Nonlinear Monte Carlo Schemes for Counterparty Risk on Credit Derivatives

Stéphane Crépey and Tuyet Mai Nguyen

Abstract Two nonlinear Monte Carlo schemes, namely, the linear Monte Carlo expansion with randomization of Fujii and Takahashi (Int J Theor Appl Financ 15(5):1250034(24), 2012 [9], Q J Financ 2(3), 1250015(24), 2012, [10]) and the marked branching diffusion scheme of Henry-Labordère (Risk Mag 25(7), 67–73, 2012, [13]), are compared in terms of applicability and numerical behavior regarding counterparty risk computations on credit derivatives. This is done in two dynamic copula models of portfolio credit risk: the dynamic Gaussian copula model and the model in which default dependence stems from joint defaults. For such high-dimensional and nonlinear pricing problems, more standard deterministic or simulation/regression schemes are ruled out by Bellman's "curse of dimensionality" and only purely forward Monte Carlo schemes can be used.

1 Introduction

Counterparty risk is a major issue since the global credit crisis and the ongoing European sovereign debt crisis. In a bilateral counterparty risk setup, counterparty risk is valued as the so-called credit valuation adjustment (CVA), for the risk of default of the counterparty, and debt valuation adjustment (DVA), for own default risk. In such a setup, the classical assumption of a locally risk-free funding asset used for both investing and unsecured borrowing is no longer sustainable. The proper accounting of the funding costs of a position leads to the funding valuation adjustment (FVA). Moreover, these adjustments are interdependent and must be computed jointly

S. Crépey (✉) · T.M. Nguyen
Laboratoire de Mathématiques et Modélisation, Université d'Évry
Val d'Essonne, 91037 Évry Cedex, France
e-mail: stephane.crepey@univ-evry.fr

T.M. Nguyen
e-mail: tuyetmai.nguyen@univ-evry.fr

through a global correction dubbed total valuation adjustment (TVA). The pricing equation for the TVA is nonlinear due to the funding costs. It is posed over a random time interval determined by the first default time of the two counterparties. To deal with the corresponding backward stochastic differential equation (BSDE), a first reduced-form modeling approach has been proposed in Crépey [3], under a rather standard immersion hypothesis between a reference (or market) filtration and the full model filtration progressively enlarged by the default times of the counterparties. This basic immersion setup is fine for standard applications, such as counterparty risk on interest rate derivatives. But it is too restrictive for situations of strong dependence between the underlying exposure and the default risk of the two counterparties, such as counterparty risk on credit derivatives, which involves strong adverse dependence, called wrong-way risk (for some insights of related financial contexts, see Fujii and Takahashi [11], Brigo et al. [2]). For this reason, an extended reduced-form modeling approach has been recently developed in Crépey and Song [4–6]. With credit derivatives, the problem is also very high-dimensional. From a numerical point of view, for high-dimensional nonlinear problems, only purely forward simulation schemes can be used. In Crépey and Song [6], the problem is addressed by the linear Monte Carlo expansion with randomization of Fujii and Takahashi [9, 10].

In the present work, we assess another scheme, namely the marked branching diffusion approach of Henry-Labordère [13], which we compare with the previous one in terms of applicability and numerical behavior. This is done in two dynamic copula models of portfolio credit risk: the dynamic Gaussian copula model and the dynamic Marshall–Olkin model in which default dependence stems from joint defaults.

The paper is organized as follows. Sections 2 and 3 provide a summary of the main pricing and TVA BSDEs that are derived in Crépey and Song [4–6]. Section 4 exposes two nonlinear Monte Carlo schemes that can be considered for solving these in high-dimensional models, such as the portfolio credit models of Sect. 5. Comparative numerics in these models are presented in Sect. 6. Section 7 concludes.

2 Prices

2.1 Setup

We consider a netted portfolio of OTC derivatives between two defaultable counterparties, generally referred to as the contract between a bank, the perspective of which is taken, and its counterparty. After having bought the contract from its counterparty at time 0, the bank sets up a hedging, collateralization (or margining), and funding portfolio. We call the funder of the bank a third party, possibly composed in practice of several entities or devices, insuring funding of the bank's strategy. The funder, assumed default-free for simplicity, plays the role of lender/borrower of last resort after the exhaustion of the internal sources of funding provided to the bank through its hedge and collateral.

For notational simplicity we assume no collateralization. All the numerical considerations, our main focus in this work, can be readily extended to the case of collateralized portfolios using the corresponding developments in Crépey and Song [6]. Likewise, we assume hedging in the simplest sense of replication by the bank and we consider the case of a fully securely funded hedge, so that the cost of the hedge of the bank is exactly reflected by the wealth of its hedging and funding portfolio.

We consider a stochastic basis $(\Omega, \mathscr{G}_T, \mathscr{G}, \mathbb{Q})$, where $\mathscr{G} = (\mathscr{G}_t)_{t \in [0,T]}$ is interpreted as a risk-neutral pricing model on the primary market of the instruments that are used by the bank for hedging its TVA. The reference filtration $\mathscr{F}$ is a subfiltration of $\mathscr{G}$ representing the counterparty risk-free filtration, not carrying any direct information about the defaults of the two counterparties. The relation between these two filtrations will be pointed out in the condition (C) introduced later. We denote by:

- $\mathbb{E}_t$, the conditional expectation under $\mathbb{Q}$ given $\mathscr{G}_t$,
- r, the risk-free short rate process, with related discount factor $\beta_t = e^{-\int_0^t r_s ds}$,
- T, the maturity of the contract,
- τ_b and τ_c, the default time of the bank and of the counterparty, modeled as $\mathscr{G}$ stopping times with $(\mathscr{G}, \mathbb{Q})$ intensities γ^b and γ^c,
- $\tau = \tau_b \wedge \tau_c$, the first-to-default time of the two counterparties, also a $\mathscr{G}$ stopping time, with intensity γ such that $\max(\gamma^b, \gamma^c) \leq \gamma \leq \gamma^b + \gamma^c$,
- $\bar{\tau} = \tau \wedge T$, the effective time horizon of our problem (there is no cashflow after $\bar{\tau}$),
- D, the contractual dividend process,
- $\Delta = D - D_-$, the jump process of D.

2.2 Clean Price

We denote by P the reference (or clean) price of the contract ignoring counterparty risk and assuming the position of the bank financed at the risk-free rate r, i.e. the $\mathscr{G}$ conditional expectation of the future contractual cash-flows discounted at the risk-free rate r. In particular,

$$\beta_t P_t = \mathbb{E}_t \left[\int_t^{\bar{\tau}} \beta_s dD_s + \beta_{\bar{\tau}} P_{\bar{\tau}} \right], \ \forall t \in [0, \bar{\tau}]. \tag{1}$$

We also define $Q_t = P_t + \mathbb{1}_{\{t = \tau < T\}} \Delta_\tau$, so that Q_τ represents the clean value of the contract inclusive of the promised dividend at default (if any) Δ_τ, which also belongs to the "debt" of the counterparty to the bank (or vice versa depending on the sign of Q_τ) in case of default of a party. Accordingly, at time τ (if $< T$), the close-out cash-flow of the counterparty to the bank is modeled as

$$\mathscr{R} = \mathbb{1}_{\{\tau = \tau_c\}} \left(R_c Q_\tau^+ - Q_\tau^- \right) - \mathbb{1}_{\{\tau = \tau_b\}} \left(R_b Q_\tau^- - Q_\tau^+ \right) - \mathbb{1}_{\{\tau_b = \tau_c\}} Q_\tau, \tag{2}$$

where R_b and R_c are the recovery rates of the bank and of the counterparty to each other.

2.3 All-Inclusive Price

Let Π be the all-inclusive price of the contract for the bank, including the cost of counterparty risk and funding costs. Since we assume a securely funded hedge (in the sense of replication) and no collateralization, the amounts invested and funded by the bank at time t are respectively given by Π_t^- and Π_t^+. The all-inclusive price Π is the discounted conditional expectation of all effective future cash flows including the contractual dividends before τ, the cost of funding the position prior to time τ and the terminal cash flow at time τ. Hence,

$$\beta_t \Pi_t = \mathbb{E}_t \left[\int_t^{\bar{\tau}} \beta_s \mathbb{1}_{s < \tau} dD_s - \int_t^{\bar{\tau}} \beta_s \bar{\lambda}_s \Pi_s^+ ds + \beta_{\bar{\tau}} \mathbb{1}_{\tau < T} \mathscr{R} \right], \tag{3}$$

where $\bar{\lambda}$ is the funding spread over r of the bank toward the external funder, i.e. the bank borrows cash from its funder at rate $r + \bar{\lambda}$ (and invests cash at the risk-free rate r). Since the right hand side in (3) depends also on Π, (3) is in fact a backward stochastic differential equation (BSDE). Consistent with the no arbitrage principle, the gain process on the hedge is a $\mathbb{Q}$ martingale, which explains why it does not appear in (3).

3 TVA BSDEs

The total valuation adjustment (TVA) process Θ is defined as

$$\Theta = Q - \Pi. \tag{4}$$

In this section we review the main TVA BSDEs that are derived in Crépey and Song [4–6]. Three BSDEs are presented. These three equations are essentially equivalent mathematically. However, depending on the underlying model, they are not always amenable to the same numerical schemes or the numerical performance of a given scheme may differ between them.

3.1 Full TVA BSDE

By taking the difference between (1) and (3), we obtain

$$\beta_t \Theta_t = \mathbb{E}_t \left[\int_t^{\bar{\tau}} \beta_s fva_s(\Theta_s) ds + \beta_{\bar{\tau}} \mathbb{1}_{\tau < T} \xi \right], \ \forall t \in [0, \bar{\tau}], \tag{5}$$

where $fva_t(\vartheta) = \bar{\lambda}_t(P_t - \vartheta)^+$ is the funding coefficient and where

$$\xi = Q_\tau - \mathscr{R} = \mathbb{1}_{\{\tau=\tau_c\}}(1 - R_c)(P_\tau + \Delta_\tau)^+ - \mathbb{1}_{\{\tau=\tau_b\}}(1 - R_b)(P_\tau + \Delta_\tau)^- \quad (6)$$

is the exposure at default of the bank. Equivalent to (5), the "full TVA BSDE" is written as

$$\Theta_t = \mathbb{E}_t\left[\int_t^{\bar{\tau}} f_s(\Theta_s)ds + \mathbb{1}_{\tau<T}\xi\right], \ 0 \le t \le \bar{\tau}, \quad (\mathrm{I})$$

for the coefficient $f_t(\vartheta) = fva_t(\vartheta) - r_t\vartheta$.

3.2 Partially Reduced TVA BSDE

Let $\hat{\xi}$ be a $\mathscr{G}$-predictable process, which exists by Corollary 3.23 2 in He et al. [12], such that $\hat{\xi}_\tau = \mathbb{E}[\xi|\mathscr{G}_{\tau-}]$ on $\tau < \infty$ and let $\bar{f}$ be the modified coefficient such that

$$\bar{f}_t(\vartheta) + r_t\vartheta = \underbrace{\gamma_t\hat{\xi}_t}_{cdva_t} + \underbrace{\bar{\lambda}_t(P_t - \vartheta)^+}_{fva_t(\vartheta)}. \quad (7)$$

As easily shown (cf. [4, Lemma 2.2]), the full TVA BSDE (I) can be simplified into the "partially reduced BSDE"

$$\bar{\Theta}_t = \mathbb{E}_t\left[\int_t^{\bar{\tau}} \bar{f}_s(\bar{\Theta}_s)ds\right], \ 0 \le t \le \bar{\tau}, \quad (\mathrm{II})$$

in the sense that if Θ solves (I), then $\bar{\Theta} = \Theta\mathbb{1}_{[0,\tau)}$ solves (II), while if $\bar{\Theta}$ solves (II), then the process Θ defined as $\bar{\Theta}$ before $\bar{\tau}$ and $\Theta_{\bar{\tau}} = \mathbb{1}_{\tau<T}\xi$ solves (I). Note that both BSDEs (I) and (II) are $(\mathscr{G}, \mathbb{Q})$ BSDEs posed over the random time interval $[0, \bar{\tau}]$, but with the terminal condition ξ for (I) as opposed to a null terminal condition (and a modified coefficient) for (II).

3.3 Fully Reduced TVA BSDE

Let

$$\hat{f}_t(\vartheta) = \bar{f}_t(\vartheta) - \gamma_t\vartheta = cdva_t + fva_t(\vartheta) - (r_t + \gamma_t)\vartheta.$$

Assume the following conditions, which are studied in Crépey and Song [4–6]:

Condition (C). There exist:

(C.1) a subfiltration $\mathscr{F}$ of $\mathscr{G}$ satisfying the usual conditions and such that $\mathscr{F}$ semi-martingales stopped at τ are $\mathscr{G}$ semimartingales,

(C.2) a probability measure $\mathbb{P}$ equivalent to $\mathbb{Q}$ on $\mathscr{F}_T$ such that any $(\mathscr{F}, \mathbb{P})$ local martingale stopped at $(\tau-)$ is a $(\mathscr{G}, \mathbb{Q})$ local martingale on $[0, T]$,

(C.3) an $\mathscr{F}$ progressive "reduction" $\widetilde{f}_t(\vartheta)$ of $\hat{f}_t(\vartheta)$ such that $\int_0^{\cdot} \hat{f}_t(\vartheta)dt = \int_0^{\cdot} \widetilde{f}_t(\vartheta)dt$ on $[0, \bar{\tau}]$.

Let $\widetilde{\mathbb{E}}_t$ denote the conditional expectation under $\mathbb{P}$ given $\mathscr{F}_t$. It is shown in Crépey and Song [4–6]) that the full TVA BSDE (I) is equivalent to the following "fully reduced BSDE":

$$\widetilde{\Theta}_t = \widetilde{\mathbb{E}}_t \left[\int_t^T \widetilde{f}_s(\widetilde{\Theta}_s)ds \right], \quad t \in [0, T], \tag{III}$$

equivalent in the sense that if Θ solves (I), then the "$\mathscr{F}$ optional reduction" $\widetilde{\Theta}$ of Θ ($\mathscr{F}$ optional process that coincides with Θ before τ) solves (III), while if $\widetilde{\Theta}$ solves (III), then $\Theta = \widetilde{\Theta} \mathbb{1}_{[0,\tau)} + \mathbb{1}_{[\tau]} \mathbb{1}_{\tau < T} \xi$ solves (I).

Moreover, under mild assumptions (see e.g. Crépey and Song [6, Theorem 4.1]), one can easily check that $\bar{f}_t(\vartheta)$ in (7) (resp. $\widetilde{f}_t(\vartheta)$) satisfies the classical BSDE monotonicity assumption

$$\big(\bar{f}_t(\vartheta) - \bar{f}_t(\vartheta')\big)(\vartheta - \vartheta') \leq C(\vartheta - \vartheta')^2$$

(and likewise for $\widetilde{f}$), for some constant C. Hence, by classical BSDE results nicely surveyed in Kruse and Popier [14, Sect. 2 (resp. 3)], the partially reduced TVA BSDE (II), hence the equivalent full TVA BSDE (I) (resp. the fully reduced BSDE (III)), is well-posed in the space of $(\mathscr{G}, \mathbb{Q})$ (resp. $(\mathscr{F}, \mathbb{P})$) square integrable solutions, where well-posedness includes existence, uniqueness, comparison and BSDE standard estimates.

3.4 Marked Default Time Setup

In order to be able to compute $\gamma \hat{\xi}$ in $\bar{f}$, we assume that τ is endowed with a mark e in a finite set E, in the sense that

$$\tau = \min_{e \in E} \tau_e, \tag{8}$$

where each τ_e is a stopping time with intensity γ_t^e such that $\mathbb{Q}(\tau_e \neq \tau_{e'}) = 1$, $e \neq e'$, and

$$\mathscr{G}_\tau = \mathscr{G}_{\tau-} \vee \sigma(\varepsilon),$$

where $\varepsilon = \mathrm{argmin}_{e \in E} \tau_e$ yields the "identity" of the mark. The role of the mark is to convey some additional information about the default, e.g. to encode wrong-way and gap risk features. The assumption of a finite set E in (8) ensures tractability of the setup. In fact, by Lemma 5.1 in Crépey and Song [6], there exists $\mathscr{G}$-predictable processes $\widetilde{P}_t^e$ and $\widetilde{\Delta}_t^e$ such that

$$P_\tau = \widetilde{P}_\tau^e \text{ and } \Delta_\tau = \widetilde{\Delta}_\tau^e \text{ on the event } \{\tau = \tau_e\}.$$

Assuming further that $\tau_b = \min_{e \in E_b} \tau_e$ and $\tau_c = \min_{e \in E_c} \tau_e$, where $E = E_b \cup E_c$ (not necessarily a disjoint union), one can then take on $[0, \bar{\tau}]$:

$$\gamma_t \hat{\xi}_t = (1 - R_c) \sum_{e \in E_c} \gamma_t^e \left(\widetilde{P}_t^e + \widetilde{\Delta}_t^e\right)^+ - (1 - R_b) \sum_{e \in E_b} \gamma_t^e \left(\widetilde{P}_t^e + \widetilde{\Delta}_t^e\right)^-,$$

where the two terms have clear respective CVA and DVA interpretation. Hence, (7) is rewritten, on $[0, \bar{\tau}]$, as

$$\bar{f}_t(\vartheta) + r_t \vartheta = \underbrace{(1 - R_c) \sum_{e \in E_c} \gamma_t^e \left(\widetilde{P}_t^e + \widetilde{\Delta}_t^e\right)^+}_{\text{CVA coefficient } (cva_t)} - \underbrace{(1 - R_b) \sum_{e \in E_b} \gamma_t^e \left(\widetilde{P}_t^e + \widetilde{\Delta}_t^e\right)^-}_{\text{DVA coefficient } (dva_t)} \qquad (9)$$
$$+ \underbrace{\bar{\lambda}_t (P_t - \vartheta)^+}_{\text{FVA coefficient } (fva_t(\vartheta))}.$$

If the functions $\widetilde{P}_t^e$ and $\widetilde{\Delta}_t^e$ above not only exist, but can be computed explicitly (as will be the case in the concrete models of Sects. 5.1 and 5.2), once stated in a Markov setup where

$$\bar{f}_t(\vartheta) = \bar{f}(t, X_t, \vartheta), \ t \in [0, T], \qquad (10)$$

for some $(\mathscr{G}, \mathbb{Q})$ jump diffusion X, then the partially reduced TVA BSDE (II) can be tackled numerically. Similarly, once stated in a Markov setup where

$$\widetilde{f}_t(\vartheta) = \widetilde{f}(t, \widetilde{X}_t, \vartheta), \ t \in [0, T], \qquad (11)$$

for some $(\mathscr{F}, \mathbb{P})$ jump diffusion $\widetilde{X}$, then the fully reduced TVA BSDE (III) can be tackled numerically.

4 TVA Numerical Schemes

4.1 Linear Approximation

Our first TVA approximation is obtained replacing Θ_s by 0 in the right hand side of
(I), i.e.

$$\Theta_0 \approx \mathbb{E}\left[\int_0^{\bar{\tau}} f_s(0)ds + \mathbb{1}_{\tau<T}\xi\right] = \mathbb{E}\left[\int_0^{\bar{\tau}} \bar{\lambda}_s P_s^+ ds + \mathbb{1}_{\tau<T}\xi\right]. \tag{12}$$

We then approximate the TVA by standard Monte-Carlo, with randomization of
the integral to reduce the computation time (at the cost of a small increase in the
variance). Hence, introducing an exponential time ζ of parameter μ, i.e. a random
variable with density $\phi(s) = \mathbb{1}_{s\geq 0}\,\mu\,e^{-\mu s}$, we have

$$\mathbb{E}\left[\int_0^{\bar{\tau}} f_s(0)ds\right] = \mathbb{E}\left[\int_0^{\bar{\tau}} \phi(s)\frac{1}{\mu}e^{\mu s}f_s(0)ds\right] = \mathbb{E}\left[\mathbb{1}_{\zeta<\bar{\tau}}\frac{e^{\mu\zeta}}{\mu}f_\zeta(0)\right]. \tag{13}$$

We can use the same technic for (II) and (III), which yields:

$$\Theta_0 = \bar{\Theta}_0 \approx \mathbb{E}\left[\int_0^{\bar{\tau}} \bar{f}_s(0)ds\right] = \mathbb{E}\left[\mathbb{1}_{\zeta<\bar{\tau}}\frac{e^{\mu\zeta}}{\mu}\bar{f}_\zeta(0)\right], \tag{14}$$

$$\Theta_0 = \widetilde{\Theta}_0 \approx \widetilde{\mathbb{E}}\left[\int_0^{T} \widetilde{f}_s(0)ds\right] = \widetilde{\mathbb{E}}\left[\mathbb{1}_{\zeta<T}\frac{e^{\mu\zeta}}{\mu}\widetilde{f}_\zeta(0)\right]. \tag{15}$$

4.2 Linear Expansion and Interacting Particle
Implementation

Following Fujii and Takahashi [9, 10], we can introduce a perturbation parameter ε
and the following perturbed form of the fully reduced BSDE (III):

$$\widetilde{\Theta}_t^\varepsilon = \widetilde{\mathbb{E}}_t\left[\int_t^T \varepsilon\widetilde{f}_s(\widetilde{\Theta}_s^\varepsilon)ds\right], \quad t \in [0, T], \tag{16}$$

where $\varepsilon = 1$ corresponds to the original BSDE (III). Suppose that the solution of
(16) can be expanded in a power series of ε:

$$\widetilde{\Theta}_t^\varepsilon = \widetilde{\Theta}_t^{(0)} + \varepsilon\widetilde{\Theta}_t^{(1)} + \varepsilon^2\widetilde{\Theta}_t^{(2)} + \varepsilon^3\widetilde{\Theta}_t^{(3)} + \cdots. \tag{17}$$

The Taylor expansion of f at $\widetilde{\Theta}^{(0)}$ reads

$$\widetilde{f}_t(\widetilde{\Theta}_t^\varepsilon) = \widetilde{f}_t(\widetilde{\Theta}_t^{(0)}) + (\varepsilon \widetilde{\Theta}_t^{(1)} + \varepsilon^2 \widetilde{\Theta}_t^{(2)} + \cdots) \partial_\vartheta \widetilde{f}_t(\widetilde{\Theta}_t^{(0)})$$
$$+ \frac{1}{2}(\varepsilon \widetilde{\Theta}_t^{(1)} + \varepsilon^2 \widetilde{\Theta}_t^{(2)} + \cdots)^2 \partial_{\vartheta^2}^2 \widetilde{f}_t(\widetilde{\Theta}_t^{(0)}) + \cdots$$

Collecting the terms of the same order with respect to ε in (16), we obtain $\widetilde{\Theta}_t^{(0)} = 0$, due to the null terminal condition of the fully reduced BSDE (III), and

$$\widetilde{\Theta}_t^{(1)} = \widetilde{\mathbb{E}}_t \left[\int_t^T \widetilde{f}_s(\widetilde{\Theta}_s^{(0)}) ds \right],$$
$$\widetilde{\Theta}_t^{(2)} = \widetilde{\mathbb{E}}_t \left[\int_t^T \widetilde{\Theta}_s^{(1)} \partial_\vartheta \widetilde{f}_s(\widetilde{\Theta}_s^{(0)}) ds \right], \tag{18}$$
$$\widetilde{\Theta}_t^{(3)} = \widetilde{\mathbb{E}}_t \left[\int_t^T \widetilde{\Theta}_s^{(2)} \partial_\vartheta \widetilde{f}_s(\widetilde{\Theta}_s^{(0)}) ds \right],$$

where the third order term should contain another component based on $\partial_{\vartheta^2}^2 \widetilde{f}$. But, in our case, $\partial_{\vartheta^2}^2 \widetilde{f}$ involves a Dirac measure via the terms $(P_t - \vartheta)^+$ in $fva_t(\vartheta)$, so that we truncate the expansion to the term $\widetilde{\Theta}_t^{(3)}$ as above. If the nonlinearity in (III) is sub-dominant, one can expect to obtain a reasonable approximation of the original equation by setting $\varepsilon = 1$ at the end of the calculation, i.e.

$$\widetilde{\Theta}_0 \approx \widetilde{\Theta}_0^{(1)} + \widetilde{\Theta}_0^{(2)} + \widetilde{\Theta}_0^{(3)}.$$

Carrying out a Monte Carlo simulation by an Euler scheme for every time s in a time grid and integrating to obtain $\widetilde{\Theta}_0^{(1)}$ would be quite heavy. Moreover, this would become completely unpractical for the higher order terms that involve iterated (multivariate) time integrals. For these reasons, Fujii and Takahashi [10] have introduced a particle interpretation to randomize and compute numerically the integrals in (18), which we call the FT scheme. Let η_1 be the interaction time of a particle drawn independently as the first jump time of a Poisson process with an arbitrary intensity $\mu > 0$ starting from time $t \geq 0$, i.e., η_1 is a random variable with density

$$\phi(t, s) = \mathbb{1}_{s \geq t} \, \mu \, e^{-\mu(s-t)}. \tag{19}$$

From the first line in (18), we have

$$\widetilde{\Theta}_t^{(1)} = \widetilde{\mathbb{E}}_t \left[\int_t^T \phi(t, s) \frac{e^{\mu(s-t)}}{\mu} \widetilde{f}_s(\widetilde{\Theta}_s^{(0)}) ds \right] = \widetilde{\mathbb{E}}_t \left[\mathbb{1}_{\eta_1 < T} \frac{e^{\mu(\eta_1-t)}}{\mu} \widetilde{f}_{\eta_1}(\widetilde{\Theta}_{\eta_1}^{(0)}) \right]. \tag{20}$$

Similarly, the particle representation is available for the higher order. By applying the same procedure as above, we obtain

$$\widetilde{\Theta}_t^{(2)} = \widetilde{\mathbb{E}}_t \left[\mathbb{1}_{\eta_1 < T} \widetilde{\Theta}_{\eta_1}^{(1)} \frac{e^{\mu(\eta_1 - t)}}{\mu} \partial_\vartheta \widetilde{f}_{\eta_1}(\widetilde{\Theta}_{\eta_1}^{(0)}) \right],$$

where $\widetilde{\Theta}_{\eta_1}^{(1)}$ can be computed by (20). Therefore, by using the tower property of conditional expectations, we obtain

$$\widetilde{\Theta}_t^{(2)} = \widetilde{\mathbb{E}}_t \left[\mathbb{1}_{\eta_2 < T} \frac{e^{\mu(\eta_2 - \eta_1)}}{\mu} \widetilde{f}_{\eta_2}(\widetilde{\Theta}_{\eta_2}^{(0)}) \frac{e^{\mu(\eta_1 - t)}}{\mu} \partial_\vartheta \widetilde{f}_{\eta_1}(\widetilde{\Theta}_{\eta_1}^{(0)}) \right], \tag{21}$$

where η_1, η_2 are the two consecutive interaction times of a particle randomly drawn with intensity μ starting from t. Similarly, for the third order, we get

$$\widetilde{\Theta}_t^{(3)} = \widetilde{\mathbb{E}}_t \left[\mathbb{1}_{\eta_3 < T} \frac{e^{\mu(\eta_3 - \eta_2)}}{\mu} \widetilde{f}_{\eta_3}(\widetilde{\Theta}_{\eta_3}^{(0)}) \frac{e^{\mu(\eta_2 - \eta_1)}}{\mu} \partial_\vartheta \widetilde{f}_{\eta_2}(\widetilde{\Theta}_{\eta_2}^{(0)}) \frac{e^{\mu(\eta_1 - t)}}{\mu} \partial_\vartheta \widetilde{f}_{\eta_1}(\widetilde{\Theta}_{\eta_1}^{(0)}) \right], \tag{22}$$

where η_1, η_2, η_3 are consecutive interaction times of a particle randomly drawn with intensity μ starting from t. In case $t = 0$, (20), (21) and (22) can be simplified as

$$\widetilde{\Theta}_0^{(1)} = \widetilde{\mathbb{E}} \left[\mathbb{1}_{\zeta_1 < T} \frac{e^{\mu \zeta_1}}{\mu} \widetilde{f}_{\zeta_1}(\widetilde{\Theta}_{\zeta_1}^{(0)}) \right]$$

$$\widetilde{\Theta}_0^{(2)} = \widetilde{\mathbb{E}} \left[\mathbb{1}_{\zeta_1 + \zeta_2 < T} \frac{e^{\mu \zeta_1}}{\mu} \partial_\vartheta \widetilde{f}_{\zeta_1}(\widetilde{\Theta}_{\zeta_1}^{(0)}) \frac{e^{\mu \zeta_2}}{\mu} \widetilde{f}_{\zeta_1 + \zeta_2}(\widetilde{\Theta}_{\zeta_1 + \zeta_2}^{(0)}) \right]$$

$$\widetilde{\Theta}_0^{(3)} = \widetilde{\mathbb{E}} \left[\mathbb{1}_{\zeta_1 + \zeta_2 + \zeta_3 < T} \frac{e^{\mu \zeta_1}}{\mu} \partial_\vartheta \widetilde{f}_{\zeta_1}(\widetilde{\Theta}_{\zeta_1}^{(0)}) \frac{e^{\mu \zeta_2}}{\mu} \partial_\vartheta \widetilde{f}_{\zeta_1 + \zeta_2}(\widetilde{\Theta}_{\zeta_1 + \zeta_2}^{(0)}) \frac{e^{\mu \zeta_3}}{\mu} \widetilde{f}_{\zeta_1 + \zeta_2 + \zeta_3}(\widetilde{\Theta}_{\zeta_1 + \zeta_2 + \zeta_3}^{(0)}) \right] \tag{23}$$

where ζ_1, ζ_2, ζ_3 are the elapsed time from the last interaction until the next interaction, which are independent exponential random variables with parameter μ.

Note that the pricing model is originally defined with respect to the full stochastic basis $(\mathscr{G}, \mathbb{Q})$. Even in the case where there exists a stochastic basis $(\mathscr{F}, \mathbb{Q})$ satisfying the condition (C), $(\mathscr{F}, \mathbb{Q})$ simulation may be nontrivial. Lemma 8.1 in Crépey and Song [6] allows us to reformulate the $\mathbb{Q}$ expectations in (23) as the following $\mathbb{Q}$ expectations, with $\bar{\Theta}^{(0)} = 0$:

$$\widetilde{\Theta}_0^{(1)} = \bar{\Theta}_0^{(1)} = \mathbb{E} \left[\mathbb{1}_{\zeta_1 < \bar{\tau}} \frac{e^{\mu \zeta_1}}{\mu} \bar{f}_{\zeta_1}(\bar{\Theta}_{\zeta_1}^{(0)}) \right]$$

$$\widetilde{\Theta}_0^{(2)} = \bar{\Theta}_0^{(2)} = \mathbb{E} \left[\mathbb{1}_{\zeta_1 + \zeta_2 < \bar{\tau}} \frac{e^{\mu \zeta_1}}{\mu} \partial_\vartheta \bar{f}_{\zeta_1}(\bar{\Theta}_{\zeta_1}^{(0)}) \frac{e^{\mu \zeta_2}}{\mu} \bar{f}_{\zeta_1 + \zeta_2}(\bar{\Theta}_{\zeta_1 + \zeta_2}^{(0)}) \right]$$

$$\widetilde{\Theta}_0^{(3)} = \bar{\Theta}_0^{(3)} = \mathbb{E} \Big[\mathbb{1}_{\zeta_1 + \zeta_2 + \zeta_3 < \bar{\tau}} \frac{e^{\mu \zeta_1}}{\mu} \partial_\vartheta \bar{f}_{\zeta_1}(\bar{\Theta}_{\zeta_1}^{(0)}) \frac{e^{\mu \zeta_2}}{\mu} \partial_\vartheta \bar{f}_{\zeta_1 + \zeta_2}(\bar{\Theta}_{\zeta_1 + \zeta_2}^{(0)})$$

$$\times \frac{e^{\mu \zeta_3}}{\mu} \bar{f}_{\zeta_1 + \zeta_2 + \zeta_3}(\bar{\Theta}_{\zeta_1 + \zeta_2 + \zeta_3}^{(0)}) \Big], \tag{24}$$

which is nothing but the FT scheme applied to the partially reduced BSDE (II). The tractability of the FT schemes (23) and (24) relies on the nullity of the terminal condition of the related BSDEs (III) and (II), which implies that $\bar{\Theta}^{(0)} = \widetilde{\Theta}^{(0)} = 0$. By contrast, an FT scheme would not be practical for the full TVA BSDE (5) with terminal condition $\xi \neq 0$. Also note that the first order in the FT scheme (23) (resp. (24)) is nothing but the linear approximation (15) (resp. (14)).

4.3 Marked Branching Diffusion Approach

Based on an old idea of McKean [16], the solution $u(t_0, x_0)$ to a PDE

$$\partial_t u + \mathscr{L} u + \mu(F(u) - u) = 0, \quad u(T, x) = \Psi(x), \tag{25}$$

where $\mathscr{L}$ is the infinitesimal generator of a strong Markov process X and $F(y) = \sum_{k=0}^{d} a_k y^k$ is a polynomial of order d, admits a probabilistic representation in terms of a random tree $\mathscr{T}$ (branching diffusion). The tree starts from a single particle ("trunk") born from (t_0, x_0). Subsequently, every particle born from a node (t, x) evolves independently according to the generator $\mathscr{L}$ of X until it dies at time $t' = (t + \zeta)$ in a state x', where ζ is an independent μ-exponential time (one for each particle). Moreover, in dying, a particle gives birth to an independent number of k' new particles starting from the node (t', x'), where k' is drawn in the finite set $\{0, 1, \ldots, d\}$ with some fixed probabilities $p_0, p_1, \ldots, p_d$. The marked branching diffusion probabilistic representation reads

$$
\begin{aligned}
u(t_0, x_0) &= \mathbb{E}_{t_0, x_0}\left[\prod_{\{\text{inner nodes } (t,x,k) \text{ of } \mathscr{T}\}} \frac{a_k}{p_k} \prod_{\{\text{states } x \text{ of particles alive at } T\}} \Psi(x) \right] \\
&= \mathbb{E}_{t_0, x_0}\left[\prod_{k=0}^{d} \left(\frac{a_k}{p_k}\right)^{n_k} \prod_{l=1}^{\nu} \Psi(x_l) \right],
\end{aligned} \tag{26}
$$

where n_k is the number of branching with k descendants up on $(0, T)$ and ν is the number of particles alive at T, with corresponding locations $x_1, \ldots, x_\nu$.

The marked branching diffusion method of Henry-Labordère [13] for CVA computations, dubbed PHL scheme henceforth, is based on the idea that, by approximating y^+ by a well-chosen polynomial $F(y)$, the solution to the PDE

$$\partial_t u + \mathscr{L} u + \mu(u^+ - u) = 0, \quad u(T, x) = \Psi(x), \tag{27}$$

can be approximated by the solution to the PDE (25), hence by (26). We want to apply this approach to solve the TVA BSDEs (I), (II) or (III) for which, instead of fixing the approximating polynomial $F(y)$ once for all in the simulations, we need a state-dependent polynomial approximation to $g_t(y) = (P_t - y)^+$ (cf. (7)) in

a suitable range for y. Moreover, (I) and (II) are BSDEs with random terminal time $\bar{\tau}$, equivalently written in a Markov setup as Cauchy–Dirichlet PDE problems, as opposed to the pure Cauchy problem (27). Hence, some adaptation of the method is required. We show how to do it for (II), after which we directly give the algorithm in the similar case of (I) and in the more classical (pure Cauchy) case of (III). Assuming τ given in terms of a $(\mathscr{G}, \mathbb{Q})$ Markov factor process X as $\tau = \inf\{t > 0 : X_t \notin \mathscr{D}\}$ for some domain $\mathscr{D}$, the Cauchy–Dirichlet PDE used for approximating the partially reduced BSDE (II) reads:

$$(\partial_t + \mathscr{A})\bar{u} + \mu\left(\bar{F}(\bar{u}) - \bar{u}\right) = 0 \text{ on } [0, T] \times \mathscr{D}, \quad \bar{u}(t, x) = 0 \text{ for } t = T \text{ or } x \notin \mathscr{D},$$
(28)

where $\mathscr{A}$ is the generator of X and $\bar{F}_{t,x}(y) = \sum_{k=0}^{d} \bar{a}_k(t, x) y^k$ is such that

$$\mu(\bar{F}_{t,x}(y) - y) \approx \bar{f}(t, x, y), \text{ i.e. } \bar{F}_{t,x}(y) \approx \frac{\bar{f}(t, x, y)}{\mu} + y.$$
(29)

Specifically, in view of (9), one can set

$$\bar{F}_{t,x}(y) = \frac{1}{\mu}\left(cdva(t, x) + \bar{\lambda}pol\big(P(t, x) - y\big) - ry\right) + y = \sum_{k=0}^{d} \bar{a}_k(t, x) y^k,$$
(30)

where $pol(r)$ is a d-order polynomial approximation of r^+ in a suitable range for r. The marked branching diffusion probabilistic representation of $\bar{u}(t_0, x_0) \in \mathscr{D}$ involves a random tree $\overline{\mathscr{T}}$ made of nodes and "particles" between consecutive nodes as follows. The tree starts from a single particle (trunk) born from the root (t_0, x_0). Subsequently, every particle born from a node (t, x) evolves independently according to the generator $\mathscr{L}$ of X until it dies at time $t' = (t + \zeta)$ in a state x', where ζ is an independent μ-exponential time. Moreover, in dying, if its position x' at time t' lies in $\mathscr{D}$, the particle gives birth to an independent number of k' new particles starting from the node (t', x'), where k' is drawn in the finite set $\{0, 1, \ldots, d\}$ with some fixed probabilities $p_0, p_1, \ldots, p_d$. Figure 1 describes such a random tree in case $d = 2$. The first particle starts from the root (t_0, x_0) and dies at time t_1, generating two new particles. The first one dies at time t_{11} and generates a new particle, who dies at time $t_{111} > T$ without descendant. The second one dies at time t_{12} and generates two new particles, where the first one dies at time t_{121} without descendant and the second one dies at time t_{122} outside the domain $\mathscr{D}$, hence also without descendant. The blue points represent the inner nodes, the red points the outer nodes and the green points the exit points of the tree out of the time–space domain $[0, T] \times \mathscr{D}$.

The marked branching diffusion probabilistic representation of $\bar{u}$ is written as

$$\bar{u}(t_0, x_0) = \mathbb{E}_{t_0, x_0}\left[\mathbb{1}_{\overline{\mathscr{T}} \subset [0,T] \times \mathscr{D}} \prod_{\{\text{inner nodes } (t,x,k) \text{ of } \overline{\mathscr{T}}\}} \frac{\bar{a}_k(t, x)}{p_k}\right], \quad (t_0, x_0) \in [0, T] \times \mathscr{D}.$$
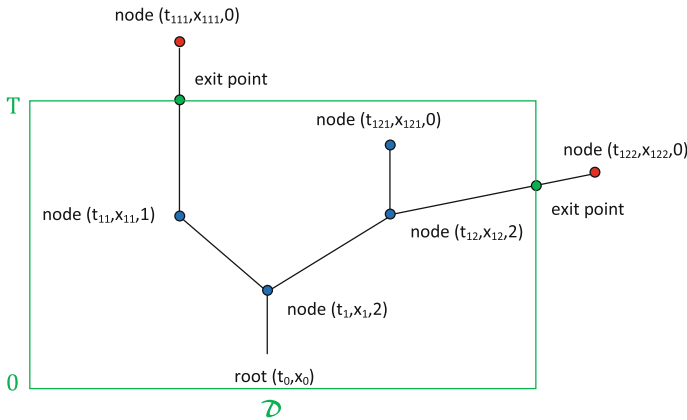(31)

Fig. 1 PHL random tree

Note that (31) is unformal at that stage, where we did not justify whether the PDE (28) has a solution $\bar{u}$ and in which sense. In fact, the following result could be used for proving that the function $\bar{u}$ defined in the first line is a viscosity solution to (28).

Proposition 1 *Denoting by $\bar{u}$ the function defined by the right hand side in* (31) *(assuming integrability of the integrand on the domain $[0, T] \times \mathscr{D}$), the process $Y_t = \bar{u}(t, X_t), 0 \leq t \leq \bar{\tau}$, solves the BSDE associated with the Cauchy–Dirichlet PDE* (28)*, namely*

$$Y_t = \mathbb{E}_t \left[\int_t^{\bar{\tau}} \mu\left(\bar{F}_{s,X_s}(Y_s) - Y_s\right) ds \right], \quad t \in [0, \bar{\tau}] \tag{32}$$

(which, in view of (29), approximates the partially reduced BSDE (II), so that $Y \approx \bar{\Theta}$ provided Y is square integrable).

Proof Let (t_1, x_1, k_1) be the first branching point in the tree rooted at $(0, X_0)$ and let $\overline{\mathscr{T}}_j$ denote k_1 independent trees of the same kind rooted at (t_1, x_1). By using the independence and the strong Markov property postulated for X, we obtain

$$\bar{u}(t, X_t) = \sum_{k_1=0}^{d} \mathbb{E}_{t,X_t} \left[\mathbb{1}_{t_1 < T} p_{k_1} \frac{a_{k_1}(t_1, x_1)}{p_{k_1}} \right.$$

$$\times \prod_{j=1}^{k_1} \mathbb{E}_{t_1, x_1} \left[\mathbb{1}_{\overline{\mathscr{T}}_j \subset [0,T] \times \mathscr{D}} \prod_{\{\text{inner node } (s,x,k) \text{ of } \overline{\mathscr{T}}_j\}} \frac{a_k(s, x)}{p_k} \right] \right]$$

$$= \mathbb{E}_{t,X_t} \left[\mathbb{1}_{t_1 < T} \sum_{k_1=0}^{d} a_{k_1}(t_1, x_1) \prod_{j=1}^{k_1} \mathbb{E}_{t_1, x_1} \left[\mathbb{1}_{\overline{\mathscr{T}}_j \subset [0,T] \times \mathscr{D}} \prod_{\{\text{inner node } (s,x,k) \text{ of } \overline{\mathscr{T}}_j\}} \frac{a_k(s, x)}{p_k} \right] \right]$$

$$= \mathbb{E}_{t,X_t} \left[\mathbb{1}_{t_1 < T} \sum_{k_1=0}^{d} a_{k_1}(t_1, x_1) \prod_{j=1}^{k_1} \bar{u}(t_1, x_1) \right]$$

$$= \mathbb{E}_{t,X_t} \left[\mathbb{1}_{t_1 < T} \bar{F}_{t_1,x_1}(\bar{u}(t, X_t^{t_1,x_1})) \right]$$

$$= \mathbb{E}_{t,X_t} \left[\int_t^{\bar{\tau}} \mu(s) e^{-\int_t^s \mu(u)du} \bar{F}_{s,X_s^{t,x}}(\bar{u}(s, X_s^{t,x}))ds \right], \ 0 \le t \le \bar{\tau},$$

i.e. $Y_t = \bar{u}(t, X_t)$ solves (32). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\Box$

If $\mathbb{1}_{\tau < T}\xi$ is given as a deterministic function $\Psi(\tau, X_\tau)$, then a similar approach (using the same tree $\overline{\mathcal{T}}$) can be applied to the full BSDE (I) in terms of the Cauchy–Dirichlet PDE

$$(\partial_t + \mathscr{A})u + \mu (F(u) - u) = 0 \text{ on } [0, T] \times \mathscr{D}, \quad u(t, x) = \Psi(t, x) \text{ for } t = T \text{ or } x \notin \mathscr{D}, \tag{33}$$

where $F_{t,x}(y) = \sum_{k=0}^d a_k(t, x)y^k$ is such that

$$\mu(F_{t,x}(y) - y) \approx f(t, x, y), \text{ i.e. } F_{t,x}(y) \approx \frac{f(t, x, y)}{\mu} + y.$$

This yields the approximation formula alternative to (31):

$$\Theta_0 \approx \mathbb{E} \left[\prod_{\{\text{inner node } (t,x,k) \text{ of } \overline{\mathcal{T}}\}} \frac{a_k(t, x)}{p_k} \prod_{\{\text{exit point } (t,x) \text{ of } \overline{\mathcal{T}}\}} \Psi(t, x) \right], \tag{34}$$

where an exit point of $\overline{\mathcal{T}}$ means a point where a branch of the tree leaves for the first time the time–space domain $[0, T] \times \mathscr{D}$. Last, regarding the $(\mathscr{F}, \mathbb{Q})$ reduced BSDE (III), assuming an $(\mathscr{F}, \mathbb{Q})$ Markov factor process $\widetilde{X}$ with generator $\widetilde{\mathscr{A}}$ and domain $\mathscr{D}$, we can apply a similar approach in terms of the Cauchy PDE

$$(\partial_t + \widetilde{\mathscr{A}})\widetilde{u} + \mu \left(\widetilde{F}_{t,x}(\widetilde{u}) - \widetilde{u} \right) = 0 \text{ on } [0, T] \times \mathscr{D}, \quad \widetilde{u}(t, x) = 0 \text{ for } t = T \text{ or } x \notin \mathscr{D}, \tag{35}$$

where $\widetilde{F}_{t,x}(y) = \sum_{k=0}^d \widetilde{a}_k(t, x)y^k$ is such that

$$\mu(\widetilde{F}_{t,x}(y) - y) \approx \widetilde{f}(t, x, y), \text{ i.e. } \widetilde{F}_{t,x}(y) \approx \frac{\widetilde{f}(t, x, y)}{\mu} + y.$$

We obtain

$$\Theta_0 = \widetilde{\Theta}_0 \approx \widetilde{\mathbb{E}} \left[\mathbb{1}_{\widetilde{\mathcal{T}} \subset [0,T] \times \mathscr{D}} \prod_{\text{inner node } (t,x,k) \text{ of } \widetilde{\mathcal{T}}} \frac{\widetilde{a}_k(t, x)}{p_k} \right], \tag{36}$$

where $\widetilde{\mathcal{T}}$ is the branching tree associated with the Cauchy PDE (35) (similar to $\overline{\mathcal{T}}$ but for the generator $\widetilde{\mathscr{A}}$).

5 TVA Models for Credit Derivatives

Our goal is to apply the above approaches to TVA computations on credit derivatives referencing the names in $N^\star = \{1, \ldots, n\}$, for some positive integer n, traded between the bank and the counterparty respectively labeled as -1 and 0. In this section we briefly survey two models of the default times τ_i, $i \in N = \{-1, 0, 1, \ldots, n\}$, that will be used for that purpose with $\tau_b = \tau_{-1}$ and $\tau_c = \tau_0$, namely the dynamic Gaussian copula (DGC) model and the dynamic Marshall–Olkin copula (DMO) model. For more details the reader is referred to [8, Chaps. 7 and 8] and [6, Sects. 6 and 7].

5.1 Dynamic Gaussian Copula TVA Model

5.1.1 Model of Default Times

Let there be given a function $\varsigma(\cdot)$ with unit L^2 norm on $\mathbb{R}_+$ and a multivariate Brownian motion $\mathbf{B} = (B^i)_{i \in N}$ with pairwise constant correlation $\rho \geq 0$ in its own completed filtration $\mathscr{B} = (\mathscr{B}_t)_{t \geq 0}$. For each $i \in N$, let h_i be a continuously differentiable increasing function from $\mathbb{R}_+^*$ to $\mathbb{R}$, with $\lim_0 h_i(s) = -\infty$ and $\lim_{+\infty} h_i(s) = +\infty$, and let

$$\tau_i = h_i^{-1}(\varepsilon_i), \text{ where } \varepsilon_i = \int_0^{+\infty} \varsigma(u) dB_u^i. \tag{37}$$

Thus the $(\tau_i)_{i \in N}$ follow the standard Gaussian copula model of Li [15], with correlation parameter ρ and with marginal survival function $\Phi \circ h_i$ of τ_i, where Φ is the standard normal survival function. In particular, these τ_i do not intersect each other. In order to make the model dynamic as required by counterparty risk applications, the model filtration $\mathscr{G}$ is given as the Brownian filtration $\mathscr{B}$ progressively enlarged by the τ_i, i.e.

$$\mathscr{G}_t = \mathscr{B}_t \vee \bigvee_{i \in N} \left(\sigma(\tau_i \wedge t) \vee \sigma(\{\tau_i > t\}) \right), \forall t \geq 0, \tag{38}$$

and the reference filtration $\mathscr{F}$ is given as $\mathscr{B}$ progressively enlarged by the default times of the reference names, i.e.

$$\mathscr{F}_t = \mathscr{B}_t \vee \bigvee_{i \in N^\star} \left(\sigma(\tau_i \wedge t) \vee \sigma(\{\tau_i > t\}) \right), \forall t \geq 0. \tag{39}$$

As shown in Sect. 6.2 of Crépey and Song [6], for the filtrations $\mathscr{G}$ and $\mathscr{F}$ as above, there exists a (unique) probability measure $\mathbb{P}$ equivalent to $\mathbb{Q}$ such that the condition (C) holds. For every $i \in N$, let

$$m_t^i = \int_0^t \varsigma(u) dB_u^i, \ k_t^i = \tau_i \mathbb{1}_{\{\tau_i \leq t\}},$$

and let $\mathbf{m}_t = (m_t^i)_{i \in N}$, $\mathbf{k}_t = (k_t^i)_{i \in N}$, $\widetilde{\mathbf{k}}_t = (\mathbb{1}_{i \in N^\star} k_t^i)_{i \in N}$. The couple $X_t = (\mathbf{m}_t, \mathbf{k}_t)$ (resp. $\widetilde{X}_t = (\mathbf{m}_t, \widetilde{\mathbf{k}}_t)$) plays the role of a $(\mathscr{G}, \mathbb{Q})$ (resp. $(\mathscr{F}, \mathbb{P})$) Markov factor process in the dynamic Gaussian copula (DGC) model.

5.1.2 TVA Model

A DGC setup can be used as a TVA model for credit derivatives, with mark $i = -1, 0$ and $E_b = \{-1\}$, $E_c = \{0\}$. Since there are no joint defaults in this model, it is harmless to assume that the contract promises no cash-flow at τ, i.e., $\Delta_\tau = 0$, so that $Q_\tau = P_\tau$. By [8, Propositions 7.3.1 p. 178 and 7.3.3 p. 181], in the case of vanilla credit derivatives on the reference names, namely CDS contracts and CDO tranches (cf. (47)), there exists a continuous, explicit function $\widetilde{P}_i$ such that

$$P_\tau = \widetilde{P}_i(\tau, \mathbf{m}_\tau, \mathbf{k}_{\tau-}), \tag{40}$$

or $\widetilde{P}_\tau^i$ in a shorthand notation, on the event $\{\tau = \tau_i\}$. Hence, (9) yields

$$\bar{f}_t(\vartheta) + r_t \vartheta = (1 - R_c)\gamma_t^0 (\widetilde{P}_t^0)^+ - (1 - R_b)\gamma_t^{-1}(\widetilde{P}_t^{-1})^- + \bar{\lambda}_t (P_t - \vartheta)^+, \quad \forall t \in [0, \bar{\tau}].$$

Assume that the processes r and $\bar{\lambda}$ are given before τ as continuous functions of (t, X_t), which also holds for P in the case of vanilla credit derivatives on names in N. Then the coefficients $\bar{f}$ and in turn $\widetilde{f}$ are deterministically given in terms of the corresponding factor processes as

$$\bar{f}_t(\vartheta) = \bar{f}(t, X_t, \vartheta), \quad \widetilde{f}_t(\vartheta) = \widetilde{f}(t, \widetilde{X}_t, \vartheta),$$

so that we are in the Markovian setup where the FT and the PHL schemes are valid and, in principle, applicable.

5.2 Dynamic Marshall–Olkin Copula TVA Model

The above dynamic Gaussian copula model allows dealing with TVA on CDS contracts. But a Gaussian copula dependence structure is not rich enough for ensuring a proper calibration to CDS and CDO quotes at the same time. If CDO tranches are also present in a portfolio, a possible alternative is the following dynamic Marshall–Olkin (DMO) copula model, also known as the "common shock" model.

5.2.1 Model of Default Times

We define a family $\mathcal{Y}$ of "shocks", i.e. subsets $Y \subseteq N$ of obligors, usually consisting of the singletons $\{-1\}, \{0\}, \{1\}, \ldots, \{n\}$, and a few "common shocks" $I_1, I_2, \ldots, I_m$ representing simultaneous defaults. For $Y \in \mathcal{Y}$, the shock time η_Y is defined as an i.i.d. exponential random variable with parameter γ_Y. The default time of obligor i in the common shock model is then defined as

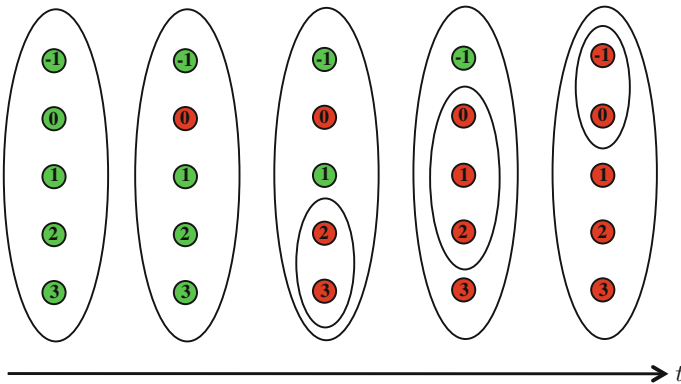$$\tau_i = \min_{Y \in \mathcal{Y}, i \in Y} \eta_Y. \tag{41}$$

Example 1 Figure 2 shows one possible default path in a common-shock model with $n = 3$ and $\mathcal{Y} = \{\{-1\}, \{0\}, \{1\}, \{2\}, \{3\}, \{2, 3\}, \{0, 1, 2\}, \{-1, 0\}\}$. The inner oval shows which shocks happened and caused the observed default scenarios at successive default times.

The full model filtration $\mathcal{G}$ is defined as

$$\mathcal{G}_t = \bigvee_{Y \in \mathcal{Y}} \left(\sigma(\eta_Y \wedge t) \vee \sigma(\{\eta_Y > t\}) \right), \ \forall t \geq 0.$$

Letting $\mathcal{Y}_\circ = \{Y \in \mathcal{Y}; \ -1, 0 \notin Y\}$, the reference filtration $\mathcal{F}$ is given as

$$\mathcal{F}_t = \bigvee_{Y \in \mathcal{Y}_\circ} \left(\sigma(\eta_Y \wedge t) \vee \sigma(\{\eta_Y > t\}) \right), \ t \geq 0.$$

Fig. 2 One possible default path in the common-shock model with $n = 3$ and $\mathcal{Y} = \{\{-1\}, \{0\}, \{1\}, \{2\}, \{3\}, \{2, 3\}, \{0, 1, 2\}, \{-1, 0\}\}$

As shown in Sect. 7.2 of Crépey and Song [6], in the DMO model with $\mathscr{G}$ and $\mathscr{F}$ as above, the condition (C) holds for $\mathbb{P} = \mathbb{Q}$. Let $J^Y = \mathbb{1}_{[0,\eta_Y)}$. Similar to $(\mathbf{m}, \mathbf{k})$ (resp. $(\mathbf{m}, \widetilde{\mathbf{k}})$) in the DGC model, the process

$$X = (J^Y)_{Y \in \mathscr{Y}} \ (\text{resp. } \widetilde{X} = (\mathbb{1}_{Y \in \mathscr{Y}_\circ} J^Y)_{Y \in \mathscr{Y}}) \tag{42}$$

plays the role of a $(\mathscr{G}, \mathbb{Q})$ (resp. $(\mathscr{F}, \mathbb{Q})$) Markov factor in the DMO model.

5.2.2 TVA Model

A DMO setup can be used as a TVA model for credit derivatives, with

$$E_b = \mathscr{Y}_b := \{Y \in \mathscr{Y}; \ -1 \in Y\}, \ E_c = \mathscr{Y}_c := \{Y \in \mathscr{Y}; \ 0 \in Y\}, \ E = \mathscr{Y}_\bullet := \mathscr{Y}_b \cup \mathscr{Y}_c$$

and

$$\tau_b = \tau_{-1} = \min_{Y \in \mathscr{Y}_b} \eta_Y, \ \tau_c = \tau_0 = \min_{Y \in \mathscr{Y}_c} \eta_Y,$$

hence

$$\tau = \min_{Y \in \mathscr{Y}_\bullet} \eta_Y, \ \gamma = \mathbb{1}_{[0,\tau)}\widetilde{\gamma} \text{ with } \widetilde{\gamma} = \sum_{Y \in \mathscr{Y}_\bullet} \gamma_Y. \tag{43}$$

By [8, Proposition 8.3.1 p. 205], in the case of CDS contracts and CDO tranches, for every shock $Y \in \mathscr{Y}$ and process $U = P$ or Δ, there exists a continuous, explicit function $\widetilde{U}_Y$ such that

$$U_\tau = \widetilde{U}_Y(\tau, X_{\tau-}), \tag{44}$$

or $\widetilde{U}_\tau^Y$ in a shorthand notation, on the event $\{\tau = \eta_Y\}$. The coefficient $\bar{f}_t(\vartheta)$ in (9) is then given, for $t \in [0, \bar{\tau}]$, by

$$\begin{aligned}
\bar{f}_t(\vartheta) + r_t \vartheta = (1 - R_c) \sum_{Y \in \mathscr{Y}_c} \gamma_t^Y \left(\widetilde{P}_t^Y + \widetilde{\Delta}_t^Y\right)^+ - (1 - R_b) \sum_{Y \in \mathscr{Y}_b} \gamma_t^Y \left(\widetilde{P}_t^Y + \widetilde{\Delta}_t^Y\right)^- \\
+ \bar{\lambda}_t (P_t - \vartheta)^+.
\end{aligned} \tag{45}$$

Assuming that the processes r and $\bar{\lambda}$ are given before τ as continuous functions of (t, X_t), which also holds for P in case of vanilla credit derivatives on the reference names, then

$$\bar{f}_t(\vartheta) = \bar{f}(t, X_t, \vartheta), \widetilde{f}_t(\vartheta) = \bar{f}_t(\vartheta) - \widetilde{\gamma}\vartheta = \widetilde{f}(t, \widetilde{X}_t, \vartheta) \tag{46}$$

(cf. (43)), so that we are again in a Markovian setup where the FT and the PHL schemes are valid and, in principle, applicable.

5.3 Strong Versus Weak Dynamic Copula Model

However, one peculiarity of the TVA BSDEs in our credit portfolio models is that, even though full and reduced Markov structures have been identified, which is required for justifying the validity of the FT and/or PHL numerical schemes, and the corresponding generators $\mathscr{A}$ or $\tilde{\mathscr{A}}$ can be written explicitly, the Markov structures are too heavy for being of any practical use in the numerics. Instead, fast and exact simulation and clean pricing schemes are available based on the dynamic copula structures.

Moreover, in the case of the DGC model, we lose the Gaussian copula structure after a branching point in the PHL scheme. In fact, as visible in [8, Formula (7.7) p. 175], the DGC conditional multivariate survival probability function is stated in terms of a ratio of Gaussian survival probability functions, which is explicit but does not simplify into a single Gaussian survival probability function. It is only in the DMO model that the conditional multivariate survival probability function, which arises as a ratio of exponential survival probability functions (see [8, Formula (8.11) p. 197 and Sect. 8.2.1.1]), simplifies into a genuine exponential survival probability function. Hence, the PHL scheme is not applicable in the DGC model.

The FT scheme based on (III) is not practical either because the Gaussian copula structure is only under $\mathbb{Q}$ and, again, the (full or reduced) Markov structures are not practical. In the end, the only practical scheme in the DGC model is the FT scheme based on the partially reduced BSDE (II). Eventually, it is only in the DMO model that the FT and the PHL schemes are both practical and can be compared numerically.

6 Numerics

For the numerical implementation, we consider stylized CDS contracts and protection legs of CDO tranches corresponding to dividend processes of the respective form, for $0 \le t \le T$:

$$D_t^i = \big((1 - R_i)\mathbb{1}_{t \ge \tau_i} - S_i(t \wedge \tau_i)\big)Nom_i$$
$$D_t = \Big(\big((1 - R)\sum_{j \in N}\mathbb{1}_{t \ge \tau_j} - (n + 2)a\big)^+ \wedge (n + 2)(b - a)\Big)Nom, \qquad (47)$$

where all the recoveries R_i and R (resp. nominals Nom_i and Nom) are set to 40 % (resp. to 100). The contractual spreads S_i of the CDS contracts are set such that the corresponding prices are equal to 0 at time 0. Protection legs of CDO tranches, where the attachment and detachment points a and b are such that $0 \le a \le b \le 100$ %, can also be seen as CDO tranches with upfront payment. Note that credit derivatives traded as swaps or with upfront payment coexist since the crisis. Unless stated otherwise, the following numerical values are used:

$$r = 0, R_b = 1, R_c = 40\%, \bar{\lambda} = 100 \text{ bp} = 0.01, \mu = \frac{2}{T}, m = 10^4.$$

6.1 Numerical Results in the DGC Model

First we consider DGC random times τ_i defined by (37), where the function h_i is chosen so that τ_i follows an exponential distribution with parameter γ_i (which in practice can be calibrated to a related CDS spread or a suitable proxy). More precisely, let Φ and Ψ_i be the survival functions of a standard normal distribution and an exponential distribution with intensity γ_i. We choose $h_i = \Phi^{-1} \circ \Psi_i$, so that (cf. (37))

$$\mathbb{Q}(\tau_i \geq t) = \mathbb{Q}\left(\Psi_i^{-1}\left(\Phi\left(\varepsilon_i\right)\right) \geq t\right) = \mathbb{Q}\left(\Phi\left(\varepsilon_i\right) \leq \Psi_i(t)\right) = \Psi_i(t),$$

for $\Phi\left(\varepsilon_i\right)$ has a standard uniform distribution. Moreover, we use a function $\varsigma(\cdot)$ in (37) constant before a time horizon $\bar{T} > T$ and null after $\bar{T}$, so that $\varsigma(0) = \frac{1}{\sqrt{\bar{T}}}$ (given the constraint that $v^2(0) = \int_0^\infty \varsigma^2(s) ds = 1$) and, for $t \leq \bar{T}$,

$$v^2(t) = \int_t^\infty \varsigma^2(s) ds = \frac{\bar{T} - t}{\bar{T}}, \quad m_t^i = \int_0^t \varsigma(u) dB_u^i = \frac{1}{\sqrt{\bar{T}}} B_t^i, \quad \int_0^\infty \varsigma(u) dB_u^i = \frac{1}{\sqrt{\bar{T}}} B_{\bar{T}}^i.$$

In the case of the DGC model, the only practical TVA numerical scheme is the FT scheme (24) based on the partially reduced BSDE (II), which can be described by the following steps:

1. Draw a time ζ_1 following an exponential law of parameter μ. If $\zeta_1 < T$, then simulate $\mathbf{m}_{\zeta_1} = (\frac{1}{\sqrt{\bar{T}}} B_{\zeta_1}^i)_{l \in N} \sim \mathcal{N}(0, \frac{\zeta_1}{\bar{T}} I_n(1, \rho))$, where $I_n(1, \rho)$ is a $n \times n$ matrix with diagonal equal to 1 and all off-diagonal entries equal to ρ, and go to Step 2. Otherwise, go to Step 4.
2. Draw a second time ζ_2, independent from ζ_1, following an exponential law of parameter μ. If $\zeta_1 + \zeta_2 < T$, then obtain the vector $\mathbf{m}_{\zeta_1+\zeta_2}$ as $\mathbf{m}_{\zeta_1} + (\mathbf{m}_{\zeta_1+\zeta_2} - \mathbf{m}_{\zeta_1})$, where $\mathbf{m}_{\zeta_1+\zeta_2} - \mathbf{m}_{\zeta_1} = (\frac{1}{\sqrt{\bar{T}}}(B_{\zeta_1+\zeta_2}^i - B_{\zeta_1}^i))_{l \in N} \sim \mathcal{N}(0, \frac{\zeta_2}{\bar{T}} I_n(1, \rho))$, and go to Step 3. Otherwise, go to Step 4.
3. Draw a third time ζ_3, independent from ζ_1 and ζ_2, following an exponential law of parameter μ. If $\zeta_1 + \zeta_2 + \zeta_3 < T$, then obtain the vector $\mathbf{m}_{\zeta_1+\zeta_2+\zeta_3}$ as $\mathbf{m}_{\zeta_1+\zeta_2} + (\mathbf{m}_{\zeta_1+\zeta_2+\zeta_3} - \mathbf{m}_{\zeta_1+\zeta_2})$, where $\mathbf{m}_{\zeta_1+\zeta_2+\zeta_3} - \mathbf{m}_{\zeta_1+\zeta_2} = (\frac{1}{\sqrt{\bar{T}}}(B_{\zeta_1+\zeta_2+\zeta_3}^i - B_{\zeta_1+\zeta_2}^i))_{l \in N} \sim \mathcal{N}(0, \frac{\zeta_3}{\bar{T}} I_n(1, \rho))$. Go to Step 4.

4. Simulate the vector $\mathbf{m}_{\bar{T}}$ from the last simulated vector $\mathbf{m}_t$ ($t = 0$ by default) as $\mathbf{m}_t + (\mathbf{m}_{\bar{T}} - \mathbf{m}_t)$, where $\mathbf{m}_{\bar{T}} - \mathbf{m}_t = (\frac{1}{\sqrt{\bar{T}}}(B^i_{\bar{T}} - B^i_t))_{i \in N} \sim \mathcal{N}(0, \frac{\bar{T}-t}{\bar{T}} I_n(1, \rho))$. Deduce $(B^i_{\bar{T}})_{i \in N}$, hence $\tau_i = \Psi_i^{-1} \circ \Phi\left(\frac{1}{\sqrt{\bar{T}}}B^i_{\bar{T}}\right)$, $i \in N$, and in turn the vectors $\mathbf{k}_{\zeta_1}$ (if $\zeta_1 + \zeta_2 + \zeta_3 < T$), $\mathbf{k}_{\zeta_1 + \zeta_2}$ (if $\zeta_1 + \zeta_2 < T$) and $\mathbf{k}_{\zeta_1 + \zeta_2 + \zeta_3}$ (if $\zeta_1 + \zeta_2 + \zeta_3 < T$).
5. Compute $\bar{f}_{\zeta_1}, \bar{f}_{\zeta_1 + \zeta_2}$, and $\bar{f}_{\zeta_1 + \zeta_2 + \zeta_3}$ for the three orders of the FT scheme.

We perform TVA computations on CDS contracts with maturity $T = 10$ years, choosing for that matter $\bar{T} = T + 1 = 11$ years, hence $\varsigma = \frac{\mathbb{1}_{[0,11]}}{\sqrt{11}}$, for $\rho = 0.6$ unless otherwise stated. Table 1 displays the contractual spreads of the CDS contracts used in these experiments. In Fig. 3, the left graph shows the TVA on a CDS on name 1, computed in a DGC model with $n = 1$ by FT scheme of order 1 to 3, for different levels of nonlinearity represented by the value of the unsecured borrowing spread $\bar{\lambda}$. The right graph shows similar results regarding a portfolio comprising one CDS contract per name $i = 1, \ldots, 10$. The time-0 clean value of the default leg of the CDS in case $n = 1$, respectively the sum of the ten default legs in case $n = 10$, is 4.52, respectively 40.78 (of course $P_0 = 0$ in both cases by definition of fair contractual spreads). Hence, in relative terms, the TVA numbers visible in Fig. 3 are quite high, much greater for instance than in the cases of counterparty risk on interest rate derivatives considered in Crépey et al. [7]. This is explained by the wrong-way risk feature of the DGC model, namely the default intensities of the surviving names and the value of the CDS protection spike at defaults in this model. When $\bar{\lambda}$ increases (for $\bar{\lambda} = 0$ that's a case of linear TVA where FT higher order terms equal 0), the second (resp. third) FT term may represent in each case up to 5–10% of the first

Table 1 Time-0 bp CDS spreads of names -1 (the bank), 0 (the counterparty) and of the reference names 1 to n used when $n = 1$ (*left*) and $n = 10$ (*right*)

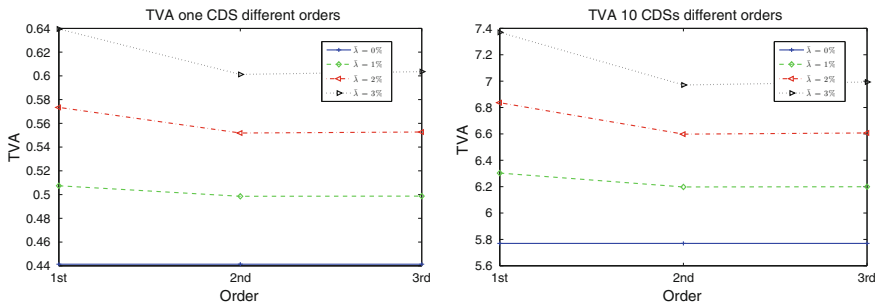| i	-1	0	1	i	-1	0	1	2	3	4	5	6	7	8	9	10
S_i	36	41	47	S_i	39	40	47	36	41	48	54	54	27	30	36	50

Fig. 3 *Left* DGC TVA on one CDS computed by FT scheme of order 1–3, for different levels of nonlinearity (unsecured borrowing spread $\bar{\lambda}$). *Right* similar results regarding the portfolio of CDS contracts on ten names
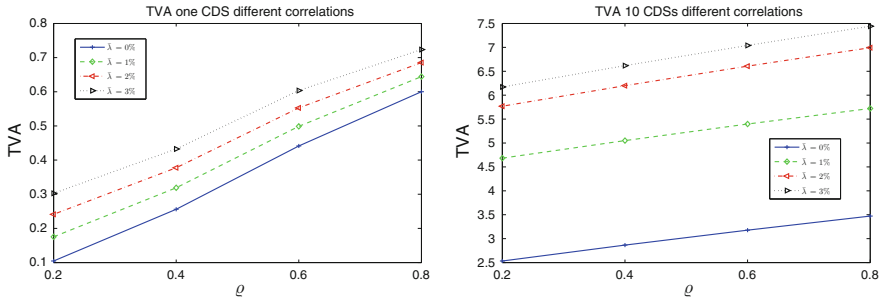
Fig. 4 *Left* TVA on one CDS computed by FT scheme of order 3 as a function of the DGC correlation parameter ρ. *Right* similar results regarding a portfolio of CDS contracts on ten different names
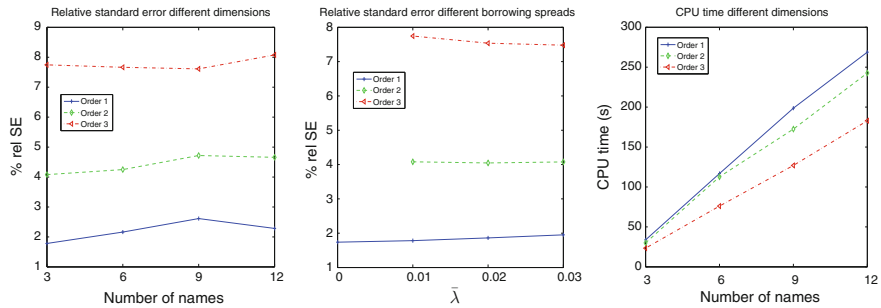

Fig. 5 *Left* the % relative standard errors of the different orders of the expansions do not explode with the number of names ($\bar{\lambda} = 100$ bp). *Middle* the % relative standard errors of the different orders of the expansions do not explode with the level of nonlinearity represented by the unsecured borrowing spread $\bar{\lambda}$ ($n = 1$). *Right* since FT terms are computed by purely forward Monte Carlo schemes, their computation times are linear in the number of names ($\bar{\lambda} = 100$ bp)

(resp. second) FT term, from which we conclude that the first FT term can be used as a first order linear estimate of the TVA, with a nonlinear correction that can be estimated by the second FT term.

In Fig. 4, the left graph shows the TVA on one CDS computed by FT scheme of order 3 as a function of the DGC correlation parameter ρ, with other parameters set as before. The right graph shows the analogous results regarding the portfolio of ten CDS contracts. In both cases, the TVA numbers increase (roughly linearly) with ρ, including for high values of ρ, as desirable from the financial interpretation point of view, whereas it has been noted in Brigo and Chourdakis [1] (see the blue curve in Fig. 1 of the ssrn version of the paper) that for high levels of the correlation between names, other models may show some pathological behaviors.

In Fig. 5, the left graph shows that the errors, in the sense of the relative standard errors (% rel. SE), of the different orders of the FT scheme do not explode with the dimension (number of credit names that underlie the CDS contracts). The middle graph, produced with $n = 1$, shows that the errors do not explode with the level of nonlinearity represented by the unsecured borrowing spread $\bar{\lambda}$. Consistent with

the fact that the successive FT terms are computed by purely forward Monte Carlo schemes, their computation times are essentially linear in the number of names, as visible in the right graph.

To conclude this section, we compare the linear approximation (14) corresponding to the first FT term in (24) (FT1 in Table 2) with the linear approximations (12)–(13) (LA in Table 2). One can see from Table 2 that the LA and FT1 estimates are consistent (at least in the sense of their 95 % confidence intervals, which always intersect each other). But the LA standard errors are larger than the FT1 ones. In fact, using the formula for the intensity γ of τ in FT1 can be viewed as a form of variance reduction with respect to LA, where τ is simulated. Of course, for $\bar{\lambda} \neq 0$ (case of the right tables where $\bar{\lambda} = 3\,\%$), both linear approximations are biased as compared with the complete FT estimate (with nonlinear correction, also shown in Table 2), particularly in the high dimensional case with 10 CDS contracts (see the bottom panels in Table 2). Figure 6 completes these results by showing the LA, FT1

Table 2 LA, FT1 and FT estimates: 1 CDS (*top*) and 10 CDSs (*bottom*), with parameters $\bar{\lambda} = 0\,\%$, $\rho = 0.8$ (*left*) and $\bar{\lambda} = 3\,\%$, $\rho = 0.6$ (*right*)

Method	TVA	95% CI	Rel. SE	Method	TVA	95% CI	Rel. SE
LA	0.65	[0.57, 0.73]	6.08 %	LA	0.66	[0.60, 0.72]	4.39%
FT1	0.61	[0.59, 0.63]	1.66%	FT1	0.62	[0.59, 0.64]	1.96%
FT	0.60	[0.58, 0.62]	1.64 %	FT	0.60	[0.58, 0.63]	1.84%

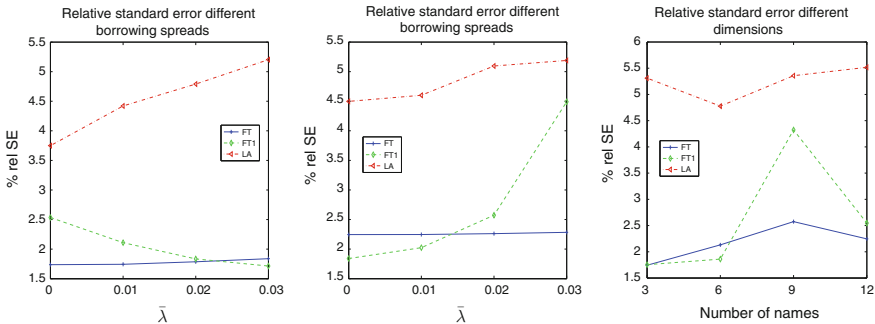| Method	TVA	95% CI	Rel. SE	Method	TVA	95% CI	Rel. SE
LA	6.17	[5.43, 6.92]	6.03%	LA	6.81	[6.16, 7.45]	4.76%
FT1	6.24	[5.77, 6.72]	3.78%	FT1	7.82	[7.39, 8.25]	2.73%
FT	6.17	[5.66, 6.68]	4.15%	FT	6.99	[6.67, 7.31]	2.28%

Fig. 6 The % relative standard errors of the different schemes do not explode with the level of nonlinearity represented by the unsecured borrowing spread $\bar{\lambda}$. *Left* 1 CDS. *Middle* 10 CDSs. *Right* the % relative standard errors of the different schemes (LA, FT1, FT in figures) do not explode with the number of names ($\bar{\lambda} = 100$ bp, $\rho = 0.6$)

and FT standard errors computed for different levels of nonlinearity and different dimensions.

Summarizing, in the DGC model, the PHL is not practical. The FT scheme based on the partially reduced TVA BSDE (II) gives an efficient way of estimating the TVA. The nonlinear correction with respect to the linear approximations (14) or (15) amounts up to 5 % in relative terms, depending on the unsecured borrowing spread $\bar{\lambda}$.

6.2 Numerical Results in the DMO Model

In the DMO model, the FT scheme (18) for the fully reduced BSDE (23) can be implemented through following steps:

1. Simulate the time η_Y of each (individual or joint) shock following an independent exponential law of parameter γ_Y, $Y \in \mathscr{Y}$, then retrieve the τ_i through the formula (41).
2. Draw a time ζ_1 following an exponential law of parameter μ. If $\zeta_1 < T$, compare the default time of each name with ζ_1 to obtain the reduced Markov factor $\widetilde{X}_{\zeta_1}$ as of (42) and in turn $\widetilde{f}_{\zeta_1}$ as of (45)–(46), then go to Step 3. Otherwise stop.
3. Draw a second time ζ_2 following an independent exponential law of parameter μ. If $\zeta_1 + \zeta_2 < T$, compare the default time τ_i of each name with $\zeta_1 + \zeta_2$ to obtain the Markov factor $\widetilde{X}_{\zeta_1+\zeta_2}$ and $\widetilde{f}_{\zeta_1+\zeta_2}$ then go to Step 4. Otherwise stop.
4. Draw a third time ζ_3 following an independent exponential law of parameter μ. If $\zeta_1 + \zeta_2 + \zeta_3 < T$, compare the default time of each name with $\zeta_1 + \zeta_2 + \zeta_3$ to obtain the Markov factor $\widetilde{X}_{\zeta_1+\zeta_2+\zeta_3}$ and $\widetilde{f}_{\zeta_1+\zeta_2+\zeta_3}$.

We can also consider the PHL scheme (31) based on the partially reduced BSDE (II) with

$$\mathscr{D} = \{x = (x^Y)_{Y\in\mathscr{Y}} \in \{0, 1\}^{\mathscr{Y}} \text{ such that } x^Y = 1 \text{ for } Y \in \mathscr{Y}_\bullet\}.$$

To simulate the random tree $\overline{\mathscr{T}}$ in (31), we follow the approach sketched before (31) where, in order to evolve X according to the DMO generator $\mathscr{A}$ during a time interval ζ, a particle born from a node $x = (j_Y)_{Y\in\mathscr{Y}} \in \{0, 1\}^{\mathscr{Y}}$ at time t, all one needs is, for each Y such that $j_Y = 1$, draw an independent exponential random variable η_Y of parameter γ_Y and then set $x' = (j_Y \mathbb{1}_{[0,\eta_Y)}(\zeta))_{Y\in\mathscr{Y}}$. Rephrasing in more algorithmic terms:

1. To simulate the random tree $\overline{\mathscr{T}}$ under the expectation in (31), we repeat the following step (generation of particles, or segments between consecutive nodes of the tree) until a generation of particles dies without children:

> For each node $(t, x = (j_Y)_{Y\in\mathscr{Y}}, k)$ issued from the previous generation of particles (starting with the root-node $(0, X_0, k = 1)$), for each of the k new particles, indexed by

l, issued from that node, simulate an independent exponential random variable ζ_l and set

$$(t'_l, x'_l, k'_l) = (t + \zeta_l, (j_Y \mathbb{1}_{[0, \eta^l_Y)}(\zeta_l))_{Y \in \mathscr{Y}}, \mathbb{1}_{x'_l \in \mathscr{D}} \nu_l),$$

where, for each l, the η^l_Y are independent exponential-γ_Y random draws and ν_l is an independent draw in the finite set $\{0, 1, \ldots, d\}$ with some fixed probabilities $p_0, p_1, \ldots, p_d$.

2. To compute the random variable Φ under the expectation in (31), we loop over the nodes of the tree $\overline{\mathscr{T}}$ thus constructed (if $\overline{\mathscr{T}} \subset [0, T] \times \mathscr{D}$, otherwise $\Phi = 0$ in the first place) and we form the product in (31), where the $\bar{a}_k(t, x)$ are retrieved as in (30).

The PHL schemes (34) based on the full BSDE (I) or (36) based on the fully reduced BSDE (III) can be implemented along similar lines.

We perform TVA computations in a DMO model with $n = 120$, for individual shock intensities taken as $\gamma_{\{i\}} = 10^{-4} \times (100 + i)$ (increasing from ∼100 bps to 220 bps as i increases from 1 to 120) and four nested groups of common shocks $I_1 \subset I_2 \subset I_3 \subset I_4$, respectively consisting of the riskiest 3, 9, 21 and 100 % (i.e. all) names, with respective shock intensities $\gamma_{I_1} = 20$ bp, $\gamma_{I_2} = 10$ bp, $\gamma_{I_3} = 6.67$ bp and $\gamma_{I_4} = 5$ bp. The counterparty (resp. the bank) is taken as the eleventh (resp. tenth) safest name in the portfolio. In the model thus specified, we consider CDO tranches with upfront payment, i.e. credit protection bought by the bank from the counterparty at time 0, with nominal 100 for each obligor, maturity $T = 2$ years and attachment (resp. detachment) points are 0, 3 and 14 % (resp. 3 %, 14 % and 100 %). The respective value of P_0 (upfront payment) for the equity, mezzanine and senior tranche is 229.65, 5.68, and 2.99. Accordingly, the ranges of approximation chosen for $pol(y) \approx y^+$ in the respective PHL schemes are 250, 200, and 10. We use polynomial approximation of order $d = 4$ with $(p_0, p_1, p_2, p_3, p_4) = (0.5, 0.3, 0.1, 0.09, 0.01)$. We set $\mu = 0.1$ in all PHL schemes and $\mu = 2/T = 0.2$ in all FT schemes.
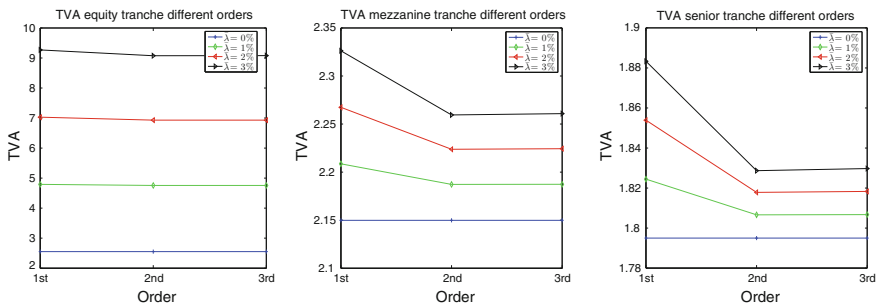

Fig. 7 TVA on CDO tranches with 120 underlying names computed by FT scheme of order 1–3 for different levels of nonlinearity (unsecured borrowing basis $\bar{\lambda}$). *Left* equity tranche. *Middle* mezzanine tranche. *Right* senior tranche. Originally published in Crépey and Song [6]. Published with kind permission of © Springer-Verlag Berlin Heidelberg 2016. All Rights Reserved. This figure is subject to copyright protection and is not covered by a Creative Commmons License

Table 3 FT, PHL, $\overline{\text{PHL}}$ and $\widetilde{\text{PHL}}$ schemes applied to the equity (*top*), mezzanine (*middle*), and senior (*bottom*) tranche, for the parameters $\bar{\lambda} = 0\,\%$, $\lambda_{I_j} = 60bp/j$ (*left*) or $\bar{\lambda} = 3\,\%$, $\lambda_{I_j} = 20bp/j$ (*right*)

Method	TVA	95% CI	Rel. SE	Method	TVA	95% CI	Rel. SE
FT	3.13	[3.10 , 3.16]	0.48 %	FT	9.08	[9.00 , 9.16]	0.46 %
$\widetilde{\text{PHL}}$	3.07	[2.87 , 3.28]	3.35 %	$\widetilde{\text{PHL}}$	9.05	[8.40 , 9.70]	3.58 %
$\overline{\text{PHL}}$	3.16	[2.94 , 3.37]	3.37 %	$\overline{\text{PHL}}$	9.28	[8.63 , 9.94]	3.51 %
PHL	2.53	[2.13 , 2.94]	8.02%	PHL	12.59	[6.92 , 18.27]	22.54%

Method	TVA	95% CI	Rel. SE	Method	TVA	95% CI	Rel. SE
FT	6.43	[6.33 , 6.53]	0.75 %	FT	2.29	[2.25 , 2.32]	0.77 %
$\widetilde{\text{PHL}}$	6.34	[5.93 , 6.75]	3.22 %	$\widetilde{\text{PHL}}$	2.51	[2.35 , 2.67]	3.17 %
$\overline{\text{PHL}}$	6.34	[5.93 , 6.75]	3.25 %	$\overline{\text{PHL}}$	2.68	[2.52 , 2.85]	3.12 %
PHL	4.86	[2.84 , 6.89]	20.82%	PHL	1.93	[0.79 , 3.08]	29.57%

Method	TVA	95% CI	Rel. SE	Method	TVA	95% CI	Rel. SE
FT	5.32	[5.24 , 5.40]	0.75 %	FT	1.83	[1.80 , 1.86]	0.78 %
$\widetilde{\text{PHL}}$	5.24	[4.90 , 5.58]	3.22 %	$\widetilde{\text{PHL}}$	1.80	[1.69 , 1.92]	3.13 %
$\overline{\text{PHL}}$	5.25	[4.90 , 5.58]	3.25 %	$\overline{\text{PHL}}$	1.87	[1.75 , 1.99]	3.11 %
PHL	4.01	[2.32 , 5.70]	21.03%	PHL	1.36	[0.41 , 2.31]	35.05%

Figure 7 shows the TVA computed by the FT scheme (23) based on the fully reduced BSDE (III), for different levels of nonlinearity (unsecured borrowing basis $\bar{\lambda}$). We observe that, in all cases, the third order term is negligible. Hence,
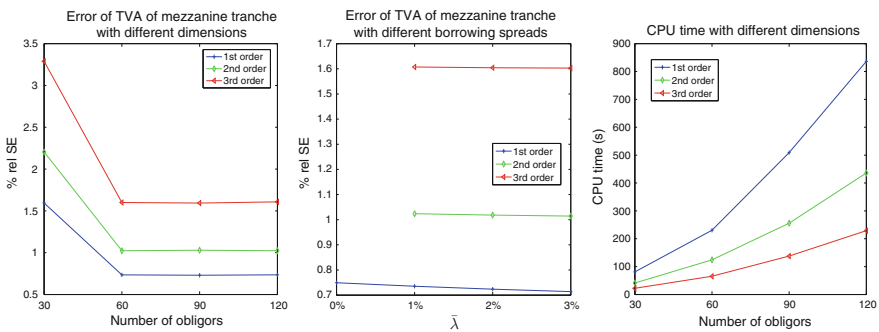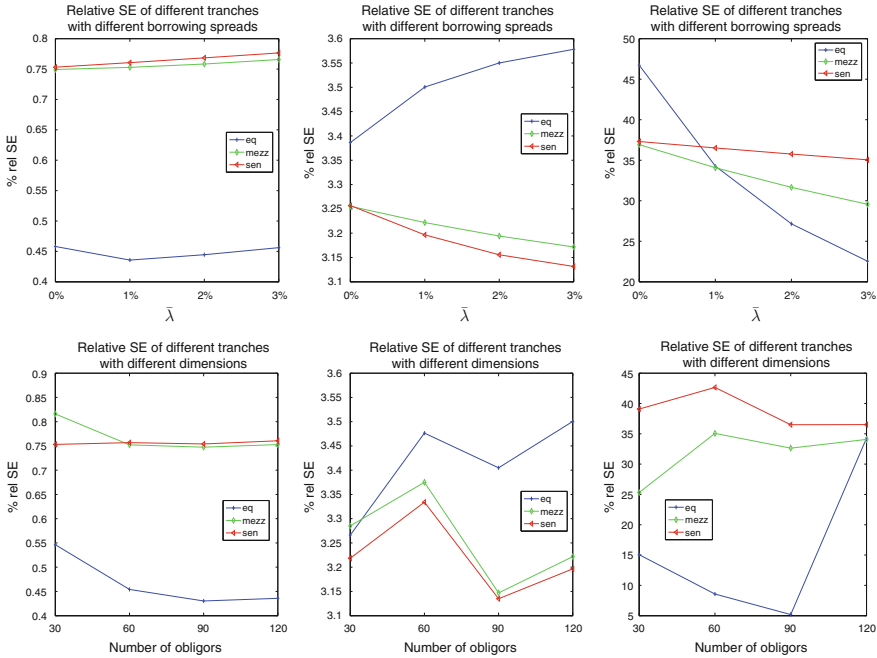

Fig. 8 Analog of Fig. 5 for the CDO tranche of Fig. 7 in the DMO model ($\bar{\lambda} = 0.01$). Originally published in Crépey and Song [6]. Published with kind permission of © Springer-Verlag Berlin Heidelberg 2016. All Rights Reserved. This figure is subject to copyright protection and is not covered by a Creative Commmons License

Fig. 9 *Bottom* the % relative standard errors do not explode with the number of names ($\bar{\lambda} = 100$ bp). *Top* the % relative standard errors do not explode with the level of nonlinearity represented by the unsecured borrowing spread $\bar{\lambda}$ ($n = 120$). *Left* FT scheme. *Middle* $\widetilde{\mathrm{PHL}}$ scheme. *Right* PHL scheme

in further FT computations, we only compute the orders 1 (linear part) and 2 (nonlinear correction) (Fig. 8). Table 3 compares the results of the above FT scheme (23) based on the fully reduced BSDE (III) with those of the PHL schemes (36) based on (III) again ($\widetilde{\mathrm{PHL}}$ in the tables), (31) based on the partially reduced BSDE (II) ($\overline{\mathrm{PHL}}$ in the tables) and (34) based on the full BSDE (I) (PHL in the tables), for the three CDO tranches and two sets of parameters. The three PHL schemes are of course slightly biased, but the first two, based on the BSDEs with null terminal condition (III) or (II), exhibit much less variance than the third one, based on the full BSDE with terminal condition ξ. This is also visible in Fig. 9 (note the different scales of the y axes going from left to right in the picture), which also shows that, for any of these schemes, the relative standard errors do not explode with the level of nonlinearity or the number of reference names in the CDO (the results for the $\overline{\mathrm{PHL}}$ scheme are not shown on the figure as very similar to those of the $\widetilde{\mathrm{PHL}}$ scheme). In comparing the TVA values on the left and the right hand side of Table 3, we see that the intensities of the common shocks, which play a role similar to the correlation ρ in the DGC model, have a more important impact on the higher tranches (mezzanine and senior tranche), whereas the equity tranche is more sensitive to the level of the unsecured borrowing spread $\bar{\lambda}$.

7   Conclusion

Under mild assumptions, three equivalent TVA BSDEs are available. The original
"full" BSDE (I) is stated with respect to the full model filtration $\mathscr{G}$ and the original
pricing measure $\mathbb{Q}$. It does not involve the intensity γ of the counterparty first-to-
default time τ. The partially reduced BSDE (II) is also stated with respect to $(\mathscr{G}, \mathbb{Q})$
but it involves both τ and γ. The fully reduced BSDE (III) is stated with respect to a
smaller "reference filtration" $\mathscr{F}$ and it only involves γ. Hence, in principle, the full
BSDE (I) should be preferred for models with a "simple" τ whereas the fully reduced
BSDE (III) should be preferred for models with a "simple" γ. But, in nonimmersive
setups, the fully reduced BSDE (III) is stated with respect to a modified probability
measure $\mathbb{P}$. Even though switching from $(\mathscr{G}, \mathbb{Q})$ to $(\mathscr{F}, \mathbb{P})$ is transparent in terms of
the generator of related Markov factor processes, this can be an issue in situations
where the Markov structure is important in the theory to guarantee the validity of the
numerical schemes, but is not really practical from an implementation point of view.
This is for instance the case with the credit portfolio models that we use for illustrative
purposes in our numerics, where the Markov structure that emerges from the dynamic
copula model is too heavy and it is only the copula features that can be used in the
numerics—copula features under the original stochastic basis $(\mathscr{G}, \mathbb{Q})$, which do not
necessarily hold under a reduced basis $(\mathscr{F}, \mathbb{P})$ (especially when $\mathbb{P} \neq \mathbb{Q}$). As for the
partially reduced BSDE (II), as compared with the full BSDE (I), its interest is its
null terminal condition, which is key for the FT scheme as recalled below. But of
course (II) can only be used when one has an explicit formula for γ.

For nonlinear and very high-dimensional problems such as counterparty risk on
credit derivatives, the only feasible numerical schemes are purely forward simu-
lation schemes, such as the linear Monte Carlo expansion of Fujii and Takahashi
[9, 10] or the branching particles scheme of Henry–Labordère [13], respectively
dubbed "FT scheme" and "PHL scheme" in the paper. In our setup, the PHL scheme
involves a nontrivial and rather sensitive fine-tuning for finding a polynomial in ϑ
that approximates the terms $(P_t - \vartheta)^{\pm}$ in $fva_t(\vartheta)$ in a suitable range for ϑ. This fine-
tuning requires a preliminary knowledge on the solution obtained by running another
approximation (linear approximation or FT scheme) in the first place. Another lim-
itation of the PHL scheme in our case is that it is more demanding than the FT
scheme in terms of the structural model properties that it requires. Namely, in our
credit portfolio problems, both a Markov structure and a dynamic copula are required
for the PHL scheme. But, whereas a "weak" dynamic copula structure in the sense
of simulation and forward pricing by copula means is sufficient for the FT scheme,
a dynamic copula in the stronger sense that the copula structure is preserved in the
future is required in the case of the PHL scheme. This strong dynamic copula prop-
erty is satisfied by our common-shock model but not in the Gaussian copula model.
In conclusion, the FT schemes applied to the partially or fully reduced BSDEs (II)
or (III) (a null terminal condition is required so that the full BSDE (I) is not eligible
for this scheme) appear as the method of choice on these problems.

An important message of the numerics is that, even for realistically high levels of nonlinearity, i.e. an unsecured borrowing spread $\bar{\lambda} = 3\%$, the third order FT correction was always found negligible and the second order FT correction less than 5–10% of the first order, linear FT term. In conclusion, a first order FT term can be used for obtaining "the best linear approximation" to our problem, whereas a nonlinear correction, if wished, can be computed by a second order FT term.

References

1. Brigo, D., Chourdakis, K.: Counterparty risk for credit default swaps: impact of spread volatility and default correlation. Int. J. Theor. Appl. Financ. **12**(7), 1007–1026 (2008)
2. Brigo, D., Capponi, A., Pallavicini, A.: Arbitrage-free bilateral counterparty risk valuation under collateralization and application to credit default swaps. Math. Financ. **24**(1), 125–146 (2014)
3. Crépey, S.: Bilateral counterparty risk under funding constraints - Part II: CVA. Math. Financ. **25**(1), 23–50 (2012)
4. Crépey, S., Song, S.: Invariant times. hal-01088940v1 (2014)
5. Crépey, S., Song, S.: BSDEs of counterparty risk. Stoch. Process. Appl. **125**(8), 3023–3052 (2014)
6. Crépey, S., Song, S.: Counterparty risk and funding: immersion and beyond. Finance Stoch. **20**(4), 901–930 (2016)
7. Crépey, S., Gerboud, R., Grbac, Z., Ngor, N.: Counterparty risk and funding: the four wings of the TVA. Int. J. Theor. Appl. Financ. **16**(2), 1350006(31 pp.) (2013)
8. Crépey, S., Bielecki, T.R., Brigo, D.: Counterparty Risk and Funding - A Tale of Two Puzzles. Chapman and Hall/CRC Financial Mathematics Series. Chapman and Hall/CRC, Boca Raton (2014)
9. Fujii, M., Takahashi, A.: Analytical approximation for non-linear FBSDEs with perturbation scheme. Int. J. Theor. Appl. Financ. **15**(5), 1250034(24) (2012)
10. Fujii, M., Takahashi, A.: Perturbative expansion of FBSDE in an incomplete market with stochastic volatility. Q. J. Financ. **2**(3), 1250015(24) (2012)
11. Fujii, M., Takahashi, A.: Collateralized credit default swaps and default dependence: implications for the central counterparties. J. Credit Risk **3**(8), 97–113 (2012)

12. He, S.-W., Wang, J.-G., Yan, J.-A.: Semimartingale Theory and Stochastic Calculus. CRC Press Boca Raton (1992)
13. Henry-Labordère, P.: Cutting CVA's complexity. Risk Mag. **25**(7), 67–73 (2012)
14. Kruse, T., Popier, A.: BSDEs with jumps in a general filtration. arXiv:1412.4622 (2014)
15. Li, D.: On default correlation: a copula function approach. J. Fixed Income **9**(4), 43–54 (2000)
16. McKean, H.P.: Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. Commun. Pure Appl. Math. **28**(3), 323–331 (1975)