

ESAIM: PROCEEDINGS AND SURVEYS, February 2019, Vol. 65, p. 1-26

B. Bouchard, J.-F. Chassagneux, F. Delarue, E. Gobet and J. Lelong, Editors

NUMERICAL APPROXIMATIONS OF MCKEAN ANTICIPATIVE BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS ARISING IN INITIAL MARGIN REQUIREMENTS *

A. AGARWAL¹, S. DE MARCO², E. GOBET³, J. G. LÓPEZ-SALAS⁴, F. NOUBIAGAIN⁵
AND A. ZHOU⁶

Abstract. We introduce a new class of anticipative backward stochastic differential equations with a dependence of McKean type on the law of the solution, that we name MKABSDE. We provide existence and uniqueness results in a general framework with relatively general regularity assumptions on the coefficients. We show how such stochastic equations arise within the modern paradigm of derivative pricing where a central counterparty (CCP) requires the members to deposit variation and initial margins to cover their exposure. In the case when the initial margin is proportional to the Conditional Value-at-Risk (CVaR) of the contract price, we apply our general result to define the price as a solution of a MKABSDE. We provide several linear and non-linear simpler approximations, which we solve using different numerical (deterministic and Monte-Carlo) methods.

Keywords: non-linear pricing, CVaR initial margins, anticipative BSDE, weak non-linearity.

MSC2000: 60H30, 65C05, 65C30

* This work has been carried out during the 6-weeks summerschool CEMRACS, July-August 2017, with the support of the *Chaire Risques Financiers* (Ecole Nationale des Ponts et Chaussées, Ecole Polytechnique, Société Générale, Sorbonne Université - with the partnership of the *Fondation du Risque*).

¹ Adam Smith Business School, University of Glasgow, University Avenue, G12 8QQ Glasgow, United Kingdom. Email: ankush.agarwal@glasgow.ac.uk. This author research was conducted while at CMAP, Ecole Polytechnique, and is part of the *Chaire Risque Financiers*.

² Centre de Mathématiques Appliquées (CMAP), Ecole Polytechnique, Route de Saclay, 91128 Palaiseau Cedex, France. Email: stefano.de-marco@polytechnique.edu

³ CMAP, Ecole Polytechnique. Email: emmanuel.gobet@polytechnique.edu. This author research is part of the *Chaire Risque Financiers* of the *Fondation du Risque* and of the *Finance for Energy Market Research Centre*.

⁴ Department of Mathematics, Faculty of Informatics, Universidade da Coruña, Campus de Elviña s/n, 15071 - A Coruña, Spain. Email: jose.lsalas@udc.es. This author research was conducted while at CMAP, Ecole Polytechnique, and is part of the *Chaire Risque Financiers*.

⁵ This author research was conducted while at Laboratoire Manceau de Mathématiques, Le Mans Université, France. Email: larissafanny@yahoo.fr

⁶ This author research was conducted while at Université Paris-Est, CERMICS (ENPC), 77455, Marne-la-Vallée, France. Email: alexandre.zhou@enpc.fr

Résumé. Nous introduisons une nouvelle famille d'équations différentielles stochastiques rétrogrades anticipatives ayant une dépendance par rapport à la loi de la solution, que nous appelons MKABSDE. Ces équations apparaissent dans le contexte moderne de la valorisation de dérivés en présence d'appels de marge de la part d'une chambre de compensation. Nous démontrons un résultat d'existence et unicité sous des hypothèses relativement faibles sur les coefficients de l'équation. Dans le cas où les appels de marge sont proportionnels à la VaR conditionnelle (CVaR) du prix du contrat, notre résultat général entraîne l'existence et unicité pour le prix en tant que solution d'une MKABSDE. Nous considérons plusieurs approximations linéaires et non-linéaires de cette équation, que nous abordons avec différentes méthodes numériques.

1. INITIAL MARGIN AND MCKEAN ANTICIPATIVE BSDE (MKABSDE)

1.1. Financial context and motivation

The paradigm of linear risk-neutral pricing of financial contracts has changed in the last few years, influenced by the regulators. Nowadays, for several type of trades banking institutions have to post collateral to a central counterparty (CCP, also called clearing house) in order to secure their positions. On a daily basis, the CCP imposes to each member to post a certain amount computed according to the estimated exposure of their contracts. The variation margin deposit corresponds to a collateral that compensates the daily fluctuations of the market value of a contract, while the initial margin deposit is intended to reduce the gap risk, which is the possible mark-to-market loss encountered during the liquidation period upon default of the contract's counterpart (see [CBB14, Chapter 3], or the Basel Committee on Banking Supervision document [Bas15], for more details). In this work we focus only on the initial margin requirement (IM for short), and we investigate how it affects the valuation and hedging of the contract. As stated in [Bas15, p.11 3(d)], “*IM protects the transacting parties from the potential future exposure that could arise from future changes in the mark-to-market value of the contract during the time it takes to close out and replace the position in the event that one or more counterparties default. The amount of initial margin reflects the size of the potential future exposure. It depends on a variety of factors, [...] the expected duration of the contract closeout and replacement period, and can change over time.*” In this work, we will consider IM deposits that are proportional to the Conditional Value-at-Risk (CVaR) of the contract value over a future period of length Δ (typically $\Delta = 1$ week or 10 days, standing for the closeout and replacement period). We focus on CVaR rather than Value-at-Risk (VaR) due to its pertinent mathematical properties; it is indeed well established that CVaR is a coherent risk measure whereas VaR is not [ADEH99].

We make some distinctions in our analysis according to the way the *contract value* is computed in the presence of IM. While [Bas15] refers to a mark-to-market value of the contract that can be seen as an *exogenous* value, we investigate the case where this value is *endogenous* and is given by the value of the hedging portfolio including the additional IM costs. A similar setting is considered in Nie and Rutkowski [NR16]. By doing so, we introduce a new non-linear pricing rule, that is: the value of the hedging portfolio V_t together with its hedging component π_t solve a stochastic equation including a term depending on the law of the solution (due to the CVaR). We justify that this problem can be seen as a new type of anticipative Backward Stochastic Differential Equation (BSDE) with McKean interaction [McK66]. From now on, we refer to this kind of equation as MKABSDE, standing for McKean Anticipative BSDE; Section 1.2 below gives a toy example of such a model. We derive stability estimates for these MKABSDEs, under general Lipschitz conditions, and prove existence and uniqueness results. In Section 3, we verify that these results can be applied to a general complete Itô market [KS98], when accounting for IM requirements. Then, we derive some approximations based on classical non-linear BSDEs whose purpose is to quantify the impact of choosing the reference price for the IM as exogenous or endogenous, and to compare with the case without IM. Essentially, in Theorem 3.1 we prove that the hedging portfolio with exogenous or endogenous reference price for the IM coincide up to order 1 in Δ when Δ is small

(which is compatible with Δ equal to a few days), while the difference between valuation *with* or *without* IM correction has a size of order $\sqrt{\Delta}$. Finally, Section 4 is devoted to numerical experiments: we solve the different approximating BSDEs using finite difference methods in dimension 1, and nested Monte Carlo and regression Monte Carlo methods in higher dimensions.

1.2. An example of anticipative BSDE with dependence in law

We start with a simple financial example with IM requirements, in the case of a single tradable asset. A more general version with a multidimensional Itô market will be studied in Section 3. Let us assume that the price of a tradable asset, denoted S , evolves accordingly to a geometric Brownian motion

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (1.1)$$

where $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}^+$ and W is an one-dimensional Brownian motion.

In the classical financial setting (see, for example, [MR05]), consider the situation where a trader wants to sell a European option with maturity $T > 0$ and payoff $\Phi(S_T)$, and to hedge it dynamically with risky and riskless assets S and S^0 , where $S_t^0 = e^{rt}$ for $t \in [0, T]$ and r is a risk-free interest rate. We denote by (V, π) , the value of the self-financing portfolio and π the amount of money invested in the risky asset, respectively. In order to ensure the replication of the payoff at maturity, the couple (V, π) should solve the following stochastic equation

$$\begin{cases} dV_t &= r(V_t - \pi_t) dt + \pi_t \frac{dS_t}{S_t}, \quad t \in [0, T], \\ V_T &= \Phi(S_T). \end{cases} \quad (1.2)$$

Eq (1.2) is a BSDE since the terminal condition of V is imposed. Because all the coefficients are linear in V and π , (1.2) is a linear BSDE (see [EPQ97] for a broad overview on BSDEs and their applications in finance).

Accounting for IM requirement will introduce an additional cost in the above self-financing dynamics. We assume that the required deposit is proportional to the CVaR of the portfolio over Δ days (typically $\Delta = 10$ days) at the risk-level α (typically $\alpha = 99\%$). The funding cost for this deposit is determined by an interest rate R .¹ Therefore, the IM cost can be modelled as an additional term in the dynamics of the self-financing portfolio as

$$dV_t = (r(V_t - \pi_t) - R \mathbf{CVaR}_{\mathcal{F}_t}^\alpha(V_t - V_{t+\Delta})) dt + \pi_t \frac{dS_t}{S_t}, \quad (1.3)$$

where the CVaR of a random variable L , conditional on the underlying sigma-field \mathcal{F}_t at time t , is defined by (see [RU00])

$$\mathbf{CVaR}_{\mathcal{F}_t}^\alpha(L) = \inf_{x \in \mathbb{R}} \mathbb{E} \left[\frac{(L - x)^+}{1 - \alpha} + x \middle| \mathcal{F}_t \right]. \quad (1.4)$$

Since $V_{t+\Delta}$ may be meaningless as t gets close to T , in (1.3) one should consider $V_{(t+\Delta) \wedge T}$ instead. Rewriting (1.3) in integral form together with the replication constraint, we obtain a BSDE

$$V_t = \Phi(S_T) + \int_t^T (-r(V_s - \pi_s) - \mu\pi_s + R \mathbf{CVaR}_{\mathcal{F}_s}^\alpha(V_s - V_{(s+\Delta) \wedge T})) ds - \int_t^T \pi_s \sigma dW_s, \quad t \in [0, T]. \quad (1.5)$$

The conditional CVaR term is anticipative and non-linear in the sense of McKean [McK66], for it involves the law of future variations of the portfolio conditional to the knowledge of the past. This is an example of McKean Anticipative BSDE, which we study in broader generality in Section 2.

Coming back to the financial setting, (V, π) stands for a valuation rule which treats the IM adjustment as endogenous (in the sense that CVaR is computed on V itself). One could alternatively consider that CVaR is

¹This interest rate corresponds to the difference of a funding rate minus the interest rate paid by the CCP for the deposit, typically $R \approx 3\%$

related to an exogenous valuation (the so-called *mark-to-market*), for instance the one due to (1.2) (assuming that (1.2) models the market evolution of the option price). Later in Section 3, we give quantitative error bounds between these different valuation rules. Without advocating one with respect to the other, we rather compare their values and estimate (theoretically and numerically) how well one of their output prices approximates the others. As a consequence, these results may serve as a support for banks and regulators for improving risk management and margin requirement rules.

1.3. Literature review on anticipative BSDEs and comparison with our contribution

BSDEs were introduced by Pardoux and Peng [PP90]. Since then, the theoretical properties of BSDEs with different generators and terminal conditions have been extensively studied. The link between Markovian BSDEs and partial differential equations (PDEs) was studied in [PP92]. Under some smoothness assumptions, [PP92] established that the solution of the Markovian BSDE corresponds to the solution of a semi-linear parabolic PDE. In addition, several applications in finance have been proposed, in particular by El Karoui and co-authors [EPQ97] who considered the application to European option pricing in the constrained case. In fact, [EPQ97] showed that, under some constraints on the hedging strategy, the price of a contingent claim is given by the solution of a non-linear convex BSDE.

Recently, a new class of BSDEs called anticipated² BSDEs (ABSDEs for short) was introduced by Peng and Yang [PY09]. The main feature of this class is that the generator includes not only the value of the solution at the present, but also at a future date. In [PY09] the existence, uniqueness and a comparison theorem for the solution is provided under a kind of Lipschitz condition which depends on the conditional expectation. One can also find more general formulations of ABSDE in Cheridito and Nam [CN17]. As in the case of classical BSDEs, the question of weakening the Lipschitz condition considered in [PY09] has been tackled by Yang and Elliott [YE13], who extended the existence theorem for ABSDEs from Lipschitz to continuous coefficients, and proved that the comparison theorem for anticipated BSDEs still holds. They also established a minimal solution.

At the same time, Buckdahn and Imkeller [BI09] introduced the so-called time-delayed BSDEs (see also Delong and Imkeller [DI09, DI10]). As opposed to the ABSDEs of [PY09], in this case the generator depends on the values of the solution at the present and at past dates, weighted with a time delay function. Assuming that the generator satisfies a certain kind of Lipschitz assumption depending on a probability measure, Delong and Imkeller [DI10] proved the existence and uniqueness of a solution for a sufficiently small time horizon or for a sufficiently small Lipschitz constant of the generator. These authors also showed that, when the generator is independent of y and for a small delay, existence and uniqueness hold for an arbitrary Lipschitz constant. Later, Delong and Imkeller [DI12] provided an application of time-delayed BSDEs to problems of pricing and hedging, and portfolio management. This work focuses on participating contracts and variable annuities, which are worldwide life insurance products with capital protections, and on claims based on the performance of an underlying investment portfolio.

More recently, Crépey *et al.* [CESS17] have worked in a setting which is close to the problem we tackle here, introducing an application of ABSDEs to the problem of computing different types of valuation adjustments (XVAs) for derivative prices, related to funding ($X=F$), capital ($X=K$) and credit risk ($X=C$). In particular, they focus on the case where the variation margins of an OTC contract can be funded directly with the capital of the bank involved in the trade, giving rise to different terms in the portfolio evolution equation. The connection of economic capital and funding valuation adjustment leads to an ABSDE, whose anticipated part consists of a conditional risk measure over the martingale part of the portfolio on a future time period. These authors have showed that the system of ABSDEs for the FVA and the KVA processes is well-posed. Mathematically, the existence and uniqueness of the solution to the system is established together with the convergence of Picard iterations.

Motivated by the dynamics of the self-financing portfolio (1.5), we consider a type of ABSDEs (McKean ABSDEs) where the generator depends on the value of the solution, but also on the law of the future trajectory, possibly up to maturity, in analogy with [CESS17]. We state a priori estimates on the differences between the

²We equivalently use the word anticipated or anticipative in this work.

solutions of two such MKABSDEs. Based on these estimates, we derive existence and uniqueness results via a fixed-point theorem. Comparing again with the close work [CESS17], one main difference is that the time horizon Δ in the IM setting is around one week, as apposed to one year for economic capital in [CESS17]. Consequently, the IM framework naturally lends itself to an asymptotic analysis as Δ goes to zero, giving rise to the approximating non linear or linear (non anticipative) BSDEs for which we provide error estimates and explore different numerical methods.

2. A GENERAL MCKEAN ANTICIPATIVE BSDE

In order to give meaning to (1.5) and to more general (multidimensional) cases such as (3.2) below, we now introduce a general mathematical setup for studying existence and uniqueness of solutions.

2.1. Notation

Let $T > 0$ be the finite time horizon and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a d -dimensional Brownian motion, where $d \geq 1$. We denote $(\mathcal{F}_t)_{t \in [0, T]}$ the filtration generated by W , completed with the \mathbb{P} -null sets of \mathcal{F} . Let $t \in [0, T]$, $\beta \geq 0$ and $m \in \mathbb{N}^*$. We will make use of the following notations:

- For any $a = (a_1, \dots, a_m) \in \mathbb{R}^m$, $|a| = \sqrt{\sum_{i=1}^m a_i^2}$.
- Given a process $(x_s)_{s \in [0, T]}$, we set $x_{t:T} := (x_s)_{s \in [t, T]}$.
- $L_T^2(\mathbb{R}^m) = \{\mathbb{R}^m\text{-valued } \mathcal{F}_T\text{-measurable random variables } \xi \text{ such that } \mathbb{E}[|\xi|^2] < \infty\}$.
- $\mathbb{H}_{\beta, T}^2(\mathbb{R}^m) = \{\mathbb{R}^m\text{-valued and } \mathcal{F}\text{-adapted stochastic processes } \varphi \text{ such that } \mathbb{E}\left[\int_0^T e^{\beta t} |\varphi_t|^2 dt\right] < \infty\}$. For $\varphi \in \mathbb{H}_{\beta, T}^2(\mathbb{R}^m)$, we define $\|\varphi\|_{\mathbb{H}_{\beta, T}^2} = \sqrt{\mathbb{E}\left[\int_0^T e^{\beta t} |\varphi_t|^2 dt\right]}$.
- $\mathbb{S}_{\beta, T}^2(\mathbb{R}^m) = \{\text{Continuous processes } \varphi \in \mathbb{H}_{\beta, T}^2(\mathbb{R}^m) \text{ such that } \mathbb{E}\left[\sup_{t \in [0, T]} e^{\beta t} |\varphi_t|^2\right] < \infty\}$. For $\varphi \in \mathbb{S}_{\beta, T}^2(\mathbb{R}^m)$, we define $\|\varphi\|_{\mathbb{S}_{\beta, T}^2} = \sqrt{\mathbb{E}\left[\sup_{t \in [0, T]} e^{\beta t} |\varphi_t|^2\right]}$.

Note that $\mathbb{H}_{\beta, T}^2(\mathbb{R}^m) = \mathbb{H}_{0, T}^2(\mathbb{R}^m)$ and $\mathbb{S}_{\beta, T}^2(\mathbb{R}^m) = \mathbb{S}_{0, T}^2(\mathbb{R}^m)$, for any $\beta \geq 0$. The additional degree of freedom given by the parameter β in the definition of the space norm will be useful when deriving a priori estimates (see Lemma 2.2).

2.2. Main result

Our aim is to find a pair of processes $(Y, Z) \in \mathbb{S}_{0, T}^2(\mathbb{R}) \times \mathbb{H}_{0, T}^2(\mathbb{R}^d)$ satisfying

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, \Lambda_s(Y_{s:T})) ds - \int_t^T Z_s dW_s, \quad t \in [0, T], \quad (2.1)$$

for a certain mapping $\Lambda_t(\cdot)$ to be defined below. We call Equation (2.1) McKean Anticipative BSDE (MKABSDE) with parameters (f, Λ, ξ) . In order to obtain existence and uniqueness of solutions, we require that the mappings f and Λ satisfy some suitable Lipschitz properties (specified below), and that the terminal condition ξ be square integrable.

Assumption (S). For any $y, z, \lambda \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$, $f(\cdot, y, z, \lambda)$ is a \mathcal{F} -adapted stochastic process with values in \mathbb{R} and there exists a constant $C_f > 0$ such that almost surely, for all $(s, y_1, z_1, \lambda_1), (s, y_2, z_2, \lambda_2) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$,

$$|f(s, y_1, z_1, \lambda_1) - f(s, y_2, z_2, \lambda_2)| \leq C_f (|y_1 - y_2| + |z_1 - z_2| + |\lambda_1 - \lambda_2|).$$

Moreover, $\mathbb{E}\left[\int_0^T |f(s, 0, 0, 0)|^2 ds\right] < \infty$.

Assumption (A). For any $X \in \mathbb{S}_{0,T}^2(\mathbb{R})$, $(\Lambda_t(X_{t:T}))_{t \in [0,T]}$ defines a stochastic process that belongs to $\mathbb{H}_{0,T}^2(\mathbb{R})$. There exist a constant $C_\Lambda > 0$ and a family of positive measures $(\nu_t)_{t \in [0,T]}$ on \mathbb{R} such that for every $t \in [0, T]$, ν_t has support included in $[t, T]$, $\nu_t([t, T]) = 1$, and for any $y^1, y^2 \in \mathbb{S}_{0,T}^2(\mathbb{R})$, we have

$$|\Lambda_t(y_{t:T}^1) - \Lambda_t(y_{t:T}^2)| \leq C_\Lambda \mathbb{E} \left[\int_t^T |y_s^1 - y_s^2| \nu_t(ds) \middle| \mathcal{F}_t \right], dt \otimes d\mathbb{P} \text{ a.e. .}$$

Note that, under Assumption (A), for every $\beta \geq 0$ and every continuous path $x : [0, T] \rightarrow \mathbb{R}$ we have

$$\int_0^T e^{\beta s} \int_s^T |x_u| \nu_s(du) ds \leq T \sup_{t \in [0, T]} e^{\beta t} |x_t|.$$

We will say that a function \tilde{f} (resp. a mapping $\tilde{\Lambda}$) satisfies Assumption (S) (resp. (A)) if the assumption holds for the choice $f = \tilde{f}$ (resp. $\Lambda = \tilde{\Lambda}$). We can now give the main result of this section.

Theorem 2.1. Under Assumptions (S) and (A), for any terminal condition $\xi \in L_T^2(\mathbb{R})$ the BSDE (2.1) has a unique solution $(Y, Z) \in \mathbb{S}_{0,T}^2(\mathbb{R}) \times \mathbb{H}_{0,T}^2(\mathbb{R}^d)$.

2.3. Proof of Theorem 2.1

The proof uses classical arguments. We first establish apriori estimates in the same spirit as in [EPQ97] on the solutions to the BSDE. Then for a suitable constant $\beta \geq 0$, we use Picard's fixed point method in the space $\mathbb{S}_{\beta,T}^2(\mathbb{R}) \times \mathbb{H}_{\beta,T}^2(\mathbb{R}^d)$ to obtain existence and uniqueness of a solution to Equation (2.1).

Lemma 2.2. Let $(Y^1, Z^1), (Y^2, Z^2) \in \mathbb{S}_{0,T}^2(\mathbb{R}) \times \mathbb{H}_{0,T}^2(\mathbb{R}^d)$ be solutions to MKABSDE (2.1) associated respectively to the parameters (f^1, Λ^1, ξ^1) and (f^2, Λ^2, ξ^2) . We assume that f^1 satisfies Assumption (S) and that Λ^1 satisfies Assumption (A). Let us define $\delta Y := Y^1 - Y^2$, $\delta Z := Z^1 - Z^2$, $\delta \xi := \xi^1 - \xi^2$. Finally, let us define for $s \in [0, T]$,

$$\delta_2 f_s = f^1(s, Y_s^2, Z_s^2, \Lambda^2(Y_{s:T}^2)) - f^2(s, Y_s^2, Z_s^2, \Lambda^2(Y_{s:T}^2)), \quad \text{and} \quad \delta_2 \Lambda_s = \Lambda_s^1(Y_{s:T}^2) - \Lambda_s^2(Y_{s:T}^2).$$

Then there exists a constant $C > 0$ such that for $\mu > 0$, we have for β large enough

$$\begin{aligned} \|\delta Y\|_{\mathbb{S}_{\beta,T}^2}^2 &\leq C \left(e^{\beta T} \mathbb{E} [|\delta \xi|^2] + \frac{1}{\mu^2} \left(\|\delta_2 f\|_{\mathbb{H}_{\beta,T}^2}^2 + C_{f^1} \|\delta_2 \Lambda\|_{\mathbb{H}_{\beta,T}^2}^2 \right) \right), \\ \|\delta Z\|_{\mathbb{H}_{\beta,T}^2}^2 &\leq C \left(e^{\beta T} \mathbb{E} [|\delta \xi|^2] + \frac{1}{\mu^2} \left(\|\delta_2 f\|_{\mathbb{H}_{\beta,T}^2}^2 + C_{f^1} \|\delta_2 \Lambda\|_{\mathbb{H}_{\beta,T}^2}^2 \right) \right). \end{aligned}$$

Proof. The proof is based on similar arguments used in [EPQ97]. Let us use the decomposition:

$$\begin{aligned} &|f^1(s, Y_s^1, Z_s^1, \Lambda_s^1(Y_{s:T}^1)) - f^2(s, Y_s^2, Z_s^2, \Lambda_s^2(Y_{s:T}^2))| \\ &\leq |f^1(s, Y_s^1, Z_s^1, \Lambda_s^1(Y_{s:T}^1)) - f^1(s, Y_s^2, Z_s^2, \Lambda_s^2(Y_{s:T}^2))| \\ &\quad + |f^1(s, Y_s^2, Z_s^2, \Lambda_s^2(Y_{s:T}^2)) - f^2(s, Y_s^2, Z_s^2, \Lambda_s^2(Y_{s:T}^2))| \\ &\leq C_{f^1} (|\delta Y_s| + |\delta Z_s| + |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^2(Y_{s:T}^2)|) + |\delta_2 f_s| \\ &\leq C_{f^1} (|\delta Y_s| + |\delta Z_s| + |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^1(Y_{s:T}^2)| + |\delta_2 \Lambda_s|) + |\delta_2 f_s|. \end{aligned}$$

By Itô's lemma on the process $t \rightarrow e^{\beta t} |\delta Y_t|^2$, where $\beta \geq 0$, and using the previous inequality, we have that

$$e^{\beta t} |\delta Y_t|^2 + \beta \int_t^T e^{\beta s} |\delta Y_s|^2 ds + \int_t^T e^{\beta s} |\delta Z_s|^2 ds$$

$$\begin{aligned}
&= e^{\beta T} |\delta\xi|^2 + 2 \int_t^T e^{\beta s} \delta Y_s (f^1(s, Y_s^1, Z_s^1, \Lambda_s^1(Y_{s:T}^1)) - f^2(s, Y_s^2, Z_s^2, \Lambda_s^2(Y_{s:T}^2))) ds - 2 \int_t^T e^{\beta s} \delta Y_s \delta Z_s dW_s \\
&\leq e^{\beta T} |\delta\xi|^2 + 2 \int_t^T e^{\beta s} |\delta Y_s| (C_{f^1} (|\delta Y_s| + |\delta Z_s| + |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^1(Y_{s:T}^2)| + |\delta_2 \Lambda_s|) + |\delta_2 f_s|) ds \\
&\quad - 2 \int_t^T e^{\beta s} \delta Y_s \delta Z_s dW_s. \tag{2.2}
\end{aligned}$$

Applying Young's inequality with $\lambda, \mu \neq 0$, we have

$$\begin{aligned}
&2|\delta Y_s| (C_{f^1} (|\delta Z_s| + |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^1(Y_{s:T}^2)| + |\delta_2 \Lambda_s|) + |\delta_2 f_s|) \\
&\leq \frac{C_{f^1}}{\lambda^2} |\delta Z_s|^2 + \lambda^2 C_{f^1} |\delta Y_s|^2 + \frac{C_{f^1}}{\lambda^2} |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^1(Y_{s:T}^2)|^2 + \lambda^2 C_{f^1} |\delta Y_s|^2 \\
&\quad + \frac{C_{f^1}}{\mu^2} |\delta_2 \Lambda_s|^2 + \mu^2 C_{f^1} |\delta Y_s|^2 + \frac{1}{\mu^2} |\delta_2 f_s|^2 + \mu^2 |\delta Y_s|^2 \\
&\leq (\mu^2 + C_{f^1}(\mu^2 + 2\lambda^2)) |\delta Y_s|^2 + \frac{C_{f^1}}{\lambda^2} |\delta Z_s|^2 + \frac{C_{f^1}}{\lambda^2} |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^1(Y_{s:T}^2)|^2 \\
&\quad + \frac{C_{f^1}}{\mu^2} |\delta_2 \Lambda_s|^2 + \frac{1}{\mu^2} |\delta_2 f_s|^2.
\end{aligned}$$

Then plug this bound into (2.2) to get

$$\begin{aligned}
&e^{\beta t} |\delta Y_t|^2 + \beta \int_t^T e^{\beta s} |\delta Y_s|^2 ds + \int_t^T e^{\beta s} |\delta Z_s|^2 ds \\
&\leq e^{\beta T} |\delta\xi|^2 + (\mu^2 + C_{f^1}(2 + \mu^2 + 2\lambda^2)) \int_t^T e^{\beta s} |\delta Y_s|^2 ds + \frac{C_{f^1}}{\lambda^2} \int_t^T e^{\beta s} |\delta Z_s|^2 ds \\
&\quad + \frac{C_{f^1}}{\lambda^2} \int_t^T e^{\beta s} |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^1(Y_{s:T}^2)|^2 ds + \frac{C_{f^1}}{\mu^2} \int_t^T e^{\beta s} |\delta_2 \Lambda_s|^2 ds + \frac{1}{\mu^2} \int_t^T e^{\beta s} |\delta_2 f_s|^2 ds \\
&\quad - 2 \int_t^T e^{\beta s} \delta Y_s \delta Z_s dW_s. \tag{2.3}
\end{aligned}$$

Choosing $\lambda^2 > C_{f^1}$ and

$$\beta \geq \mu^2 + C_{f^1}(2 + \mu^2 + 2\lambda^2), \tag{2.4}$$

we get from (2.3) that

$$\begin{aligned}
\mathbb{E} \left[\int_t^T e^{\beta s} |\delta Z_s|^2 ds \right] &\leq \frac{\lambda^2}{\lambda^2 - C_{f^1}} \mathbb{E} \left[e^{\beta T} |\delta\xi|^2 + \frac{C_{f^1}}{\lambda^2} \int_t^T e^{\beta s} |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^1(Y_{s:T}^2)|^2 ds \right] \\
&\quad + \frac{\lambda^2}{\lambda^2 - C_{f^1}} \mathbb{E} \left[\frac{C_{f^1}}{\mu^2} \int_t^T e^{\beta s} |\delta_2 \Lambda_s|^2 ds + \frac{1}{\mu^2} \int_t^T e^{\beta s} |\delta_2 f_s|^2 ds \right].
\end{aligned}$$

Here we have used that the stochastic integral in (2.3) is a true martingale, by invoking $\delta Y \in \mathbb{S}_{0,T}^2(\mathbb{R})$, $\delta Z \in \mathbb{H}_{0,T}^2(\mathbb{R}^d)$, computations similar to (2.7) and a localization procedure. From (2.3) we also have that

$$\mathbb{E} \left[\sup_{t \in [0, T]} e^{\beta t} |\delta Y_t|^2 + \left(1 - \frac{C_{f^1}}{\lambda^2}\right) \int_0^T e^{\beta s} |\delta Z_s|^2 ds \right]$$

$$\begin{aligned} &\leq \mathbb{E} \left[e^{\beta T} |\delta \xi|^2 + \frac{C_{f^1}}{\lambda^2} \int_0^T e^{\beta s} |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^1(Y_{s:T}^2)|^2 ds + \frac{C_{f^1}}{\mu^2} \int_0^T e^{\beta s} |\delta_2 \Lambda_s|^2 ds + \frac{1}{\mu^2} \int_0^T e^{\beta s} |\delta_2 f_s|^2 ds \right. \\ &\quad \left. + 2 \sup_{t \in [0, T]} \left| \int_t^T e^{\beta s} \delta Y_s \delta Z_s dW_s \right| \right]. \end{aligned} \quad (2.5)$$

As Λ^1 satisfies Assumption (A), Jensen's inequality yields

$$\begin{aligned} \mathbb{E} \left[\int_0^T e^{\beta s} |\Lambda_s^1(Y_{s:T}^1) - \Lambda_s^1(Y_{s:T}^2)|^2 ds \right] &\leq C_\Lambda^2 \mathbb{E} \left[\int_0^T e^{\beta s} \int_s^T |\delta Y_u|^2 \nu_s(du) ds \right] \\ &\leq TC_\Lambda^2 \mathbb{E} \left[\sup_{t \in [0, T]} e^{\beta t} |\delta Y_t|^2 \right]. \end{aligned} \quad (2.6)$$

By the Burkholder-Davis-Gundy inequality, there exists a positive constant C_1 such that

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \int_t^T e^{\beta s} \delta Y_s \delta Z_s dW_s \right| \right] &\leq C_1 \mathbb{E} \left[\left(\int_0^T e^{2\beta s} |\delta Y_s|^2 |\delta Z_s|^2 ds \right)^{1/2} \right] \\ &\leq C_1 \mathbb{E} \left[\left(\sup_{s \in [0, T]} e^{\beta s} |\delta Y_s|^2 \right)^{1/2} \left(\int_0^T e^{\beta s} |\delta Z_s|^2 ds \right)^{1/2} \right]. \end{aligned} \quad (2.7)$$

Therefore, by Young's inequality with $\gamma > 0$, we have

$$\begin{aligned} 2\mathbb{E} \left[\sup_{t \in [0, T]} \left| \int_t^T e^{\beta s} \delta Y_s \delta Z_s dW_s \right| \right] &\leq \frac{C_1}{\gamma^2} \mathbb{E} \left[\sup_{t \in [0, T]} e^{\beta t} |\delta Y_t|^2 \right] + \gamma^2 C_1 \mathbb{E} \left[\int_0^T e^{\beta s} |\delta Z_s|^2 ds \right] \\ &\leq \frac{C_1}{\gamma^2} \mathbb{E} \left[\sup_{t \in [0, T]} e^{\beta t} |\delta Y_t|^2 \right] + \frac{\gamma^2 C_1 \lambda^2}{\lambda^2 - C_{f^1}} \mathbb{E} \left[e^{\beta T} |\delta \xi|^2 + \frac{C_{f^1}}{\mu^2} \int_t^T e^{\beta s} |\delta_2 \Lambda_s|^2 ds \right. \\ &\quad \left. + \frac{1}{\mu^2} \int_t^T e^{\beta s} |\delta_2 f_s|^2 ds \right] + \frac{\gamma^2 C_1 \lambda^2}{\lambda^2 - C_{f^1}} \frac{C_{f^1}}{\lambda^2} TC_\Lambda^2 \mathbb{E} \left[\sup_{t \in [0, T]} e^{\beta t} |\delta Y_t|^2 \right]. \end{aligned} \quad (2.8)$$

Combining the inequalities (2.5)–(2.8) leads to

$$\begin{aligned} &\left(1 - \frac{C_1}{\gamma^2} - \frac{TC_{f^1} C_\Lambda^2}{\lambda^2} - \frac{TC_{f^1} C_1 C_\Lambda^2 \gamma^2}{\lambda^2 - C_{f^1}} \right) \mathbb{E} \left[\sup_{t \in [0, T]} e^{\beta t} |\delta Y_t|^2 \right] + \left(1 - \frac{C_{f^1}}{\lambda^2} \right) \mathbb{E} \left[\int_0^T e^{\beta s} |\delta Z_s|^2 ds \right] \\ &\leq \left(1 + \frac{\gamma^2 C_1 \lambda^2}{\lambda^2 - C_{f^1}} \right) \left(\mathbb{E} [e^{\beta T} |\delta \xi|^2] + \frac{1}{\mu^2} \mathbb{E} \left[C_{f^1} \int_0^T e^{\beta s} |\delta_2 \Lambda_s|^2 ds + \int_0^T e^{\beta s} |\delta_2 f_s|^2 ds \right] \right). \end{aligned}$$

Let us define a continuous function Γ by setting

$$\Gamma(\gamma, \lambda) = 1 - \frac{C_1}{\gamma^2} - \frac{TC_{f^1} C_\Lambda^2}{\lambda^2} - \frac{TC_{f^1} C_1 C_\Lambda^2 \gamma^2}{\lambda^2 - C_{f^1}}$$

for any $\gamma > 0$ and any $\lambda > 0$ with $\lambda^2 > C_{f^1}$. Observe that if we set $\gamma(\lambda) = \sqrt{\lambda}$ with $\lambda > 0$, we have $\lim_{\lambda \rightarrow \infty} \Gamma(\gamma(\lambda), \lambda) = 1$, so that there exist λ, γ large enough such that $\Gamma(\gamma, \lambda) > 0$. For such a choice of γ and

λ , we obtain the announced result with the constant

$$C = \frac{1 + \frac{C_1 \gamma^2 \lambda^2}{\lambda^2 - C_{f^1}}}{\min\left(\Gamma(\gamma, \lambda), 1 - \frac{C_{f^1}}{\lambda^2}\right)}.$$

Recall that β is chosen according to the value of λ (see inequality (2.4)). \square

Proof of Theorem 2.1. We use the previous apriori estimates in the case where (Y^1, Z^1) and (Y^2, Z^2) solve respectively the BSDEs

$$\begin{aligned} Y_t^1 &= \xi + \int_t^T f_s(U_s^1, V_s^1, \Lambda_s(U_{s:T}^1)) ds - \int_t^T Z_s^1 dW_s, \\ Y_t^2 &= \xi + \int_t^T f_s(U_s^2, V_s^2, \Lambda_s(U_{s:T}^2)) ds - \int_t^T Z_s^2 dW_s. \end{aligned}$$

Here, $(U^1, V^1), (U^2, V^2) \in \mathbb{S}_{0,T}^2(\mathbb{R}) \times \mathbb{H}_{0,T}^2(\mathbb{R}^d)$ are given processes. Therefore $f_s(U_s^1, V_s^1, \Lambda_s(U_{s:T}^1))$ and $f_s(U_s^2, V_s^2, \Lambda_s(U_{s:T}^2))$ define processes in $\mathbb{H}_{0,T}^2(\mathbb{R})$ owing to the Assumptions (S) and (A). Therefore, the existence and uniqueness of (Y^1, Z^1) and (Y^2, Z^2) in $\mathbb{S}_{0,T}^2(\mathbb{R}) \times \mathbb{H}_{0,T}^2(\mathbb{R}^d)$ as solutions of BSDEs is classical (see [EPQ97, Theorem 2.1, Proposition 2.2]).

In addition, the process $Y^1 - Y^2$ is then solution to the BSDE $Y_t^1 - Y_t^2 = \int_t^T \delta_2 f_s ds - \int_t^T (Z_s^1 - Z_s^2) dW_s$, where the driver $\delta_2 f_s = f_s(U_s^1, V_s^1, \Lambda_s(U_{s:T}^1)) - f_s(U_s^2, V_s^2, \Lambda_s(U_{s:T}^2))$ does not depend on Y_s^1 nor Y_s^2 . Using Lemma 2.2 for $C_f = 0$ and $\mu > 0$, we have that for $\beta > 0$ large enough,

$$\|\delta Y\|_{\mathbb{S}_{\beta,T}^2}^2 + \|\delta Z\|_{\mathbb{H}_{\beta,T}^2}^2 \leq \frac{C}{\mu^2} \|\delta_2 f\|_{\mathbb{H}_{\beta,T}^2}^2.$$

Moreover,

$$\begin{aligned} \|\delta_2 f\|_{\mathbb{H}_{\beta,T}^2}^2 &= \mathbb{E} \left[\int_0^T e^{\beta s} |f_s(U_s^1, V_s^1, \Lambda_s(U_{s:T}^1)) - f_s(U_s^2, V_s^2, \Lambda_s(U_{s:T}^2))|^2 ds \right] \\ &\leq 3C_f^2 \mathbb{E} \left[\int_0^T e^{\beta s} (|\delta U_s|^2 + |\delta V_s|^2 + |\Lambda_s(U_{s:T}^1) - \Lambda_s(U_{s:T}^2)|^2) ds \right]. \end{aligned}$$

As we have that $\|\delta U\|_{\mathbb{H}_{\beta,T}^2}^2 \leq T \|\delta U\|_{\mathbb{S}_{\beta,T}^2}^2$, and

$$\mathbb{E} \left[\int_0^T e^{\beta s} |\Lambda_s(U_{s:T}^1) - \Lambda_s(U_{s:T}^2)|^2 ds \right] \leq C_\Lambda^2 \mathbb{E} \left[\int_0^T e^{\beta s} \int_s^T |\delta U_u|^2 \nu_s(du) ds \right] \leq TC_\Lambda^2 \|\delta U\|_{\mathbb{S}_{\beta,T}^2}^2.$$

We obtain that

$$\|\delta Y\|_{\mathbb{S}_{\beta,T}^2}^2 + \|\delta Z\|_{\mathbb{H}_{\beta,T}^2}^2 \leq \frac{3CC_f^2}{\mu^2} \left((TC_\Lambda^2 + T) \|\delta U\|_{\mathbb{S}_{\beta,T}^2}^2 + \|\delta V\|_{\mathbb{H}_{\beta,T}^2}^2 \right).$$

We now choose $\mu^2 > 3CC_f^2(TC_\Lambda^2 + T + 1)$, and obtain that for β large enough, the mapping $\phi : (U, V) \rightarrow (Y, Z)$ is a contraction in the space $\mathbb{S}_{\beta,T}^2(\mathbb{R}) \times \mathbb{H}_{\beta,T}^2(\mathbb{R}^d)$. Hence, we get existence and uniqueness of a solution to the BSDE (2.1). \square

3. THE CASE OF CVAR INITIAL MARGINS

In this section, we apply the previous results on MKABSDE to a generalization of equation (1.5). We will consider a multi-dimensional setting (several assets S^i), a more general dynamic model and a more general terminal condition: we refer to the following section for precise definitions. Beyond usual existence and uniqueness results, our aim is to analyse related approximations of the option price and delta, obtained when the CVaR is evaluated using Gaussian expansions (justified as $\Delta \rightarrow 0$, see Theorem 3.1).

3.1. Well-posedness of the problem

Let us consider a general Itô market with d tradable assets [KS98, Chapter 1]. The riskless asset S^0 (money account) follows the dynamics $\frac{dS_t^0}{S_t^0} = r_t dt$, and we have d risky assets (S^1, \dots, S^d) following

$$\frac{dS_t^i}{S_t^i} = \mu_t^i dt + \sum_{j=1}^d \sigma_t^{ij} dW_t^j, \quad S_0^i = s_0^i \in \mathbb{R}, \quad 1 \leq i \leq d. \quad (3.1)$$

The processes $r, \mu := (\mu^i)_{1 \leq i \leq d}, \sigma := (\sigma^{ij})_{1 \leq i, j \leq d}$ are \mathcal{F} -adapted stochastic processes with values respectively in \mathbb{R}, \mathbb{R}^d , and the set of matrices of size $d \times d$. Moreover, we assume that $dt \otimes d\mathbb{P}$ a.e., the matrix σ_t is invertible and the processes r and $\sigma^{-1}(\mu - r\mathbf{1})$ are uniformly bounded, where we define the column vector $\mathbf{1} := (1, \dots, 1)^\top \in \mathbb{R}^d$. For a path-dependent payoff ξ paid at maturity T , the dynamics of the hedging portfolio (V, π) with CVaR initial margin requirement (over a period $\Delta > 0$) is given by

$$V_t = \xi + \int_t^T (-r_s V_s + \pi_s (r_s \mathbf{1} - \mu_s) + R \mathbf{CVaR}_{\mathcal{F}_s}^\alpha (V_s - V_{(s+\Delta) \wedge T})) ds - \int_t^T \pi_s \sigma_s dW_s.$$

Here π is a row vector whose i th coordinate consists of the amount invested in the i th asset. The derivation is analogous to that of Section 1.2. This equation rewrites, in terms of the variables $(V, Z = \pi\sigma)$,

$$V_t = \xi + \int_t^T (-r_s V_s + Z_s \sigma_s^{-1} (r_s \mathbf{1} - \mu_s) + R \mathbf{CVaR}_{\mathcal{F}_s}^\alpha (V_s - V_{(s+\Delta) \wedge T})) ds - \int_t^T Z_s dW_s. \quad (3.2)$$

Existence and uniqueness for the MKABSDE above are a consequence of Theorem 2.1.

Corollary 3.1. *For any square integrable terminal condition ξ , the CVaR initial margin problem (3.2) is well posed with a unique solution $(V, Z) \in \mathbb{S}_{\beta, T}^2(\mathbb{R}) \times \mathbb{H}_{\beta, T}^2(\mathbb{R}^d)$ for any $\beta \geq 0$.*

Note that, when $R = 0$, the coefficients r_t, μ_t and σ_t are deterministic functions of time and $\xi = \Phi(S_T)$, (3.2) is solved by $V_t = v(t, S_t)$, where S follows (3.1) and v is the solution to the classical (multi-dimensional) Black-Scholes pricing equation with interest rate r_t , volatility parameter σ_t and payoff function Φ . As usual, the function v does not depend on the drift parameter μ , while the portfolio process $V_t = v(t, S_t)$ (aka the MtM of the trade) does, via the asset price S . In general, the solution (V, Z) to (3.2) will still depend on μ , even if we cannot identify an analogue of the function v in the general case.

Proof. We check that Assumptions S and A are satisfied for the problem (3.2). The driver of the BSDE has the form

$$f(t, v, z, \lambda) = -r_t v + z \sigma_t^{-1} (r_t \mathbf{1} - \mu_t) + \lambda, \quad t \geq 0, \quad v, \lambda \in \mathbb{R}, \quad z \in \mathbb{R}^d,$$

and we introduce the functional

$$\Lambda_t(X_{t:T}) := R \mathbf{CVaR}_{\mathcal{F}_t}^\alpha (X_t - X_{(t+\Delta) \wedge T}) = R \inf_{x \in \mathbb{R}} \mathbb{E} \left[\frac{(X_t - X_{(t+\Delta) \wedge T} - x)^+}{1 - \alpha} + x \mid \mathcal{F}_t \right], \quad t \in [0, T], \quad X \in \mathbb{S}_{0, T}^2(\mathbb{R}).$$

Since r and $\sigma^{-1}(\mu - r\mathbf{1})$ are uniformly bounded, f clearly satisfies Assumption (S). We now check that Λ satisfies Assumption (A). For $X \in \mathbb{S}_{0,T}^2(\mathbb{R})$ and $x \in \mathbb{R}$, we have

$$\mathbb{E} [X_t - X_{(t+\Delta)\wedge T} | \mathcal{F}_t] \leq \inf_{x \in \mathbb{R}} \mathbb{E} \left[\frac{(X_t - X_{(t+\Delta)\wedge T} - x)^+}{1 - \alpha} + x \middle| \mathcal{F}_t \right] \leq \mathbb{E} \left[\frac{(X_t - X_{(t+\Delta)\wedge T})^+}{1 - \alpha} \middle| \mathcal{F}_t \right], \quad (3.3)$$

where for the left hand side (l.h.s.) we use the fact that $\frac{(z-x)^+}{1-\alpha} + x \geq z$ for $\alpha \in (0, 1)$ and $z, x \in \mathbb{R}$, and for the right hand side (r.h.s.), we bound the infimum from above with the value taken at $x = 0$. Since it is easy to check that both the l.h.s. and the r.h.s. of (3.3) belong to $\mathbb{H}_{0,T}^2$, we conclude that $\Lambda_t(X) \in \mathbb{H}_{0,T}^2(\mathbb{R})$. Now, let $X^1, X^2 \in \mathbb{S}_{0,T}^2$. We have

$$\begin{aligned} |\Lambda_t(X^1) - \Lambda_t(X^2)| &\leq R \mathbb{E} \left[\left| \frac{X_t^1 - X_t^2 - (X_{(t+\Delta)\wedge T}^1 - X_{(t+\Delta)\wedge T}^2)}{1 - \alpha} \right| \middle| \mathcal{F}_t \right] \\ &\leq \frac{2R}{1 - \alpha} \mathbb{E} \left[\int_t^T |X_s^1 - X_s^2| \nu_t(ds) \middle| \mathcal{F}_t \right], \end{aligned}$$

where, for the first inequality, we use the fact that $|\inf_{x \in \mathbb{R}} g^1(x) - \inf_{x \in \mathbb{R}} g^2(x)| \leq \sup_{x \in \mathbb{R}} |g^1(x) - g^2(x)|$ for any couple of functions $g^1, g^2 : \mathbb{R} \rightarrow \mathbb{R}$ and the 1-Lipschitz property of the positive part function, and for the second inequality, we have set $\nu_t := \frac{1}{2}(\delta_t + \delta_{(t+\Delta)\wedge T})$, where δ_u is the Dirac measure on $\{u\}$. We conclude that Assumption (A) holds with $C_\Lambda = \frac{2R}{1-\alpha}$. We finally apply Theorem 2.1 to obtain the claim. \square

3.2. Approximation by standard BSDEs when $\Delta \ll 1$

The numerical solution of (3.2) is challenging in full generality. In fact, it is a priori more difficult than solving a standard BSDE, for which we can employ, for example, regression Monte-Carlo methods (see e.g. [GT16] and references therein). In this work, we take advantage of the fact that Δ is small (recall $\Delta =$ one week or 10 days) in order to provide handier approximations of (V, Z) defined in terms of standard non-linear or linear BSDEs. Below, we define these different BSDEs and provide the related error estimates.

At the lowest order in the parameter $\sqrt{\Delta}$, formally we have that, conditionally to \mathcal{F}_s ,

$$V_s - V_{(s+\Delta)\wedge T} \approx - \int_s^{(s+\Delta)\wedge T} Z_u dW_u \stackrel{(d)}{=} -|Z_s| \sqrt{(s+\Delta) \wedge T - s} \times G,$$

where we freeze the process Z at current time $s \in [0, T]$, and $G \stackrel{(d)}{=} \mathcal{N}(0, 1)$ is independent from \mathcal{F}_s . This is an approximation of CVaR using the ‘‘Delta’’ portion of the portfolio (in the spirit of [GHS00, Section 2]). Plugging this approximation into (3.2), and defining

$$C_\alpha := \mathbf{CVaR}^\alpha(\mathcal{N}(0, 1)) = \frac{e^{-x^2/2}}{(1 - \alpha)\sqrt{2\pi}} \Bigg|_{x=\mathcal{N}^{-1}(\alpha)}, \quad (3.4)$$

we obtain a standard non-linear BSDE

$$V_t^{NL} = \xi + \int_t^T \left(-r_s V_s^{NL} + Z_s^{NL} \sigma_s^{-1} (r_s \mathbf{1} - \mu_s) + R C_\alpha \sqrt{(s+\Delta) \wedge T - s} |Z_s^{NL}| \right) ds - \int_t^T Z_s^{NL} dW_s. \quad (3.5)$$

Seeing V^{NL} as a function of the parameter Δ appearing in the driver, we follow the expansion procedure in [GP15] and perform an expansion at the orders 0 and 1 w.r.t. $\sqrt{\Delta}$. We obtain two linear BSDEs, respectively

(V^{BS}, Z^{BS}) and (V^L, Z^L) , where

$$V_t^{BS} = \xi + \int_t^T (-r_s V_s^{BS} + Z_s^{BS} \sigma_s^{-1} (r_s \mathbf{1} - \mu_s)) ds - \int_t^T Z_s^{BS} dW_s, \quad (3.6)$$

$$V_t^L = \xi + \int_t^T \left(-r_s V_s^L + Z_s^L \sigma_s^{-1} (r_s \mathbf{1} - \mu_s) + RC_\alpha \sqrt{(s + \Delta) \wedge T - s} |Z_s^{BS}| \right) ds - \int_t^T Z_s^L dW_s. \quad (3.7)$$

Let us comment on these different models:

- The simplest equation is (V^{BS}, Z^{BS}) , corresponding to the usual linear valuation rule [EPQ97, Theorem 1.1] without IM requirement. When the model is a one-dimensional geometric Brownian motion and $\xi = (S_T - K)_+$, the solution is given by the usual Black-Scholes formula.
- The second simplest equation is (V^L, Z^L) where the IM cost is computed using the “Delta” of an exogenous reference price given by the simpler pricing rule (V^{BS}, Z^{BS}) without IM. This is still a linear BSDE, but the numerical simulation requires to know Z^{BS} in order to evaluate (V^L, Z^L) . We will use a nested Monte-Carlo procedure in our experiments.
- The third equation is (3.5), where the IM cost is computed using the “Delta” of the endogenous price (V^{NL}, Z^{NL}) itself.

Existence and uniqueness of a solution to the BSDEs (3.5), (3.6) and (3.7) are direct consequences of [PP90], as the respective drivers satisfy standard Lipschitz properties and the processes r and $\sigma^{-1}(r\mathbf{1} - \mu)$ are bounded.

Proposition 3.1. *The BSDEs (3.5), (3.6) and (3.7) have a unique solution in the L_2 -space $\mathbb{S}_{0,T}^2 \times \mathbb{H}_{0,T}^2$, whose norm is uniformly bounded in $\Delta \leq T$.*

The main result of this section is the following theorem.

Theorem 3.1. *Define the L_2 time-regularity index of Z^{NL} by*

$$\mathcal{E}^{NL}(\Delta) := \frac{1}{\Delta} \mathbb{E} \left[\int_0^T \int_t^{(t+\Delta) \wedge T} |Z_s^{NL} - Z_t^{NL}|^2 ds dt \right]. \quad (3.8)$$

We always have $\sup_{0 < \Delta \leq T} \mathcal{E}^{NL}(\Delta) < +\infty$. Moreover, there exist constants $K_1, K_2, K_3 > 0$, independent from Δ , such that

$$\|V^L - V^{BS}\|_{\mathbb{S}_{0,T}^2}^2 + \|Z^L - Z^{BS}\|_{\mathbb{H}_{0,T}^2}^2 \leq K_1 \Delta, \quad (3.9)$$

$$\|V^{NL} - V^L\|_{\mathbb{S}_{0,T}^2}^2 + \|Z^{NL} - Z^L\|_{\mathbb{H}_{0,T}^2}^2 \leq K_2 \Delta^2, \quad (3.10)$$

$$\|V - V^{NL}\|_{\mathbb{S}_{0,T}^2}^2 + \|Z - Z^{NL}\|_{\mathbb{H}_{0,T}^2}^2 \leq K_3 \Delta (\Delta + \mathcal{E}^{NL}(\Delta)). \quad (3.11)$$

In addition, we have

$$\mathcal{E}^{NL}(\Delta) = O(\Delta), \quad (3.12)$$

and thus $\|V - V^{NL}\|_{\mathbb{S}_{0,T}^2}^2 + \|Z - Z^{NL}\|_{\mathbb{H}_{0,T}^2}^2 = O(\Delta^2)$ provided that the additional sufficient conditions below are fulfilled:

- (i) the terminal condition is a Lipschitz functional of S , that is, $\xi = \Phi(S_{0:T})$ for some functional Φ satisfying

$$|\Phi(x_{0:T}) - \Phi(x'_{0:T})| \leq C_\Phi \sup_{t \in [0,T]} |x_t - x'_t|,$$

for any continuous paths $x, x' : [0, T] \rightarrow \mathbb{R}^d$;

- (ii) the coefficients r, σ, μ are constant.

Estimate (3.12) follows from [Zha04]. Let us remark that the tools used in [Zha04] to prove (3.12), and consequently the estimate (3.12) itself, should also hold if we assume (i) and the following more general conditions:

- (iii) the processes r, σ, μ are Markovian, i.e. $r_t = \hat{r}(t, S_t)$, $\sigma_t^{ij} = \hat{\sigma}^{ij}(t, S_t)$ and $\mu_t^i = \hat{\mu}^i(t, S_t)$ for some deterministic functions $\hat{r}, \hat{\mu}^i, \hat{\sigma}^{ij}$;
- (iv) the functions $x \rightarrow \hat{\mu}^i(x)x_i$, $x \rightarrow \hat{\sigma}^{ij}(x)x_i$ are globally Lipschitz in $(t, x) \in [0, T] \times \mathbb{R}^d$, for any $1 \leq i, j \leq d$;
- (v) the functions \hat{r} and $\hat{\sigma}^{-1}(\hat{r}\mathbf{1} - \hat{\mu})$ are globally Lipschitz in $(t, x) \in [0, T] \times \mathbb{R}^d$.

Remark 3.1. When $\mathcal{E}^{NL}(\Delta) = O(\Delta)$, then (V, Z) and (V^{NL}, Z^{NL}) are close to each other: precisely, the squared error is of order Δ^2 . Overall, the error estimates (3.9)-(3.10)-(3.11) show that, on the one hand, there is a significant difference between valuation with or without initial margin cost: the size of the corrections $\|V^L - V^{BS}\|_{\mathbb{S}_{0,T}^2}$ and $\|Z^L - Z^{BS}\|_{\mathbb{H}_{0,T}^2}$ is of order $\sqrt{\Delta}$ (in (3.9)). On the other hand, the other valuation rules yield comparable values when $\Delta \ll 1$, since the other corrections ((3.10) and (3.11)) have size of order Δ .

3.3. Proof of Theorem 3.1

▷ *Estimate on $\mathcal{E}^{NL}(\Delta)$.* We start with a deterministic inequality. For any positive function Ψ and any $\beta \geq 0$, we have

$$\int_0^T e^{\beta t} \left(\int_t^{(t+\Delta) \wedge T} \Psi_s ds \right) dt \leq \Delta \int_0^T e^{\beta s} \Psi_s ds. \quad (3.13)$$

Indeed the left hand side of (3.13) can be written as

$$\int_0^T \int_0^T e^{\beta t} \Psi_s \mathbf{1}_{t \leq s \leq (t+\Delta) \wedge T} ds dt = \int_0^T \Psi_s \left(\int_0^T e^{\beta t} \mathbf{1}_{t \leq s \leq (t+\Delta) \wedge T} dt \right) ds, \quad (3.14)$$

from which (3.13) follows easily. Using $(a+b)^2 \leq 2a^2 + 2b^2$ and (3.13) with $\beta = 0$ gives

$$\mathcal{E}^{NL}(\Delta) \leq \frac{2}{\Delta} \mathbb{E} \left[\int_0^T \int_t^{(t+\Delta) \wedge T} (|Z_s^{NL}|^2 + |Z_t^{NL}|^2) ds dt \right] \leq 4 \mathbb{E} \left[\int_0^T |Z_t^{NL}|^2 dt \right], \quad (3.15)$$

and the last term is uniformly bounded in Δ (Proposition 3.1).

We now derive finer estimates that reveal the L_2 time-regularity of Z^{NL} under the additional assumptions (i)-(ii). In this Markovian setting, we know that Z^{NL} has a càdlàg version (see [Zha04, Remark (ii) after Lemma 2.5]). Then, introduce the time grid $t_i = i\Delta$ for $0 \leq i \leq n := \lfloor \frac{T}{\Delta} \rfloor$ and $t_{n+1} = T$. We claim that

$$\mathcal{E}^{NL}(\Delta) \leq 4 \sum_{i=0}^n \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_s^{NL} - Z_{t_i}^{NL}|^2 + |Z_s^{NL} - Z_{t_{i+1}}^{NL}|^2 ds \right]. \quad (3.16)$$

With this result at hand, the estimate (3.12) follows directly from an application of [Zha04, Theorem 3.1]. In order to show (3.16), denote $\varphi^-(s)$ and $\varphi^+(s)$ the points on the grid before and after s , and write

$$\begin{aligned} \int_0^T \int_t^{(t+\Delta) \wedge T} |Z_s^{NL} - Z_t^{NL}|^2 ds dt &\leq 2 \sum_{i=0}^n \int_{t_i}^{t_{i+1}} \int_t^{(t+\Delta) \wedge T} (|Z_s^{NL} - Z_{t_{i+1}}^{NL}|^2 + |Z_t^{NL} - Z_{t_{i+1}}^{NL}|^2) ds dt \\ &\leq 2 \int_0^T \int_t^{(t+\Delta) \wedge T} (|Z_s^{NL} - Z_{\varphi^-(s)}^{NL}|^2 + |Z_s^{NL} - Z_{\varphi^+(s)}^{NL}|^2) ds dt + 2\Delta \int_0^T |Z_t^{NL} - Z_{\varphi^+(t)}^{NL}|^2 dt \\ &\leq 4\Delta \left(\int_0^T |Z_t^{NL} - Z_{\varphi^-(t)}^{NL}|^2 dt + \int_0^T |Z_t^{NL} - Z_{\varphi^+(t)}^{NL}|^2 dt \right) \end{aligned}$$

where we have used (3.13) with $\beta = 0$. The inequality (3.16) follows.

▷ *Proof of (3.9)*. This error estimate is related to the difference of two linear BSDEs. The drivers of V^{BS} and V^L are respectively $f^{BS}(s, y, z, \lambda) = -r_s y + z \sigma_s^{-1}(r_s \mathbf{1} - \mu_s)$, and

$$f^L(s, y, z, \lambda) = -r_s y + z \sigma_s^{-1}(r_s \mathbf{1} - \mu_s) + RC_\alpha \sqrt{(s + \Delta) \wedge T - s} |Z_s^{BS}|,$$

for $s \in [0, T]$, $v, \lambda \in \mathbb{R}$ and $z \in \mathbb{R}^d$, hence

$$(f^L - f^{BS})(s, y, z, \lambda) = RC_\alpha \sqrt{(s + \Delta) \wedge T - s} |Z_s^{BS}|.$$

By Lemma 2.2, we obtain that for $\mu > 0$, β large enough and $K_1 = \frac{C}{\mu^2} (RC_\alpha)^2 \|Z^{BS}\|_{\mathbb{H}_{\beta, T}^2}$,

$$\|V^L - V^{BS}\|_{\mathbb{S}_{0, T}^2}^2 + \|Z^L - Z^{BS}\|_{\mathbb{H}_{0, T}^2}^2 \leq \|V^L - V^{BS}\|_{\mathbb{S}_{\beta, T}^2}^2 + \|Z^L - Z^{BS}\|_{\mathbb{H}_{\beta, T}^2}^2 \leq K_1 \Delta.$$

We are done with (3.9).

▷ *Proof of (3.10)*. Then, as $\xi \in L^2$, as the processes $r, \sigma^{-1}(\mu - r\mathbf{1})$ are bounded and as the non-linear term

$$t, z \in [0, T] \times \mathbb{R}^d \rightarrow RC_\alpha \sqrt{(t + \Delta) \wedge T - t} |z|$$

is Lipschitz in the variable z , uniformly in time, we obtain Inequality (3.10) as an application of [GP15, Theorem 2.4], for which assumptions **H.1** – **H.3** are satisfied.

▷ *Proof of (3.11)*. Using computations similar to those in the proof of Lemma 2.2, we obtain existence of a constant $C > 0$ such that for $\mu > 0$ and β large enough,

$$\|V - V^{NL}\|_{\mathbb{S}_{\beta, T}^2}^2 + \|Z - Z^{NL}\|_{\mathbb{H}_{\beta, T}^2}^2 \leq \frac{C}{\mu^2} \left\| \mathbf{CVaR}_{\mathcal{F}_t}^\alpha \left(V^{NL} - V_{(\cdot + \Delta) \wedge T}^{NL} \right) - \mathbf{CVaR}_{\mathcal{F}_t}^\alpha \left(- \int_{\cdot}^{(\cdot + \Delta) \wedge T} Z^{NL} dW_s \right) \right\|_{\mathbb{H}_{\beta, T}^2}^2.$$

As the CVaR function is subadditive [RU00], we have that given A, B two random variables, and $t \in [0, T]$, $\mathbf{CVaR}_{\mathcal{F}_t}^\alpha(A) \leq \mathbf{CVaR}_{\mathcal{F}_t}^\alpha(B) + \mathbf{CVaR}_{\mathcal{F}_t}^\alpha(A - B)$. Inverting the roles of A and B , we obtain that

$$\begin{aligned} 0 \leq | \mathbf{CVaR}_{\mathcal{F}_t}^\alpha(A) - \mathbf{CVaR}_{\mathcal{F}_t}^\alpha(B) | &\leq \max(\mathbf{CVaR}_{\mathcal{F}_t}^\alpha(A - B), \mathbf{CVaR}_{\mathcal{F}_t}^\alpha(B - A)) \\ &\leq \frac{1}{1 - \alpha} \left(\mathbb{E} \left[(A - B)^+ \mid \mathcal{F}_t \right] + \mathbb{E} \left[(B - A)^+ \mid \mathcal{F}_t \right] \right) = \frac{\mathbb{E} [|A - B| \mid \mathcal{F}_t]}{1 - \alpha}, \end{aligned}$$

where for the last inequality, we have used that for $U \in \{A - B, B - A\}$, $\inf_{x \in \mathbb{R}} \mathbb{E} \left[\frac{(U - x)^+}{1 - \alpha} + x \mid \mathcal{F}_t \right] \leq \mathbb{E} \left[\frac{U^+}{1 - \alpha} \mid \mathcal{F}_t \right]$. We then have that

$$| \mathbf{CVaR}_{\mathcal{F}_t}^\alpha(A) - \mathbf{CVaR}_{\mathcal{F}_t}^\alpha(B) |^2 \leq \frac{1}{(1 - \alpha)^2} \mathbb{E} \left[(A - B)^2 \mid \mathcal{F}_t \right].$$

Setting, for $t \in [0, T]$, $A_t = - \int_t^{(t + \Delta) \wedge T} Z_t^{NL} dW_s$, $B_t = V_t^{NL} - V_{(t + \Delta) \wedge T}^{NL}$ and using the previous inequality, we obtain

$$\|V - V^{NL}\|_{\mathbb{S}_{\beta, T}^2}^2 + \|Z - Z^{NL}\|_{\mathbb{H}_{\beta, T}^2}^2 \leq \frac{C}{\mu^2 (1 - \alpha)^2} \mathbb{E} \left[\int_0^T e^{\beta t} \left(V_t^{NL} - V_{(t + \Delta) \wedge T}^{NL} + \int_t^{(t + \Delta) \wedge T} Z_t^{NL} dW_s \right)^2 dt \right].$$

Now writing

$$\begin{aligned} V_t^{NL} - V_{(t+\Delta)\wedge T}^{NL} + \int_t^{(t+\Delta)\wedge T} Z_t^{NL} dW_s \\ = \int_t^{(t+\Delta)\wedge T} \left(-r_s V_s^{NL} + Z_s^{NL} \sigma_s^{-1} (r_s \mathbf{1} - \mu_s) + R C_\alpha \sqrt{(s+\Delta) \wedge T - s} |Z_s^{NL}| \right) ds \\ - \int_t^{(t+\Delta)\wedge T} (Z_s^{NL} - Z_t^{NL}) dW_s =: \Pi_1(t) - \Pi_2(t), \end{aligned}$$

we obtain $\|V - V^{NL}\|_{\mathbb{S}_{\beta,T}^2}^2 + \|Z - Z^{NL}\|_{\mathbb{H}_{\beta,T}^2}^2 \leq \frac{2C}{\mu^2(1-\alpha)^2} \mathbb{E} \left[\int_0^T e^{\beta t} (\Pi_1^2(t) + \Pi_2^2(t)) dt \right]$. By Jensen's inequality and the inequality (3.13), we get

$$\begin{aligned} \mathbb{E} \left[\int_0^T e^{\beta t} \Pi_1^2(t) dt \right] &\leq 3\Delta \mathbb{E} \left[\int_0^T e^{\beta t} \int_t^{(t+\Delta)\wedge T} \left((-r_s V_s^{NL})^2 + (Z_s^{NL} \sigma_s^{-1} (r_s \mathbf{1} - \mu_s))^2 + (R C_\alpha \sqrt{\Delta} |Z_s^{NL}|)^2 \right) ds dt \right] \\ &\leq 3\Delta^2 \left(|r|_\infty^2 + |\sigma^{-1}(r\mathbf{1} - \mu)|_\infty^2 + (R C_\alpha)^2 \right) \mathbb{E} \left[\int_0^T e^{\beta t} (|V_t^{NL}|^2 + |Z_t^{NL}|^2) dt \right]. \end{aligned}$$

Now invoking the uniform estimate in Proposition 3.1, we finally obtain $\mathbb{E} \left[\int_0^T e^{\beta t} (\Pi_1(t))^2 dt \right] \leq \tilde{K}_1 \Delta^2$ for some positive constant \tilde{K}_1 . Moreover, using Ito's isometry, we obtain

$$\mathbb{E} \left[\int_0^T e^{\beta t} \Pi_2^2(t) dt \right] = \mathbb{E} \left[\int_0^T e^{\beta t} \int_t^{(t+\Delta)\wedge T} |Z_s^{NL} - Z_t^{NL}|^2 ds dt \right] \leq e^{\beta T} \Delta \mathcal{E}^{NL}(\Delta).$$

Gathering all the previous arguments, we get (3.11). The proof of Theorem 3.1 is now complete. \square

4. NUMERICAL EXAMPLES

In the absence of numerical methods to estimate the solution of the McKean Anticipative BSDE (3.2) in full generality, we rather solve numerically the BSDE approximations (3.5) or (3.7) as discussed in Section 3.2. For this purpose, when the dimension d is greater than one, we use the Stratified Regression Multistep-forward Dynamical Programming (SRMDP) scheme developed in [GLTV16]. In our numerical tests in this section, we set the coefficients μ_t, σ_t and r_t of the model (3.1) to be constant (multi-dimensional geometric Brownian motion).

Note that the linear BSDE (3.7) is solved by $V_t^L = v^L(t, S_t)$, where S follows (3.1) and v is the solution to the (linear) PDE

$$\partial_t v^L(t, x) + (rx \partial_x + \frac{1}{2} \sigma^2 x^2 \partial_{xx}) v^L(t, x) + R C_\alpha \sqrt{(t+\Delta) \wedge T - t} \sigma |\delta^{BS}(t, x)| x - r v^L(t, x) = 0, \quad (4.1)$$

and $v^L(T, x) = \Phi(x)$, where δ^{BS} is the option's delta in the Black-Scholes model with interest rate r and volatility parameter σ . Section 4.3 is precisely devoted to the numerical solution of this PDE via an appropriate Nested Monte Carlo method. In particular, the initial value $V_0^L = v^L(0, S_0)$ does not depend on the drift parameter μ . The same comment applies to the non-linear BSDE (3.5): this equation is solved by $V_t^{NL} = v^{NL}(t, S_t)$, where v^{NL} is the solution of the (non-linear) PDE (4.2)-(4.3), which, once again, does not depend on μ . When the initial portfolio values V_0^L or V_0^{NL} are evaluated via a Monte Carlo method that requires to simulate the underlying process S , the error of the algorithm will, in general, display a dependence with respect to the drift parameter μ . For simplicity, we take $\mu^i = r$ in our experiments.

Observe that setting $R = 0$ reduces the original BSDE to the linear equation (3.6). The (explicit) solution in this case will serve as a reference value in order to assess the impact of Initial Margins.

4.1. Finite difference method for (V^{NL}, Z^{NL}) in dimension 1

In order to check the validity of our results, we first obtain a benchmark when $d = 1$ by solving the semi-linear parabolic PDE related to the BSDE (3.5) when $\xi = \Phi(S_T)$, see [PR14]. By an application of Itô's lemma, the semi-linear PDE is given by

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} + C_\alpha R \sigma \sqrt{(t+\Delta) \wedge T - t} \left| S \frac{\partial V}{\partial S} \right| - rV = 0, \quad (t, S) \in [0, T) \times \mathbb{R}^+, \quad (4.2)$$

$$V(T, S) = \Phi(S), \quad S \in \mathbb{R}^+, \quad (4.3)$$

and $(V_t^{NL}, Z_t^{NL}) = (V(t, S_t), \frac{\partial V}{\partial S}(t, S_t)\sigma S_t)$.

Remark 4.1. *If $\Phi(S) = \max(S - K, 0)$ or $\max(K - S, 0)$ for some $K > 0$, i.e., either a call or a put option payoff, we expect the gradient $\frac{\partial V}{\partial S}$ to have a constant sign. In such a case, the PDE (4.2)–(4.3) becomes linear and in fact has an explicit solution, given by a Black-Scholes formula with time-dependent continuous dividend yield $d(t) = -C_\alpha R \sigma \sqrt{(t+\Delta) \wedge T - t} \operatorname{sign}(\frac{\partial V}{\partial S})$.*

We use a classical finite difference methods to solve (4.2)–(4.3) (see, for example, [AP05]). First, we perform a change of variable, $x = \ln S$, so that the PDE can be rewritten in the following form for the function $v(t, x) := V(t, e^x)$:

$$\frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 \frac{\partial^2 v}{\partial x^2} + \left(r - \frac{1}{2}\sigma^2\right) \frac{\partial v}{\partial x} + C_\alpha R \sigma \sqrt{(t+\Delta) \wedge T - t} \left| \frac{\partial v}{\partial x} \right| - rv = 0, \quad (t, x) \in [0, T) \times \mathbb{R}, \quad (4.4)$$

$$v(T, x) = \Phi(e^x), \quad x \in \mathbb{R}. \quad (4.5)$$

We denote the finite difference domain by $D = [0, T] \times [x_{\min}, x_{\max}]$ with $-\infty < x_{\min} < x_{\max} < \infty$. The domain D is approximated with a uniform mesh $\mathcal{D} = \{(t^n, x_i) : n = 0, 1, \dots, N, i = 0, 1, \dots, M\}$, where $t^n := n\Delta t$ and $x_i := x_{\min} + i\Delta x$. Here, for N time intervals, $\Delta t = T/N$ and $\Delta x = (x_{\max} - x_{\min})/M$ for M spatial steps. Furthermore, we denote $v(t^n, x_i) = v_i^n$. Next, consider the following finite difference derivative approximations under the well-known ω -scheme, i.e., we replace v_i^n by $\omega v_i^n + (1-\omega)v_i^{n+1}$, where $\omega \in [0, 1]$ is a constant parameter, such that

$$\begin{aligned} \frac{\partial v}{\partial t}(t^n, x_i) &\approx \frac{v_i^{n+1} - v_i^n}{\Delta t}, & \frac{\partial v}{\partial x}(t^n, x_i) &\approx \omega \frac{v_{i+1}^n - v_i^n}{\Delta x} + (1-\omega) \frac{v_{i+1}^{n+1} - v_i^{n+1}}{\Delta x}, \\ \frac{\partial^2 v}{\partial x^2}(t^n, x_i) &\approx \omega \frac{v_{i+1}^n - 2v_i^n + v_{i-1}^n}{(\Delta x)^2} + (1-\omega) \frac{v_{i+1}^{n+1} - 2v_i^{n+1} + v_{i-1}^{n+1}}{(\Delta x)^2}. \end{aligned}$$

The choice $\omega = 0.5$ corresponds to Crank-Nicolson method. We also “linearize” the non-linear term by treating it as explicit, i.e., at any time t_n we take $\left| \frac{\partial v}{\partial x}(t^n, x_i) \right| \approx \left| \frac{\partial v}{\partial x}(t^{n+1}, x_i) \right|$.

The substitution of finite difference derivative approximations in (4.4)–(4.5) along with the “linearization” step, leads to the following tridiagonal linear system at each time step $n = N-1, \dots, 0$ which can be solved by Thomas algorithm [YG73]: $Av^n = b^{n+1}$, with nonzero coefficients of the tridiagonal matrix $A = (a_{i,j})$ given by

$$\begin{aligned} a_{0,0} &= 1, & a_{0,1} &= 0, & a_{M,M-1} &= 0, & a_{M,M} &= 1, \\ a_{i,i} &= 1 + 2\theta\omega + \kappa\omega + \rho\omega, & a_{i,i+1} &= -\theta\omega - \kappa\omega, & a_{i-1,i} &= -\theta\omega, & i &= 1, \dots, M-1, \end{aligned}$$

and the time dependent vector b^{n+1} as:

$$b_0^{n+1} = v_0^{n+1}, \quad b_M^{n+1} = v_M^{n+1}, \quad b_i^{n+1} = \theta(1-\omega)v_{i-1}^{n+1} + (1-2\theta(1-\omega) - \kappa(1-\omega) - \rho(1-\omega))v_i^{n+1} \\ + (\theta(1-\omega) + \kappa(1-\omega))v_{i+1}^{n+1} + \beta^n |v_{i+1}^{n+1} - v_i^{n+1}|, \quad i = 1, \dots, M-1,$$

where v_0^{n+1} , v_M^{n+1} are given by the boundary conditions and the remaining constants are defined as below

$$\theta = \frac{\sigma^2 \Delta t}{2(\Delta x)^2}, \quad \kappa = \frac{(r - \frac{1}{2}\sigma^2)\Delta t}{\Delta x}, \quad \rho = r\Delta t, \quad \beta^n = \frac{C_\alpha R \sigma \Delta t \sqrt{(t_n + \Delta) \wedge T - t_n}}{\Delta x}.$$

The i th coordinate of vector v^n is the approximation of the value $v(t^n, x_i)$.

We set the model parameters as $T = 1$, $\sigma = 0.25$, $r = 0.02$, $\alpha = 0.99$ and $\Delta = 0.02$ (1 week) and consider three different options – call, put and butterfly, for different strikes. We set $R = 0.02$ when accounting for IM and $R = 0$ otherwise. The space domain we consider for the finite difference scheme is $[\ln(10^{-6}), \ln(4K)]$, while for the SRMDP regression algorithm, we fix the space domain to be $[-5, 5]$. Furthermore, for the finite difference scheme, the number of steps are $N = 10^3$ and $M = 10^6$. For the **LPO** version of the SRMDP algorithm, we consider 2800 hypercubes, 50 time steps, and 2500 simulations per hypercube. In Figure 1 and Table 1, we present the results for implied volatilities, prices and deltas of call options with different strikes. First, we compute the values using the classical Black-Scholes formula (the curve labeled B-S $R = 0$) in order to assess the impact of taking into account the IM. Next, we solve the non-linear BSDE (3.5) using the three methods presented above: the exact Black-Scholes formula when IM is considered as a time-dependent dividend yield (labeled BS $R = 0.02$) (see Remark 4.1), the finite difference method (FD) and the SRMDP algorithm. For the last method, we compute 95% confidence intervals for the price and the delta of the options. The results for put options are given in Figure 2 and Table 2. In all cases, we observe that the IM has a significant impact on the implied volatility of option prices (around 20-30 basis points for standard values of the volatility). Finally, we consider butterfly options with payoff function

$$\Phi(S_T) = (S_T - (K - 2))^+ - 2(S_T - K)^+ + (S_T - (K + 2))^+ \quad (4.6)$$

in Figure 3 and Table 3. This derivative product involves three options with different strikes – a long position in the butterfly consists in buying a call option with strike $K - 2$, a call with strike $K + 2$, and selling two call options with strike K . Note that the first derivative of the price of a butterfly option (the option delta) is expected to take values of both signs, therefore an explicit Black-Scholes formula is not available any more when IM is taken into account the ($R = 0.02$ here). For all the payoffs above, we observe that the SRMDP algorithm provides good accuracy when compared to the exact values or finite difference estimates.

As a side remark, we note that, due to the non-linear pricing rule V^{NL} , the call and put prices that we compute do not satisfy the put-call parity relation. As a consequence, the implied volatilities extracted from these call and put prices do not coincide, as Figures 1 and 2 show. We analyse more in detail this situation in the appendix. In particular, we explain why, in this context, the implied volatility smile is decreasing with respect to strike for calls and increasing for puts.

4.2. Variance reduction for solving (V^{NL}, Z^{NL}) using (V^{BS}, Z^{BS})

In order to assess the impact of using $R > 0$ on the solution of the BSDE (3.5), in the case of European call and put options in one dimension, it is better to solve the BSDE difference $(V_t^{DF}, Z_t^{DF}) = (V_t^{NL} - V_t^{BS}, Z_t^{NL} - Z_t^{BS})$ which has a reduced variance in the algorithm. Note that for a call option

$$Z_t^{BS} = \sigma S_t \Phi(d_1), \quad d_{1,t} = \frac{\ln(S_t/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

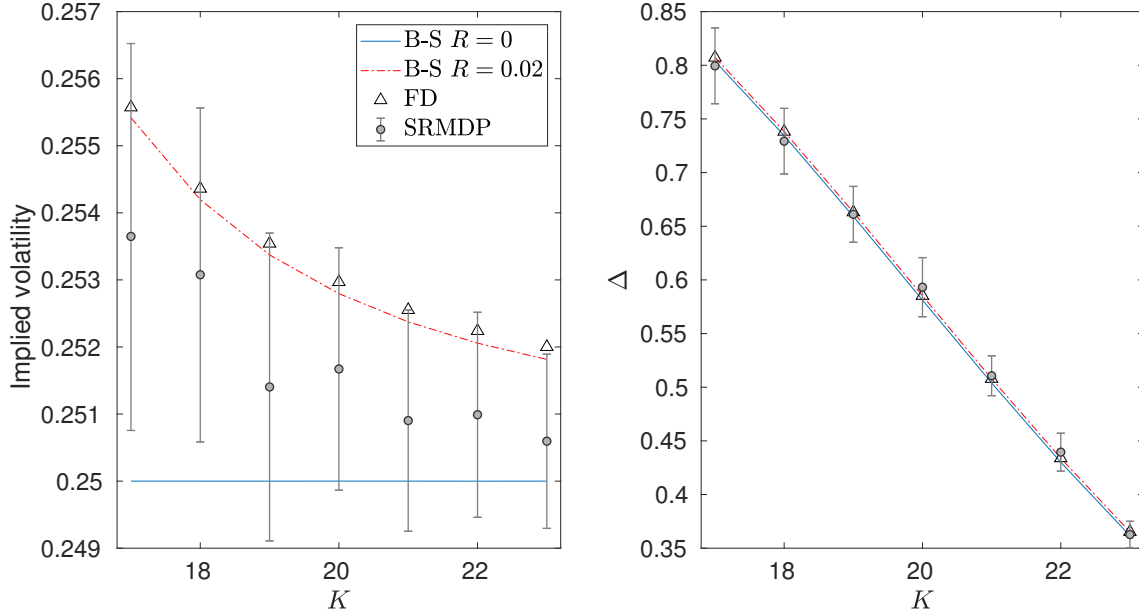


FIGURE 1. Implied volatility and delta for call options with spot value $S_0 = 20$ and different strikes K

K	B-S $R = 0$		B-S $R = 0.02$		FD	SRMDP	SRMDP	
	B-S(0, S_0)	$\Delta(0, S_0)$	B-S(0, S_0)	$\Delta(0, S_0)$			95% CI $V_0^{NL}(S_0)$	95% CI $Z_0^{NL}(S_0)/(\sigma S_0)$
17	3.9534	0.8037	3.9835	0.8073	3.9844	0.8072	[3.9575, 3.9897]	[0.7641, 0.8347]
18	3.2795	0.7345	3.3071	0.7383	3.3082	0.7382	[3.2833, 3.3161]	[0.6986, 0.7598]
19	2.6863	0.6592	2.7111	0.6631	2.7123	0.6630	[2.6797, 2.7134]	[0.6350, 0.6871]
20	2.1741	0.5812	2.1959	0.5852	2.1973	0.5852	[2.1730, 2.2012]	[0.5656, 0.6207]
21	1.7398	0.5039	1.7587	0.5079	1.7601	0.5078	[1.7338, 1.7601]	[0.4920, 0.5292]
22	1.3777	0.4301	1.3939	0.4338	1.3953	0.4338	[1.3734, 1.3975]	[0.4218, 0.4571]
23	1.0805	0.3617	1.0941	0.3651	1.0954	0.3652	[1.0752, 1.0946]	[0.3503, 0.3750]

TABLE 1. Price and delta for call options with spot value $S_0 = 20$ and different strikes K .

where Φ is the standard Gaussian cumulative distribution function. Therefore $|Z_t^{NL}| = |Z_t^{DF} + \sigma S_t \Phi(d_{1,t})|$. Then, the BSDE for the difference (V_t^{DF}, Z_t^{DF}) in the case of a call option³ is given by:

$$V_t^{DF} = 0 + \int_t^T \left(-rV_s^{DF} + C_\alpha R \sqrt{(s + \Delta) \wedge T - s} |Z_s^{DF} + \sigma S_s \Phi(d_{1,s})| \right) ds - \int_t^T Z_s^{DF} dW_s.$$

In Table 4, the BSDE (V_t^{DF}, Z_t^{DF}) is solved for several call and put options using the SRMDP algorithm. Besides, exact solutions (ES) are computed through the difference between Black-Scholes formula where IM's contribution is considered as a time-dependent dividend yield and the classical Black-Scholes formula with $R = 0$.

³For a put option an analogous BSDE can be written taking into account that $Z_t^{BS} = \sigma S_t (\Phi(d_{1,t}) - 1)$.

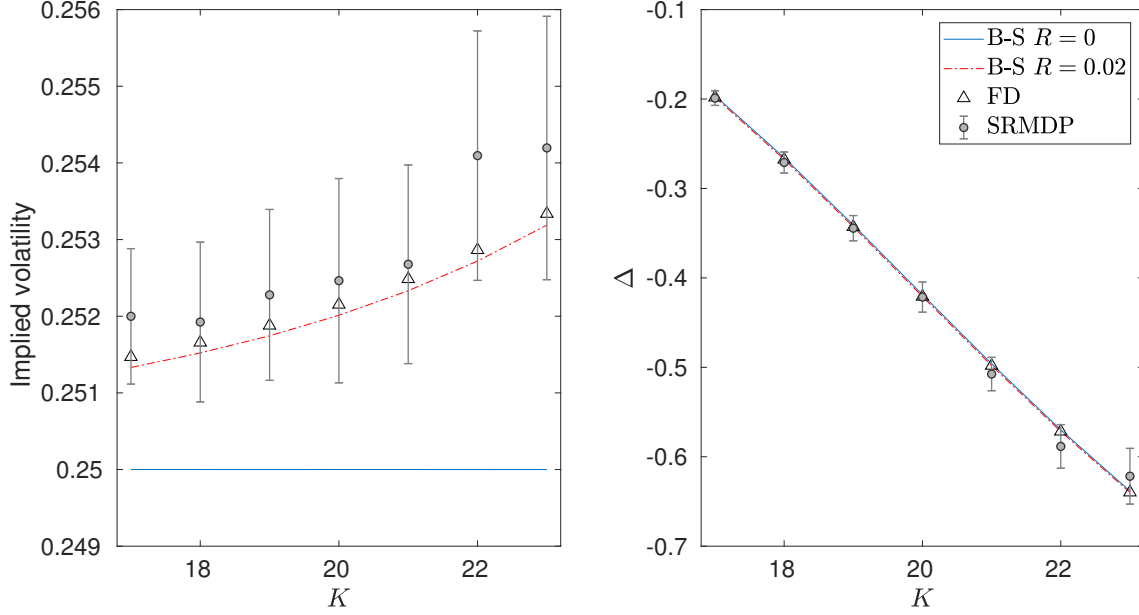


FIGURE 2. Implied volatility and delta for put options with spot value $S_0 = 20$ and different strikes K .

K	B-S $R = 0$		B-S $R = 0.02$		FD		SRMDP	
	B-S(0, S_0)	$\Delta(0, S_0)$	B-S(0, S_0)	$\Delta(0, S_0)$	$V(0, S_0)$	$\nabla V(0, S_0)$	95% CI $V_0^{NL}(S_0)$	95% CI $Z_0^{NL}(S_0)/(\sigma S_0)$
17	0.6168	-0.1963	0.6241	-0.1980	0.6249	-0.1981	[0.6229, 0.6328]	[-0.2071, -0.1908]
18	0.9231	-0.2655	0.9331	-0.2675	0.9340	-0.2676	[0.9289, 0.9426]	[-0.2827, -0.2592]
19	1.3101	-0.3408	1.3229	-0.3429	1.3239	-0.3430	[1.3186, 1.3350]	[-0.3585, -0.3303]
20	1.7781	-0.4188	1.7938	-0.4209	1.7949	-0.4209	[1.7869, 1.8077]	[-0.4383, -0.4046]
21	2.3240	-0.4961	2.3426	-0.4981	2.3438	-0.4981	[2.3350, 2.3557]	[-0.5262, -0.4888]
22	2.9421	-0.5699	2.9635	-0.5718	2.9646	-0.5717	[2.9615, 2.9871]	[-0.6128, -0.5641]
23	3.6251	-0.6383	3.6490	-0.6400	3.6501	-0.6398	[3.6436, 3.6695]	[-0.6529, -0.5906]

TABLE 2. Price and delta for put options with spot value $S_0 = 20$ and different strikes K .

Once again these tests allow us to demonstrate that SRMDP algorithm provides accurate results in one dimension.

4.3. Nested Monte Carlo for computing (V^L, Z^L) in dimension 1

As discussed in Section 3.2, we can further approximate the solution of the non-linear BSDE V^{NL} by the linear BSDE V^L with external source term $Z = Z^{BS}$. In this case, we have an explicit representation for Z_t^{BS} , given by

$$Z_t^{BS} = \partial_s v^{BS}(t, S_t) \sigma S_t,$$

where $v^{BS}(t, s) := \mathbb{E}[e^{-r(T-t)} \Phi(S_T) | S_t = s]$. We can use the likelihood ratio method of Broadie and Glasserman [BG96] (see also the automatic differentiation formula in [Nua06, Proposition 6.2.1]) to rewrite the derivative

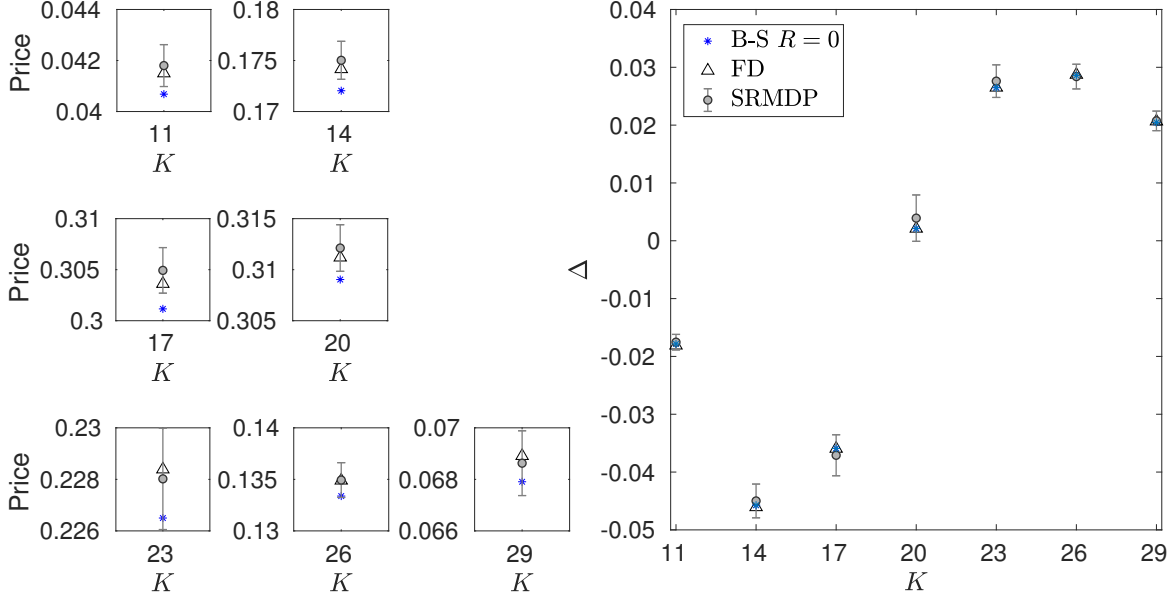


FIGURE 3. Price and delta for the butterfly option (4.6) with spot value $S_0 = 20$ and different strikes K .

K	B-S $R = 0$		FD		SRMDP	
	B-S(0, S_0)	$\Delta(0, S_0)$	$V(0, S_0)$	$\nabla V(0, S_0)$	95% CI $V_0^{NL}(S_0)$	95% CI $Z_0^{NL}(S_0)/(\sigma S_0)$
11	0.0407	-0.0178	0.0415	-0.0181	[0.0410, 0.0426]	[-0.0189, -0.0162]
14	0.1720	-0.0457	0.1742	-0.0461	[0.1732, 0.1769]	[-0.0479, -0.0421]
17	0.3012	-0.0359	0.3036	-0.0359	[0.3027, 0.3072]	[-0.0406, -0.0336]
20	0.3090	0.0021	0.3112	0.0021	[0.3098, 0.3144]	[-0.0001, 0.0079]
23	0.2265	0.0265	0.2284	0.0265	[0.2261, 0.2300]	[0.0248, 0.0304]
26	0.1334	0.0286	0.1349	0.0287	[0.1333, 0.1366]	[0.0263, 0.0305]
29	0.0679	0.0205	0.0689	0.0207	[0.0674, 0.0699]	[0.0190, 0.0224]

TABLE 3. Price and delta for the butterfly option (4.6) with spot value $S_0 = 20$ and different strikes K .

in terms of an expectation, getting

$$\begin{aligned}
 Z_t^{BS}(s) &= \frac{\partial}{\partial s} \mathbb{E}[e^{-r(T-t)} \Phi(S_T) | S_t = s] \sigma s = \mathbb{E} \left[e^{-r(T-t)} \Phi(S_T) \frac{W_T - W_t}{\sigma S_t (T-t)} \middle| S_t = s \right] \sigma s \\
 &= \mathbb{E} \left[e^{-r(T-t)} (\Phi(S_T) - \Phi(S_t)) \frac{W_T - W_t}{T-t} \middle| S_t = s \right]. \quad (4.7)
 \end{aligned}$$

The additional conditionally centered term $-\Phi(S_t) \frac{W_T - W_t}{T-t}$ that we have introduced in the last expression plays the role of a control variate, allowing to reduce the variance in the Monte-Carlo simulation. Recall that, in the linear BSDE (3.7) with source term Z^{BS} , we have $V_t = v^L(t, S_t)$, where v^L solves the PDE (4.1). The classical

	ES: B-S $R = 0.02$ – B-S $R = 0$		SRMDP	
	B-S(0, S_0)	$\Delta(0, S_0)$	95% CI	95% CI
			$V_0^{DF}(S_0)$	$Z_0^{DF}(S_0)/(\sigma S_0)$
Call, $K = 17$	0.0302	0.0036	[0.0302, 0.0304]	[0.0033, 0.0037]
Call, $K = 20$	0.0218	0.0040	[0.0218, 0.0219]	[0.0038, 0.0042]
Call, $K = 23$	0.0136	0.0035	[0.0136, 0.0137]	[0.0033, 0.0036]
Put, $K = 17$	0.0074	-0.0017	[0.0074, 0.0075]	[-0.0020, -0.0016]
Put, $K = 20$	0.0157	-0.0021	[0.0157, 0.0158]	[-0.0023, -0.0019]
Put, $K = 23$	0.0239	-0.0016	[0.0239, 0.0240]	[-0.0018, -0.0015]

TABLE 4. SRMDP algorithm for the BSDE (V_t^{DF}, Z_t^{DF}) .

Feynman-Kac representation of the solution to this PDE yields

$$V_0^L = \mathbb{E} \left[e^{-rT} \Phi(S_T) + \int_0^T e^{-rs} \left(R C_\alpha |Z_s^{BS}| \sqrt{(s + \Delta) \wedge T - s} \right) ds \right],$$

which we can rewrite as

$$V_0^L = \mathbb{E} \left[e^{-rT} \Phi(S_T) + T e^{-rU} R C_\alpha |Z_U^{BS}| \sqrt{(U + \Delta) \wedge T - U} \right], \quad (4.8)$$

where U is an independent random variable uniformly distributed on $[0, T]$. We introduce this additional randomization so to avoid any discretisation of the integral $\int_0^T ds$ – which would introduce a bias. Taking into account that Z_0^L is the derivative of V_0^L with respect to the risky asset S_0 times the diffusion coefficient of S computed at time 0, we apply again automatic differentiation formula (see [Nua06, Proposition 6.2.1]) on (4.8) to obtain

$$Z_0^L = \mathbb{E} \left[e^{-rT} \Phi(S_T) \frac{W_T}{T} + T e^{-rU} \left(R C_\alpha |Z_U^{BS}| \sqrt{(U + \Delta) \wedge T - U} \right) \frac{W_U}{U} \right]. \quad (4.9)$$

We solve the linear BSDE V^L for different payoffs (call, put and butterfly option) with a Nested Monte-Carlo algorithm (Nested MC), and compare again with finite difference method as in the previous sections. The Nested MC algorithm consists in approximating the outer expectation in (4.8) and in (4.9) with an empirical mean over M i.i.d. samples, and the inner expectation (4.7) with an empirical mean over N i.i.d. samples. More precisely, the Nested MC estimator for V_0^L is

$$\begin{aligned} \widehat{V}_0^L &= \frac{1}{M} \sum_{m=1}^M \left[e^{-rT} \Phi(S_0 e^{(r-\sigma^2/2)T + \sigma\sqrt{T}X^m}) \right. \\ &\left. + T e^{-rT} R C_\alpha \left| \frac{1}{N} \sum_{n=1}^N \left(\Phi \left(S_m e^{(r-\frac{\sigma^2}{2})(T-U^m) + \sigma\sqrt{T-U^m}Y^{m,n}} \right) - \Phi(S^m) \right) \frac{Y^{m,n}}{\sqrt{T-U^m}} \right| \sqrt{(U^m + \Delta) \wedge T - U^m} \right], \end{aligned}$$

where $S^m = S_0 e^{(r-\sigma^2/2)U^m + \sigma\sqrt{U^m}X^m}$, and $(X^m)_{m=1, \dots, M}$, $(Y^{m,n})_{m=1, \dots, M}^{n=1, \dots, N}$ are independent standard normal random variables. More details and an analysis of the bias and variance of analogous Nested MC estimators can be found in [GJ10] and [GH18]. Note that the problem is eventually non-linear, due to the presence of the absolute value function applied to the inner expectation. In our tests, we use $N = 100$ and $M = 100000$. The results are presented in Table 5. In this one-dimensional example, we observe a good agreement between the Nested MC algorithm and the finite difference solution, which we use as a benchmark. While the finite difference can typically hardly go beyond dimension $d = 2$ or $d = 3$, the Nested MC provides a competitive method to

estimate (V^L, Z^L) in higher dimensions. Moreover, comparing the values in Table 5 with the corresponding lines in Tables 1-2-3, one can observe that the non-linear price and delta V^{NL}, Z^{NL} are well approximated by their linear counterparts V^L, Z^L , as predicted by Theorem 3.1.

	FD		Nested MC	
	$V(0, S_0)$	$\nabla V(0, S_0)$	95% CI $V_0^L(S_0)$	95% CI $Z_0^L(S_0)/(\sigma S_0)$
Call, $K = 17$	3.9843	0.8072	[3.9796, 3.9856]	[0.8059, 0.8082]
Call, $K = 20$	2.1971	0.5852	[2.1931, 2.1979]	[0.5843, 0.5862]
Call, $K = 23$	1.0953	0.3653	[1.0924, 1.0958]	[0.3644, 0.3659]
Put, $K = 17$	0.6249	-0.1981	[0.6233, 0.6249]	[-0.1983, -0.1977]
Put, $K = 20$	1.7950	-0.4209	[1.7937, 1.7966]	[-0.4216, -0.4205]
Put, $K = 23$	3.6502	-0.6398	[3.6468, 3.6511]	[-0.6407, -0.6393]
Butterfly, $K = 11$	0.0414	-0.0181	[0.0412, 0.0415]	[-0.0181, -0.0180]
Butterfly, $K = 20$	0.3112	0.0021	[0.3112, 0.3119]	[0.0021, 0.0022]
Butterfly, $K = 29$	0.0689	0.0206	[0.0686, 0.0690]	[0.0206, 0.0207]

TABLE 5. Nested MC algorithm for the BSDE (V_t^L, Z_t^L) .

4.4. Basket options in higher dimensions

In this section we solve the non-linear BSDE in high dimensions using SRMDP algorithm. In this setting, traditional full grid methods like finite difference are not able to tackle the problem for dimension greater than 3.

We consider call option on a basket of d assets where the asset process is modelled by multi-dimensional geometric Brownian motion with constant correlation $\rho_{ij} = \rho = 0.75$ for $i \neq j$ and constant volatility $\sigma_0 = 0.25$. The full-rank volatility matrix σ in model (3.1) is then given by

$$\sigma\sigma^\top = \Sigma \text{ where } \Sigma := (\Sigma_{ij})_{1 \leq i, j \leq d} \text{ with } \Sigma_{ij} = \sigma_0^2 \rho, i \neq j \text{ and } \Sigma_{ii} = \sigma_0^2.$$

Then, $A_0 := \left(((\sigma^1 S_0^1)^\top, \dots, (\sigma^d S_0^d)^\top)^\top \right)^{-1}$ where σ^i is the i th row of σ . The payoff is given by

$$\Phi(S_T^1, \dots, S_T^d) = \left(\sum_i p^i S_T^i - K \right)^+.$$

The option expiration is set to $T = 1$ year and the interest rate $r = 0.02$. We suppose that weights $p_i = \frac{1}{d}$ for all i . The strike price K equals 20 and the initial values of the assets $S_0 = (S_0^1, \dots, S_0^d)$ are specified in Table 6. The rest of the model parameters are the same as earlier. In this table, we present prices and deltas for different basket options with several underlyings. In the first column, classical crude Monte Carlo values are shown (MC $R = 0$, IM was not considered). In the second column SRMDP values are displayed taking into account IM. In order to cope with the curse of dimensionality, here we used the **LP1** version of the SRMDP algorithm, with 500, 200, 50, 20 hypercubes and 1000, 1500, 2000, 2500 simulations per hypercube, for $d = 2, 3, 4, 5$ respectively, 5 time steps were considered. With respect to Table 6, in the absence of analytical solutions to compare with, we can only assert that Monte Carlo numerical results taking into account IM are close to those without IM, as seen before with the same Monte Carlo solvers executed over one dimensional problems.

S_0	MC ($R = 0$)		SRMDP ($R = 0.02$)	
	95% CI $V_0^{BS}(S_0)$	95% CI $Z_0^{BS}(S_0)A_0$	95% CI $V_0^{NL}(S_0)$	95% CI $Z_0^{NL}(S_0)A_0$
(18, 20)	[1.5102, 1.5113]	[-0.0685, -0.0682] [0.6237, 0.6245]	[1.5015, 1.5468]	[-0.0772, -0.0649] [0.6297, 0.6556]
(18, 20, 22)	[2.0067, 2.0081]	[-0.4676, -0.4671] [0.3813, 0.3817] [0.7435, 0.7443]	[1.9915, 2.0447]	[-0.4756, -0.4641] [0.3873, 0.4167] [0.7725, 0.7882]
(16, 18, 20, 22)	[1.4470, 1.4481]	[-0.6589, -0.6582] [0.1062, 0.1064] [0.4334, 0.4338] [0.6093, 0.6100]	[1.4677, 1.5090]	[-0.6689, -0.6182] [0.0962, 0.1374] [0.4234, 0.4628] [0.5943, 0.6310]
(16, 18, 20, 22, 24)	[1.9672, 1.9676]	[-1.0467, -1.0455] [-0.0855, -0.0852] [0.3342, 0.3347] [0.5655, 0.5662] [0.7039, 0.7047]	[1.9928, 2.0692]	[-1.0767, -1.0242] [-0.1155, -0.0752] [0.3042, 0.3467] [0.5355, 0.5762] [0.6839, 0.7167]

TABLE 6. Prices and deltas for the basket call option.

A. APPENDIX. ON THE IMPLIED VOLATILITY SMILES FROM CALL AND PUT OPTIONS WITH IM CORRECTION

We analyse more in detail the behavior of the implied volatility smiles obtained from the non-linear call and put option prices in Figures 1 and 2. As already pointed out at the end of Section 4.1, these prices do not satisfy put-call parity, which explains why the implied volatilities from calls and puts do not coincide.

Under the non-linear pricing rule V^{NL} , the price of a call option is

$$\begin{aligned}
V_0^{NL} = C(T, K) &:= \mathbb{E} \left[e^{-rT} \left(S_0 e^{rT - \int_0^T d^{\text{Call}}(t) dt + \sigma \sqrt{T} G - \frac{1}{2} \sigma^2 T} - K \right)^+ \right] \\
&= F^T(r, d^{\text{Call}}) e^{-rT} \mathbb{E} \left[\left(e^{\sigma \sqrt{T} G - \frac{1}{2} \sigma^2 T} - \frac{K}{F^T(r, d^{\text{Call}})} \right)^+ \right] \\
&:= F^T(r, d^{\text{Call}}) e^{-rT} C_{\text{BS}} \left(\frac{K}{F^T(r, d^{\text{Call}})}, \sigma \sqrt{T} \right)
\end{aligned}$$

where $F^T(r, d) = S_0 e^{rT - \int_0^T d(t) dt}$ is the usual forward price of S with interest rate r and dividend rate $d(t)$, and we denote $C_{\text{BS}}(X, v)$ the normalized BS call price with moneyness X and total implied volatility v (so that $C_{\text{BS}}(0, v) = 1$). In our case,

$$\begin{aligned}
d^{\text{Call}}(t) &= -C_\alpha R \sigma \sqrt{(t + \Delta) \wedge T - t}, \\
d^{\text{Put}}(t) &= C_\alpha R \sigma \sqrt{(t + \Delta) \wedge T - t}.
\end{aligned}$$

How do we imply the implied volatility σ^{Call} ? Using the same reference Black-Scholes price that leads to a constant smile when $R = 0$ (as in Figures 1-2): that is, we impose

$$\begin{aligned}
C(T, K) &= F^T(r, d^{\text{Call}}) e^{-rT} C_{\text{BS}} \left(\frac{K}{F^T(r, d^{\text{Call}})}, \sigma \sqrt{T} \right) \\
&= F^T(r, d = 0) e^{-rT} C_{\text{BS}} \left(\frac{K}{F^T(r, d = 0)}, \sigma^{\text{Call}}(K) \sqrt{T} \right) = \text{RefCallPrice}(T, K) \quad (\text{A.1})
\end{aligned}$$

and analogously for $\sigma^{\text{Put}}(\cdot)$. Equation (A.1) is of the form

$$C_{\text{BS}}\left(X, \sigma^{\text{Call}}(XF_1)\sqrt{T}\right) = \frac{F_2}{F_1} C_{\text{BS}}\left(X \frac{F_1}{F_2}, \sigma\sqrt{T}\right) = \mathbb{E}\left[\left(\frac{F_2}{F_1} e^{\sigma\sqrt{T} - \frac{1}{2}\sigma^2 T} - X\right)^+\right], \quad X \geq 0,$$

with $F_2 = F^T(r, d^{\text{Call}}) > F_1 = F^T(r, 0)$. This already shows that:

- (1) There is a smile: $\sigma^{\text{Call}}(\cdot)$ is not constant.
- (2) $\sigma^{\text{Call}}(K) \geq \sigma$ for all K , since $F_2 \geq F_1$.

The same conclusions (1) and (2) also hold for $\sigma^{\text{Put}}(\cdot)$, for which we have $d^{\text{Put}} \geq 0$, hence $F^T(r, d^{\text{Put}}) < F^T(r, 0)$.

Why are the implied volatilities of calls and puts with IM different? Because the reference price satisfies the following parity relation

$$\text{RefCallPrice}(T, K) - \text{RefPutPrice}(T, K) = e^{-rT} (F^T(r, 0) - K)$$

while the call and put prices with IM satisfy

$$\begin{aligned} C(T, K) - P(T, K) &= e^{-rT} \left[F(d^{\text{Call}}) C_{\text{BS}}(d^{\text{Call}}) - F(d^{\text{Put}}) P_{\text{BS}}(d^{\text{Put}}) \right] \\ &= e^{-rT} (F(d^{\text{Call}}) - K) \\ &\quad + e^{-rT} \left[F(d^{\text{Call}}) (P_{\text{BS}}(d^{\text{Call}}) - P_{\text{BS}}(d^{\text{Put}})) \right. \\ &\quad \left. + P_{\text{BS}}(d^{\text{Put}}) (F(d^{\text{Call}}) - F(d^{\text{Put}})) \right], \end{aligned}$$

where we have slightly simplified notations by setting $F(d^{\text{Call}}) = F^T(r, d^{\text{Call}})$, $F(d^{\text{Put}}) = F^T(r, d^{\text{Put}})$, and $C_{\text{BS}}(d) = C_{\text{BS}}\left(\frac{K}{F^T(r, d)}, \sigma\sqrt{T}\right)$, $P_{\text{BS}}(d) = P_{\text{BS}}\left(\frac{K}{F^T(r, d)}, \sigma\sqrt{T}\right)$. Note that the right hand side is not an affine function of K .

We can now assess the sign of the ATMF skew $\partial_K \sigma(K)|_{K=F^T(r, 0)}$. Taking derivatives with respect to K in (A.1), we get

$$\begin{aligned} \mathbb{P}\left(S_T^{r, d^{\text{Call}}, \sigma} > K\right) &= F^T(r, 0) \left[\partial_K C_{\text{BS}}\left(\frac{K}{F^T(r, 0)}, \sigma^{\text{Call}}(K)\sqrt{T}\right) + \partial_\sigma C_{\text{BS}}\left(\frac{K}{F^T(r, 0)}, \sigma^{\text{Call}}(K)\sqrt{T}\right) \partial_K \sigma^{\text{Call}}(K) \right] \\ &= -\mathbb{P}\left(S_T^{r, d=0, \sigma^{\text{Call}}(K)} > K\right) + K\sqrt{T} \phi\left(d_2\left(\frac{K}{F^T(r, 0)}, \sigma^{\text{Call}}(K)\sqrt{T}\right)\right) \partial_K \sigma^{\text{Call}}(K), \end{aligned}$$

or yet

$$\sqrt{T} \partial_K \sigma^{\text{Call}}(K) = \frac{-\mathbb{P}\left(S_T^{r, d^{\text{Call}}, \sigma} > K\right) + \mathbb{P}\left(S_T^{r, d=0, \sigma^{\text{Call}}(K)} > K\right)}{K \phi\left(d_2\left(\frac{K}{F^T(r, 0)}, \sigma^{\text{Call}}(K)\sqrt{T}\right)\right)} \quad (\text{A.2})$$

where $\phi(\cdot)$ is the standard normal density. Note that $\partial_K \sigma^{\text{Call}}(K)$ has the sign of the numerator $\text{Num}^{\text{CallSkew}}$ in (A.2). Setting

$$q(d, \sigma) := \mathbb{P}\left(S_T^{r, d, \sigma} > K\right) = N\left(d_2\left(\frac{K}{F^T(r, d)}, \sigma\sqrt{T}\right)\right) = N\left(\frac{\ln \frac{F^T(r, d)}{K}}{\sigma\sqrt{T}} - \frac{1}{2}\sigma\sqrt{T}\right)$$

we have

$$\partial_\sigma q(d, \sigma) = \sqrt{T} \phi(d_2) \left(-\frac{\ln \frac{F^T(r, d)}{K}}{\sigma^2 T} - \frac{1}{2} \right), \quad \partial_d q(d, \sigma) = \frac{1}{\sigma \sqrt{T}} \phi(d_2) \frac{\partial_d F^T(r, d)}{F^T(r, d)} = -\frac{T}{\sigma \sqrt{T}} \phi(d_2)$$

the last derivative being taken in the case of a constant dividend rate d (which is our case up to time $t \leq T - \Delta$). We can write the numerator in (A.2) as

$$\begin{aligned} \text{Num}^{\text{CallSkew}} &= q(d = 0, \sigma^{\text{Call}}(K)) - q(d^{\text{Call}}, \sigma) \\ &= q(d = 0, \sigma^{\text{Call}}(K)) - q(d = 0, \sigma) + q(d = 0, \sigma) - q(d^{\text{Call}}, \sigma) =: \text{Num}_1^{\text{CallSkew}} + \text{Num}_2^{\text{CallSkew}}. \end{aligned}$$

We now consider the ATMF point, that is $K = F^T(r, 0)$. Since, in this case,

$$\partial_\sigma q(d = 0, \sigma) = -\frac{1}{2} \sqrt{T} \phi(d_2) < 0,$$

we have $\text{Num}_1^{\text{CallSkew}} < 0$, recalling that $\sigma^{\text{Call}}(K) > \sigma$. Moreover, since $\partial_d q(d, \sigma) < 0$, we also have $\text{Num}_2^{\text{CallSkew}} < 0$, recalling that $0 > d^{\text{Call}}$. Overall, we have obtained

$$\sqrt{T} \partial_K \sigma^{\text{Call}}(K)|_{K=F^T(r, 0)} < 0.$$

REFERENCES

- [ADEH99] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.
- [AP05] Y. Achdou and O. Pironneau. *Computational Methods for Option Pricing*. SIAM series, Frontiers in Applied Mathematics, Philadelphia, 2005.
- [Bas15] Basel Committee on Banking Supervision. Margin requirements for non-centrally-cleared derivatives. <http://www.bis.org/bcbs/publ/d317.pdf>, 2015.
- [BG96] M. Broadie and P. Glasserman. Estimating security price derivatives using simulation. *Management science*, 42(2):269–285, 1996.
- [BI09] R. Buckdahn and P. Imkeller. Backward stochastic differential equations with time delayed generator. *preprint*, 2009.
- [CBB14] S. Crépey, T. Bielecki, and D. Brigo. *Counterparty risk and funding: A tale of two puzzles*. Chapman & Hall CRC, 2014.
- [CESS17] S. Crépey, R. Élie, W. Sabbagh, and S. Song. When capital is a funding source: The XVA anticipated BSDEs. *preprint*, 2017.
- [CN17] P. Cheredito and K. Nam. BSE’s, BSDE’s and fixed point problems. *The Annals of Probability*, 45(6A):3795–3828, 2017.
- [DI09] L. Delong and P. Imkeller. On Malliavin’s differentiability of time delayed BSDEs driven by Brownian motions and Poisson random measures. *preprint*, 2009.
- [DI10] L. Delong and P. Imkeller. Backward Stochastic Differential Equations with time delayed generators - Results and counterexamples. *The Annals of Probability*, 20(4):1512–1536, 2010.
- [DI12] L. Delong and P. Imkeller. Applications of time-delayed backward stochastic differential equations to pricing, hedging and portfolio management in insurance and finance. *Applicationes Mathematicae*, 39:463–488, 2012.
- [EPQ97] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7:1–71, 1997.
- [GH18] M. Giles and A.-L. Haji-Ali. Multilevel nested simulation for efficient risk estimation. *arXiv:1802.05016*, 2018.
- [GHS00] P. Glasserman, P. Heidelberger, and P. Shahabuddin. Variance reduction techniques for estimating Value-at-Risk. *Management Science*, 46:1349–1364, 2000.
- [GJ10] M. B. Gordy and S. Juneja. Nested simulation in portfolio risk measurement. *Management Science*, 56(10):1833–1848, 2010.
- [GLTV16] E. Gobet, J. G. López-Salas, P. Turkedjiev, and C. Vázquez. Stratified Regression Monte-Carlo Scheme for Semilinear PDEs and BSDEs with Large Scale Parallelization on GPUs. *SIAM Journal on Scientific Computing*, 38(6):C652–C677, 2016.
- [GP15] E. Gobet and S. Pagliarani. Analytical approximations of BSDEs with nonsmooth driver. *SIAM Journal on Financial Mathematics*, 6(1):919–958, 2015.

- [GT16] E. Gobet and P. Turkedjiev. Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Math. Comp.*, 85(299):1359–1391, 2016.
- [KS98] I. Karatzas and S.E. Shreve. *Methods of mathematical finance*. Springer-Verlag, New York, 1998.
- [McK66] H.P. McKean. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA*, 56:1907–1911, 1966.
- [MR05] M. Musiela and M. Rutkowski. *Martingale methods in financial modelling*. Springer Verlag, second edition, 2005.
- [NR16] T. Nie and M. Rutkowski. A BSDE approach to fair bilateral pricing under endogenous collateralization. *Finance and Stochastics*, 20(4):855–900, Oct 2016.
- [Nua06] D. Nualart. *The Malliavin Calculus and Related Topics*. Probability and Its Applications. Springer Verlag, second edition, 2006.
- [PP90] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equations. *Systems & Control Letters*, 14(1):55–61, 1990.
- [PP92] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equation. *Lecture Notes in CIS, Springer-Verlag*, 176:200–217, 1992.
- [PR14] E. Pardoux and A. Rascanu. *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*, volume 69 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, 2014.
- [PY09] S. Peng and Z. Yang. Anticipated backward stochastic differential equations. *The Annals of Probability*, 37(3):877–902, 2009.
- [RU00] R.T. Rockafellar and S. Uryasev. Optimization of conditional Value-at-Risk. *Journal of risk*, 2:21–42, 2000.
- [YE13] Z. Yang and R.J. Elliott. Anticipated backward stochastic differential equations with continuous coefficients. *Communications on Stochastic Analysis*, 7(2):303–319, 2013.
- [YG73] D.M. Young and R.T. Gregory. *A survey on Numerical Mathematics*, volume 1. Dover Publications, 1973.
- [Zha04] J. Zhang. A numerical scheme for BSDEs. *Ann. Appl. Probab.*, 14(1):459–488, 02 2004.