Monte Carlo Variance Reduction by conditioning for pricing with underlying a continuous-time finite state Markov process

Juan Miguel Montes
Department of Mathematics
LMU University Munich
montes.jmiguel@gmail.com*

Valentina Prezioso
Market Risk Management
Intesa Sanpaolo, Milan
valentinaprezioso@gmail.com†

Wolfgang J.Runggaldier
Department of Mathematics
University of Padova
runggal@math.unipd.it‡

June 14, 2014

Abstract

We consider the pricing of derivatives when the evolution of the underlying is given by a continuous time finite-state Markov chain. We present a semi-analytic approach that consists in: i) simulating the number of transitions of the underlying up to a given time horizon, ii) computing via an explicit analytic formula the derivative price for each simulated number of transitions and iii) approximating the actual price by the empirical average over the values computed in ii). This corresponds to a Monte Carlo approach with variance reduction by conditioning and, with respect to a plain Monte Carlo, it leads thus to a smaller variance in addition to more precise values. The method is in particular applied to path dependent derivatives and numerical results are presented and discussed.

Keywords: Derivative pricing, path dependent derivatives, continuous time Markov chains, Monte Carlo simulation, variance reduction by conditioning.

AMS classification: 91G20, 91G60, 60J28, 65C05.

1 Introduction

Traditional pricing formulae concern market models either in continuous or discrete time. There are however situations, in which it is more natural to model the evolution of the underlying as a continuous time Markov chain (CTMC) that combines features of both, continuous as well as

*Postal address: Mathematisches Institut, LMU Muenchen, Theresienstrasse 39, D-80333 Munich, Germany

†Market Risk Management, Intesa Sanpaolo, Piazza Paolo Ferrari 10, 20121 Milan, Italy. The views expressed here are those of the author and do not represent the opinions of her employer. They are not responsible for any use that may be made of these contents in any circumstance.

‡Dipartimento di Matematica, Università di Padova, Via Trieste 63, 35121-Padova, Italy.

discrete time models; in fact, while being a model in continuous time, it has an embedded Markov chain that corresponds to a discrete time one. A CTMC has the following relevant features:

i) it jumps at random points in time;

ii) the number of jumps in a fixed time interval is random.

Market models based on CTMC may turn out to be useful models in the following specific situations:

a) In the pricing of bonds and interest rate derivatives, where the "underlying" can be considered as given by the short rate of interest. This covers, similarly, all those situations where the "underlying" is a rate (e.g. exchange rate) or an intensity (e.g. default intensity, see Remark 2.1 below). Since the actual evolution of rates or intensities may follow a pure jump process, a CMTC model appears to be more natural than a diffusion-type model or even a jump-diffusion model. For the latter the computations may in fact be rather involved (see e.g. section 6 in [2], see also [12]).

b) In the context of small time scale models, where prices vary by tick size at random times in reaction to trading or to the arrival of significant new information. Here the jumps are naturally more frequent than in the case of a).

c) A CTMC modeling might also intervene in situations, where the evolution of the underlying is modeled by a Poisson process or a market point process model with a finite number of marks, to which a (deterministic or stochastic) evolution process is superimposed. (For a related situation see e.g. [2] where, as mentioned under a) above, in section 6 a pricing methodology is presented for derivatives having as underlying the short rate that satisfies an affine jump-diffusion model; the solution is given in explicit form, but is difficult to compute in practice). As an example consider an Insurance context with the risk process as underlying; the latter is piecewise deterministic (premium payments) with the random part consisting of a market point process, where the jumps are due to the arrival of a claim (see e.g. [1], [14]). Another example may be in Credit Risk when the underlying is affected by rating transitions or when a sequence of partial defaults (multiple defaults) may occur, whereby, at each rating transition or default time, the value of the firm changes by a random amount (see [16]). For a specific application of CTMCs to Credit Risk see Remark 2.1 below.

In this paper we shall consider a generic underlying process X_t, of which the evolution is given by a CTMC and that may be multivariate and/or time inhomogeneous. We derive a method that allows for an actual computation of the price also of path dependent claims. It is a mixture of an analytic expression and a Monte Carlo (MC) simulation and corresponds to an MC method with variance reduction by conditioning. As we shall illustrate in section 5 below where we consider barrier and lookback options as examples, our method extends rather straightforwardly to path dependent options; such an extension is not easily achieved by other models and approaches. In fact, while for a Black and Scholes model there exist closed formulas for barrier options, for more general diffusion models the pricing of path dependent options becomes more involved, even if one uses a Monte Carlo simulation approach; for the latter a preliminary time discretization is required not only for the underlying itself but in general also for possible supplementary state variables summarizing the path dependence (see e.g. section 6.2.3 in [8]).

Although for the computation of the expectation in the pricing formulas of the various claims in our CTMC market model one can always resort to a plain Monte Carlo (MC) simulation, a plain MC approach has various drawbacks as it will also appear from the numerical results below: in addition to a possibly large variance, it may also lead to biased results, unless one performs an extremely large number of simulation runs. The MC-based approach proposed in this paper and that we shall call "conditional MC" approach, results from first conditioning on the number $\nu_{t,T}$ of transitions of the underlying CTMC process X_t between the evaluation time t of the claim and its maturity T. We shall show that the conditional expectation, given $\nu_{t,T}$, allows for an explicit analytic solution also in the case of claims that are not simple claims and this is by itself one of the other contributions of the paper. In this way there remains only to possibly simulate $\nu_{t,T}$ (an explicit, but cumbersome formula for computing analytically the distribution of $\nu_{t,T}$ can be found in [13] for X_t scalar; see also Remark 4.8). The fact that a fair portion of what is computed by simulation in plain MC is here computed explicitly, makes it intuitively clear that one can thus obtain more precise results.

The outline of the paper is as follows. In section 2 we describe more specifically our CTMC market model and formulate the problem more explicitly. The basic idea of the proposed pricing approach is then illustrated in section 3 , while in section 4 we describe the details of the approach. In section 5 we then show how to apply the conditional MC approach to the pricing of barrier options and of path dependent claims and in section 6 we present numerical results and comparisons. For comparison purposes, and only for this purpose, in subsection 6.1 we consider a CTMC model for the short rate of interest that is obtained by a specific quantization approach from an affine diffusion model of the short rate considered as a benchmark. Indeed, for affine diffusion models bond prices admit an explicit analytical expression and this allows for a comparison of the validity of our approach in spite of the approximation due to the quantization. In all the remaining part of the paper CTMCs are considered as a model in itself.

2 The model and the problem formulation

Let X_t be a CTMC with values in $\{x^1, \cdots, x^N\}$ and transition intensity matrix Q that we assume for the moment to be time homogeneous. Below we shall occasionally identify x^i with i, $(i = 1, \cdots, N)$. By a transition intensity matrix Q we mean here the matrix, where the off-diagonal elements $q_{i,j}$ represent the transition rates from state i to state j, while the diagonal elements $q_{i,i}$ are equal to zero. This is different from the infinitesimal generator matrix, where $q_{i,i} = -\sum_{j \neq i} q_{i,j}$, while $q_{i,j}$ remain the same for $i \neq j$ (notice that the information content is the same in both types of matrices).

For a given maturity T let there be given a claim that, in order to better describe our approach, we take here as a simple claim of the following form (the extension to more general, in particular path dependent, claims is then given in section 5).

$$H = H(X_T) = \left[H(x^1), \cdots, H(x^N)\right]' \quad \text{where } ' \text{ denotes transposition} \tag{1}$$

and which, given the finite state assumption for X_t, can be represented as a vector. Denoting by τ_n the random time at which the $n-$th transition of the chain X_t takes place, put, for simplicity, $X_n := X_{\tau_n}$ so that $X_s = X_n$ for $s \in [\tau_n, \tau_{n+1})$. Assume also that the short rate of interest r_t is related to the underlying X_t in the sense that $r_t = r(X_t)$ so that r_t can undergo a change only at

the time points τ_n and we let $r_n := r_{\tau_n} = r(X_n) \in \{r^1, \cdots, r^N\}$. Let $\tilde{P} \sim P$ be an equivalent (to the physical measure P) martingale measure that is used for pricing and that will typically result from a calibration to the market. The arbitrage-free price at $t < T$ of the claim H, when $X_t = x^i$ is then

$$\Pi_i(t) = E^{\tilde{P}} \left\{ \exp\left[-\int_t^T r_s ds \right] H \mid X_t = x^i \right\} . \tag{2}$$

Denoting by $\nu_t := \sup\{n \mid \tau_n \leq t\}$ the number of transitions of X_t up to a given time t so that $\nu_{t,T} = \nu_T - \nu_t$, one may rewrite (2) as

$$\begin{aligned} \Pi_i(t) &= E^{\tilde{P}} \Big\{ \exp[r_t(t - \tau_{\nu_t}) \exp\left[-\sum_{j=\nu_t}^{\nu_T - 1} r_j(\tau_{j+1} - \tau_j) - r_T(T - \tau_{\nu_T}) \right] H(X_T) \mid X_t = i \Big\} \\ &= \exp[r_t(t - \tau_{\nu_t})]\, E^{\tilde{P}} \left\{ \exp\left[-\sum_{j=\nu_t}^{\nu_T - 1} r_j(\tau_{j+1} - \tau_j) - r_T(T - \tau_{\nu_T}) \right] H(X_T) \mid X_t = i \right\} \end{aligned} \tag{3}$$

where we have used the fact that τ_{ν_t} is known at time t and so $\exp[r_t(t - \tau_{\nu_t})]$ can be taken as being deterministic. Since we may without loss of generality assume $t = \tau_{\nu_t}$, our problem can now be described as follows:

Problem: Compute

$$V_{H,t,T}(X_t) = E^{\tilde{P}} \left\{ \exp\left[-\sum_{j=\nu_t}^{\nu_T - 1} r_j(\tau_{j+1} - \tau_j) - r_{\nu_T}(T - \tau_{\nu_T}) \right] H(X_T) \mid X_t \right\} \tag{4}$$

where $V_{H,t,T}(X_t)$ denotes the vector with components $V_{H,t,T}(X_t)\mathbf{1}_{\{X_t = x^i\}}$ that occasionally we shall also denote by $V_{H,t,T}(X_t)_{|X_t = x^i}$ and where we may also write

$$H(\cdot) := H_0(\cdot) = \sum_{i=1}^N w_i^0 \mathbf{1}_{\{\cdot = x^i\}},\ x^i \in E,\ \ w_i^0 \in \mathbb{R} \tag{5}$$

with w_i^0 representing the value $H(x^i)$. Notice that $V_{H,t,T}(X_t)_{|X_t = x^i}$ corresponds to $\Pi_i(t)$ in (2) and (3) and we shall use this more complex symbol below whenever we want to make explicit all the ingredients for the price $\Pi_i(t)$.

A theory of financial markets, when the underlying follows a time homogeneous CTMC has been developed in [11] according to which the price $\Pi_i(t)$ of a simple claim H can be computed as

$$\Pi_i(t) = \left[\exp\{(Q - R)\,(T - t)\}\, H \right]_i \tag{6}$$

where $[z]_i$ denotes the $i-$th component of the vector z, R is the diagonal $(N \times N)-$matrix with elements r^i $(i = 1, \cdots, N)$ and Q is assumed here to be time homogeneous. The main purpose of the present study is to extend the basic theory and to provide an efficient semianalytic method to compute the price $\Pi_i(t)$ for these extensions. The extensions concern basically path dependent claims that in many cases can be reduced to simple claims at the expense however of augmenting the dimension of the underlying and occasionally, like in the case of barrier options, leading to a time inhomogeneous process even when starting from a time homogeneous one. In the time

inhomogeneous case the explicit formula (6) cannot be applied and it becomes cumbersome to apply it also in the time homogeneous case when the underlying is multivariate. Some of the extensions and the corresponding solutions for our Problem will be discussed in Section 5 below. Here we add the following remark concerning possible extensions to a Credit Risk setup.

Remark 2.1. *A multivariate CTMC may also arise in defaultable bond pricing, where the underlying is $X_t = r_t$ and the price is given by*

$$\Pi(t) = \mathbf{1}_{\{\tau > t\}} E^{\tilde{P}} \left\{ \exp\left[-\int_t^T (r_s + \lambda_s)\, ds \right] \mid \mathcal{F}_t \right\} \tag{7}$$

with τ denoting the default time and λ_t the default intensity. The pair $(r_t,\, \lambda_t)$ may then be taken to form a bivariate CTMC, i.e. $X_t = r_t, , Y_t = \lambda_t$. To take correlation into account, one may put $r_t = r(X_t, Z_t)$ and $\lambda_t = \lambda(Y_t, Z_t)$ with X_t, Y_t, Z_t three independent CTMCs. One can extend this further to a multicurve setup for the term structure of interest rates as it was introduced after the financial crisis $2008 - 2010$ *and where λ_t can more generally be considered as a short rate spread due not only to credit risk but also other risks such as liquidity (see e.g.[7]). Notice finally that the pricing of barrier options, as described for CTMCs in subsection 5.1 below, can be conveniently adapted also to the pricing of credit risky products within the structural approach.*

We close this section by mentioning the notation that we shall use for the time inhomogeneous and multivariate case. In the time inhomogeneous case, instead of a fixed transition intensity matrix Q, we shall consider a sequence $Q(n)$ where n refers to the $n-$th transition time τ_n of the underlying X_t. For the multivariate case let us take the special case of a bivariate CTMC (X_t, Y_t) with $X_t \in \{x^1, \cdots, x^N\}$, $Y_t \in \{y^1, \cdots, y^M\}$. Letting τ_n denote the $n-$th jump time of the pair (X_t, Y_t), assume again that the short rate r_t changes only at the time points τ_n, i.e. assume that $r_n := r_{\tau_n} = r(X_n, Y_n)$ with $(X_n, Y_n) = (X_{\tau_n}, Y_{\tau_n})$. For the (time inhomogeneous) multivariate transition intensity matrix we shall put

$$Q(n) = \left\{ q_{(i,h),(j,k)}(n) \right\}_{\substack{i,j=1,\cdots,N \\ h,k=1,\cdots,M}}$$

3 The underlying methodology

As mentioned in the Introduction, in the general multivariate and time inhomogeneous cases one can always resort to a Monte Carlo (MC) simulation to compute the expectation in (2) or, more specifically, in (4), but this leads to various drawbacks. The "conditional MC" approach proposed in this paper results from first conditioning on the number $\nu_{t,T}$ of transitions of the chain X_t between t and T, namely by rewriting the expression in (2) as

$$\begin{aligned} \Pi_i(t) &= E^{\tilde{P}} \left\{ e^{-\int_t^T r_s ds} H(X_T) \mid X_t = i \right\} \\ &= E^{\tilde{P}} \left\{ E^{\tilde{P}} \left\{ e^{-\int_t^T r_s ds} H(X_T) \mid \nu_{t,T}, X_t = i \right\} \mid X_t = i \right\} . \end{aligned} \tag{8}$$

We shall show that the inner expression allows for an explicit analytic computation also in the case of claims that are not simple claims and X_t is time inhomogeneous.

First recall that a plain MC approach consists in simulating the successive transition times τ_n of the chain X_t and the values X_n of X_t at τ_n and then averaging over the values obtained in each simulation run for the argument in the expectation of the right hand side of (4). To be precise, consider for a moment again a time homogeneous chain with transition intensity matrix Q. Putting $q_i := \sum_{i \neq j} q_{i,j}$, one has that, if $X_n = X_{\tau_n} = x^i$, then the inter-jump times $\tau_{n+1} - \tau_n$ are exponentially distributed with parameter q_i and the transition probability of the embedded chain, namely the probability that $X_{\tau_{n+1}} = x^j \neq x^i$ is given by $p_{i,j} = \frac{q_{i,j}}{q_i}$ implying, as it should be, that $p_{i,i} = 0$. Given these values for q_i and $p_{i,j}$, one can then simulate the successive values of τ_n and of the corresponding X_n.

Our problem is now to compute $V_{H,t,T}(X_t)$ in (4) that we shall call *Prototype product* since various more general derivatives can be obtained either as particular cases or as linear combinations of prototype products. This is in particular the case when $X_t = r_t$ and one has to deal with interest rate derivatives where e.g. Caps and Swaptions can be expressed as linear combinations of prototype products (see [13]). Notice also the analogy between the prototype product and Arrow-Debreu prices.

As already mentioned, we shall compute (4) by conditioning first on the number of jumps $\nu_{t,T}$, i.e. in line with (8) we shall compute a conditional price $V^{\nu_{t,T}}_{H,t,T}(X_t)$ according to

$$V^{\nu_{t,T}}_{H,t,T}(X_t) = E^{\tilde{P}} \left\{ \exp \left[- \sum_{j=\nu_t}^{\nu_T - 1} r_j(\tau_{j+1} - \tau_j) - r_{\nu_T}(T - \tau_{\nu_T}) \right] H(X_T) \mid X_t, \nu_{t,T} \right\} \tag{9}$$

so that

$$V_{H,t,T}(X_t) = E^{\tilde{P}} \left\{ V^{\nu_{t,T}}_{H,t,T}(X_t) \mid X_t \right\} . \tag{10}$$

It follows that, if we are able to compute in an exact analytic way $V^{\nu_{t,T}}_{H,t,T}(X_t)$ in (9), then we may compute the expectation in (10) by MC simulations of only $\nu_{t,T}$. It turns out that the exact analytic computation of $V^{\nu_{t,T}}_{H,t,T}(X_t)$ in (9) is made difficult by the presence of the random variables r_{ν_T} and τ_{ν_T} both in the sum as well as in the last term of the exponential in (9) thus preventing the use of a double conditioning to separate the expectation of the exponential of the sum from that of the last term. As we shall show below, it is however possible to compute in an exact analytic way both an upper and a lower bound on $V^{\nu_{t,T}}_{H,t,T}(X_t)$, which we shall denote by $\bar{V}^{\nu_{t,T}}_{H,t,T}(X_t)$ and $\underline{V}^{\nu_{t,T}}_{H,t,T}(X_t)$ respectively and that are defined as follows. First, let

$$V^{\nu_{t,T},0}_{H,t,T}(X_t) = E^{\tilde{P}} \left\{ \exp \left[- \sum_{j=\nu_t}^{\nu_T - 1} r_j(\tau_{j+1} - \tau_j) \right] H(X_T) \mid X_t, \nu_{t,T} \right\} \tag{11}$$

and

$$V^{\nu_{t,T},1}_{H,t,T}(X_t) = E^{\tilde{P}} \left\{ \exp \left[- \sum_{j=\nu_t}^{\nu_T} r_j(\tau_{j+1} - \tau_j) \right] H(X_T) \mid X_t, \nu_{t,T} \right\} \tag{12}$$

denoting, furthermore, by $V^{\nu_{t,T},0}_{H,t,T}(\underline{x})$ and $V^{\nu_{t,T},1}_{H,t,T}(\underline{x})$ the vectors with components $V^{\nu_{t,T},0}_{H,t,T}(X_t)_{|X_t = x^i}$ and $V^{\nu_{t,T},0}_{H,t,T}(X_t)_{|X_t = x^i}$ respectively. Notice that in (11) the exponential does not include the term $-r_{\nu_T}(T - \tau_{\nu_T})$ that is however present in (9) and this makes the argument in the expectation in the right hand of (11) a.s. larger than the corresponding one in (9). On the other hand, the last

term $-r_{\nu_T}(\tau_{\nu_T+1}-\tau_{\nu_T})$ in the sum on the right hand side of (12) is a.s. smaller than $-r_{\nu_T}(T-\tau_{\nu_T})$ and this causes the argument in the expectation in (12) to be a.s. smaller than the corresponding one in (9). Then, put

$$\bar{V}^{\nu_{t,T},T}_{H,t,T}(X_t) := \max\left[V^{\nu_{t,T},T,0}_{H,t,T}(X_t), V^{\nu_{t,T},T,1}_{H,t,T}(X_t)\right] \tag{13}$$

$$\underline{V}^{\nu_{t,T},T}_{H,t,T}(X_t) := \min\left[V^{\nu_{t,T},T,0}_{H,t,T}(X_t), V^{\nu_{t,T},T,1}_{H,t,T}(X_t)\right] \tag{14}$$

considering, analogously to $V^{\nu_{t,T},T,i}_{H,t,T}(\underline{x})$, $i=0,1$, also here the vectors $\bar{V}^{\nu_{t,T},T}_{H,t,T}(\underline{x})$ and $\underline{V}^{\nu_{t,T},T}_{H,t,T}(\underline{x})$. We also recall that $X_T = X_{\nu_T}$ in (11) and $X_T = X_{\nu_T+1}$ in (12).

Remark 3.1. *Since for each $X_t = x^i$ the argument of the expectation in the right hand side of (12) is a.s. smaller than the corresponding one in (11), one might wonder why we did not define the upper and lower bounds directly as $V^{\nu_{t,T},T,0}_{H,t,T}(X_t)$ and $V^{\nu_{t,T},T,1}_{H,t,T}(X_t)$ respectively. The reason is that it is the norm of the entire vector $V^{\nu_{t,T},T,1}_{H,t,T}(\underline{x})$ determined according to (12) that is guaranteed to be smaller than or equal to that determined according to (11), but the individual components of $V^{\nu_{t,T},T,1}_{H,t,T}(\underline{x})$ may not necessarily be smaller than those of $V^{\nu_{t,T},T,0}_{H,t,T}(\underline{x})$. In fact, in our computations the individual components of $V^{\nu_{t,T},T,0}_{H,t,T}(\underline{x})$ and $V^{\nu_{t,T},T,1}_{H,t,T}(\underline{x})$ turned out to exhibit an initial oscillatory behavior. It is however always the case that the actual price $V_{H,t,T}(X_t)$ in (4) (see also (10)) belongs to the interval with extreme points given by the corresponding values of $\bar{V}^{\nu_{t,T},T}_{H,t,T}(X_t)$ and $\underline{V}^{\nu_{t,T},T}_{H,t,T}(X_t)$.*

Instead of computing analytically the exact value of $V^{\nu_{t,T},T}_{H,t,T}(X_t)$ in (9), we shall determine analytically the approximation given by the midpoint between $\bar{V}^{\nu_{t,T},T}_{H,t,T}(X_t)$ and $\underline{V}^{\nu_{t,T},T}_{H,t,T}(X_t)$, namely

$$V^{\nu_{t,T},T}_{H,t,T}(X_t) \sim \tilde{V}^{\nu_{t,T},T}_{H,t,T}(X_t) := \frac{1}{2}\left(\bar{V}^{\nu_{t,T},T}_{H,t,T}(X_t) + \underline{V}^{\nu_{t,T},T}_{H,t,T}(X_t)\right) \tag{15}$$

which, combined with (10), leads to the approximation

$$V_{H,t,T}(X_t) \sim E^{\tilde{P}}\left\{\tilde{V}^{\nu_{t,T},T}_{H,t,T}(X_t)\right\} := \frac{1}{2}E^{\tilde{P}}\left\{\bar{V}^{\nu_{t,T},T}_{H,t,T}(X_t) + \underline{V}^{\nu_{t,T},T}_{H,t,T}(X_t)\right\} \tag{16}$$

where the expectation is with respect to $\nu_{t,T}$.

Remark 3.2. *Concerning the accuracy of the approximation in (15) and (16) it will be shown in Corollary 4.7 below that the norm of the difference $\bar{V}^{n}_{H,t,T}(\underline{x}) - \underline{V}^{n}_{H,t,T}(\underline{x})$ tends to zero for n tending to infinity so that the approximation can be expected to be rather precise in situations where there are many transitions of X_t. This is also in line with the fact that, as can be easily shown, the norm of the difference $\bar{V}^{\nu_{t,T},T}_{H,t,T}(\underline{x}) - \underline{V}^{\nu_{t,T},T}_{H,t,T}(\underline{x})$ is bounded from above by $\bar{\rho} := \max_{i\le N}\frac{r_i}{q_i}$, which is in general a small value, especially if the transitions of X_t are frequent. In section 6 we perform some numerical tests to assess the accuracy of the approximation of $V_{H,t,T}(X_t)$ by $E^{\tilde{P}}\left\{\tilde{V}^{\nu_{t,T},T}_{H,t,T}(X_t)\right\}$.*

The expectation with respect to $\nu_{t,T}$ in (16) will be computed by MC simulations. In order to simulate just the number of jumps $\nu_{t,T}$, for each simulation run we cannot avoid determining also the successive values of τ_n and X_n, but we do not have to record them. In the next section we shall now present our approach to determine in an analytically exact way the value of $\tilde{V}^{\nu_{t,T},T}_{H,t,T}(X_t)$ for each given value of $\nu_{t,T}$.

4 The algorithm to implement our methodology

4.1 Preliminaries

Given t and ν_t (recall that, see description following (3), we had assumed without loss of generality that $t = \nu_t$ and so $X_t = X_{\nu_t}$) fix an $M \in \mathbb{N}$ and define the sequence of functions $H_n(\cdot)$, $n = 0, \cdots, M$ recursively as follows: $H_0(\cdot)$ is given by the Prototype payoff, namely (see (5)

$$H_0(\cdot) = \sum_{i=1}^{N} w_i^0 \mathbf{1}_{\{\cdot = x^i\}} \tag{17}$$

and, for $0 < n \leq M$,

$$H_n(X_{\nu_t+M-n}) = E^{\tilde{P}} \left\{ e^{-r_{\nu_t+M-n}(\tau_{\nu_t+M-n+1} - \tau_{\nu_t+M-n})} H_{n-1}(X_{\nu_t+M-n+1}) \mid X_{\nu_t+M-n} \right\} . \tag{18}$$

This definition has been inspired by an approach described in [6], where the authors derive the analog in discrete time of the continuous time affine term structure models, and here we use the fact that X_t is Markov and that, given X_{ν_t}, the distribution of the inter-arrival time $\tau_{\nu_t+1} - \tau_{\nu_t}$ depends only on X_{ν_t}. Notice also that from (18) it follows that

$$H_M(X_{\nu_t}) = E^{\tilde{P}} \left\{ \exp \left[- \sum_{m=\nu_t}^{\nu_t+M-1} r_m(\tau_{m+1} - \tau_m) \right] H_0(X_{\nu_t+M}) \mid X_{\nu_t} \right\} . \tag{19}$$

Since X_t takes a finite number of possible values, just as with $H(\cdot) = H_0(\cdot)$, also each of the $H_n(x)$ can be represented as a vector. More precisely, letting $\underline{x} = [x^1, \cdots, x^N]'$, we have

$$H_n(\underline{x}) = [w_1^n, \cdots, w_N^n]' \tag{20}$$

with w_i^n representing the value $H_n(x^i)$.

Remark 4.1. *Notice the immediate relationship between the vectors $H_n(\underline{x})$ defined recursively in (18) above and the basic quantities of our MC-with-conditioning approach, namely the vectors $V_{H,t,T}^{\nu_{t,T},0}(\underline{x})$ and $V_{H,t,T}^{\nu_{t,T},1}(\underline{x})$ with components defined for each value of $X_t = x^i$, $i = 1, \cdots, N$ according to (11) and (12) respectively. From (11), (12), the equality $H(X_T) = H_0(X_T)$, the representation of $H_M(\cdot)$ in (19), and recalling that $\nu_{t,T} = \nu_T - \nu_t$, we have in fact that*

$$V_{H,t,T}^{\nu_{t,T},0}(\underline{x}) = H_{\nu_{t,T}}(\underline{x}) \quad ; \quad V_{H,t,T}^{\nu_{t,T},1}(\underline{x}) = H_{\nu_{t,T}+1}(\underline{x}) . \tag{21}$$

This requires the sequence $H_n(\cdot)$ to be computed for a sufficiently large value of M so that there is always an $n \leq M$ for each value of $\nu_{t,T}$ that might occur. In practice (see the Algorithm in subsection 4.2 below) one takes M to be the largest value of $\nu_{t,T}$ that has occurred in the various simulation runs.

From the definitions of $\bar{V}_{H,t,T}^{\nu_{t,T}}(\underline{x})$ and $\underline{V}_{H,t,T}^{\nu_{t,T}}(\underline{x})$ with components defined for each value of $X_t = x^i, i = 1, \cdots, N$ according to (13) and (14), from that of $\tilde{V}_{H,t,T}^{\nu_{t,T}}(X_t)$ in (15), from the equalities (21), and recalling that $X_T = X_{\nu_T}$ in (11) and $X_T = X_{\nu_T+1}$ in (12), one obtains immediately the following corollary where, in line with the other analogous vectors defined previously, $\tilde{V}_{H,t,T}^{\nu_{t,T}}(\underline{x})$ denotes the vector with components $\tilde{V}_{H,t,T}^{\nu_{t,T}}(X_t)_{|X_t=x^i}$, $i = 1, \cdots, N$.

Corollary 4.2. *We have*

$$\bar{V}^{\nu_{t,T}}_{H,t,T}(\underline{x}) = \max\left[H_{\nu_{t,T}}(\underline{x}), H_{\nu_{t,T}+1}(\underline{x})\right] \quad ; \quad \underline{V}^{\nu_{t,T}}_{H,t,T}(\underline{x}) = \min\left[H_{\nu_{t,T}}(\underline{x}), H_{\nu_{t,T}+1}(\underline{x})\right]$$

$$\tilde{V}^{\nu_{t,T}}_{H,t,T}(\underline{x}) = \tfrac{1}{2}\left(\bar{V}^{\nu_{t,T}}_{H,t,T}(\underline{x}) + \underline{V}^{\nu_{t,T}}_{H,t,T}(\underline{x})\right) = \tfrac{1}{2}\left(H_{\nu_{t,T}}(\underline{x}) + H_{\nu_{t,T}+1}(\underline{x})\right) .$$

This Corollary shows clearly the relevance of the functions $H_n(\cdot)$ computed recursively in (18).

4.2 Derivation of the Algorithm

We shall now derive an easily implementable procedure to analytically compute the functions $H_n(\cdot)$. Considering the general case, where the transition intensity of the process X_t is given as a sequence of $Q-$matrices $Q(n) = \{q_{i,j}(n)\}$, define the sequence of matrices $\widetilde{Q}(n)$ as

$$\widetilde{Q}(n) = (\tilde{q}_{i,j}(n))_{1\le i,j\le N} \quad \text{with} \quad \tilde{q}_{i,j}(n) = \begin{cases} \frac{q_{i,j}(n)}{r^i + q_i(n)} & i \neq j \\ 0 & i = j \end{cases} \tag{22}$$

where $r^i = r(X)$ when $X = x^i$ and $q_i(n) = \sum_{j\neq i} q_{i,j}(n) = \sum_{i,j} q_{i,j}(n)$. We have now

Proposition 4.3. *Starting from the given $H_0(\cdot)$ (see (5)), the functions $H_n(\cdot)$ in (18) can be computed recursively by the following matrix multiplication*

$$H_n(\underline{x}) = \widetilde{Q}(n)\, H_{n-1}(\underline{x})\,. \tag{23}$$

Proof. Fixing a generic n, by the representation of $H_n(\underline{x})$ in (20) we have to prove that

$$[w^n_1, \cdots, w^n_N]' = \widetilde{Q}(n)\,[w^{n-1}_1, \cdots, w^{n-1}_N]'$$

For this purpose it suffices to show that for a generic $i \in \{1, \cdots, N\}$ we have

$$w^n_i = \sum_{j=1}^{N} \frac{q_{i,j}(n)}{r^i + q_i(n)}\, w^{n-1}_j \tag{24}$$

where, due to the fact that $q_{i,i} = 0$, the sum extends actually only over the $j \neq i$. We next have, see (18)

$$\begin{aligned} w^n_i &= H_n(X_{\nu_t+M-n} = x^i) \\ &= E^{\tilde{P}}\left\{e^{-r_{\nu_t+M-n}(\tau_{\nu_t+M-n+1}-\tau_{\nu_t+M-n})} H_{n-1}(X_{\nu_t+M-n+1}) \mid X_{\nu_t+M-n} = x^i\right\} \\ &= E^{\tilde{P}}\left\{e^{-r^i(\tau_{\nu_t+M-n+1}-\tau_{\nu_t+M-n})} \textstyle\sum_{j=1}^N w^{n-1}_j \mathbf{1}_{\{X_{\nu_t+M-n+1}=x^j\}} \mid X_{\nu_t+M-n} = x^i\right\} \\ &= \textstyle\sum_{j=1}^N w^{n-1}_j E^{\tilde{P}}\left\{e^{-r^i(\tau_{\nu_t+M-n+1}-\tau_{\nu_t+M-n})} \mathbf{1}_{\{X_{\nu_t+M-n+1}=x^j\}} \mid X_{\nu_t+M-n} = x^i\right\} \end{aligned} \tag{25}$$

where we have used the fact that, at a generic transition time τ_n, we have $r_n = r^i$ if $X_n = x^i$. We have also used the definition of w^{n-1}_j according to (20) with $n-1$ replacing n.

Next, by the general properties of CTMCs we have that, conditional on X_{ν_t+M-n}, the inter-arrival time $\tau_{\nu_t+M-n+1} - \tau_{\nu_t+M-n}$ is independent of $X_{\nu_t+M-n+1}$. Recalling furthermore that, for $X_{\nu_t+M-n} = x^i$, the distribution of $\tau_{\nu_t+M-n+1} - \tau_{\nu_t+M-n}$ is (negative) exponential with parameter q_i, we obtain

$$
\begin{aligned}
&E^{\tilde{P}}\left\{e^{-r^i(\tau_{\nu_t+M-n+1}-\tau_{\nu_t+M-n})}\mathbf{1}_{\{X_{\nu_t+M-n+1}=x^j\}} \mid X_{\nu_t+M-n}=x^i\right\} \\
&= E^{\tilde{P}}\left\{e^{-r^i(\tau_{\nu_t+M-n+1}-\tau_{\nu_t+M-n})} \mid X_{\nu_t+M-n}=x^i\right\} E^{\tilde{P}}\left\{\mathbf{1}_{\{X_{\nu_t+M-n+1}=x^j\}} \mid X_{\nu_t+M-n}=x^i\right\} \\
&= \int_0^\infty e^{-r^i u} q_i(n) e^{-q_i(n)u} du\, \tilde{P}\left\{X_{\nu_t+M-n+1}=x^j \mid X_{\nu_t+M-n}=x^i\right\} \\
&= \tfrac{q_i(n)}{r^i+q_i(n)}\, p_{i,j}(n) = \tfrac{q_{i,j}(n)}{r^i+q_i(n)}
\end{aligned}
\tag{26}
$$

where in the last passage we have used the fact that one has that $p_{i,j}(n) = \frac{q_{i,j}(n)}{q_i(n)}$. Combining (25) with (26) we obtain (24) and thus the statement of the Proposition. ∎

Notice that (23) implies $H_n(\underline{x}) = \widetilde{Q}(n)\widetilde{Q}(n-1)\cdots\widetilde{Q}(1)H_0(\underline{x})$ which in the time homogeneous case becomes $H_n(\underline{x}) = \widetilde{Q}^n H_0(\underline{x})$. Notice also that in the bivariate (multivariate) case, putting $\underline{z} = (\underline{x}, \underline{y})'$ with $\underline{x} = (x^1, \cdots, x^N)$, $\underline{y} = (y^1, \cdots, y^M)$ we have $H_0(\underline{z}) = [w_1^0, \cdots, w_{N\cdot M}^0]'$ and $H_n(\underline{z}) = \widetilde{Q}(n)H_{n-1}(\underline{z})$ where

$$
\widetilde{Q}(n) = \left\{\frac{q_{(i,h),(j,k)}(n)}{r^{i,h}+q_{i,h}(n)}\right\}_{\substack{i,j=1,\cdots,N \\ h,k=1,\cdots,M}}
$$

with $r^{i,h} = r(X,Y)$ when $X = x^i$ and $Y = y^h$ and $q_{i,h}(n) = \sum_{j\neq i, k\neq h} q_{(i,h),(j,k)}(n)$.

The computation of $\tilde{V}_{H,t,T}^{\nu_{t,T}}(X_t)$ by recursive matrix multiplication and then that of the expectation $E^{\tilde{P}}\left\{\tilde{V}_{H,t,T}^{\nu_{t,T}}(X_t)\right\}$ by simulating $\nu_{t,T}$ forms the backbone of our (hybrid) MC method with conditioning. It is based on the rather immediate next Proposition, which in fact follows from (16), Proposition 4.3 and Corollary 4.2.

Proposition 4.4. *The value of $E^{\tilde{P}}\left\{\tilde{V}_{H,t,T}^{\nu_{t,T}}(X_t)\right\}$, by which (see (16)) we determine the price $V_{H,t,T}(X_t)$ of the Prototype product, is given by*

$$
E^{\tilde{P}}\left\{\tilde{V}_{H,t,T}^{\nu_{t,T}}(X_t)\right\}_{|X_t=x^i} = \frac{1}{2}\, E^{\tilde{P}}\left\{\left[\left(1+\widetilde{Q}(\nu_{t,T}+1)\right)\widetilde{Q}(\nu_{t,T})\cdots\widetilde{Q}(1)H_0(\underline{x})\right]_i\right\} \tag{27}
$$

where $[z]_i$ denotes the $i-$th component of the vector z and where the expectation is with respect to $\nu_{t,T}$. In the time homogeneous case the above expression reduces to

$$
E^{\tilde{P}}\left\{\tilde{V}_{H,t,T}^{\nu_{t,T}}(X_t)\right\}_{|X_t=x^i} = \frac{1}{2}\, E^{\tilde{P}}\left\{\left[\left(1+\widetilde{Q}\right)\widetilde{Q}^{\nu_{t,T}}H_0(\underline{x})\right]_i\right\}. \tag{28}
$$

Remark 4.5. *The expressions in (28) can be further simplified if $\widetilde{Q}$ is diagonalizable (see [13]).*

Based on the above Proposition 4.4, our conditional MC approach (hybrid MC) can now be synthesized in the following

Algorithm:

i) Simulate a sufficiently large number of realizations of the random variable $\nu_{t,T}$.

ii) Record the maximum value, say M, of $\nu_{t,T}$ obtained during the simulations and determine the empirical distribution of $\nu_{t,T}$ derived from the simulations.

iii) Compute recursively the values of $\Big(1+\widetilde{Q}(n+1)\Big)\,\widetilde{Q}(n)\cdots\widetilde{Q}(1)H_0(\underline{x})$ for $n=1$ up to $n=M$.

iv) Determine the average of the values computed in iii) with respect to the empirical distribution of $\nu_{t,T}$ determined in ii).

Applications of this approach are discussed in the next section 5 and in section 6 we then present numerical results and comparisons.

4.3 Accuracy

As already mentioned in Remark 3.2, the accuracy of the approximation of the exact price $V_{H,t,T}(X_t)$ in (4) by the $E^{\tilde{P}}\left\{\tilde{V}^{\nu_{t,T}}_{H,t,T}(X_t)\right\}$, where $\tilde{V}^{\nu_{t,T}}_{H,t,T}(X_t)$ is the midpoint between the upper and lower bounds $\bar{V}^{\nu_{t,T}}_{H,t,T}(X_t)$ and $\underline{V}^{\nu_{t,T}}_{H,t,T}(X_t)$ respectively, can also be seen as a consequence of Corollary 4.7 below. For this purpose consider the operator, acting on $\mathbb{R}^N$ with values in $\mathbb{R}^N$, that is associated to the matrix $\widetilde{Q}(n)$ defining the recursions (23). It is given by the expectation operator in (18) which, for the generic $i-$th component of $H_n(\underline{x})$, can be expressed as (recall that we had assumed $r^i = r(x^i)$)

$$\widetilde{Q}(n)H_{n-1}(X_{n-1})_{|X_{n-1}=x^i} = E^{\tilde{P}}\left\{e^{-r^i\mathcal{I}^i}H(X_n)\mid X_{n-1}=x^i\right\} \tag{29}$$

with $\mathcal{I}^i$ denoting a (negative) exponential random variable with parameter q_i. We have the following

Proposition 4.6. *The operator $\widetilde{Q}(n)$ in (29) is a contraction operator in $\mathbb{R}^N$ with contraction constant*

$$\gamma := \max_{i\le N, n\in\mathbb{N}} \frac{q_i(n)}{r^i+q_i(n)} < 1\,.$$

.

Proof. By Jensen's inequality we have

$$\|\widetilde{Q}(n)H(\underline{x}) - \widetilde{Q}(n)\bar{H}(\underline{x})\| \le \sup_{i,n} E^{\tilde{P}}\left\{e^{-r^i\,\mathcal{I}^i}\right\}\,\|H(\underline{x})-\bar{H}(\underline{x})\|$$

where

$$\sup_{i,n} E^{\tilde{P}}\left\{e^{-r^i\,\mathcal{I}^i}\right\} = \sup_{i,n}\int_0^\infty e^{-r^i s}q_i(n)e^{-q_i(n)s}ds = \max_{i,n}\frac{q_i(n)}{r^i+q_i(n)}\,.$$

∎

This Proposition 4.6 as well as Corollary 4.2 lead immediately to the following

Corollary 4.7. *By the contraction property of the operator associated with $\widetilde{Q}(n)$ we have that*

$$\lim_{n\to\infty} \|\bar{V}^n_{H,t,T}(\underline{x}) - \underline{V}^n_{H,t,T}(\underline{x})\| = 0$$

Remark 4.8. *As already mentioned, instead of determining, on the basis of (27), the value of $E^{\tilde{P}}\left\{\tilde{V}^{\nu_{t,T}}_{H,t,T}(X_t)\right\}$ by MC simulations of $\nu_{t,T}$, one might compute this expectation in a fully analytical way as*

$$E^{\tilde{P}}\left\{\tilde{V}^{\nu_{t,T}}_{H,t,T}(X_t)\right\}_{|X_t=x^i} = \frac{1}{2}\sum_{n=0}^{\infty}\left[\left(1+\widetilde{Q}(n+1)\right)\widetilde{Q}(n)\cdots\widetilde{Q}(1)H_0(\underline{x})\right]_i \tilde{P}(\nu_{t,T}=n \mid X_t=x^i)\,.$$

The difficulties for an actual use of this formula consist in the infinite sum and the probability distribution of $\nu_{t,T}$. Concerning the infinite sum notice that, since (see Proposition 4.6) the operator associated to $\widetilde{Q}(n)$ is contracting, for the actual computations one may truncate the infinite sum thereby introducing an approximation that can be made arbitrarily precise provided the truncation is chosen to be sufficiently large. On the other hand, the probability distribution of $\nu_{t,T}$ can in fact be determined explicitly, however the corresponding procedure is rather cumbersome. Details for the case of a scalar X_t can be found in [13].

5 Applications of the conditional MC approach to the pricing of path dependent options

To give an idea of the wide applicability of our approach, in particular in the time inhomogeneous and multivariate cases, we show here how to apply it for the pricing of some path dependent claims, for which the direct formula (6) cannot be used. We start in subsection 5.1 with the case of barrier options (knock-out options), for which formula (6) cannot be used even if the underlying is time homogeneous. For the majority of the path dependent claims, in order to apply our approach, we have to first transform them into simple claims, where the underlying then becomes typically multivariate, possibly also time inhomogeneous. As an example, in section 5.2 we consider lookback options. Further examples, in particular the case of Asian options, are described in [10].

5.1 Barrier options (knock-out options)

Barrier options include various kinds of derivatives, in particular they include also credit risky derivatives in the context of the structural approach. As an example we consider here a specific case, namely knock-out options.

We consider the case when an option with underlying X_t is knocked out as soon as X_t reaches or falls below a level L. Using a notation corresponding to (5), assume that for the "background" (not knocked out) option we have

$$\bar{H}(\cdot) = \bar{H}_0(\cdot) = \sum_{i=1}^{N} \bar{w}^0_i \mathbf{1}_{\{\cdot=x^i\}},\ x^i \in E,\ \ w^0_i \in \mathbb{R}\,. \tag{30}$$

Assume furthermore that the values x^i are arranged in increasing order of magnitude and put $\ell := \min[i \in \{1, \cdots, N\} \mid x^i > L]$. For the knock-out option we may then start from

$$H(X_t) = H_0(X_T) = \sum_{i=1}^{N} \bar{w}_i^0 \mathbf{1}_{\{X_T = x^i, i \geq \ell\}} := \sum_{i=1}^{N} w_i^0 \mathbf{1}_{\{X_T = x^i\}} \tag{31}$$

having put $w_i^0 := \bar{w}_i^0 \mathbf{1}_{\{i \geq \ell\}}$.

In order to be able to apply our conditional MC approach, we want also here to obtain a relation of the form (23) for a suitable $\widetilde{Q}(n)$. We have now the rather immediate

Proposition 5.1. *Starting from* $H_0(\cdot) = \sum_{i=1}^{N} \bar{w}_i^0 \mathbf{1}_{\{\cdot = x^i, i \geq \ell\}} := \sum_{i=1}^{N} w_i^0 \mathbf{1}_{\{\cdot = x^i\}}$ *with* $w_i^0 := \bar{w}_i^0 \mathbf{1}_{\{i \geq \ell\}}$ *we have, for* $n \leq \nu_T$, *that* $H_n(\cdot) = \sum_{i=1}^{N} w_i^n \mathbf{1}_{\{\cdot = x^i\}}$, *where* $w^n = [w_1^n, \cdots, w_N^n]'$ *are given recursively by*

$$w^n = I_\ell \widetilde{Q}(n) w^{n-1} \tag{32}$$

with I_ℓ *a unit matrix having the first* ℓ *rows equal to zero and, as before,*

$$\widetilde{Q}(n) = \left\{ \frac{q_{i,j}(n)}{r_i + q_i(n)} \right\}_{i,j=1,\cdots,N} \tag{33}$$

As a consequence of Proposition 5.1 we may restrict attention to an $(N - \ell)-$vector $\tilde{w}^n$ for which

$$\tilde{w}_i^0 = w_i^0 := \bar{w}_i^0 \mathbf{1}_{\{i \geq \ell\}} \quad \text{and} \quad \tilde{w}^n = \widetilde{Q}_\ell(n) \tilde{w}^{n-1} \tag{34}$$

where $\widetilde{Q}_\ell(n)$ is the $(N - \ell) \times (N - \ell)$ sub matrix of $\widetilde{Q}^n$ formed by the last $N - \ell$ rows and columns and we have the equivalent representations

$$H_n(X_{\nu_T - n}) = \sum_{i=1}^{N} w_i^n \mathbf{1}_{\{X_{\nu_T - n} = x^i\}} = \sum_{i=1}^{N-\ell} \tilde{w}_i^n \mathbf{1}_{\{X_{\nu_T - n} = x^i\}} \,. \tag{35}$$

Notice the importance here of having the recursive relation (32) even if the underlying is time-homogeneous. Notice also that, in the case of barrier options, the time in-homogeneity arises not only if the underlying X_t is time in-homogeneous, but also if the barrier L is time varying.

5.2 Lookback options

A general form for the claim of a lookback option is

$$H_T = \left(X_T - g(X_0^T)\right)^+ \tag{36}$$

where $g(\cdot)$ is a measurable function of the generic trajectory $X_0^t := (X_0, \cdots, X_t)$ for $t \leq T$ of the process X_t for which, recalling that τ_n denote the transition times of X_t, we assume

$$g(X_0^{\tau_n}) = G(X_{\tau_n}, g(X_0^{\tau_{n-1}})) \quad \text{for some measurable } G(\cdot, \cdot) \,. \tag{37}$$

We now show how, with X_t a CTMC, one can transform this claim into a simple claim with a bivariate underlying CTMC that in certain cases may also turn out to be time inhomogeneous. This will then allow our approach to be applied.

Put $Y_t := g(X_0^t)$ and notice that, for $t \in [0,T]$, the process Y_t also takes a finite number of possible values; denote them by $h = 1, \cdots, M$. More importantly, Y_t can make a transition only at the transition times of X_t and it can be easily seen that (X_t, Y_t) forms a bivariate CTMC. Furthermore, we have that $H_T = (X_T - Y_T)^+$ and we also recall that $(X_T, Y_T) = (X_{\nu_T}, Y_{\nu_T})$. To deal with the chain (X_t, Y_t) and to price H_T, we need to derive the transition intensity matrix Q for (X_t, Y_t), for which we have

Proposition 5.2. *Given a CTMC X_t with transition intensity matrix $Q = \{q_{i,j}\}_{i,j=1,\cdots,N}$, the chain (X_t, Y_t) has*

$$\left\{q_{(i,h),(j,k)}(n)\right\}_{\substack{i,j=1,\cdots,N \\ h,k=1,\cdots,M}}$$

with $q_{(i,h),(j,k)} = q_{i,j}\mathbf{1}_{\{G(j,h)=k\}}$

Proof. Recall first that, if for a scalar CTMC X_t the $Q-$matrix is $Q = \{q_{i,j}\}$, then the transition probabilities of the embedded chain X_n are $p_{i,j} = \frac{q_{i,j}}{q_i}$ with $q_i = \sum_{j \neq i} q_{i,j}$ $(q_{i,i} = p_{i,i} = 0)$. Viceversa, given $p_{i,j}$, there are various possible $q_{i,j}$ that lead to the same $p_{i,j}$. They differ by the choice of q_i since we have $q_{i,j} = q_i p_{i,j}$. Given that in our case Y_t jumps exactly when X_t does, we may put

$$q_{(i,h)} = \sum_{j,k} q_{(i,h),(j,k)} = q_i \quad \forall h = 1, \cdots, M \tag{38}$$

and notice that at a generic τ_n the process X_t actually leaves the current state, while Y_t may jump to itself. To conclude, it thus suffices to construct $p_{(i,h),(j,k)}$. Recalling that we had put $X_n = X_{\tau_n}, Y_n = Y_{\tau_n}$, we have

$$\begin{aligned}
p_{(i,h),(j,k)} &:= P\{X_{n+1} = j, Y_{n+1} = k \mid X_n = i, Y_n = h\} \\
&= P\{X_{n+1} = j, G(X_{n+1}, Y_n) = k \mid X_n = i, Y_n = h\} \\
&= P\{G(X_{n+1}, Y_n) = k \mid X_{n+1} = j, X_n = i, Y_n = h\} \\
&\qquad \cdot P\{X_{n+1} = j \mid X_n = i, Y_n = h\} \\
&= \mathbf{1}_{\{G(j,h)=k\}} P\{X_{n+1} = j \mid X_n = i\} = \mathbf{1}_{\{G(j,h)=k\}} p_{i,j}
\end{aligned} \tag{39}$$

from which

$$q_{(i,h),(j,k)} = p_{(i,h),(j,k)} \cdot q_i = q_{i,j}\mathbf{1}_{\{G(j,h)=k\}}\,. \tag{40}$$

∎

Example 1. As an example consider the standard case where

$$Y_t = g(X_0^t) := \min_{s \leq t} X_s\,. \tag{41}$$

Notice that in this case Y_t has the same finite number of possible values as X_t and the assumptions on $g(\cdot)$ are satisfied since $G(X_{\tau_n}, g(X_0^{\tau_{n-1}})) = \min\left[X_{\tau_n}, \min_{s \leq \tau_{n-1}} X_s\right] = g(X_0^{\tau_n})$. The relation

(39) particularizes into $p_{(i,h),(j,k)} = \mathbf{1}_{\{G(j,h)=k\}} p_{i,j} = \mathbf{1}_{\{\min\{j,h\}=k\}} p_{i,j}$ which, with the states x^i in increasing order of magnitude, implies that

$$p_{(i,h),(j,k)} = \begin{cases} p_{ik} & \text{if} \quad k < h \\ p_{ij} & \text{if} \quad k = h,\ j \geq k \\ 0 & \text{if} \quad k > h \end{cases} = \begin{cases} \frac{q_{ik}}{q_i} & \text{if} \quad k < h \\ \frac{q_{ij}}{q_i} & \text{if} \quad k = h,\ j \geq k \\ 0 & \text{if} \quad k > h \end{cases} \tag{42}$$

and, consequently,

$$q_{(i,h),(j,k)} = p_{(i,h),(j,k)} \cdot q_i = \begin{cases} q_{ik} & \text{if} \quad k < h \\ q_{ij} & \text{if} \quad k = h,\ j \geq k \\ 0 & \text{if} \quad k < h \end{cases} \tag{43}$$

Notice that the chain (X_t, Y_t), as described above, is multivariate but still time homogeneous so that the pricing of the claim could be performed by the explicit analytic formula (6), although it is more complex due to the increased dimensionality. If, however, one would consider a claim of the form $H_T = (X_T - g(X_{T-\sigma}^T))^+$ for a given $0 < \sigma < T$, then our process Y_t may be defined as taking a suitable fixed value for $t \leq T - \sigma$ and thereafter evolves as in the case of $g(X_0^T)$ by letting $t = 0$ correspond to $t = T - \sigma$. In the specific case of Example 1, the process Y_t may then be defined as

$$Y_t = \begin{cases} \max_i x^i & \text{for} \quad t \leq T - \sigma \\ \\ \min_{T-\sigma<s\leq t} X_s & \text{for} \quad t > T - \sigma \end{cases} \tag{44}$$

It is then quite evident that the chain (X_t, Y_t) is not anymore time homogeneous, even if X_t is. The transition intensity matrix is then a sequence $Q(n)$ that has a certain expression for all n such that $\tau_n \leq T - \sigma$ and another one for those n for which $\tau_n > T - \sigma$. Notice also that the change from one expression to the other one depends on τ_n and therefore on the individual trajectory of X_t. Formula (6) then does not apply anymore, but our conditional MC approach (see steps i) to iv) of *Algorithm* in subsection 4.2) appears to be particularly appropriate since the expression to be averaged, namely $\left[1 + \widetilde{Q}(n+1)\right] \widetilde{Q}(n) \cdots \widetilde{Q}(1) H_0(\underline{x})$ can be computed separately for each individual simulated trajectory.

6 Numerical results and comparisons

The purpose of this section is to implement numerically our suggested approach and thereby to show that, with respect to a plain MC, in the MC with conditioning the variance is indeed reduced and the results themselves are more precise without increasing the computational complexity. To apply our method, we have first to specify a transition intensity matrix Q. For actual applications, this matrix would have to be calibrated to actual market data. Here the purpose is however that of providing some test examples and for this it may be convenient to choose the matrix freely, possibly also different matrices as we do below for the case of lookback options.

While calibration is beyond the scope of this paper, we still want to point out that, for an actual calibration to become feasible, one would have to choose specific patterns of $Q-$matrices, parametrized by a small number of parameters, and so calibrate just these parameters. We also mention that, for a time homogeneous X_t it is possible to set up a filtering approach to estimate

the values of the $Q-$matrix, either by combined filtering and parameter estimation, or by filtering and EM-parameter estimation (see e.g.[5], see also [10]).

The section is structured as follows. In subsection 6.1 we start by discussing the pricing of zero-coupon bonds as this will allow us also to assess the quality of our approach. Zero-coupon bond prices can in fact be obtained by an exact formula in the continuous time case when the short rate evolves as an affine diffusion process. By discretizing in space the affine diffusion we obtain a CTMC, to which we then apply our approach. We shall show that, in spite of the approximation due to the spatial dicretization, the prices are close to one another. We may thus be confident that also in other situations, where a comparison with a benchmark is not anymore possible, our approach performs well. In this subsection 6.1 we use a "Kushner-type" approximation according to [9]; other spatial discretization/quantization methods may also be used, in particular optimal quantization methods according to [3] (for specific financial application of optimal quantization see also [4], [15]). Referring to [13], we report prices of zero-coupon bonds that are computed according to our MC with conditioning and we compare them with the exact values of the continuous-time counterpart, with those obtained from the analytical formula (6) (to allow for this comparison we consider a time homogeneous case), and also with the values obtained from other computational methods, namely: plain MC, a recombining binomial tree model applied to the continuous-time counterpart and the algorithm described in [6] with the discrete time Markov chain obtained via a deterministic time discretization. In subsection 6.2 we numerically test our approach for the pricing of barrier options (knock-out options) by comparing the prices computed according to our conditional MC with those computed via plain MC. In subsection 6.3 we present analogous numerical tests for the case of lookback options (for the case of Asian options we refer to [10]). For this we consider a time homogeneous model and this allows for a comparison also with the price computed according to the exact analytic formula (6). We perform this test for two different $Q-$matrices, of which one induces more frequent transitions of the chain. To conclude section 6, in subsection 6.3 we comparatively discuss the computational complexity and and the induced computation times of our approach. One of the criteria that we use for comparison in subsections 6.2 and 6.3 is the empirical variance per number of samples, which we call the *estimator variance rate* (EVR).

Remark 6.1 (Estimator variance rate)**.** *Given n MC samples $C_i\,, (i = 1, \cdots, n)$ for a generic random variable C and letting $\hat{C}_n := \frac{1}{n}\sum_{i=1}^{n} C_i$ be the MC estimate for the mean of C, as estimator variance we intend the following empirical variance $\widehat{Var}(\hat{C}_n) = \frac{1}{n-1}\sum_{i=1}^{n}(C_i - \hat{C}_n)^2$. The estimator variance rate that we shall use as performance criterion in several diagrams below corresponds then to $EVR = \frac{1}{n}\widehat{Var}(\hat{C}_n)$.*

6.1 Zero Coupon Bonds

The pricing of zero coupon bonds corresponds to our setup by putting $X_t = r_t$ and $H(X_T) \equiv 1$. To define the CTMC for $X_t = r_t$, we have to choose a transition intensity matrix Q. In order to have an example that allows for an easy comparison with exact results, we choose X_t as given by a space discretization, in the spirit of [9], of well-known affine diffusion models for the short rate of interest r_t, for which the exact price can be computed analytically. We report here results from [13], where the CTMC $X_t = r_t$ is obtained from a space discretization according to [9] of the CIR model for the corresponding continuous-time short rate $\tilde{r}_t$, namely

$$d\tilde{r}_t = k(\theta - \tilde{r}_t)dt + \sigma\sqrt{\tilde{r}_t}\, dW_t\,, \quad \tilde{r}_0 = \bar{r} \tag{45}$$

The bond prices obtained from the MC method with conditioning are compared not only with the theoretically exact price $\widetilde{P}(t,T)$ for model (45), but also with the prices obtained by other methods, including plain MC. While in the present paper we consider the midpoint between the upper and lower bounds as the value to be computed analytically in our (hybrid) MC method with conditioning (see (15) and Corollary 4.2), in [13] the authors consider just the upper bound that in general does not differ much from the midpoint. In the Tables below, prices for various maturities T and different initial values for the short rate $\tilde{r}_0 = r_0 = \bar{r}$, as well as for different values of the parameters in (45) are reported from [13]. As already mentioned, $\widetilde{P}(t,T)$ denotes the exact bond price for model (45), while the other reported prices are: $P_{RBT}(t,T)$ the bond price obtained from a recombining binomial tree model, $P_{FZ}(t,T)$ the price computed for a time discretization of (45) and computed according to the recursive pricing method described in [6], $P_{Exp}(t,T)$ the price computed by formula (6) (recall that Q is here time homogeneous), $P_{pl}(t,T)$ the price computed by plain MC and $P_c^{ub}(t,T)$ the upper bound on the price computed by MC with conditioning.

The following parameters were used: $t = 0$ years, $T = 0.5$, 2 and 5 years. The number of MC simulations M and the RBT steps $\bar{M}$ were taken to be 500, which, while it is a small number with respect to what is typical, was enough because a very fine space discretization was chosen. In Table 1 and Table 2 the numerical results are presented when the values of the initial spot rate $\bar{r}$ and the mean-reversion constant θ are of the order of one hundredth; in Table 3 the values of $\bar{r}$ and θ are of the order of one tenth.

T(years)	0.5	2	5	0.5	2	5
$\bar{r}(=r^i)$	0.01	0.01	0.01	0.02	0.02	0.02
k	0.01	0.01	0.01	0.02	0.02	0.02
θ	0.8	0.8	0.8	0.5	0.5	0.5
σ	0.1	0.1	0.1	0.05	0.05	0.05
$\widetilde{P}(t,T)$	0.995014	0.980244	0.951462	0.990051	0.960821	0.905046
$P_{RBT}(t,T)$	0.995042	0.980302	0.951556	0.99007	0.960898	0.905226
$P_{FZ}(t,T)$	0.995014	0.980244	0.951463	0.990051	0.960821	0.905046
$P_{Exp}(t,T)$	0.995012	0.979568	0.947174	0.990051	0.960821	0.905047
$P_c^{ub}(t,T)$	0.995024	0.980276	0.951621	0.990143	0.960734	0.905318

Table 1: Bond prices $\widetilde{P}(t,T)$, $P_{RBT}(t,T)$, $P_{FZ}(t,T)$, $P_{Exp}(t,T)$ and $P_c^{ub}(t,T)$ $(\mathbf{M} = \bar{\mathbf{M}} = 500)$

T(years)	0.5	2	5	0.5	2	5
$\bar{r}(=r^i)$	0.03	0.03	0.03	0.02	0.02	0.02
k	0.03	0.03	0.03	0.02	0.02	0.02
θ	1.1	1.1	1.1	1.2	1.2	1.2
σ	0.1	0.1	0.1	0.1	0.1	0.1
$\widetilde{P}(t,T)$	0.985116	0.941861	0.861094	0.990052	0.960849	0.905072
$P_{RBT}(t,T)$	0.985146	0.941974	0.86135	0.990072	0.960926	0.905251
$P_{FZ}(t,T)$	0.985116	0.941861	0.861094	0.990052	0.960849	0.905072
$P_{Exp}(t,T)$	0.985116	0.941841	0.861042	0.990052	0.960738	0.904656
$P_c^{ub}(t,T)$	0.985128	0.941968	0.861319	0.990059	0.95647	0.90193

Table 2: Bond prices $\widetilde{P}(t,T)$, $P_{RBT}(t,T)$, $P_{FZ}(t,T)$, $P_{Exp}(t,T)$ and $P_c^{ub}(t,T)$ $(\mathbf{M} = \bar{\mathbf{M}} = 500)$

T(years)	0.5	0.5	0.5	0.5
$\widetilde{r}(=r^i)$	0.01	0.02	0.03	0.02
k	0.8	0.5	1.1	1.2
θ	0.01	0.02	0.03	0.02
σ	0.1	0.05	0.1	0.1
$\widetilde{P}(t,T)$	0.995014	0.990051	0.985116	0.990052
$P_{RBT}(t,T)$	0.995042	0.99007	0.985146	0.990072
$P_{FZ}(t,T)$	0.995014	0.990051	0.985116	0.990052
$P_{Exp}(t,T)$	0.995012	0.990051	0.985116	0.990052
$P_{pl}(t,T)$	0.995012	0.990051	0.985067	0.989930
$P_c(t,T)$	0.981921	0.990059	0.985041	0.989885

Table 3: Bond prices $\widetilde{P}(t,T)$, $P_{RBT}(t,T)$, $P_{FZ}(t,T)$, $P_{Exp}(t,T)$, $P_{pl}(t,T)$, $P_c(t,T)$ $(\mathbf{M}=\bar{\mathbf{M}}=500)$

6.2 Barrier options (knock-out options)

Following the pricing approach in section 5.1 for barrier options, we are able to compare the barrier options price as well as the EVR for plain MC and MC with conditioning. With $N=5$, we define the following Q matrix:

$$Q=\begin{bmatrix} -190 & 30 & 25 & 50 & 85 \\ 5 & -185 & 140 & 25 & 15 \\ 5 & 135 & -230 & 70 & 20 \\ 5 & 40 & 165 & -230 & 20 \\ 5 & 10 & 45 & 60 & -120 \end{bmatrix}$$

For the other parameters, we set the (non-knocked-out) terminal payoff $H(X_T)$ as $H(X_T)=[1,1,1,1,1]^\top$, the *Down-and-Out* barrier level $L=x^2$ $(x\in\{x^1,\cdots,x^5\})$, the interest rate values r_t as $r_t=r(X_t)=[0.01,0.015,0.02,0.025,0.05]$, the initial state $X_0=x^4$ again with $x\in\{x^1,\cdots,x^5\}$, initial time $t=0$ and terminal time $T=1$ year.

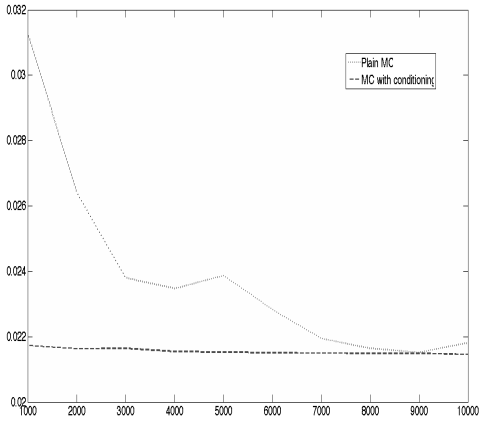

Figure 1: Estimator for barrier options using Plain MC (dotted line) and MC with conditioning (dashed line) over total number of MC simulations. Final estimated price is $Bar_{pl}(t,T)=0.0218$, *std. dev*= 0.144, and $Bar_c(t,T)=0.0215$, *std. dev*= 0.00547, $n=10000$.

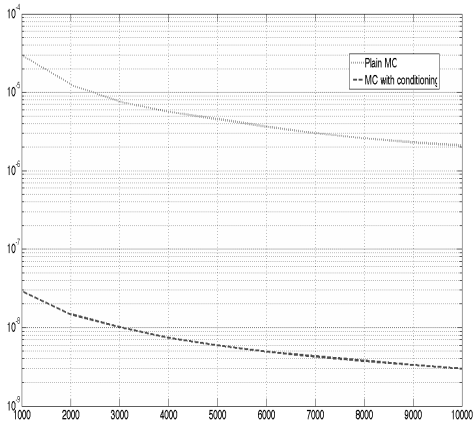

Figure 2: The EVR for barrier options using Plain MC (dotted line) and MC with conditioning (dashed line), shown on a $\log_{10}$ y-axis vs. the total number of samples. For $n=10000$ the Plain MC EVR is 2.0860×10^{-6} while the MC-with-conditioning variance rate is 2.994×10^{-9}.

Down-and-Out barrier options are computed using these specified parameters. Figure 1 shows a graph of the estimator with plain MC (dotted line) and MC with conditioning (dashed line), over the number of MC simulations, from 1000 to 10,000. Figure 2 shows the EVR of plain MC (dotted line) and MC with conditioning (dashed line); the y-axis is given in $\log_{10}$ scale to improve readability. It is clear that MC with conditioning performs better for barrier options pricing compared to plain MC using EVR as performance criterion.

6.3 Lookback options

We perform the MC pricing for lookback options as discussed in subsection 5.2 considering the case where $g(X_0^t) := \min_{s\leq t} X_s$ (Example 1) assuming, furthermore, that $\sigma = T$. This implies that, if the chain X_t is time-homogeneous, also the bivariate chain (X_t, Y_t) is and has thus a time homogeneous transition intensity matrix. Since the claim $H_T = (X_T - Y_T)^+$ is furthermore a simple claim, it is possible to compute the price also according to the analytic formula (6). We thus computed lookback prices for a plain MC namely $(\widetilde{LB}(t,T))$, for MC with conditioning $(LB(t,T))$ and according to formula (6) $(\bar{LB}(t,T))$ for different sets of parameters. We report here just two of the tests that we performed, which however show two quite different situations that may occur.

In both cases, we assumed a state space $E = [0.8, 0.9, 1.0, 1.1, 1.2]$ with $N = 5$; initial state $X_0 = 1.1$, and maturity $T = 2$ years. For Test 1 and Test 2 we used the following Q-matrices:

$$Q = \begin{bmatrix} -1440 & 360 & 360 & 360 & 360 \\ 7.2 & -28.8 & 7.2 & 7.2 & 7.2 \\ 0.72 & 0.72 & -2.88 & 0.72 & 0.72 \\ 2.52 & 2.52 & 2.52 & -10.08 & 2.52 \\ 480 & 480 & 480 & 480 & -1920 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} -0.8 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.7\bar{3} & -2.5\bar{3} & 0.\bar{6} & 0.6 & 0.5\bar{3} \\ 0.02 & 0.02 & -0.08 & 0.02 & 0.02 \\ 0.6 & 0.5\bar{3} & 0.\bar{6} & -2.4\bar{6} & 0.\bar{6} \\ 0.4 & 0.4 & 0.\bar{3} & 0.4 & -1.5\bar{3} \end{bmatrix}$$

respectively, where the over-bar denotes repeated decimals. Notice that the basic difference in the two cases consists in the fact that, in the first case, we obtain more frequent jumps/transitions of X_t.

The results of the tests are reported in the various Figures where for plain MC and the hybrid MC with conditioning we show the price, that is its averaged value over the various simulation runs, together with a confidence interval given as a multiple of the empirical standard deviation, which is given by $\sqrt{n\,EVR}$.

For Test 1, the plain MC estimate of the lookback price is $\widetilde{LB}(t,T) = 0.216839$ (std. dev $\tilde{\delta} = 0.128226$) (Figure 3), for the MC-with-conditioning the estimated price is $LB(t,T) = 0.194265$ (std. dev $\delta = 0.011603$) (Figure 4), while $\bar{LB}(t,T) = 0.194707$. Note that $\widetilde{LB}(t,T) - \tilde{\delta} < LB(t,T) - \delta < LB(t,T) + \delta < \widetilde{LB}(t,T) + \tilde{\delta}$. Figure 3 and Figure 4 show the graphs of the empirical mean as a function of the total iteration count of the MC simulation, together with a plot of the 1.5 standard deviations interval for $\widetilde{LB}(t,T)$ and $LB(t,T)$ in Test 1, respectively. For comparison, each of the Figures includes also the level corresponding to $\bar{LB}(t,T)$. For Test 1, Figure 5 is a graph of the EVR over the total number of MC simulations for the plain MC estimator (solid line) and the MC-with-conditioning estimator (dashed line). Here we see again that the MC-with-conditioning performs better under the EVR criterion.

For Test 2, where the jump intensities are considerably lower, the lookback price with the plain MC estimator is $\widetilde{LB}(t,T) = 0.147932$ (std. dev $\tilde{\delta} = 0.118035$) (Figure 6), the MC-with-conditioning price has empirical mean $LB(t,T) = 0.118050$ (std. dev $\delta = 0.037996$)(Figure 7) and

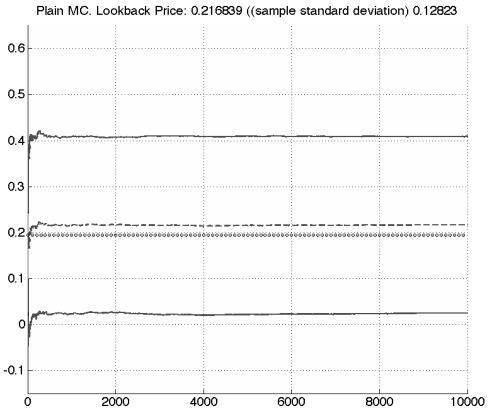

Figure 3: Plain MC estimator for lookback options showing the 1.5 standard deviation interval (Test 1). *lookback price* $\widetilde{LB}(0,2) = 0.216839$ (dashed line), *sample std dev* $= 0.128226$, $\bar{LB}(t,T) = 0.194707$ (dotted line), $n = 10000$

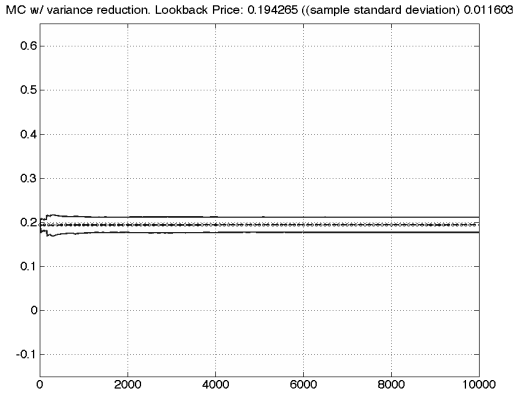

Figure 4: MC-with-Conditioning estimator for lookback options showing the 1.5 standard deviation interval (Test 1). *lookback price* $LB(0,2) = 0.194265$ (dashed line), *sample std dev* $= 0.011603$, $\bar{LB}(t,T) = 0.194707$ (dotted line), $n = 10000$

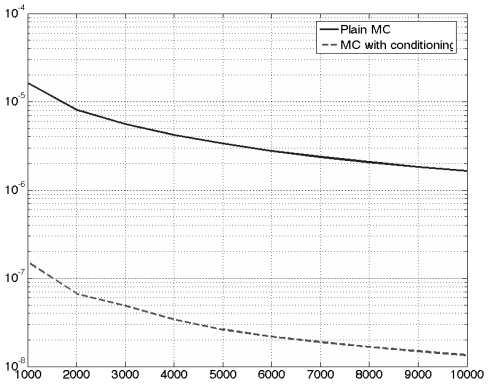

Figure 5: EVR for lookback options by Plain MC (solid line) and MC with conditioning (dashed line), over the total number of MC samples (Test 1). Y-axis in $\log_{10}$ scale. For $n = 10000$ the Plain MC EVR is 1.644×10^{-6} while the MC-with-conditioning EVR is 1.346×10^{-8}.

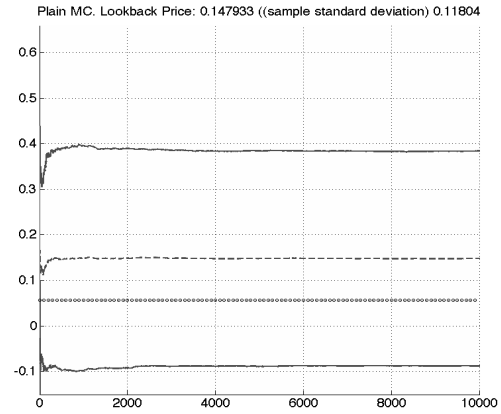

Figure 6: Plain MC estimator for lookback options showing the 2 standard deviation interval (Test 2). *lookback price* $= 0.147932$ (dashed line), *sample std dev* $= 0.118036$, $\bar{LB}(t,T) = 0.056398$ (dotted line), $n = 10000$

$\bar{LB}(t,T) = 0.056398$. Unlike in the previous test, we now have that $\widetilde{LB}(t,T) + \tilde{\delta} < LB(t,T) - \delta$ and we need to plot the 2 standard deviation interval so that it includes $\bar{LB}(t,T)$. This may indicate that, since we have less frequent state transitions, it takes more simulations to obtain sufficiently reliable estimates. In fact, clustering problems may arise, i.e., the drawn samples in the simulation belong to just a small subset of the event space. This is a well known phenomenon that may lead to highly biased MC estimators. Indeed we find that in our MC tests, samples are sometimes not drawn over a large enough subset of the event space for the random variable $\nu_{t,T}$ -

not every possible transition count in $[t, T]$ is achieved by the simulated paths (see Figures 8 and 9 which show the empirical distribution of jump counts for test 1 and test 2 respectively). As a result, many of the possible outcomes are weighted too low, causing the empirical mean, that is the weighted average, to be unreliable (see Figures 10 and 11 which show respectively the sampled price and the theoretical price for each possible jump count).

Nevertheless we still observe better performance for MC-with-conditioning under the EVR criterion as well as a considerably smaller squared bias with respect to the value $\bar{LB}(t, T)$ calculated from (6). Figure 12 shows in fact the EVR over the number of simulations, which just as before reflects the improved performance of MC-with-conditioning. This suggests that in order to meaningfully apply MC, in particular plain MC, it must be understood what conditions on the parameters may lead to models that are susceptible to clustering, and in that case, it may be better to apply a Quasi-MC approach.

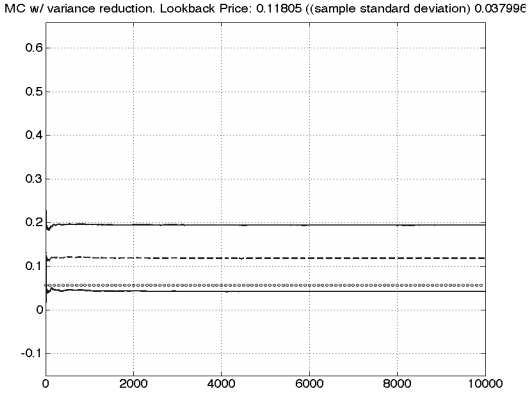

Figure 7: MC-with-Conditioning estimator for lookback options showing the 2 standard deviation interval (Test 2) *lookback price* $= 0.118$ (dashed line), *sample std dev* $= 0.0380$, $\bar{LB}(t, T) = 0.0564$ (dotted line), $n = 10000$

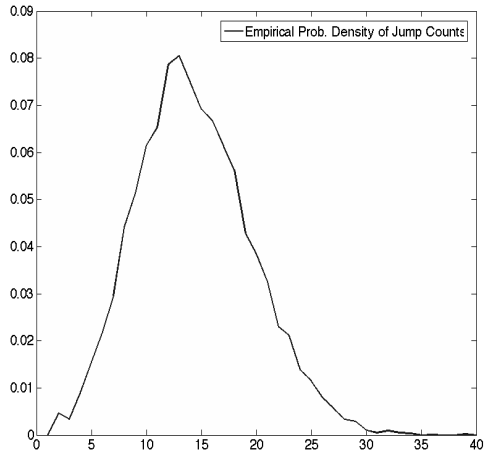

Figure 8: Empirical Distribution of Jump Counts for Test 1 samples, for $\nu = 0 \ldots 40$ jumps

6.4 Complexity and Computation Times

Table 4 shows the example computation times in running each of the various numerical experiments. They represent the computational cost to obtain the performance gain (lower variance) corresponding to each test as reported in Sections 6.1-6.3. For the MC pricing, we distinguish the time spent on the simulation of paths, which is common to the plain and the conditional MC, and the actual time for computing prices. The computations were performed using the software package *MATLAB Release 2012b* on a 2.6 GHz Intel Core i7 processor, with 16GB Memory, under an OS X Version 10.9.1 operating system. Comparing the computation times of plain MC to those of the MC with conditioning, it can be seen that the latter is slightly faster than the former in all the cases (Zero-Coupon Bonds, Barrier Options and Lookback Options). This is expected because the computation under plain MC involves several evaluations of the exponential function. It is well known that the calculation of the `exp` function in most math libraries are rather slow. Instead,

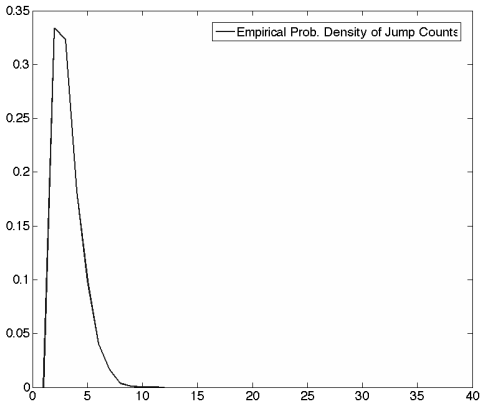

Figure 9: Empirical Distribution of Jump Counts for Test 2 samples, for $\nu = 0 \ldots 40$ jumps

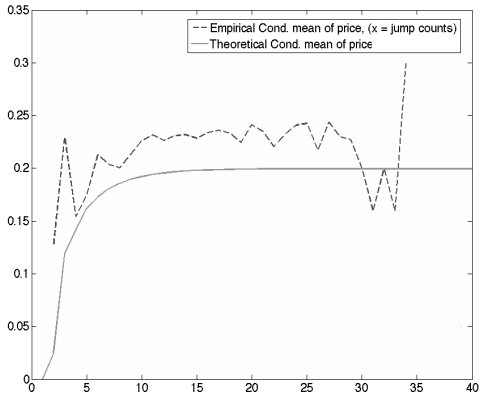

Figure 10: Conditional mean of prices per jump count (Test 1 samples). Dashed line - sample price (Plain MC); Solid line - theoretical price (MC -with-Conditioning)

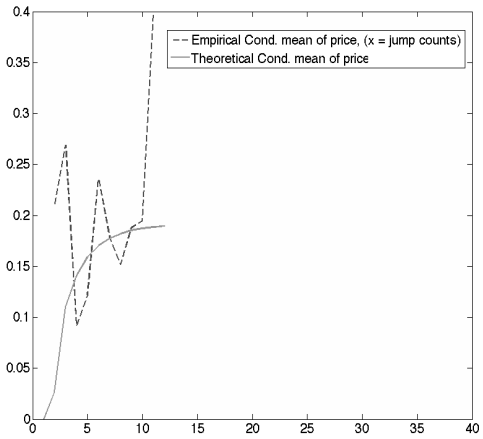

Figure 11: Conditional mean of prices per jump count (Test 2 samples). Dashed line - sample price (Plain MC); Solid line - theoretical price (MC-with-Conditioning)

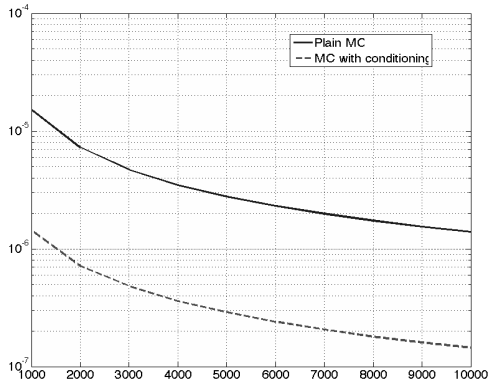

Figure 12: EVR for lookback options by Plain MC (solid line) and MC with conditioning (dashed line), over the total number of MC samples (Test 2). Y-axis in $\log_{10}$ scale. For $n = 10000$ the Plain MC EVR is 1.393×10^{-6} while the MC-with-conditioning variance rate is 1.444×10^{-7}.

under the MC with conditioning, such evaluations are not necessary, and, in fact it involves only matrix operations. Most math libraries have optimized routines for dealing with the latter.

In summary, the MC with conditioning presents a comparable, if not slightly better, method over plain MC, not only with respect to variance reduction but also with respect to a slightly lower computational complexity/cost.

Test	Simulation Time	Plain MC	MC with Conditioning
Zero-Coupon Bonds	59.57 s	0.602 s	0.503 s
Barrier Options	505.4 s	1.803 s	1.721 s
Lookback Options	15.774 s	5.611 s	2.321s

Table 4: Computation Time for the Numerical Tests

References

[1] S. Asmussen, Ruin Probabilities. World Scientific, River Edge, NJ. 2000.

[2] T. Björk, Y. Kabanov, W. J. Runggaldier, Bond market structure in the presence of marked point processes. Mathematical Finance 7 (1997), pp. 211-223.

[3] V. Bally, G.Pages and J. Printems, A stochastic quantization method for nonlinear problems, Monte Carlo Methods and Applications, 7 (2001), pp. 21-34.

[4] G. Callegaro, and A. Sagna, An application to credit risk of a hybrid Monte Carlo-Optimal quantization method, Journal of Computational Finance, 16 (2013) No. 4, pp. 123-156.

[5] R.J. Elliott, L. Aggoun, J.B. Moore, Hidden Markov Models, Estimation and Control, Springer-Verlag, 1995.

[6] D. Filipović, and J. Zabczyk, Markovian term structure models in discrete time. Ann. Appl. Probab. 7 (2002), pp. 710-729.

[7] D. Filipović, and A.B. Trolle, The term structure of interbank risk. Journal of Financial Econometrics 109 (2013), pp. 707-733.

[8] P. Glasserman, Monte Carlo Methods in Financial Engineering. Springer Verlag, New York, 2004.

[9] H. J. Kushner, and P.G. Dupuis, Numerical methods for stochastic control problems in continuous time. Springer-Verlag, 2001.

[10] J.M. Montes, Aspects of Affine Models in the Pricing of Exotic Options and in Credit Risk, Thesis, University of Padova, 2013.

[11] R. Norberg, The Markov chain market. ASTIN Bulletin Vol. 33 (2003), pp. 265-287.

[12] E. Eberlein, and W. Kluge, Exact Pricing Formulae for Caps and Swaptions in a Levy Term Structure Model. Journal of Computational Finance 9 (2006), pp. 99-125.

[13] V. Prezioso, W. Runggaldier, Interest rate derivatives pricing when the short rate is a continuous time finite state Markov process. Preprint, University of Padova (2010).

[14] R. Romera, W. Runggaldier, Ruin probabilities in a finite-horizon risk model with investment and reinsurance. Journal of Applied Probability , 49 (2012) No. 4, pp. 954-966.

[15] A. Sagna, Pricing of barrier options by marginal functional quantization method. Monte Carlo Methods and Applications, 17 (2012) No. 4, pp. 371-398.

[16] P.J. Schönbucher, Term Structure Modelling of Defaultable Bonds. Review of Derivatives Research 2 (1998), pp. 161-192.