Lihang Wang (France), Aymeric Kalife (France), Xiaolu Tan (France), Bruno Bouchard (France), Saad Mouti (France)

Understanding guaranteed minimum withdrawal benefit: a study on financial risks and rational lapse strategy

Abstract

The valuation of guaranteed minimum withdrawal benefit (GMWB) products attracts the attention of practitioners and academics both because of its long maturity and complex design properties, and also because of uncertain policyholder behaviors, such as lapse rate. In the present paper, we address the pricing problem as the valuation of a Bermudan-style option for the insurer. This evaluation approach corresponds to the price that allows the insurers to hedge the risk whatever the lapse strategy of the holder is. It is worthy to mention that so far, a historical or statistical lapse rate has generally been assumed for pricing these guarantees (see e.g. [4]). Both financial theory and past observations show that this assumption may lead to an underestimation of the risk associated to these products, the holders being rational or not. To evaluate the Bermudan-style liability, we implement two different schemes: partial differential equation (PDE) method and high-dimensional regression (HDR) method (see [11]). It is shown that the PDE method is precise for low-dimensional problems (< 3), while the HDR is more efficient when there are more than three dimensions. In the Hull and White stochastic interest rate model, the authors also show how a change of numeraire technique can be used to accelerate the numerical algorithms significantly for policies with ratchet (lookback) properties.

Keywords: GMWB, ratchet, variable annuity, rational lapse strategy, stochastic interest rates, long-term financial risk, PDE, ADI, high-dimensional regression.

Introduction

The guaranteed minimum withdrawal benefit (GMWB) riders are a recent innovation in the life insurance market. According to a report of IBBotson Associates (see [20]), the dominant sales driver for variable annuities (VA hereafter) in recent years was this kind of guarantee riders. In this paper, we focus on the GMWB rider for life.

The GMWB rider for life gives policyholders the ability to protect their retirement investments against downside market risks by allowing the annuitant the right to withdraw a fixed percentage (e.g. 4%) of the benefit base each year until death. The benefit base can either step up and be reset to the high-water mark of the account value on the rider anniversary date (Step-up policies), or can roll up with a fixed percentage annually (Roll-up policies), regardless of the market conditions. Another important feature of GMWB is that the policyholder can defer all annuities certain years after the inception. For simplicity, we assume that this “deferral period” is zero in this paper, which means that the customers begin withdrawals at the end of the first policy year. The remaining contract value at death will be paid to beneficiaries, which removes the investor concern about giving up liquidity to the heirs.

With the GMWB products, automatic annual resets are available after the contract is purchased. Thanks to these new designs on benefit bases, the guarantee protects policyholder against any nominal investment losses without losing the benefit of upside gain. In exchange for this benefit, the policyholder pays a charge fee each year. For example, suppose that Bill bought a GMWB contract with a withdrawal rate of 7%. At the end of deferral period, Bill’s account value reached historical high at 12,000, and his benefit base is also reset at 12,000. Due to good performance in the next year, the investment gains a 20% and the account value increases by about 10% (the investment gain minus the annuity paid and charge fees). Then the benefit base will be reset to a higher level so that Bill is able to receive more annuities each year thereafter.

Table 1 illustrates cash-flows of a typical GMWB rider, assuming the initial expense charge rate is 3% at the beginning of 1990. Note that the benefit base is the greater of a roll-up base (8%) and a step-up annually reset base. The guaranteed lifetime withdrawal rate is set to be 7% and a charge fee rate 1.0%. Supposing that the deferral period ends at the beginning of 1994, the policyholder will receive his first annuity at the beginning of 1995, which is equal to the withdrawal rate multiplied by the benefit base at 1994.

We observe in Table 1 that the benefit base is always higher than the account value and never decreases during all the period (thanks to the step-up feature). In addition, the guaranteed income is also compared with the net-return of 10-Y US Note. It is clear that, in our example, the annuity return of GMWB is much higher than that of long-term state notes. Moreover, GMWB can also protect contract owners from downside risk of interest rate (e.g., since 2000 in this example).

Lihang Wang, Ph.D. FX trader.
Aymeric Kalife, Affiliate Professor at University Paris-Dauphine, the Head of structuring, modeling and hedging at GIE AXA, France.
Xiaolu Tan, Affiliate Professor at University Paris-Dauphine.
Bruno Bouchard, Professor at Ceremade, University Paris Dauphine and Crest-Ensae, France.
Saad Mouti, Ph.D., Student at University Pierre et Marie Curie, France.

Table 1. Illustrations of GMWB rider

Policy year	0-Y Note	S&P 500	Account value/$	Benefit base/$	Rider fee/$	Guaranteed
	Net return	Net return				Income/$
1990			9,700	10,000	(300)	0
1991	8.4%	4,5%	10,036	10,800	(100)	0
1992	8.4%	18,9%	11,809	11,809	(118)	0
1993	8.4%	7,3%	12,549	12,597	(125)	0
1994	8.4%	9,8%	13,637	13,637	(136)	0
1995	8.4%	-2,3%	12,232	13,637	(122)	955
1996	8.4%	35.2%	15,418	15,418	(154)	955
1997	8.4%	23.6%	17,788	17,788	(178)	1079
1998	8.4%	24.7%	20,713	20,713	(207)	1245
1999	8.4%	30.5%	25,318	25,318	(253)	1450
2000	8.4%	9.0%	25,542	25,542	(256)	1772
2001	6.7%	-2.0%	22,983	25,542	(230)	1788
2002	6.7%	-17.3%	17,037	25,542	(170)	1788
2003	6.7%	-24.3%	10,982	25,542	(110)	1788
2004	6.7%	32.2%	12,584	25,542	(126)	1788
2005	6.7%	4.4%	11,223	25,542	(112)	1788
2006	6.7%	8.4%	10,252	25,542	(103)	1788
2007	6.7%	12.4%	9,615	25,542	(96)	1788
2008	6.7%	-4.2%	7,336	25,542	(73)	1788
2009	6.7%	-40.1%	2,563	25,542	(26)	1788
2010	6.7%	30.0%	1,512	25,542	(15)	1788

We observe in Table 1 that the benefit base is always higher than the account value and never decreases during all the period (thanks to the step-up feature). In addition, the guaranteed income is also compared with the net return of 10-Y US Note. It is clear that, in our example, the annuity return of GMWB is much higher than that of long-term state notes. Moreover, GMWB can also protect contract owners from downside risk of interest rate (e.g., since 2000 in this example).

Besides the financial risks related to its option features and long maturity, a major issue concerns the uncertainty on the lapse rate, i.e. the rate at which policyholders withdraw money from their accounts. In the past, historical or statistical lapse rates have generally been assumed in the literature (see e.g. [4]). This approach amounts to assuming that lapses are independent of the market conditions. Moreover, even if the above (very strong) assumption was satisfied, it would only allow to hedge the payoff in mean, without real control on the distribution of the hedging error (like bounds given by a central limit theorem). In reality, lapse rates of GMWB riders heavily depend on the account value and interest rate level. This phenomena is supported by the past observations, especially during the sub-prime crisis since 2008[1]. The dependency between lapse rate and financial market conditions is consistent with the intuition. Indeed, when the interest rate goes up, the discounted value of the insurer's annuities decreases. If the portfolio value rises at the same time, then the guarantees become worthless and more customers should surrender their policies to get back the account value. In the opposite situation, where both interest rate and account value decrease, the contract becomes to the contrary much more attractive, thus leading to a reduced lapse rate. Note that customers need not to be "rational", in the financial market sense, to adopt such behaviors. Namely, they are certainly not able to compute the optimal exercise times associated to such very complex products. However, competitors can take advantage of new market conditions to propose more attractive products, thus influencing their behavior indirectly.

In this paper, we address this question from the point of view of an insurer who wants to guarantee his ability to pay back the account value, whatever the holder's strategy is, rational or not. This means that we have to consider the embedded lapse option pricing problem as the valuation of a Bermudan-style contingent claim for the insurer. As usual, the policy price evaluated by this approach can be interpreted as the price of the claim if all policyholders use the same rational lapse strategy, which is similar to the optimal early-exercise strategy of American options. However, it is important to point out that this is only an interpretation, and that the critical aim is to make sure the lapse risk can be hedged whatever the holder's strategy is.

[1] As the market conditions were very volatile during the crisis, it was easier to observe the dependency between lapse rate and market conditions such as interest rate and account value.

This rational lapse approach, similar with the valuation problem of Bermudan-style options, is standard in financial markets (see e.g. [7]). Our main contributions is to apply it to GMWB riders, and, in particular, to highlight the critical importance of the up-front fees to protect the insurers from the wave of early termination of policies. Importantly, we allow for stochastic interest rates, which is important for products with long maturities. For simplicity, we assume that the account value has a Black and Scholes type dynamics with interest rates driven by a Hull and White one factor model. Obviously, the same methodology could be applied to more sophisticated models (including e.g. stochastic volatility or more general multi-factor term structures).

From the technical point of view, we use two different numerical methods to compute the rational lapse boundary and the early-termination premium (surrender option defined in [2]) of GMWB contracts. The first one is the Alternating-Direction Implicit Scheme (ADI, see [16], [14]). This is a pure determinist scheme allowing to solve PDEs in small dimensions (less than three). The second one is the High-Dimension Regression Scheme (HDR, see [11]). This is a Monte-Carlo based method that is more efficient in higher dimensions. In our numerical tests, these two schemes are both efficient and consistent with each other. For policies with ratchet properties, we also show how a change of numeraire technique can be used to accelerate the numerical algorithms significantly.

The reminder of this paper is organized as follows. In Section 1, we introduce the basic notations and general modeling framework for GMWB guarantees. Section 2 formulates the evaluation problem from the Bermudan-option point of view. The two numerical schemes are presented in Section 3, and the numerical tests are provided in Section 4. Final section closes with a summary of the main results.

To the best of our knowledge, it is the first time that such a study is performed in the academic literature.

1. Modeling framework

In this section, we present the modeling framework that we shall use to evaluate the liabilities of the GMWB.

From now on, we let $\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \leqslant T}, \mathbb{Q}$ denote a complete filtered probability space supporting two independent standard one dimensional Brownian motions W and $W^{\perp}$. Here $T > 0$ is a fixed time horizon. We assume that the filtration $\mathbb{F}$ is the completion of the rough filtration generated by $(W, W^{\perp})$, so that any martingale $(\mathbb{Q}, \mathbb{F})$-martingale can be represented as a stochastic integral with respect to $(W, W^{\perp})$.

1.1. Underlying asset and interest rate dynamics. During the last decade, the literature on pricing of GMWB riders has certainly evolved, but many evaluation approaches proposed (e.g. [4], [26]) are still based on the assumption of deterministic interest rates. Such an assumption is harmless in most situations since the interest rates variability is usually negligible when compared to the variability observed in equity markets. When pricing a long-maturity securities such as VA riders, however, the volatile feature of interest rates has a stronger impact on the guarantee value. In such case, it is therefore advisable to use stochastic interest rate models.

In this paper, we discuss the simple case where the short term interest rate $r = (r(t))_{t \geq 0}$ is driven by the one factor Hull and White model (see [19]), and the underlying asset $S = (S(t))_{t \geq 0}$ in which the account value is invested follows a Black and Scholes type dynamics, namely:

$$\begin{cases} dS(t) = r(t)S(t)dt + \sigma S(t)dW(t) \\ dr(t) = a(\theta(t) - r(t))dt + \sigma_r dZ(t)\,,\ Z := (1-\rho^2)^{\frac{1}{2}} W^{\perp} + \rho W \end{cases} \tag{1}$$

Here, a and σ_y are positive constants, θ is a deterministic Lebesgue-integrable function, σ is the instantaneous volatility of the asset return, and ρ is the correlation[1] between the account value and the interest rate. Note that the above financial market is complete whenever S and a zero-coupon bond with maturity T can be freely traded, and that $\mathbb{Q}$ is the only martingale (risk neutral) measure.

We formulate the complete dynamics of the account value for both roll-up riders and step-up riders in the following.

1.2. Roll-up riders. We first provide the dynamics of the account value in the case of a roll-up rider.

The anniversary dates at which the annuity payments are made and lapses can occur are denoted by

$$0 = t_0 < t_1 < \cdots < t_i < t_{i+1} < \cdots < t_n < t_{n+1} = \varpi \leq T\,,$$

where ϖ denotes the fixed maturity[2] of the policy. To simplify further the analysis, we also suppose that the policyholder is only allowed to surrender the contract at the discrete time points t_i.

[1] For VAs, the correlation is often negative, as most portfolio contains fixed income assets, such as bonds.

[2] For the GMWB rider for life, the maturity is in fact the death of the policyholder. For sake of simplicity, we suppose that the maturity is deterministic in this paper. Thus the value of a pool of GMWB riders is simply a weighted sum of the policies of different maturities according to the longevity assumption.

We assume that the account value $\breve{A}$ is fully invested in the risky asset S and that a charge fee is deducted at a rate c. This means that, between two anniversary dates t_i and t_{i+1}, it evolves according to:

$$d\breve{A}(t) = (r(t) - c)\breve{A}(t)dt + \sigma\breve{A}(t)dW(t),$$
$$t_i \leqslant t < t_{i+1}. \quad (2)$$

At the anniversary dates t_i, the account value jumps downwards (if it is not zero) due to the deduction of annuities. For roll-up riders, the annuities $(A_i)_{i\leq n}$ is determined at the inception. At each anniversary date t_i, the annuity A_i is deducted from the account value:

$$\breve{A}(t) = (\breve{A}(t-) - A_i)^+ \quad \text{for } t = t_i, i \in \{1, 2, \cdots, n\}, \quad (3)$$

where we write x^+ for $\max(x, 0)$.

Since $(r, \breve{A})$ is a Markov process, the liability process L^E associated to no early lapse can be identified to a deterministic function ℓ^E by:

$$L^E(t) := \sum_{j\geqslant 1}^{n} \mathbf{1}_{t<t_j} A_j \mathbf{E}^{\mathbb{Q}}[D_t^{t_j} \mid \mathcal{F}_t] +$$
$$+\mathbf{E}^{\mathbb{Q}}[D_t^{t_{n+1}} \breve{A}(t_{n+1}) \mid \mathcal{F}_t] = \ell^E(t, r(t), \breve{A}(t)), \quad (4)$$

where $D_{t_1}^{t_2}$ stands for the stochastic discount factor between t_1 and t_2:

$$D_{t_1}^{t_2} := \exp(-\int_{t_1}^{t_2} r(s)ds).$$

In equation (4), the first term on the right hand side is a sum of zero-coupon bond price, which can be computed explicitly in our Hull-White framework. In most cases, the second term, which can only be computed numerically, is much smaller than the first one because of the deduction of annuities during many years.

Because of the annuity payments at each anniversary of the policy purchase, the liability L^E is discontinuous at the discrete time points t_i. Combining the dynamic (3) of the account value and the definition (4) of the liability, we have at time t_i:

$$L^E(t_i-) = L^E(t_i) + A_i$$
$$\text{and} \quad \ell^E(t_i-, r, \breve{a}) = \ell^E(t_i, r, (\breve{a} - A_i)^+) + A_i. \quad (5)$$

The evaluation of the Bermudan-style liability in case of possible early lapse is deferred to the next section.

1.3. Step-up riders. The account value follows the same dynamics between the anniversary dates as in the previous case, see (2). The difference with roll-up riders is that the annuities are no more determined at the inception date but evolve according to the high-water mark of the account value $\bar{\breve{A}}(t_i-)$:

$$A_i := w\bar{\breve{A}}(t_i-), \quad (6)$$

where w is the annual withdrawal rate and:

$$\bar{\breve{A}}(t) := \sup_{s\leq m(t)} \breve{A}(s),$$
$$\text{where} \quad m(t) := \sup\{t_i, i \leq n : t_i \leqslant t\}, t \geq 0. \quad (7)$$

The complete dynamic of the account value under step-up riders is then given by:

$$\begin{cases} d\breve{A}(t) = (r(t) - c)\breve{A}(t)dt + \sigma\breve{A}(t)dW(t) & \text{for } t_i \leqslant t < t_{i+1}, i \in \{1, 2, \cdots, n\} \\ \breve{A}(t) = \max(\breve{A}(t-) - w\bar{\breve{A}}(t_i-), 0) & \text{for } t = t_i, i \in \{1, 2, \cdots, n\} \\ dr(t) = a(\theta(t) - r(t))dt + \sigma_r dZ(t) & \text{for } t \geq 0. \end{cases} \quad (8)$$

The value of the no-early lapse product, L^E, can also be identified to a deterministic function ℓ^E which now depends on the high-water mark:

$$L^E(t) := \sum_{j\geqslant 1}^{n} \mathbf{E}^{\mathbb{Q}}[\mathbf{1}_{t<t_j} D_t^{t_j} w\bar{\breve{A}}(t_j-) \mid \mathcal{F}_t] +$$
$$+\mathbf{E}^{\mathbb{Q}}[D_t^{t_{n+1}} \breve{A}(t_{n+1}) \mid \mathcal{F}_t] = \ell^E(t, r(t), \breve{A}(t), \bar{\breve{A}}(t)). \quad (9)$$

Because of the dependence with respect to the high-water mark, the evaluation issue much more complex than for roll-up riders. However, one can simplify the pricing process by apply a "change of numeraire" technique. It is based on the observation that the liability function of the step-up riders has the following scaling invariant property:

$$\ell^E(t_i-, r, \breve{a}, \bar{\breve{a}}) = \bar{\breve{a}}\, \tilde{\ell}^E(t_i-, r, \breve{a} / \bar{\breve{a}}), \quad (10)$$

where $\tilde{\ell}^E(t, r, x) := \ell^E(t, r, x, 1)$. We refer to Appendix A for the complete argument. This leads to the introduction of the positive process $(\xi(t))_{t\geqslant 0}$ defined by:

$$\xi(t) = \breve{A}(t) / \bar{\breve{A}}(t), \quad t \in [0, \bar{\omega}], \quad (11)$$

so that, recalling that the high-water mark of the account value is constant between two anniversary dates, see (7),

$$L^E(t) = \bar{\breve{A}}(t_i)\tilde{\ell}^E(t, r_t, \xi_t) \quad \text{for } t \in [t_i, t_{i+1}).$$

For the same reasons, ξ evolves in the same way as $\breve{A}$ on $[t_i, t_{i+1})$, namely:

$$d\xi(t) = (r(t) - c)\xi(t)dt + \sigma\xi(t)dW_t \quad \text{for } t \in [t_i, t_{i+1}), \tag{12}$$

while at the annuity payment point t_i, ξ has a sudden jump when $\xi(t_i-) > 0$. According to the definition of ξ and (8), we have at the time point t_i:

$$\xi(t_i) = \begin{cases} 1 & \text{for} \quad \xi(t_i-) \geqslant 1 + w \\ \max(\xi(t_i-) - w, 0) & \text{for} \quad \xi(t_i-) < 1 + w \end{cases} \tag{13}$$

with the initial condition $\xi(t_0) = 1$. We can then reduce to a two dimensional Markovian framework which considerably reduces the computation cost.

Applying the definition (11) of x and the dynamics given in (13), the liability function $\tilde{\ell}^E$ evolves according to the following equation at the annuity payment dates:

$$\tilde{\ell}^E(t_i-, r, x) = \begin{cases} (x - w)\tilde{\ell}^E(t_i, r, 1) + w & \text{for} \quad x \geqslant 1 + w \\ \tilde{\ell}^E(t_i, r, (x - w)^+) + w & \text{for} \quad 0 \leqslant x < 1 + w. \end{cases} \tag{14}$$

At the maturity t_{n+1}, we have $\tilde{\ell}^E(t_{n+1}, r, x) = x$. Note that the equation (14) can be further simplified to:

$$\tilde{\ell}^E(t_i-, r, x) = \max(x - w, 1)\tilde{\ell}^E(t_i, r, \min((x - w)^+, 1)) + w. \tag{15}$$

2. The Bermudan-option point of view

2.1. The pricing formulation. In the previous section, we have derived the price of the GMWB riders under the no-lapse assumption. If policyholders are not allowed to lapse contracts before maturity, the evaluation process of GMWB liability is similar to that of European-style contingent claims. However, in practice we can not assume the lapse rate to be zero or some other constant, as we observe that the lapse rate does change significantly in different market conditions (equity market and interest rate level) and these fluctuations of the lapse rate have notable impacts on the liability value and insurer's hedging strategy.

In this section, we discuss the pricing problem of the GMWB, with possible early lapses, from the point of view of an issuer who does not want to take any risk. If this is the case, his hedging strategy should be such that the liquidation value of the hedging portfolio is always bigger than what can be asked by the client, whatever the strategy of the client is, rational or not. If we can not control this lapse strategy, then the only way to hedge without risk consists in looking at the product as a Bermudan option, with possibly multiple exercises if multiple lapses are allowed, see e.g. [21] and [12].

In the rest of the paper, we simplify the presentation and assume that the account value can be lapsed only once, at an anniversary date, and for the total value of the account. Then, we simply face a "standard" Bermudan-type option.

For roll-up or step-up, the super-hedging price at time t of the liability, if not already exercised, in then given by:

$$L^B(t) := \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_t} \mathbf{E}^{\mathbb{Q}}\Big[\sum_{j \geq 0} 1_{t < t_j \leq \tau} D_t^{t_j} A_j + D_t^{\tau} \breve{A}(\tau) \,|\, \mathcal{F}_t\Big], \tag{16}$$

where $\mathcal{T}_t$ denotes the set of stopping times taking values in $\{t_i, i \leq n+1\} \cap [t, \infty)$.

As in the above section, we can use the fact that $(r, \breve{A}, \overline{\breve{A}})$ is a Markov process to associate the price process to a deterministic map ℓ^B satisfying

$$\ell^B(t, r(t), \breve{A}(t), \overline{\breve{A}}(t)) = L^B(t) \quad for\ t \leq \overline{\omega}. \tag{17}$$

Since the policyholder is not authorized to lapse the contract between two purchase anniversaries t_i and t_{i+1}, the process L^B of liability evolves as a $\mathbb{Q}$-martingale on each interval $[t_i, t_{i+1}[$. In particular,

$$\forall i \leqslant n, \quad \ell^B(t_i, r(t_i), \breve{A}(t_i), \overline{\breve{A}}(t_i)) = \mathbf{E}^{\mathbb{Q}}[D_i^{i+1} \ell^B(t_{i+1}-, r(t_{i+1}-), \breve{A}(t_{i+1}-), \overline{\breve{A}}(t_{i+1}-)) \,|\, \mathcal{F}_{t_i}]. \tag{18}$$

Similarly with Bermudan-style options, at discrete time points t_i, the policyholder is supposed to compare the account value with the value of the liability to decide whether to lapse or not. If the account value is bigger than the liability, policyholders surrender the contract and get back. $\breve{A}(t_i)$. Otherwise, they continue to hold the policy. Namely, the liability evolves at the anniversary date t_i as follows, depending wether it is a roll-up or a step-up rider:

$$Roll-up\ rider: \ell^B(t_i-, r, \breve{a}) = \max(\breve{a}, \ell^B(t_i, r, (\breve{a} - A_i)^+) + A_i) \tag{19}$$

$$Step-up\ rider: \ell^B(t_i-, r, \breve{a}, \overline{\breve{a}}) = \max(\breve{a}, \ell^B(t_i, r, (\breve{a} - w\overline{\breve{a}})^+, \max(\breve{a} - w\overline{\breve{a}}, \overline{\breve{a}})) + w\overline{\breve{a}}), \tag{20}$$

where in (19) we used the fact the price does not depend on the high-water mark for roll-up riders, so that we omit to write the dependence with respect to this quantity. By applying the change of variable (11) to equation (20) (see Appendix A for details), we then obtain in the case of the step-up rider:

$$\tilde{\ell}^B(t_i-,r,x)=\begin{cases}\max(x,(x-w)\tilde{\ell}^B(t_i,r,1)+w) & \text{for} \quad x\geqslant 1+w\\ \max(x,\tilde{\ell}^B(t_i,r,x-w)+w) & \text{for} \quad w\leqslant x<1+w\\ \max(x,\tilde{\ell}^B(t_i,r,0)+w) & \text{for} \quad 0\leqslant x<w\end{cases} \tag{21}$$

with $\tilde{\ell}^B(t,r,x):=\ell^B(t,r,x,1)$, so that:

$$\ell^B(t_i-,r,\breve{a},\bar{\bar{a}})=\bar{\bar{a}}\,\tilde{\ell}^B(t_i-,r,\breve{a}/\bar{\bar{a}}). \tag{22}$$

2.2. Analysis of the rational lapse strategy. Another important issue related to the pricing of the liability as a Bermudan-type option is the determination of the rational lapse strategy.

2.2.1. Roll-up riders. For roll-up riders, the annuity A_i being fixed at inception, one easily checks that the function $(t,r,\breve{a})\mapsto\ell^B(t,r,\breve{a})$ is convex and nondecreasing with respect to $\breve{a}$. In addition, according to equation (4), (16) and (17), we have $\ell^B(t,r,\breve{a})\geqslant\ell^E(t,r,\breve{a})>0$ for $(t,\breve{a})\in[0,T)\times(0,\infty)$.

Finally, as the annuities A_i of roll-up riders are all fixed at inception, it is easy to see that, for the charge fee $c > 0$, we have $\lim_{\breve{a}\to\infty}(\breve{a}-\ell^B(t_i,r,\breve{a}))>0$ for $i \leq n$. In fact, we know that at inception, $\ell^B(t_{n+1}, r(t_{n+1}), \breve{A}(t_{n+1})) = \breve{A}(t_{n+1})$. For $\breve{A}(t_{n+1}) > 0$, we have:

$$\begin{aligned}\ell^B(t_n,r(t_n),\breve{A}(t_n))&=\mathbf{E}^{\mathbb{Q}}[D_n^{n+1}\ell^B(t_{n+1}-,r(t_{n+1}-),\breve{A}(t_{n+1}-))\,|\,\mathcal{F}_{t_n}]\\ &=\mathbf{E}^{\mathbb{Q}}[D_n^{n+1}\max(\breve{A}(t_{n+1}-),\ell^B(t_{n+1},r(t_{n+1}),\breve{A}(t_{n+1}))+A_{n+1})\,|\,\mathcal{F}_{t_n}]\\ &=\mathbf{E}^{\mathbb{Q}}[D_n^{n+1}\breve{A}(t_{n+1}-)\,|\,\mathcal{F}_{t_i}]<\breve{A}(t_n),\end{aligned} \tag{23}$$

where the last inequality is a direct consequence of equation (2) in case of $c > 0$. We can prove this inequality for $i < n$ by following similar arguments.

It follows from the properties introduced above that, at the time t_i $(0 < i < n + 1)$, there exists a function $\breve{a}^*$ of the interest rate,

$$\begin{aligned}&0\leqslant\breve{a}<\breve{a}^*(t_i-,r) &&\Rightarrow \quad \ell^B(t_i-,r,\breve{a})>\breve{a}\\ &\breve{a}\geqslant\breve{a}^*(t_i-,r) &&\Rightarrow \quad \ell^B(t_i-,r,\breve{a})=\breve{a}.\end{aligned} \tag{24}$$

In this paper, $r\mapsto\breve{a}^*(t_i-,r)$ is referred to as the "critical account value" at time t_i since the policy should be lapsed as soon as the account value increases to this level at time t_i for a given interest rate level. As the interest rate is also a random variable here, the critical account value $\breve{a}^*(t_i-,r)$ is in fact a curve (a function of r) rather than a single point at time t_i. To avoid ambiguity, we call this curve the "critical surface" hereafter.

2.2.2. Step-up riders. For step-up riders, the nature of the critical surface is much more complex. As above the map $(t,r,x)\mapsto\tilde{\ell}^B(t,r,x)$ is convex and non-decreasing with respect to the x-variable. We also have $\tilde{\ell}^B(t,r,x)>\tilde{\ell}^E(t,r,x)>0$ for $(t, x)\in[0, T)\times(0, \infty)$. This allows to ensure the existence of a critical surface $r\mapsto x^*(t_i-,r)$ in three different situations:

1. $\tilde{\ell}^B(t_i,r,1)<1$: We have $\lim_{x\to\infty}[x-(x-w)\tilde{\ell}^B(t_i,r,1)-w]>0$. As we have explained, $\tilde{\ell}^B$ is a positive, convex and non-decreasing function. Therefore, when $\tilde{\ell}^B(t_i,r,1)<1$, there is a unique critical boundary $x^*(t_i-,r)$ such that:

$$\begin{aligned}&0\leqslant x<x^*(t_i-,r) &&\Rightarrow \quad \ell^B(t_i-,r,x)>x \quad \text{(not lapsed)}\\ &x\geqslant x^*(t_i-,r) &&\Rightarrow \quad \ell^B(t_i-,r,x)=x \quad \text{(lapsed)}\end{aligned} \tag{25}$$

2. $\tilde{\ell}^B(t_i,r,1)=1$: According to equation (21), $x=(x-w)\tilde{\ell}^B(t_i,r,1)+w$ for. $x \geq 1 + w$. It follows that the critical boundary $x^*(t_i,r)$ is the whole interval $[1+w,\infty)$:

$$\begin{aligned}&0\leqslant x<1+w &&\Rightarrow \quad \ell^B(t_i-,r,x)>x \quad \text{(not lapsed)}\\ &x\geqslant 1+w &&\Rightarrow \quad \ell^B(t_i-,r,x)=x \quad \text{(lapsed or not lapsed)}\end{aligned} \tag{26}$$

3. $\tilde{\ell}^B(t_i,r,1)>1$: We have $\lim_{x\to\infty}[(x-w)\tilde{\ell}^B(t_i,r,1)+w-x]>0$. This inequality gives rise to two possibilities: the map $x\mapsto(x-w)\tilde{\ell}^B(t_i,r,1)+w-x$ may cross the level 0 either twice

or never. In the former case, lapses occur if the account value ends up in the interval associated to two crossing points of the level 0. In the latter, there is no lapse whatever the account value is. Our numerical tests (see Section 4 below) support the latter case.

3. Numerical schemes

In this section, we now discuss the numerical estimation of the liability pricing function ℓ^B and the critical surface $\breve{a}^*$ (or x^*) for GMWB riders. Although many analytical approximation approaches exist in the academy literature (see e.g. [26]), most of them are not sufficiently precise for long maturities and lookback properties of the policies.

In the following, we propose two numerical methods: the PDE and the HDR schemes. As already mentioned, the PDE approach is precise for low-dimensional problems (< 3), while the HDR is more efficient when there are more than three dimensions in the pricing problem (e.g., multi-asset account value or stochastic volatility models).

3.1. The PDE approach. In this section, we transform the evaluation problem (20) of the Bermudan-style liability into a free-boundary partial differential equation, from which ℓ^B is the solution. Details on the numerical scheme used to solve this equation are given in Appendix B. The following study is divided into two parts, roll-up riders and step-up riders. In the second case, the "change of numeraire" technique is applied to reduce the number of dimensions.

3.1.1. Roll-up riders. For roll-up policies, the value of the Bermudan-style liability is a function of three variables: the time t, the interest rate r and the account value $\breve{A}$. Applying Itô's lemma and the martingale representation theorem together, we know that the value function associated to the liability is the solution of a two dimensional PDE:

$$\frac{\partial \ell^B}{\partial t} + (r-c)\breve{a}\frac{\partial \ell^B}{\partial \breve{a}} + \frac{\sigma^2 \breve{a}^2}{2}\frac{\partial^2 \ell^B}{\partial \breve{a}^2} + a(\theta - r)\frac{\partial \ell^B}{\partial r} + \frac{\sigma_r^2}{2}\frac{\partial^2 \ell^B}{\partial r^2} + \rho\sigma\sigma_r \breve{a}\frac{\partial^2 \ell^B}{\partial \breve{a}\partial r} = r\ell^B \tag{27}$$

on $\{(t,\breve{a}) : t_{i-1} \leqslant t < t_i, \breve{a} > 0\}$, subject to the boundary conditions at the time points $0 < t_i < t_{n+1}$

$$\ell^B(t_i-, r, \breve{a}) = \begin{cases} \max(\breve{a}, \ell^B(t_i, r, \breve{a} - A_i) + A_i) & \text{for } \breve{a} \geqslant A_i \\ \max(\breve{a}, \ell^B(t_i, r, 0) + A_i) & \text{for } \breve{a} < A_i \end{cases} \tag{28}$$

and at the maturity $\overline{\omega}$

$$\ell^B(t_{n+1}, r, \breve{a}) = \breve{a}.$$

More rigorously, the above holds in the sense of viscosity solutions, see e.g. [10]. Between two annuity payments, i.e. on each of the intervals $[t_{i-1}, t_i)$, the PDE (27) can be solved numerically by using the Alternating Direction Implicit (ADI)[1] scheme (see [16], [14], [18] and Appendix B). While at the discrete time points t_i, the critical lapse surface $\breve{a}^*$ can be easily found by the free-boundary constraint indicated in (28). On the boundary, we impose the zero-convexity condition[2]:

$$\frac{\partial^2 \ell^B}{\partial \breve{a}^2}\Big|_{\breve{a}\_boundary} = 0; \qquad \frac{\partial^2 \ell^B}{\partial r^2}\Big|_{r\_boundary} = 0.$$

In fact, according to [3], the precision of the final computation is not very sensible to the error on the boundaries if the solution domain of parabolic equation is large enough. So in most cases, the practitioner can choose other boundary conditions rather than the ones we propose here[3].

3.1.2. Step-up riders. For step-up policies, we evaluate $\tilde{\ell}^B(t,r,x)$ rather than $\ell^B(t,r,\breve{a},\overline{\breve{a}})$ directly.

As shown before, the Bermudan-style liability $\tilde{\ell}^B$ is a function of three variables: t, and x. In the same manner as for roll-up riders, by applying the dynamiques (12)-(13) of ξ, the evolution of $\tilde{\ell}^B$ at the annuity payment (14) and the free boundary constraint (21), we get a two dimensional PDE:

[1] In Appendix B, we show the ADI scheme in details and also verify the convergence of the numerical solution to the viscosity solution of equation (27) under certain assumptions.
[2] This assumption is based on the fact that the gamma of the liability is small on the boundary.
[3] In the specific case here, the first or second order derivative boundary condition is preferred to Dirichlet condition. As the latter could lead to significant errors on the boundary.

$$\frac{\partial\tilde{\ell}^B}{\partial t}+(r-c)x\frac{\partial\tilde{\ell}^B}{\partial x}+\frac{\sigma^2x^2}{2}\frac{\partial^2\tilde{\ell}^B}{\partial x^2}+a(\theta-r)\frac{\partial\tilde{\ell}^B}{\partial r}+\frac{\sigma_r^2}{2}\frac{\partial^2\tilde{\ell}^B}{\partial r^2}+\rho\sigma\sigma_r x\frac{\partial^2\tilde{\ell}^B}{\partial x\partial r}=r\tilde{\ell}^B \tag{29}$$

on $\{(t,x):t_{i-1}\leqslant t<t_i,x\geqslant 0\}$, subject to the boundary conditions for $0<t_i<t_{n+1}$

$$\tilde{\ell}^B(t_i-,r,x)=\begin{cases}\max(x,(x-w)\tilde{\ell}^B(t_i,r,1)+w) & \text{for } x\geqslant 1+w\\ \max(x,\tilde{\ell}^B(t_i,r,(x-w)^+)+w) & \text{for } 0\leqslant x<1+w\end{cases} \tag{30}$$

and at the maturity $\overline{\omega}$

$$\tilde{\ell}^B(t_{n+1},r,x)=x.$$

In our numerical tests, we use the ADI scheme to compute the PDE (29) numerically between two annuity payments, i.e. on each of the intervals $[t_{i-1}, t_i)$. While at each discrete time point t_i, we apply (30) to update the solution on all points of the discrete space grids. In addition, with the free boundary constraint indicated in (30), we can easily find the critical lapse surface x^*. On the boundary, we also impose the zero-convexity condition:

$$\frac{\partial^2\tilde{\ell}^B}{\partial x^2}|_{x\_boundary}=0;\ \frac{\partial^2\tilde{\ell}^B}{\partial r^2}|_{r\_boundary}=0.$$

3.2. High-dimensional regression scheme. In this paper, we always assume that the account value $\breve{A}$ evolves as a one-dimensional Wiener process. However, $\breve{A}$ is in practice a complex portfolio based on a multi-assets portfolio strategy, which may include both fix-income securities (e.g., bonds or treasuries) and stock indexes. This means that, in some cases, we need to consider the account value as a multi-dimensional Wiener process and that the evaluation of GMWB liability becomes a high-dimensional (> 3) problem.

In practice, when the number of dimension is bigger than three, purely deterministic schemes, such as finite difference methods for PDE, become inefficient. In contrast, the convergence speed of several purely Monte-Carlo based approaches (e.g. [9] or [24]) do not deteriorate too much when the number of dimension rises. Thus it is rational to apply Monte-Carlo type techniques to evaluate high-dimensional problems.

In this section, we apply the high-dimensional regression scheme (HDR) discussed in Bouchard and Warin [11] (see also [17]) to calculate the Bermudan-style liability ℓ^B (or $\tilde{\ell}^B$) of GMWB policies and the related critical surface defined earlier. Although many other Monte-Carlo based methods are available in the literature, such as [24] or [6], they are either imprecise and extremely time-consuming or difficult to implement for high dimensional problems. The reminder of this section is divided into two parts concerning successively the roll-up riders and the step-up riders cases.

3.2.1. Roll-up riders. The algorithm is based on the well-known fact that L^B satisfies the backward programming equation:

$$L^B(t_i-)=\max\{\breve{A}(t_i-),\mathbf{E}^{\mathbb{Q}}[D_{t_i}^{t_{i+1}}L^B(t_{i+1}-)\,|\,\mathcal{F}_{t_i-}]+A_i\},\quad i\leq n, \tag{31}$$

with terminal condition

$$L^B(t_{n+1})=\breve{A}(t_{n+1}).$$

It follows that:

$$L^B(t_i-)=\mathbf{E}^{\mathbb{Q}}[\sum_{j\geqslant i}1_{t_j\leq\hat{\tau}_i}D_{t_i}^{t_j}A_j+D_{t_i}^{\hat{\tau}_i}\breve{A}(\hat{\tau}_i)\,|\,\mathcal{F}_{t_i}], \tag{32}$$

where $\hat{\tau}_i$ denotes the (first) optimal stopping time after time: t_i $\hat{\tau}_i:=\inf\{t_j\geqslant t_i:L^B(t_j-)=\breve{A}(t_j-)\}$.

As pointed out in [11], the equations (31) and (32) lead to two possible algorithms (referred to as S1 and S2 hereafter) for the computation of the Bermudan liability $L^B(t_0)$ at inception. The scheme S2 is similar to the well-known Longstaff-Schwartz

(LS, see [24]) method. First, we simulate the path of $\breve{A}$ and r on $\{t_0, \ldots, t_{n+1}\}$ by using the SDE (3). In the following, we use the index $(\cdot)^{(k)}$ to denote quantities associated to the k-th simulated path. Then, we apply equation (31) backward from t_n to t_0 to estimate the liability $L^B(t_0)$ and the optimal lapse time $\hat{\tau}_0$.

Scheme $S1$: Price process computation

1. Simulate N discrete scenarios for the account value and the interest rate, noted as $\breve{A}^{(k)}(t_i)$ and $r^{(k)}(t_i)$ ($0 \le i \le n+1$ and $0 < k \le N$), according to (3).
2. Initialization: Set the Bermudan-style liability at maturity for each scenario $L_{[1]}^{B,(k)}(t_{n+1}) = \breve{A}^{(k)}(t_{n+1})$
3. Apply (31) from t_n to t_0. For $i = n$ to 0:

$$\text{if } \breve{A}^{(k)}(t_i-) = 0: L_{[1]}^{B,(k)}(t_i-) = D_i^{i+1,(k)} L_{[1]}^{B,(k)}(t_{i+1}-) + A_i$$

$$\text{if } \breve{A}^{(k)}(t_i-) > 0: L_{[1]}^{B,(k)}(t_i-) = \max\{\breve{A}^{(k)}(t_i-), \tilde{\mathbf{E}}^{\mathbb{Q}}[\tilde{D}_{t_i}^{t_{i+1}} L_{[1]}^{B}(t_{i+1}-) \,|\, \{\breve{A}, r\}^{(k)}(t_i-)] + A_i\}.$$

4. Price estimator at 0:

$$L_{[2]}^B(0) := \frac{1}{N}\sum_{k=1}^{N}[\sum_{j=0}^{\hat{\tau}_0^{(k)}} \tilde{D}_0^{t_j,(k)} A_j + \tilde{D}_0^{\hat{\tau}_0^{(k)},(k)} \breve{A}^{(k)}(\hat{\tau}_0^{(k)})].$$

In the above, we approximate $D_{t_i}^{t_j}$ by $\tilde{D}_{t_i}^{t_j} = \prod_{i \le l < j} \exp(-\frac{r(t_l) + r(t_{l+1})}{2}(t_{l+1} - t_l))$ for $i < j$, see Appendix C for more details on the simulation procedure.

The key to evaluate $L^B(0)$ by scheme S1 is to estimate the conditional expected value of the liability $\mathbf{E}^{\mathbb{Q}}[\tilde{D}_{t_i}^{t_{i+1}} L_{[1]}^B(t_{i+1}-) \,|\, \{\breve{A}, r\}^{(k)}(t_i-)]$. Here, it is approximated by a numerical estimator $\tilde{\mathbf{E}}^{\mathbb{Q}}[\cdot \,|\, \{\breve{A}, r\}^{(k)}(t_i-)]$. This estimation is done by regressing the subsequent realized cash flows from continuation ($\tilde{D}_{t_i}^{t_{i+1}} L_{[1]}^B(t_{i+1}-) + A_i$) on a set of basis functions depending on the relevant state variables ($\breve{A}^{(k)}(t_i-)$, $r^{(k)}(t_i-)$). We will introduce the estimator $\tilde{\mathbf{E}}^{\mathbb{Q}}[\cdot \,|\, \mathcal{F}_{t_i-}]$ later in full details.

In Step 3 of scheme S1, we can also identify the estimated rational lapse times $\hat{\tau}_i^{(k)}$ for each (k) scenario. At time t_i ($0 \le i \le n$), we separate all paths into two different groups, $\mathfrak{L}_i$ and its complement $\mathfrak{L}_i^c$, where $\mathfrak{L}_i := \{(k): L_{[1]}^{B,(k)}(t_i-) = \breve{A}^{(k)}(t_i-)\}$ and $\mathfrak{L}_i^c := \{(k): L_{[1]}^{B,(k)}(t_i-) > \breve{A}^{(k)}(t_i-)\}$. The stopping time $\hat{\tau}_i^{(k)}$ is the first time t_j after when t_i the k-th scenario enters into the group $\mathfrak{L}_j$ ($0 \leqslant i \le j \leqslant n+1$). Once $\hat{\tau}^{(k)}$ is recorded for each path, we can estimate our Bermudan-style liability by scheme S2 as follows:

Scheme S2: Rational lapse time estimation

1. Simulation: Use the same N simulated scenarios as in S1.
2. Initialization: Set the rational lapse time $\hat{\tau}_{n+1}^{(k)} = t_{n+1}$, for. $0 < k \le N$.
3. Backward induction: For $i = n$ to 0, $\hat{\tau}_i^{(k)} = t_i \mathbf{1}_{\{(k) \in \mathfrak{L}_i\}} + \hat{\tau}_{i+1}^{(k)} \mathbf{1}_{\{(k) \in \mathfrak{L}_i^c\}}$.

In addition, we only estimate the conditional expected value of continuation at time t_i for scenarios such that $\breve{A}^{(k)}(t_i) > 0$. As the policyholder will not lapse the contract once $\breve{A}^{(k)}$ touches 0, the Bermudan-style liability evolves like a European one. This technique was first proposed by Longstaff and Schwartz [24] to accelerate the pricing of American options.

In Bouchard and Warin (2010) the authors indicated that the following relation is true at a formal level:

$$\mathbf{E}^{\mathbb{Q}}[L_{[2]}^B(0)] \leqslant L^B(0) \leqslant \mathbf{E}^{\mathbb{Q}}[L_{[1]}^B(0)]. \tag{33}$$

If the conditional expectation estimator was satisfying $Y \in L^1 \mapsto \tilde{\mathbf{E}}^{\mathbb{Q}}[Y \,|\, \{\breve{A}, r\}^{(k)}(t_i-)] \in L^1(F_{t_i})$, then we could see the estimated optimal lapse policy as a suboptimal stopping time. This would lead to $\mathbf{E}^{\mathbb{Q}}[L_{[2]}^B(0)] \leqslant L^B(0)$. On the other hand, if we assume that the conditional expectation operators are conditionally unbiased, then a backward induction argument combined with Jensen's inequality would imply $L^B(0) \leqslant \mathbf{E}^{\mathbb{Q}}[L_{[1]}^B(0)]$. Obviously both above assumptions do not hold true and this reasoning can only be done at a formal level. In the numerical tests, we calculate both $L_{[1]}^B(0)$ and $L_{[2]}^B(0)$ to construct *confidence intervals* of the form $[L_{[2]}^B(0), L_{[1]}^B(0)]$ for the true value $L^B(0)$.

We now introduce the scheme used to calculate the conditional expected value of continuation for scenarios such that $\breve{A}^{(k)}(t_i-) > 0$. Here we use the adaptive hyper-cubes based local linear regression approach proposed in Bouchart and Warin (2010), as opposed to the global polynomial regression method developed in Longstaff and Schwartz (2004). The reason for this is that the latter can lead to some instability in the regression process, especially for high dimensional and long maturity problems (see Bouchart and Warin, 2010). Moreover, the choice of a good global polynomial basis is typically difficult, and does not allow to construct efficient payoff-free algorithms.

For our specific problem, we have two space dimensions: the interest rate r and the account value $\breve{A}$. The idea is to use, at each time step t_i, a set of functions $\psi_{d_1 d_2}$ having local hypercube support $D_{d_1 d_2}$, where the p-th dimension is cut into I_p regions, d_p = 1 to I_p, and $\{D_{d_1 d_2}\}$ is a partition of

$$[\min_{k=1,\ldots,N} r^{(k)}(t_i), \max_{k=1,\ldots,N} r^{(k)}(t_i)] \times \\ \times [\min_{k=1,\ldots,N} \breve{A}^{(k)}(t_i), \max_{k=1,\ldots,N} \breve{A}^{(k)}(t_i)].$$

On each support D_l, $l = (d_1 \times d_2)$ we define a linear function Ψ_l with 3 degrees of freedom, which are represented by a constant, and the coefficients for $\breve{A}$ and r. At each time step, we regress the future cash flows of the liability on the function Ψ_l to estimate the relevant conditional expectation. As in Bouchart and Warin (2010), we compute the hyper-cubes so that they contain approximatively the same number of simulated values. This allows to adapt automatically to the law of the underlying processes.

3.2.2. Step-up riders. Clearly, the same idea can be used to evaluate step-up riders by simulating the process ξ instead of $\breve{A}$. The main difference compared with roll-up riders is that, at each discrete time points t_i, the liability should be multiplied by the factor max $(\xi - w, 1)$.

4. Numerical experiments

We now present the numerical results obtained by using the two approaches (PDE and HDR) introduced above. Our final results show not only the consistency between these two methods, but also the efficiency and precision of both methods. We shall also see that the implicit "critical surfaces" are comparable.

In the following, the account value is supposed to evolve according to (1), for the set of parameters listed in Table 2.

Table 2. Hull-White model inputs

σ	$r(0)$	θ	a	σ_y	p
2	0.02	0.02	0.03	0.01	0

For simplicity, we also assume that the policyholder dies at the 40 th policy year ($\overline{\omega} = t_{n+1} = 40$) after the purchase of contract. Although we are focused on the liability of a single policy in the following numerical tests, the methodology we propose here can be easily extended to evaluate a GMWB policy pool by adding up policies of different maturities with a proper weight associated to given mortality rate assumptions.

4.1. Roll-up rider. For roll-up riders, we use the set of parameters given in Table 2 and Table 3. As introduced above, the annuity $A_i = w\breve{A}(0)$ is fixed at inception.

In Figure 1, we provide the numerical results obtained by the PDE scheme for different account values and different level of interest rate at inception.

Table 3. Product parameters of roll-up rider

c	w	A_i	t_{n+1}	Annuity frequency
%	3.5%	3.5% $\breve{A}(0)$	40	1/year

We observe that the value of the liability decreases when the initial interest rate rises up. This is consistent with our intuitive, since the higher the interest rate is, the less the annuities are worth.

In Figure 2, we compare the different numerical methods. For the HDR method, we used 50.000 scenarios with a time-step of 1 year. As pointed out in Bouchart and Warin (2010), the scheme S1 is biased from above while S2 is biased from below, which allows to construct a *confidence interval* of the form $[L^B_{[2]}(0), L^B_{[1]}(0)]$. We observe that this interval is very strait, around 1.5% difference. For a liability with such a long maturity, this is a significant improvement on the standard Longstaff-Schwartz approach.

In Figure 2, we also compare the value of the Bermudan-style liability with the account value at inception (red line). To be further protected from potential lapse waves or other financial risks, the insurer can ask the policyholder to pay an up-front charge fee, which equals to the difference between the liability and the account value (the asset) at inception (the difference between the blue and the red curves when $\breve{A} = \breve{A}(0) = 100$) to make sure that the balance sheet is in equilibrium initially. It is worthy to mention that once the account value (red line) rises up to the liability value (blue line) at time t_i, the rational policyholder should lapse the contract immediately.

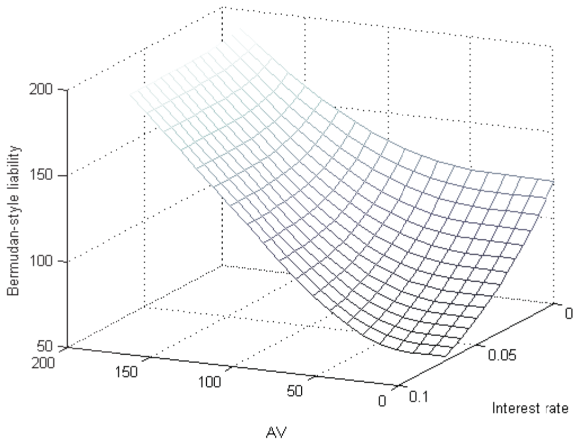

Fig. 1. Bermudan-style liability computed by PDE scheme for different account values (AV) and interest rate at inception.

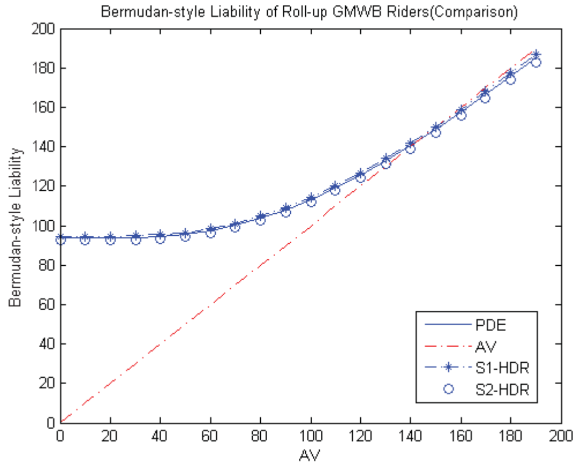

Fig. 2. Consistancy of PDE scheme and HDR scheme (S1 and S2) when $r(0) = 2\%$

4.2. Step-up rider. We now calculate the value of the Bermudan-style liability of a standard step-up rider. The principle product parameters are listed in Table (4), where the charge fees rate c and the withdrawal rate w are set to be the same as that of the roll-up rider. We also assume that the deferral period is zero and the benefit base is reset at each anniversary of the policy purchase.

Table 4. Product parameters of step-up rider at inception

c	w	$\bar{A}(0)$	t_{n+1}	Ratchet frequency	Annuity frequency
%	3.5%	$\breve{A}(0)$	40	1/year	1/year

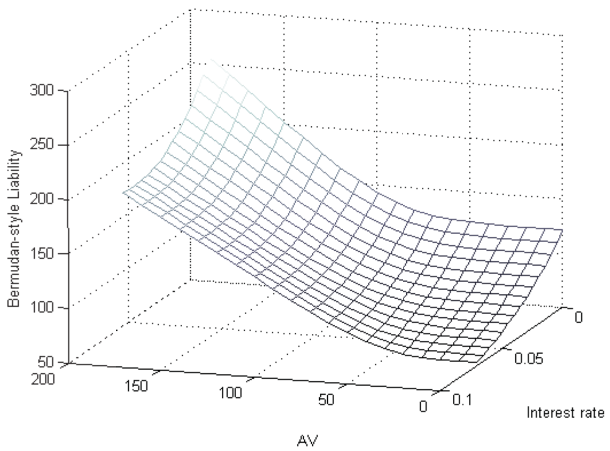

Fig. 3. Bermudan-style liability computed by PDE scheme for different AV0 ($) and $r(0)$ at inception

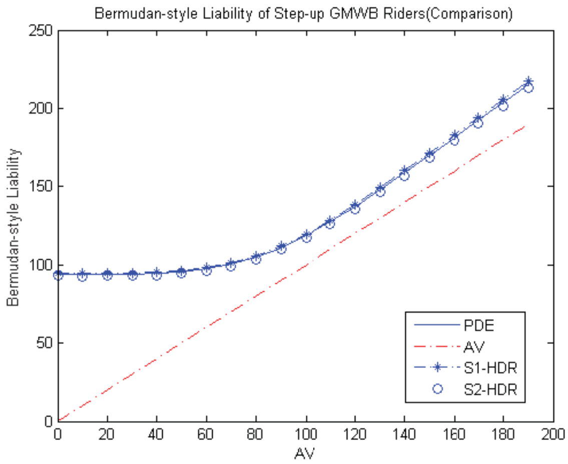

Fig. 4. Consistancy of PDE scheme and HDR scheme (S1 and S2) when $r(0) = 2\%$

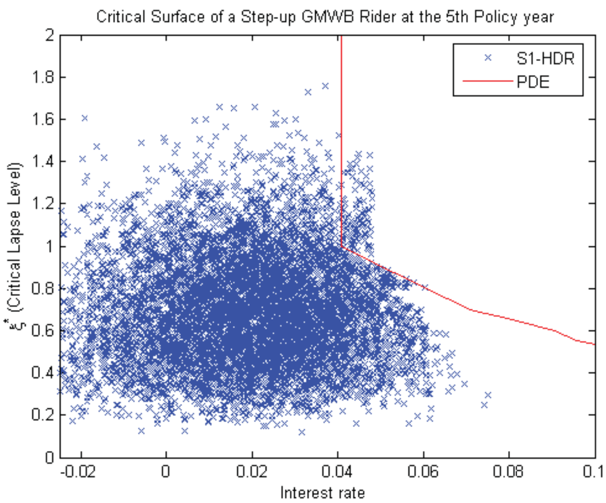

Fig. 5. Critical lapse surface at the fifth anniversary of the purchase of a step-up GMWB rider

Figure 3 shows the initial Bermudan-style liability computed by ADI scheme (implemented in C++) for different $\breve{A}$ and r at inception.

As roll-up riders, we also observe that the value of the liability decreases when the initial interest rate rises up. This is consistent with our intuitive, as the higher the interest rate is, the smaller the discount factor is and the less the annuities worth. In addition, when the account value is very small (e.g., $\breve{A} < 50$), the liability is not very sensible to the change of underlying. As in this case, the probability that the account value exceeds its high-water mark is very low and the liability value stays almost unchanged.

Figure 4 compares the numerical results computed by two methods: PDE and HDR (S1 and S2). For the HDR method used in this example, we also simulate 50000 scenarios with the step length of 1 year. As we mentioned in the previous example, the two algorithms S1 and S2 of HDR approach allow to construct a confidential interval $[L^B_{[2]}(0), L^B_{[1]}(0)]$ for the Bermudan-style liability. We observe that this interval is also very strait in Figure 4 (about 2%). As for the previous case, this is a significant improvement on the global polynomials regression approach.

In addition, the Bermudan-style liability is compared with the account value $\breve{A}$ (red line) at inception in Figure 4. To be further protected from potential lapse waves or other financial risks, the insurer can ask the policyholder to pay an up-front charge fee, which equals to the difference between the liability and the asset at inception. This up-front charge is much higher than that of roll-up riders due to the look-back property of the benefit base. It is worthy to mention that in this example, we calculate the liability of a single rider during 40 years, which may be an expensive policy for insurers. In practice, the average up-front charge fees based on a policy pool could be lower than that we show here.

Finally, we compare the "critical lapse surface" $x^*(t_i-, r)$ computed by HDR and PDE methods. According to Milevsky and Salisbury (2001) we know that at time t_i, if $\xi(t_i-) > x^*(t_i-, r_{t_i})$ and $\tilde{\ell}^B(t_i, r_{t_i}, 1) < 1$, then the rational policyholder should lapse the policy immediately. Otherwise, he should continue to hold the contract.

Figure 5 shows the estimated critical lapse surface $x^*(t_5-, r)$ of the step-up GMWB rider evaluated above.

We assume that the initial account value is 100$ and all other parameters of the policy are shown in Table 4. In Figure 5, we also record all the paths (simulated by scheme S1, blue points) corresponding to the cases where the policyholder should lapse the contract at time t_5-. It is not difficult to observe that most of these paths are located in the "lapse zone" identified by the critical lapse surface (red line) found by the PDE method. Finally, the properties of the critical surface implied by equation (25)-(26) are also verified in Figure 5.

Conclusion

In this paper, we formulate the evaluation problem of GMWB riders in a stochastic interest rate models, from the point of view of an insurer who does want to take any hedging risk. Then, two numerical methods, PDE and HDR schemes, are implemented to compute the super-hedging prices of the liability and the critical surface of these long-maturity and path-dependent products. In our numerical tests, the results obtained by these two methods are consistent with each other. Our numerical examples also indicate that some up-front charge fee may further protect the insurer to potential lapse wave or other financial risks. In addition, for step-up riders, we develop a "change of numeraire" technique to simplify and accelerate the pricing process efficiently.

References

1. J. Andreasen (1998). The Pricing of Discretely Sampled Asian and Lookback Options: A Change of Numeraire Approach, Working Paper.
2. M. Albizzati and H. Geman (1994). Interest Rate Risk Management and Valuation of the Surrender Option in Life Insurance Policies, *The Journal of Risk and Insurance*, 61 (4), pp. 616-637.
3. G. Barles, C. Daher and M. Romano (1995). Convergence of Numerical Schemes for Parabolic Equations Arising in Finance Theory, *Mathematical Models and Methods in Applied Sciences*, 5 (1), pp. 125-143.
4. D. Bauer, A. Kling and J. Russ (2006). A Universal Pricing Framework for Guaranteed Minimum Benefits in Variable Annuities, Working Paper.
5. D. Brigo and F. Mercurio (2006). Interest Rate Models: Theory and Paractice, 2nd, *Springer Finance, Heidelberg*.
6. V. Bally, G. Pages and J. Printems (2005). A Quantization Method for Pricing and Hedging Multi-dimensional American Style Options, *Mathematical Finance*, 15 (1).
7. M.J. Brennan and E.S. Schwartz (1977). The Valuation of American Put Options, *The Journal of Finance*, 32 (2), 449-462.
8. B. Bouchard (2007). Introduction to Stochastic Control of Mixed Diffusion Processes, Viscosity Solutions and Applications in Finance and Insurance. Lecture notes.

9. B. Bouchard and N. Touzi (2004). Discrete-time Approximation and Monte Carlo Simulation of Backward Stochastic Differential Equations, *Stochastic Processes and Their Applications*, 111, pp. 175-206.
10. B. Bouchard and N. Touzi (2010). Weak Dynamic Programming Principle for Viscosity Solutions, to appear in *SIAM Journal on Control and Optimization*.
11. B. Bouchard and X. Warin (2010). Monte-Carlo Valorisation of American Options: facts and new algorithms to improve existing methods, preprint submitted to Elsevier.
12. R. Carmona and N. Touzi (2008). Optimal Multiple Stopping and Valuation of Swing Options, *Mathematical Finance*, 18 (2), pp. 239-268.
13. M. Crandall, H. Ishii and P.L. Lions (1992). User's Guide to Viscosity Solutions of Second Order Partial Differential Equations, *Bull. Amer. Math. Soc.*, Vol. 27, pp. 1-67.
14. I. Craig and A. Sneyd (1988). An Alternating-Direction Implicit Scheme for Parabolic Equations with Mixed Derivatives, Comput. Math. Applie., 16 (4), pp. 341-350.
15. M. Dai, Y., K. Kwok, and J. Zhong (2007). Guaranteed minimum withdrawal benefit in variable annuities. *Mathematical Finance*.
16. J. Douglas, R. Kellogg and R. Varga (1983). Alternating Direction Iteration Methods for Space Variables, *Mathematics of Computation*, Vol. XVLI, No. 83.
17. E. Gobet, J.P. Lemor, X. Warin (2005). A Regression-Based Monte-Carlo Method to Solve Backward Stochastic Differential Equations, *Annals of Applied Probability*, 15 (3), pp. 2172-2002.
18. K.J. Hout and Foulon (2007). ADI Finite Difference Schemes for Option Pricing in the Heston Model with Correlation, Working Paper.
19. J. Hull and A. White (1990). Pricing Interest-Rate-Derivative Securities, *The Review of Financial Studies*, 3 (4), pp. 573-592.
20. IBBotson Associates (2007). Retirement Portfolio and Variable Annuity with Guaranteed Minimum Withdrawal Benefit, Working Paper.
21. I. Karatzas and S.E. Shreve (1998). Methods of Mathematical Finance, *Applications of Mathematics, Springer-Verlag*, Vol. 39.
22. T. Kluge (2002). Pricing Derivatives in Stochastic Volatility Models Using the Finite Difference Method, Thesis.
23. H.J. Kushner (1977). Probability Method for Approximations in Stochastic Control and for Elliptic Equations, Academic Press.
24. F. Longstaff and E. Schwartz (2001). Valuing American Options by Simulation: A simple least-squares, *Review of Financial Studies*, 1 (14), pp. 113-147.
25. M.A. Milevsky and T.S. Salisbury (2001). The Real Option to Lapse a Variable Annuity: Can Surrender Charges Complete the Market, Working Paper.
26. M.A. Milevsky and T.S. Salisbury (2005). Financial Valuation of guaranteed minimum withdrawal benefits, *Insurance: Mathematics and Economics*, 38, pp. 21-38.
27. S. Villeneuve and A. Zanette (2009). Comparison of Finite Difference Methods for Pricing American Options on Two Stocks, *Premia*, 9.

Appendix A. Change of numeraire

In this section, we prove the scaling property:

$$\ell(t,r,k\breve{a},k\bar{\bar{a}}) = k\ell(t,r,\breve{a},\bar{\bar{a}}),\ \forall k \in \mathbb{R}^+, \tag{34}$$

where ℓ could be ℓ^E defined in (6) or ℓ^B in (14).

Firstly, it is clear that both ℓ^E and ℓ^B satisfy (34) at the maturity t_{n+1} and $t_{n+1}-$. Once (34) holds for $t = t_{i+1}-$, it follows directly from:

$$\ell(t,r,\breve{a},\bar{\bar{a}}) = \mathbf{E}^{\mathbb{Q}}[\ell(t_{i+1}-, r(t_{i+1}-), \breve{A}(t_{i+1}-), \bar{\bar{A}}(t_{i+1}-)) \,|\, \breve{A}(t) = \breve{a}, \bar{\bar{A}}(t) = \bar{\bar{a}}], \forall t \in [t_i, t_{i+1}),$$

that (34) holds for every $t \in [t_i, t_{i+1})$. And by (20), (34) holds true for all t.

It follows that:

$$\ell(t,r,\breve{a},\bar{\bar{a}}) = \bar{\bar{a}}\tilde{\ell}(t,r,x), \tag{35}$$

where $x = \breve{a} / \bar{\bar{a}}$ and $\tilde{\ell}(t,r,x) := \ell(t,r,x,1)$.

By applying this change of numeraire technique, we can reduce the number of dimensions by one. It is worth to mention that our approach is different from the one developed by J. Andreasen (see [1]), as the latter use $\breve{a}$ rather than $\bar{\bar{a}}$ as the nominal.

Appendix B. Consistency of ADI scheme

For the ease of the reader, we explain here the ADI scheme proposed in [18][1], which is a direct generalization of the classic methods introduced in Douglas, Kellogg and Varga (1983). In the following, we consider a typical two dimensional parabolic partial differential equation,

$$-\frac{\partial u}{\partial t}-\sum_{i,j=1}^{2}a_{i,j}\frac{\partial^2 u}{\partial x_i \partial x_j}-\sum_{i=1}^{2}b_i\frac{\partial u}{\partial x_i}+cu=0 \qquad in\ \mathbb{R}^2\times[0,T), \tag{36}$$

together with some initial condition

$$u(T,\cdot)=u_T \qquad in\ \mathbb{R}^2, \tag{37}$$

where the matrix $A=\{a_{i,j}\}$ is symmetric. In the following, we assume that a comparison result holds for the above equation in the class of viscosity solutions with polynomial growth, which is easily checked under standard Lipschitz continuity assumptions on the coefficients and u_T, whenever u_T has polynomial growth. These conditions are satisfied in the application of Section 3.1.

First, we approximate the equation (36) by an equation set in a bounded smooth domain B_R :

$$-\frac{\partial u_R}{\partial t}-\sum_{i,j=1}^{2}a_{i,j}\frac{\partial^2 u_R}{\partial x_i \partial x_j}-\sum_{i=1}^{2}b_i\frac{\partial u_R}{\partial x_i}+cu_R=0 \qquad in\ [0,T)\times B_R, \tag{38}$$

with initial condition $u_R(T,\cdot)=u_T$ in B_R. We fix one of the two following boundary conditions on $[0,T)\times\partial B_R$,

$$u_R=u_T \quad or \quad \frac{\partial u_R}{\partial n}=\frac{\partial u_0}{\partial n}, \tag{39}$$

for some smooth function u_0 with first derivatives having polynomial growth, where n denotes the normal vector at the boundary. In Barles, C. Daher and M. Romano (1995), the authors show that in both cases, u_R converges to u as $R\to\infty$ and that this convergence is exponentially fast inside the domain.

The idea behind ADI technics is to approximate (38) in the form:

$$\frac{u^{n-1}-u^n}{\Delta t}-F_1u^{n-1}-F_2u^{n-1}-F_0u^{n-1}-\tilde{F}u^{n-1}=0, \tag{40}$$

for $1\leqslant n\leqslant N$, where N denotes the number of grid points in time and Δt denotes the time mesh size. The solution of the above is a vector $(u_{ij}^n)_{n,i,j}$ in $\mathbb{R}^{NIJ}$, where I and J denote the number of grid points in the two space directions. In the above, u^n stands for $(u_{ij}^n)_{ij}\in\mathbb{R}^{IJ}$ and the operators F_1, F_2, F_0 and $\tilde{F}$ denote respectively the discretization scheme of the following terms ($a_{12}=a_{21}$ as the matrix A is symmetric):

$$F_1u^n\to a_{11}\frac{\partial^2 u_R}{\partial x_1^2};\quad F_0u^n\to a_{12}\frac{\partial^2 u_R}{\partial x_1\partial x_2}+a_{12}\frac{\partial^2 u_R}{\partial x_2\partial x_1},$$

$$F_2u^n\to a_{22}\frac{\partial^2 u_R}{\partial x_2^2};\quad \tilde{F}u^n\to b_1\frac{\partial u_R}{\partial x_1}+b_2\frac{\partial u_R}{\partial x_2}-cu_R.$$

In our case of interest it takes the form:

$$u^{n-1}=(\mathrm{I}-\theta\Delta tF_2)^{-1}(\mathrm{I}-\theta\Delta tF_1)^{-1}(\mathrm{I}+(1-\theta)\Delta tF_1+(1-\theta)\Delta tF_2+\Delta tF_0+\Delta t\tilde{F}+\theta^2\Delta t^2F_1F_2)u^n, \tag{41}$$

where θ is a parameter satisfying $0\leqslant\theta\leqslant1$ and, for some space mesh sizes $\Delta x_1,\Delta x_2>0$ associated to the space grid,

$$(F_1u^n)_{i,j}=\frac{a_{11}}{\Delta x_1^2}(u_{i-1,j}^n-2u_{i,j}^n+u_{i+1,j}^n),\quad (F_2u^n)_{i,j}=\frac{a_{22}}{\Delta x_2^2}(u_{i,j-1}^n-2u_{i,j}^n+u_{i,j+1}^n), \tag{42}$$

[1] We slightly change the discretization method for the cross derivative term, which is shown explicitly later.

$$(F_0 u^n)_{i,j} = \frac{a_{12}^+}{\Delta x_1 \Delta x_2}(u_{i+1,j+1}^n + 2u_{i,j}^n + u_{i-1,j-1}^n - u_{i+1,j}^n - u_{i,j+1}^n - u_{i-1,j}^n - u_{i,j-1}^n) \tag{43}$$

$$+ \frac{a_{12}^-}{\Delta x_1 \Delta x_2}(u_{i-1,j+1}^n + 2u_{i,j}^n + u_{i+1,j-1}^n - u_{i+1,j}^n - u_{i,j+1}^n - u_{i-1,j}^n - u_{i,j-1}^n).$$

$$(\tilde{F} u^n)_{i,j} = \frac{b_1^+}{\Delta x_1}(u_{i+1,j}^n - u_{i,j}^n) + \frac{b_1^-}{\Delta x_1}(u_{i-1,j}^n - u_{i,j}^n) + \frac{b_2^+}{\Delta x_2}(u_{i,j+1}^n - u_{i,j}^n) + \frac{b_2^-}{\Delta x_2}(u_{i,j-1}^n - u_{i,j}^n) - c u_{i,j}^n, \tag{44}$$

with $a_{12}^+ = \max(0, a_{12})$ and $a_{12}^- = \max(0, -a_{12})$, see [16], [23] and [18].

According to Theorem 2.1 of Barles, C. Daher and M. Romano (1995), the solution of (40) converges as $\Delta t, \Delta x \to 0$ to the unique viscosity solution of (36), uniformly on each compact subset under three assumptions: stability, consistency and monotonicity of the scheme. In fact, it is easy to show that the stability and consistency are generally kept by ADI schemes (see Barles, C. Daher and M. Romano, 1995), and this will be the case for our cases of interest in Section 4.1. Thus we only check the monotonicity property.

We first check that the operator $(\mathrm{I} - \theta \Delta t F_1)^{-1}$ is monotone. Denote $(\hat{F}_1 u^n)_{i,j} = u_{i-1,j}^n + u_{i+1,j}^n$, then

$$(\mathrm{I} - \theta \Delta t F_1)^{-1} = [(1 + 2\frac{\theta \Delta t}{\Delta x_1^2})\mathrm{I} - (\theta \frac{\Delta t}{\Delta x_1^2})\hat{F}_1]^{-1}$$

$$= \frac{1}{1+2\alpha}[\mathrm{I} - \frac{\alpha}{1+2\alpha}\hat{F}_1]^{-1} = \frac{1}{1+2\alpha}\sum_{k=0}^{\infty}(\frac{\alpha}{1+2\alpha})^k \hat{F}_1^k, \tag{45}$$

where $\alpha = \frac{\theta \Delta t}{\Delta x_1^2}$. Since $\| \hat{F}_1 \| \leqslant 2$, we have $\| \frac{\alpha}{1+2\alpha}\hat{F}_1 \| \leqslant \frac{2\alpha}{1+2\alpha} < 1$. Therefore, $(\mathrm{I} - \theta \Delta t F_1)^{-1}$ is monotone. We can prove similarly that $(\mathrm{I} - \theta \Delta t F_2)^{-1}$ is monotone.

Finally, we need to prove the monotonicity of the operator $\mathrm{I} + (1-\theta)\Delta t F_1 + (1-\theta)\Delta t F_2 + \Delta t F_0 + \Delta t \tilde{F} + \theta^2 \Delta t^2 F_1 F_2$. Applying (42)-(43)-(44), we find that a sufficient monotonicity condition is that the following three inequalities hold true:

$$(1-\theta) a_{11}\frac{\Delta t}{\Delta x_1^2} \geqslant a_{12}\frac{\Delta t}{\Delta x_1 \Delta x_2} + 2\theta^2 a_{11} a_{22}\frac{\Delta t^2}{\Delta x_1^2 \Delta x_2^2} \tag{46}$$

$$(1-\theta) a_{22}\frac{\Delta t}{\Delta x_2^2} \geqslant a_{12}\frac{\Delta t}{\Delta x_1 \Delta x_2} + 2\theta^2 a_{11} a_{22}\frac{\Delta t^2}{\Delta x_1^2 \Delta x_2^2} \tag{47}$$

$$1 + 2a_{12}\frac{\Delta t}{\Delta x_1 \Delta x_2} \geqslant 2(1-\theta)\Delta t(\frac{a_{11}}{\Delta x_1^2} + \frac{a_{22}}{\Delta x_2^2}) + \Delta t(\frac{b_1}{\Delta x_1} + \frac{b_2}{\Delta x_2} + c). \tag{48}$$

In our case of interest, see Section 3.1., if the correlation between the account value and the interest rate (proportional with a_{12}) is not too strong, we can keep the above conditions satisfied by carefully choosing θ, Δt, Δx_1 and Δx_2.

Appendix C. Simulation algorithm

In practice, the naive Euler scheme is not very efficient to simulate the short rate process under Hull-White model. In fact, the dynamic of the instantaneous short rate r can be written as:

$$r(t) = x(t) + \phi(t), \quad r(0) = \phi(0) = r_0,$$

where $\phi(t)$ is a deterministic function calibrated to market data and the Ornstein-Uhlenbeck process x satisfies:

$$dx(t) = -a x(t) dt + \sigma_r dW(t), \quad x(0) = 0.$$

Rather than simulating r directly, we simulate the process $x(t)$ on discrete time points i as following. Define $\Delta t = t_{i+1} - t_i$ and $\Delta W_{i+1} = W(t_{i+1}) - W(t_i)$,

$$x(t_{i+1}) = x(t_i)\exp(-a\Delta t) + \sigma_r \sqrt{\frac{1-\exp(-2a\Delta t)}{2a}}\Delta W_{i+1}.$$

Once the process x is simulated, we can get the process r by simply adding the deterministic function. The discount factor is then approximated according to:

$$D_{t_i}^{t_{i+1}} \simeq \tilde{D}_{t_i}^{t_{i+1}} := \exp(-\frac{r(t_i)+r(t_{i+1})}{2}\Delta t).$$

In practice, this algorithm is very efficient. Even for the discrete step length of 1 year, the pricing result is still precise (see our numerical tests).