\right)-G_i^{\#}\left(\rho_i^{(1)}, \rho_i^{(0)}\right)\right] \\
\geq\ & \left(\rho_i^{(1)}\right)^\top E_i\left[\beta_{i+1} Y_{i+1}\right]+\rho_i^{(0)}\left(\left(r_i^*\right)^\top E_i\left[\beta_{i+1} Y_{i+1}\right]-F_i^{\#}\left(r_i^*\right)\right)-G_i^{\#}\left(\rho_i^{(1)}, \rho_i^{(0)}\right) \\
=\ & \left(\rho_i^{(1)}\right)^\top E_i\left[\beta_{i+1} Y_{i+1}\right]+\rho_i^{(0)} F_i\left(E_i\left[\beta_{i+1} Y_{i+1}\right]\right)-G_i^{\#}\left(\rho_i^{(1)}, \rho_i^{(0)}\right) \\
\geq\ & G_i\left(E_i\left[\beta_{i+1} Y_{i+1}\right], F_i\left(E_i\left[\beta_{i+1} Y_{i+1}\right]\right)\right)=Y_i .
\end{aligned}
$$

making now use of the nonnegativity of $\rho_i^{(0)}$, the duality relation (26), and (13). This establishes $E_i[\theta_i^{low}] \leq Y_i \leq E_i[\theta_i^{up}]$, for $i = j, \ldots, J$. □

B.3 Proof of Proposition 4.6

By the definition of θ^{up}, the tower property of the conditional expectation, Jensen's inequality (applied to the convex functions max and F_j), and the comparison $Y_{j+1}^{up} \geq Y_{j+1}^{low}$ (in view of Proposition B.3), we obtain

$$
\begin{aligned}
Y_j^{up} =\ & E_j\left[\left(\left(\rho_j^{(1)}\right)^\top \beta_{j+1}\right)_+ \theta_{j+1}^{up}-\left(\left(\rho_j^{(1)}\right)^\top \beta_{j+1}\right)_- \theta_{j+1}^{low}-\left(\rho_j^{(1)}\right)^\top \Delta M_{j+1}\right. \\
& \left.+\rho_j^{(0)} \max_{\iota \in\{up, low\}} F_j\left(\beta_{j+1} \theta_{j+1}^{\iota}-\Delta M_{j+1}\right)-G_j^{\#}\left(\rho_j^{(1)}, \rho_j^{(0)}\right)\right] \\
\geq\ & E_j\left[\left(\left(\rho_j^{(1)}\right)^\top \beta_{j+1}\right)_+ Y_{j+1}^{up}\right]-E_j\left[\left(\left(\rho_j^{(1)}\right)^\top \beta_{j+1}\right)_- Y_{j+1}^{low}\right] \\
& +\rho_j^{(0)} \max_{\iota \in\{up, low\}} F_j\left(E_j\left[\beta_{j+1} Y_{j+1}^{\iota}\right]\right)-G_j^{\#}\left(\rho_j^{(1)}, \rho_j^{(0)}\right) \\
\geq\ & \left(\rho_j^{(1)}\right)^\top E_j\left[\beta_{j+1} Y_{j+1}^{up}\right]+\rho_j^{(0)} \max_{\iota \in\{up, low\}} F_j\left(E_j\left[\beta_{j+1} Y_{j+1}^{\iota}\right]\right)-G_j^{\#}\left(\rho_j^{(1)}, \rho_j^{(0)}\right) \\
\geq\ & \left(\rho_j^{(1)}\right)^\top E_j\left[\beta_{j+1} Y_{j+1}^{up}\right]+\rho_j^{(0)} F_j\left(E_j\left[\beta_{j+1} Y_{j+1}^{up}\right]\right)-G_j^{\#}\left(\rho_j^{(1)}, \rho_j^{(0)}\right) \\
\geq\ & G_j\left(E_j\left[\beta_{j+1} Y_{j+1}^{up}\right], F_j\left(E_j\left[\beta_{j+1} Y_{j+1}^{up}\right]\right)\right),
\end{aligned}
$$

making, again, use of the nonnegativity of $\rho_j^{(0)}$ and (13). As in the previous proofs, the argument for the lower bound Y^{low} is essentially the same, and we skip the details.

B.4 Proof of Theorem 4.7

Let $j \in \{0, ..., J-1\}$ be fixed. Due to Proposition B.2 and the comparison principle (Comp), it only remains to show that the pair of processes Y^{up} and Y^{low} given by $Y_i^{up} = E_i[\Theta_i^{up}]$ and $Y_i^{low} = E_i[\Theta_i^{low}]$, $i = j, ..., J$, defines a super- and a subsolution to (12). The proof of this claim is similar to the proof of Proposition 4.6 but simpler.

C Proofs on Stochastic Games

C.1 Alternative Representation

We first derive an alternative representation for $\Theta_0^{low}(r,M)$ as a pathwise minimization problem. To this end, define, for some fixed control $r \in \mathcal{A}_0^F$ and martingale $M \in \mathcal{M}_D$,

$$\begin{aligned}\tilde{\Theta}_i^{low} &= \inf_{(v_j^{(1)},v_j^{(0)})\in\mathbb{R}^{D+1},\ j=i,\dots,J-1} \Bigg(w_{i,J}(v^{(1)},v^{(0)},r)\xi \\ &\quad - \sum_{j=i}^{J-1} w_{i,j}(v^{(1)},v^{(0)},r)\left(v_j^{(0)}F_j^{\#}(r_j) + (v_j^{(1)}+v_j^{(0)}r_j)^\top \Delta M_{j+1} + G_j^{\#}(v_j^{(1)},v_j^{(0)})\right)\Bigg),\end{aligned}$$

where $w_{i,j}(v^{(1)},v^{(0)},u) = \prod_{k=i}^{j-1}(v_k^{(1)}+v_k^{(0)}u_k)^\top \beta_{k+1}$. Then, $\tilde{\Theta}_J^{low} = \xi$ and, for $i = 0,\dots,J-1$,

$$\begin{aligned}\tilde{\Theta}_i^{low} &= \inf_{(v_i^{(1)},v_i^{(0)})\in\mathbb{R}^{D+1}} \ \inf_{(v_j^{(1)},v_j^{(0)})\in\mathbb{R}^{D+1},\ j=i+1,\dots,J-1} \Bigg((v_i^{(1)}+v_i^{(0)}r_i)^\top \beta_{i+1} \Bigg(w_{i+1,J}(v^{(1)},v^{(0)},r)\xi \\ &\quad - \sum_{j=i+1}^{J-1} w_{i+1,j}(v^{(1)},v^{(0)},r)\left(v_j^{(0)}F_j^{\#}(r_j) + (v_j^{(1)}+v_j^{(0)}r_j)^\top \Delta M_{j+1} + G_j^{\#}(v_j^{(1)},v_j^{(0)})\right)\Bigg) \\ &\quad - \left(v_i^{(0)}F_i^{\#}(r_i) + (v_i^{(1)}+v_i^{(0)}r_i)^\top \Delta M_{i+1} + G_i^{\#}(v_i^{(1)},v_i^{(0)})\right)\Bigg).\end{aligned}$$

The outer infimum can be taken restricted to such $(v_i^{(1)},v_i^{(0)}) \in \mathbb{R}^{D+1}$ which belong to $D_{G_i^{\#}(\omega,\cdot)}$, because the expression which is to be minimized is $+\infty$ otherwise. Then, (17) implies that the inner infimum can be interchanged with the nonnegative factor $(v_i^{(1)}+v_i^{(0)}r_i)^\top\beta_{i+1}$, which yields

$$\begin{aligned}\tilde{\Theta}_i^{low} &= \inf_{(v_i^{(1)},v_i^{(0)})\in D_{G_i^{\#}(\omega,\cdot)}} \Bigg((v_i^{(1)}+v_i^{(0)}r_i)^\top \beta_{i+1}\tilde{\Theta}_{i+1}^{low} \\ &\quad - \left(v_i^{(0)}F_i^{\#}(r_i) + (v_i^{(1)}+v_i^{(0)}r_i)^\top \Delta M_{i+1} + G_j^{\#}(v_i^{(1)},v_i^{(0)})\right)\Bigg), \qquad (40)\end{aligned}$$

and the infimum can, by the same argument as above, again be replaced by that over the whole $\mathbb{R}^{D+1}$. Further rewriting (40) using (13) shows that the recursion for $\tilde{\Theta}_i^{low}$ coincides with the one for Θ_i^{low}. Therefore,

$$\begin{aligned}\Theta_0^{low}(r,M) &= \tilde{\Theta}_0^{low} = \inf_{(v_j^{(1)},v_j^{(0)})\in\mathbb{R}^{D+1},\ j=0,\dots,J-1} \Bigg(w_J(v^{(1)},v^{(0)},r)\xi \\ &\quad - \sum_{j=0}^{J-1} w_j(v^{(1)},v^{(0)},r)\left(v_j^{(0)}F_j^{\#}(r_j) + (v_j^{(1)}+v_j^{(0)}r_j)^\top \Delta M_{j+1} + G_j^{\#}(v_j^{(1)},v_j^{(0)})\right)\Bigg).\end{aligned}$$

The alternative expression for $\Theta_0^{up}(\rho^{(1)},\rho^{(0)},M)$ in (28) can be shown in the same way.

C.2 Equilibrium Value

It remains to show that Y_0 is the equilibrium value of the two-player zero-sum game. We apply the alternative representations for Θ^{up} and Θ^{low} as pathwise optimization problems from Section C.1 as well as Theorem 4.7 (twice) in order to conclude

$$\begin{aligned}
Y_0 &= \sup_{r\in\mathcal{A}_0^F,\, M\in\mathcal{M}_D} E[\Theta_0^{low}(r,M)] \\
&= \sup_{r\in\mathcal{A}_0^F,\, M\in\mathcal{M}_D} E\Bigg[\inf_{(v_j^{(1)},v_j^{(0)})\in\mathbb{R}^{D+1},\, j=0,\dots,J-1}\Big(w_J(v^{(1)},v^{(0)},r)\xi \\
&\quad -\sum_{j=0}^{J-1} w_j(v^{(1)},v^{(0)},r)\left(v_j^{(0)}F_j^{\#}(r_j)+(v_j^{(1)}+v_j^{(0)}r_j)^\top\Delta M_{j+1}+G_j^{\#}(v_j^{(1)},v_j^{(0)})\right)\Big)\Bigg] \\
&\le \sup_{r\in\mathcal{A}_0^F,\, M\in\mathcal{M}_D}\ \inf_{(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_0^G} E\Bigg[w_J(\rho^{(1)},\rho^{(0)},r)\xi \\
&\quad -\sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},r)\Big(\rho_j^{(0)}F_j^{\#}(r_j)+(\rho_j^{(1)}+\rho_j^{(0)}r_j)^\top\Delta M_{j+1}+G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)})\Big)\Bigg] \\
&= \sup_{r\in\mathcal{A}_0^F}\ \inf_{(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_0^G} E\left[w_J(\rho^{(1)},\rho^{(0)},r)\xi-\sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},r)\Big(\rho_j^{(0)}F_j^{\#}(r_j)+G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)})\Big)\right] \\
&\le \inf_{(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_0^G}\ \sup_{r\in\mathcal{A}_0^F} E\left[w_J(\rho^{(1)},\rho^{(0)},r)\xi-\sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},r)\Big(\rho_j^{(0)}F_j^{\#}(r_j)+G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)})\Big)\right] \\
&\le \inf_{(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_0^G,\, M\in\mathcal{M}_D} E\Bigg[\sup_{(u_j)\in\mathbb{R}^D,\, j=0,\dots,J-1}\Big(w_J(\rho^{(1)},\rho^{(0)},u)\xi \\
&\quad -\sum_{j=0}^{J-1} w_j(\rho^{(1)},\rho^{(0)},u)\left(\rho_j^{(0)}F_j^{\#}(u_j)+(\rho_j^{(1)}+\rho_j^{(0)}u_j)^\top\Delta M_{j+1}+G_j^{\#}(\rho_j^{(1)},\rho_j^{(0)})\right)\Big)\Bigg] \\
&= \inf_{(\rho^{(1)},\rho^{(0)})\in\mathcal{A}_0^G,\, M\in\mathcal{M}_D} E[\Theta_0^{up}(\rho^{(1)},\rho^{(0)},M)] = Y_0.
\end{aligned}$$

Here we also applied the zero-expectation property of martingale penalties at adapted controls. Consequently, all inequalities turn into equalities, which completes the proof.

References

S. Alanko and M. Avellaneda. Reducing variance in the numerical solution of BSDEs. *Comptes Rendus Mathematique*, 351(3):135–138, 2013.

L. Andersen and M. Broadie. Primal-dual simulation algorithm for pricing multidimensional American options. *Management Science*, 50(9):1222–1234, 2004.

M. Avellaneda, A. Levy, and A. Parás. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2(2):73–88, 1995.

C. Bender and J. Steiner. Least-squares Monte Carlo for Backward SDEs. In R. Carmona, P. Del Moral, P. Hu, and N. Oudjane, editors, *Numerical Methods in Finance*, pages 257–289. Springer, 2012.

C. Bender, N. Schweizer, and J. Zhuo. A primal-dual algorithm for BSDEs. *Mathematical Finance*, Early View, DOI: 10.1111/mafi.12100, 2015.

C. Bender, C. Gärtner, and N. Schweizer. Iterative improvement of lower and upper bounds for backward SDEs. *arXiv preprint 1605.07500*, 2016.

D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume I. Athena Scientific, 3rd edition, 2005.

C. Beveridge and M. Joshi. Monte Carlo bounds for game options including convertible bonds. *Management Science*, 57(5):960–974, 2011.

D. Brigo and F. Mercurio. *Interest Rate Models – Theory and Practice: With Smile, Inflation and Credit*. Springer, 2nd edition, 2006.

D. Brigo and A. Pallavicini. Counterparty risk pricing under correlation between default and interest rates. In J. H. Miller, D. C. Edelman, and J. A. D. Appleby, editors, *Numerical Methods for Finance*, pages 63–81. Chapman and Hall/CRC, 2007.

D. Brigo, M. Morini, and A. Pallavicini. *Counterparty Credit Risk, Collateral and Funding: With Pricing Cases for All Asset Classes*. John Wiley & Sons, 2013.

D. B. Brown and M. B. Haugh. Information relaxation bounds for infinite horizon Markov decision processes. *Preprint*, 2016.

D. B. Brown and J. E. Smith. Dynamic portfolio optimization with transaction costs: Heuristics and dual bounds. *Management Science*, 57(10):1752–1770, 2011.

D. B. Brown, J. E. Smith, and P. Sun. Information relaxations and duality in stochastic dynamic programs. *Operations Research*, 58(4):785–801, 2010.

S. S. Chandramouli and M. B. Haugh. A unified approach to multiple stopping and duality. *Operations Research Letters*, 40(4):258–264, 2012.

P. Cheridito, M. Kupper, and N. Vogelpoth. Conditional analysis on $\mathbb{R}^d$. In A. H. Hamel, F. Heyde, A. Löhne, B. Rudloff, and C. Schrage, editors, *Set Optimization and Applications in Finance - The State of the Art*, pages 179–211. Springer, 2015.

S. Crépey. *Financial Modeling.* Springer, 2013.

S. Crépey, R. Gerboud, Z. Grbac, and N. Ngor. Counterparty risk and funding: The four wings of the TVA. *International Journal of Theoretical and Applied Finance*, 16(2):1350006, 2013.

S. Crépey, T. R. Bielecki, and D. Brigo. *Counterparty Risk and Funding: A Tale of Two Puzzles.* Chapman and Hall/CRC, 2014.

D. Duffie, M. Schroder, and C. Skiadas. Recursive valuation of defaultable securities and the timing of resolution of uncertainty. *Annals of Applied Probability*, 6(4):1075–1090, 1996.

N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997.

A. Fahim, N. Touzi, and X. Warin. A probabilistic numerical method for fully nonlinear parabolic PDEs. *Annals of Applied Probability*, 21(4):1322–1364, 2011.

E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, 3(4):391–412, 1999.

A. Friedman. *Partial differential equations of parabolic type.* Prentice-Hall, 1964.

P. Glasserman and B. Yu. Simulation for American options: Regression now or regression later? In H. Niederreiter, editor, *Monte Carlo and Quasi-Monte Carlo Methods 2002*, pages 213–226. Springer, 2004.

J. Guyon and P. Henry-Labordère. The uncertain volatility model: A Monte Carlo approach. *Journal of Computational Finance*, 14(3):37–71, 2011.

J. Guyon and P. Henry-Labordère. *Nonlinear Option Pricing.* Chapman and Hall/CRC, 2013.

M. Haugh and A. E. B. Lim. Linear–quadratic control and information relaxations. *Operations Research Letters*, 40(6):521–528, 2012.

M. B. Haugh and L. Kogan. Pricing American options: A duality approach. *Operations Research*, 52(2):258–270, 2004.

M. B. Haugh and C. Wang. Information relaxations and dynamic zero-sum games. *arXiv preprint 1405.4347*, 2015.

I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, 2nd edition, 1991.

I. Kharroubi, N. Langrené, and H. Pham. A numerical algorithm for fully nonlinear HJB equations: an approach by control randomization. *Monte Carlo Methods and Applications*, 20(2): 145–165, 2014.

H. Kraft and F. T. Seifried. Stochastic differential utility as the continuous-time limit of recursive utility. *Journal of Economic Theory*, 151:528–550, 2014.

C. Kühn, A. E. Kyprianou, and K. van Schaik. Pricing Israeli options: A pathwise approach. *Stochastics*, 79(1-2):117–137, 2007.

R. Lord, R. Koekkoek, and D. van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194, 2010.

T. J. Lyons. Uncertain volatility and the risk-free synthesis of derivatives. *Applied Mathematical Finance*, 2(2):117–133, 1995.

T. Nie and M. Rutkowski. A BSDE approach to fair bilateral pricing under endogenous collateralization. *Finance and Stochastics*, 20(4):855–900, 2016.

É. Pardoux. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. In *Stochastic Analysis and Related Topics VI (Geilo 1996)*, pages 79–127. Birkhäuser, 1998.

H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications.* Springer, 2009.

W. B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality.* John Wiley & Sons, 2nd edition, 2011.

N. S. Rasmussen. Control variates for Monte Carlo valuation of American options. *Journal of Computational Finance*, 9(1):84–102, 2005.

R. T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

L. C. G. Rogers. Monte Carlo valuation of American options. *Mathematical Finance*, 12(3): 271–286, 2002.

L. C. G. Rogers. Pathwise stochastic optimal control. *SIAM Journal on Control and Optimization*, 46(3):1116–1132, 2007.

J. Schoenmakers. A pure martingale dual for multiple stopping. *Finance and Stochastics*, 16(2): 319–334, 2012.

L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953.

J. M. Vanden. Exact superreplication strategies for a class of derivative assets. *Applied Mathematical Finance*, 13(1):61–87, 2006.

J. Zhang. A numerical scheme for BSDEs. *Annals of Applied Probability*, 14(1):459–488, 2004.