Valuation of Path-Dependent Interest Rate Derivatives in a Finite Difference Setup

Mikkel Svenstrup*
The Aarhus School of Business

October 2000
This version February 2001

Abstract

In this paper we study and implement a finite difference version of the augmented state variable approach proposed by Hull & White (1993) that allows for path-dependent securities. We apply the method to a class of path-dependent interest rate derivatives and consider several examples including mortgage backed securities and collateralized mortgage obligations. The efficiency of the method is assessed in a comparative study with Monte Carlo simulation and we find it to be faster for a similar accuracy.

JEL Codes: G13, G12, C19

Keywords: Path-dependent Options; Finite Difference; Mortgage Backed Securities

1 Introduction

In Hull & White (1993) a method to price path-dependent securities in trees is demonstrated to be an efficient way of handling particular path-dependent securities. The main idea is to augment the state space with additional state variables to represent movements in the past. In Wilmott, Dewynne & Howison (1993) the same technique is applied but in a more general finite differences framework to value exotic options like look-back and Asian options.

In this paper we first summarize the method for interest rate derivatives in a finite difference setup. The method allows us to handle most common features in fixed income products including particular types of path-dependencies as well as American features.

Secondly we apply the technique to other types of path-dependent securities, and we illustrate that the valuation of collateralized mortgage obligations under rational prepayments can be done in a single backward run, as opposed to the two-step procedure proposed in McConnell & Singh (1994) that employs both finite difference and Monte Carlo techniques.

The numerical results presented in Hull & White (1993) indicate that the method is faster and just as accurate as Monte Carlo simulation and that the method is not particularly sensitive to the density of the discretized augmented state space. However, our numerical results show that there are in fact large differences in the density of the augmented state space needed in order for the method to converge, but it is still at least as fast as standard Monte Carlo for similar accuracy. The examples we consider are a mortgage backed security (MBS) with a path-dependent prepayment function, collateralized mortgage obligations (CMO) such as the Interest Only (IO), the Principal Only (PO)

*The author would like to thank Tom Engsted, Svend Jakobsen, Jesper Lund, Kristian Miltersen, Rolf Poulsen and Nicki Rasmussen for helpful comments. This research was supported by ScanRate Financial Systems.

and Sequential Pay tranches, and, finally, a capped amortizing Adjustable Rate Mortgage (ARM) with a coupon that is settled as an average of historical interest rates.

In section 2 we go through the model framework. Section 3 describes the numerical implementation while section 4 contains applications of the method. Finally, we make our conclusion.

2 The Model Setup

The following exposition is based primarily on Wilmott et al. (1993), and the main difference is that we derive the fundamental partial differential equation in an interest rate model, whereas Wilmott et al. (1993) work in a Black-Scholes world.

2.1 Interest Rate Dynamics

We work in a one-factor term structure setup, with models for the instantaneous short rate r_t that can be represented by the following SDE,

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dW_t,$$

where μ and σ denote drift- and volatility functions that satisfy the usual conditions. W_t is a one-dimensional Wiener-process. This setup covers many of the most commonly used single factor models, but the technique is also applicable to multi-factor models.

Let V denote the value of an interest rate contingent claim, that is dependent on the history of the short rate. Assume that this dependency can be summarized in a z-dimensional state-vector $A \in \mathbb{R}^z$, in the following way

$$A_t = \int_0^t f(r_s, s)ds.$$

To keep notation simple we assume that $z = 1$. However, it is possible to have $z > 1$. With these specifications the value $V(t, r_t, A_t)$ of the claim is Markov with respect to the information generated by the triple (t, r_t, A_t). In other words, we assume that the value of the path-dependent security is given by the real valued function $V(t, r_t, A_t)$ defined on $\mathbb{R}_+ \times \mathcal{D}(r_t) \times \mathcal{D}(A_t)$. Here $\mathcal{D}(\cdot)$ denotes the domain for a given variable. This domain will in general depend on the specific term structure model and the definition of the state-vector. Before continuing notice that

$$dA_t = f(r_t, t)dt,$$

which means that A_t is a state variable of finite variation, and does not add further noise to the system. In particular this means that we do not need to worry about additional risk premia.

2.2 The Partial Differential Equation

A standard arbitrage argument leads to the fundamental partial differential equation for the security (the derivation can be found in appendix A.1 for completeness).

$$\begin{aligned} r_t V(t, r_t, A_t) &= \frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} \\ &\quad + (\mu(r_t, t) - \lambda(r_t, t)\sigma(r_t, t)) \frac{\partial V}{\partial r_t} + f(r_t, t)\frac{\partial V}{\partial A_t}. \end{aligned} \tag{1}$$

Here $\lambda(r_t, t)$ denotes the market price of interest rate risk. A terminal condition must be specified in order to determine a single solution to the problem, so let this be given by

$$V(T, r_T, A_T) = h(T, r_T, A_T).$$

With the appropriate boundary conditions, these equations will define the value function and must in general be solved by numerical methods. Observe that the last term in equation (1) is due to the state variable and will be zero for path-independent securities, leaving the usual term-structure equation.

2.3 Discrete Sampling

When the state-variable is updated at discrete time points, the term $\frac{\partial V}{\partial A_t} f(r_t, t)$ in the PDE found above will disappear, as $dA_t = f(r_t, t)\, dt = 0$ between sampling dates. The simplification facilitates the solution compared to the continuous sampling as discrete updates of the state-variable introduce a type of jump condition. Note that in the case of continuous sampling greater care should be taken when implementing this method, but we will not get into the details here, but refer the reader to Forsyth, Vetzal & Zvan (2000) for a rigorous treatment of the numerical aspects.

Let Φ denote the time points where the state variable is updated. By definition discretely sampled state variables remain constant between the sampling dates, and on a sampling date they should be updated through a so-called *update* rule

$$A_{t_i} = U\left(t_i, r_{t_i}, A_{t_{i-1}}\right).$$

A no arbitrage argument (Wilmott et al. (1993)) will show that a corresponding jump condition will be

$$V\left(t_i^-, r_{t_i}, A_{t_{i-1}}\right) = V\left(t_i^+, r_{t_i}, U\left(t_i, A_{t_{i-1}}, r_{t_i}\right)\right), \quad i \in \Phi. \tag{2}$$

In order to provide some intuition for the jump conditions due to discrete sampling of the state variable, consider the following example. Assume we know the current value of the state variable, and that time approaches the next sampling time. The uncertainty regarding the new value of the state variable will diminish and immediately before the fixing time we will know the new value. As the realization of the price process should be continuous when no payments are made to either side of the contract, the values immediately before and after the update should be equal.

It is worth noting that a clever choice of state variable and update rule is important for optimal use of this method. As we shall see later it is sometimes possible to exploit particular properties in a given security or the update rule to reduce the dimensionality of the solution function.

2.4 Discrete Dividends

If the security pays discrete coupons an arbitrage argument leads to jump conditions. Let Ψ denote the set of dates at which the security pays the coupons $D_i\left(t_i, r_{t_i}, A_{t_i}\right)$. Following standard notation let t_i^- and t_i^+ denote the time immediately before and after the $i'th$ payment is made, respectively. This means that the $i'th$ jump condition due to coupons is

$$V(t_i^-, r_{t_i}, A_{t_i}) = V\left(t_i^+, r_{t_i}, A_{t_i}\right) + D_i\left(t_i, r_{t_i}, A_{t_i}\right), \quad i \in \Psi. \tag{3}$$

2.5 Amortization of Principal

Another feature we must be able to incorporate is the amortization of the remaining principal P_t. If t_i is the time where Z_{t_i} units of the principal are repaid, we have

$$V\left(t_i^-, r_{t_i}, P_{t_i^-}, A_{t_i}\right) = V\left(t_i^+, r_{t_i}, P_{t_i^-} - Z_{t_i}, A_{t_i}\right) + Z_{t_i}.$$

If the amortization scheme depends on the interest rate movements it will induce a special kind of path-dependency, but in most cases these value functions have a *similarity solution* without this path-dependency. As demonstrated below, securities where the amortization Z_{t_i} is linear in the remaining principal, support this similarity reduction.

If the amortization schedule Z_{t_i} is defined as a fraction $\theta(t, r_t, A_{t_i})$ of remaining principal, i.e. $Z_{t_i} = \theta(t_i, r_{t_i}, A_{t_i}) \cdot P_{t_i^-}$, then we have the following jump condition

$$V\left(t_i^-, r_{t_i}, P_{t_i^-}, A_{t_i}\right) = V\left(t_i^+, r_{t_i}, (1 - \theta(t_i, r_{t_i}, A_{t_i})) \cdot P_{t_i^-}, A_{t_i}\right) + \theta(t_i, r_{t_i}, A_{t_i}) \cdot P_{t_i^-}.$$

For fixed income securities that are homogeneous of first degree in the remaining principal P_t[1], we can apply the similarity reduction

$$V\left(t_i^-, r_{t_i}, P_{t_i^-}, A_{t_i}\right) = (1 - \theta(t_i, r_{t_i}, A_{t_i})) \cdot V\left(t_i^+, r_{t_i}, P_{t_i^-}, A_{t_i}\right) + \theta(r_{t_i}, t_i, A_{t_i}) \cdot P_{t_i^-}$$

which implies

$$V\left(t_i^-, r_{t_i}, 1, A_{t_i}\right) = (1 - \theta(t_i, r_{t_i}, A_{t_i})) \cdot V\left(t_i^+, r_{t_i}, 1, A_{t_i}\right) + \theta(t_i, r_{t_i}, A_{t_i}) \cdot 1 \tag{4}$$

This facilitates the solution as we shall find a function of one variable less. We just need to incorporate a version of this jump condition whenever principal is redeemed. Basically, we always measure the value in terms of 100% remaining principal of the security.

3 The Numerical Solution

3.1 Transformation of the PDE

We apply a standard transformation of the interest rate state space (see e.g. Duffie (1996), Stanton & Wallace (1999) or James & Webber (2000)). Define,

$$x(r) = \frac{1}{1 + \pi r}, \ \pi > 0,$$

with inverse

$$r(x) = \frac{1 - x}{\pi x}, \ \pi > 0.$$

There are mainly two reasons that we want to transform the state space for the spot rate. First, the transformation of the PDE (1) allows us to work with the solution on a bounded space. Secondly, it enables us to increase the number of points in the most relevant part of the state space using the constant π.

Let $u(x, t) = V(r(x), t)$. We now transform the PDE (1) into an PDE in u defined on the bounded state space 0 to 1.

$$\begin{aligned}
\frac{\partial V(r,t)}{\partial r} &= \frac{\partial u(x,t)}{\partial x}\frac{\partial x}{\partial r} = u_x \frac{-\pi}{(1+\pi r)^2} = -\pi x^2 u_x, \\
\frac{\partial^2 V(r,t)}{\partial r^2} &= \pi^2 x^4 \frac{\partial^2 u(x,t)}{\partial x^2} + 2\pi^2 x^3 \frac{\partial u(x,t)}{\partial x}.
\end{aligned}$$

Substituting into (1) we obtain the following PDE for u in x, t where subscripts are short hand notation for partial derivative

$$\begin{aligned}
0 &= \frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} + \tilde{\mu}(r_t, t)\frac{\partial V}{\partial r_t} - V_t r \\
&= u_t + \tfrac{1}{2}\sigma(r(x), t)^2\left(\pi^2 x^4 u_{xx} + 2\pi^2 x^3 u_x\right) + \tilde{\mu}(r(x), t)\left(-\pi x^2 u_x\right) - u r(x) \\
&= u_t + \tfrac{1}{2}\sigma(r(x), t)^2 \pi^2 x^4 u_{xx} + \pi x^2\left(\sigma(r(x), t)^2 \pi x - \tilde{\mu}(r(x), t)\right) u_x - u r(x) \\
&= u_t(x,t) + \beta(x,t) u_{xx}(x,t) + \alpha(x,t) u_x(x,t) - r(x) u(x,t),
\end{aligned} \tag{5}$$

with terminal condition

$$u(T, x_T, A_T) = h(T, r(x_T), A_T).$$

[1] Conditions for similarity reductions must also be satisfied on the boundary as well as by the terminal function.

3.1.1 Boundary conditions.

In general we need to specify boundary conditions if we are using implicit schemes to solve parabolic PDE's. However, as described in Vetzal (1998) for interest rate models with mean reversion and constant standard deviation, the PDE above behaves more like a hyperbolic PDE due to the size of the convection term, even though it is formally parabolic. Hence, it will not only be unnecessary to use boundary conditions, it will actually be most efficient to avoid specifying them.

Unfortunately not many interesting models have constant standard deviation, so we might need to use something else. However, another boundary condition arises naturally, as also noted by Vetzal (1998), by the fact that the $-r(x)\,u$ term causes exponential decay, thereby driving u and its derivatives to zero. This means that in these cases appropriate boundary conditions could be e.g. $u = 0$, $u_x = 0$ or $u_{xx} = 0$ on the upper boundary in r space (lower boundary in x space).

3.2 The Finite Difference Schemes

The PDE in (1) will in general have to be solved numerically, and in this section we describe the finite difference solution used.

Crank-Nicolson and implicit schemes are unconditionally stable, allowing us to match any cash flow, sampling, or decision date. Furthermore, as the Crank-Nicolson scheme is second-order accurate in time, we are able to take much larger steps in the time direction. However, if the terminal condition is not differentiable in the state-variable, the conditions for the Crank-Nicolson scheme are violated, which often causes oscillations in the solution. This can often be avoided by using the pure implicit scheme for the first couple of steps, or by smoothening the payoff function (see e.g. Tavella & Randall (2000)).

Therefore, we will use what is sometimes referred to as the "delta" method, which is basically a convex combination of pure explicit and implicit schemes, with the Crank-Nicolson scheme as the special case with equal weight. This implementation facilitates shifts between different finite difference schemes, by changing the weight ω.

On the boundary we use inside approximations that are second order in space, when applying the implied boundary conditions. We refer to Appendix A.2 for further details.

3.3 Implementing an Augmented State-Variable

To fix some notation let $V^n_{s,k}$, denote the value of the security at time t_n, when the short rate is r_s, and where k denotes level of the state variable. We denote the discretization of the augmented state variable by $\mathcal{A} = \{A_0, ..., A_K\}$. At all sampling times, where the augmented state variable is updated using the update scheme, the value must satisfy the jump condition in (2)

$$V\left(t_j^-, r_{t_j}, A_{j-1}\right) = V\left(t_j^+, r_{t_j}, U\left(t_j, r_{t_j}, A_{j-1}\right)\right) = V\left(t_j^+, r_{t_j}, A_j\right).$$

However, the update function U does not necessarily take values in $\mathcal{A}$, so we will not know the exact value of $V\left(t_j^+, r_{t_j}, A_j\right)$. The basic idea in this method is to approximate it by interpolating the future values at known levels of A.

With a view to this interpolation, define the mapping function $k^*(A) : \mathbb{R} \to \{0, ..., K\}$ by

$$A_{k^*} \leq A < A_{k^*+1}.$$

That is, the mapping picks the index of the highest level of the state variable that is still less than or equal to the value A, assuming that the discretization of the state space has been done such that this is a well-defined mapping. Notice that if V is non-linear in the state variable, we get a biased estimate using linear interpolation. E.g. if V is a convex function of A, then the estimate is too high.

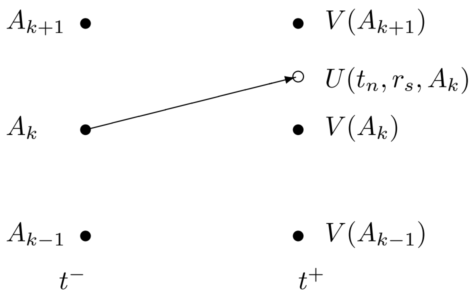

Figure 1: Illustration of the interpolation in the augmented state space.

We use either linear- or polynomial interpolation of order 2[2]. Algorithms can be found in Press, Flannery, Teukolsky & Vetterling (1989), and written in a pseudo notation we get

$$V_{s,k}^{n-} = int\left(A_k, \{A_{k^*-1}, A_{k^*}, A_{k^*+1}\}, \left\{V_{s,k^*-1}^{n+}, V_{s,k^*}^{n+}, V_{s,k^*+1}^{n+}\right\}\right).$$

It is possible to make the number of levels of the augmented state space time and state-dependent in order to minimize calculation time, as there is no need to consider levels of the state variable that are not feasible. In situations where the state variable is monotonincally increasing or decreasing, a simple example could be to use the current value as either upper or lower bound of the augmented state space.

4 Applications

The technique can be applied to a wide range of path-dependent securities. The essential part is to make a clever choice of state variables and update rules. As a complicated example Dewynne & Wilmott (n.d.) show how to value a trend based option like a "Five-times-up-and-out" using this approach.

In the following numerical analysis we use the Cox-Ingersoll-Ross model,

$$\mu(r_t, t) = \kappa\,(\mu - r_t)\,, \quad \sigma(r_t, t) = \sigma\sqrt{r_t}, \quad \lambda\,(r_t, t) = \lambda^{CIR}\sqrt{r_t}/\sigma$$

with parameters as given in table 1.

κ	μ	σ	λ^{CIR}
0.3	0.08	0.12	0

Table 1: Parameters in the CIR model

4.1 Mortgage Backed Securities

A standard mortgage backed security(MBS) is a fixed rate mortgage with an embedded option that allows the borrower to repay the remaining principal at par at any time during the life of the mortgage. This means that when refinancing rates fall, borrowers prepay their loans by taking up new loans at the prevailing market rate. Any reasonable pricing model for MBS's is designed to incorporate what is known as the burnout effect, namely that borrowers most inclined to prepay leave the mortgage pool, causing future prepayment rates to decline *ceteris paribus*.

[2] Other interpolation schemes such as cubic splines and rational interpolation have been tested without improvements.

This heterogeneity among borrowers can be modelled in basically two ways, which we will denote explicit and implicit burnout. Implicit modelling of burn out consists of summarizing the historical interest and prepayment behavior in state variables which enter directly into the prepayment function. This is also termed a path-dependent prepayment function. Early contributions in this direction were made by Schwartz & Torous (1989) and Richard & Roll (1989). Examples of explicit modelling of burnout can be found in Jakobsen (1992) and Stanton (1995). By regarding a bond in a large and heterogeneous mortgage pool as a portfolio of homogeneous sub pools, each having a path-independent prepayment function, they demonstrate that changes in the mixture of borrowers will induce a burnout pattern very similar to that of the implicit models.

When it comes to valuing MBS, Monte Carlo simulation has by some been considered superior to backward induction techniques as Monte Carlo simulation allows the prepayment model to combine the two approaches, but as shown here so do recombining lattice methods. On the other hand, Monte Carlo simulation has the serious flaw, namely that it is unable, or at least unsuitable to handle MBS's under rational prepayment behavior. Especially the fact that American or Bermudan option pricing is very hard to do by Monte Carlo simulation, means that we cannot use this approach to compute the optimal prepayment strategy. Furthermore, as mentioned earlier, the finite difference approach facilitates the task of valuing options on MBS's or CMO's as we just use backward induction.

As mentioned above we need to define the state variable and the update rule in order to make use of the method. One variable that has been applied in many prepayment models in various forms is a so called pool factor B_j, that measures the current remaining principal relative to the originally scheduled. If θ_j denotes the conditional prepayment rate, i.e. the fraction of the remaining borrowers that prepay at time t_j, we have that

$$B_j = \Pi_{i=1}^{j} (1 - \theta_i), \quad B_0 = 1.$$

The update rule U is given by

$$B_j = U(t, r_t, B_{j-1}) = B_{j-1} \cdot (1 - \theta_j),$$

Assume that the conditional prepayment rate (CPR) $\theta_j = f\left(t_j, r_{t_j}, B_{j-1}\right)$ is a function f of some explaining variables, one of them being the pool factor, making the prepayment model path-dependent. This means that at a term of notice[3], where the borrowers have to decide whether to prepay or not, we apply the jump condition

$$V(t^-, r_t, B_{j-1}) = \theta_j \cdot 1 + (1 - \theta_j) \cdot V\left(t^+, r_t, U(t, r_t, B_{j-1})\right), \tag{6}$$

measured in terms of principal at time t^-.

4.1.1 MBS: An example

As an example we consider the pricing of a 20-year annuity bond, with a fixed 8% coupon and quarterly payments, where the borrowers' behavior is described by the very simplified but path-dependent prepayment function for the conditional prepayment rate,

$$\theta_j\left(r_{t_j}, B_{j-1}\right) = \min\left((1 + 30 \cdot B_{j-1}) \cdot \left(Coupon - \left(r_{t_j} + 1\%\right)\right)^+, 100\%\right).$$

[3] Almost all mortgages have a term of notice, but in these examples we ignore these features, such that prepayment decisions are taken at the term date.

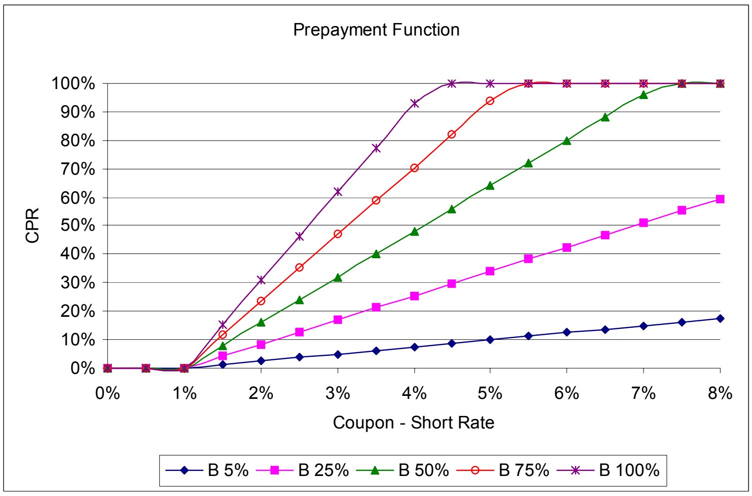

Figure 2: Illustration of the path-dependent prepayment function. CPR denotes the conditional prepayment rate and B the path-dependen burnout factor.

Table 2 illustrates the convergence of the value of the MBS for different values of the short rate as we increase the density of the discrete augmented state space using first linear and then quadratic interpolation. Monte Carlo estimates based on 40.000 paths using antithetic variates as variance reduction (a total of 80.000 paths) are given below. The right column shows the differences between the finite difference solution and the MC estimates measured in basis points. For three out of four levels of the short rate we cannot at a 95% significance level reject the hypotehesis that the MC and PDE values are equal, when the number of states is high enough. However for all practical applications the differences are not significant as they are way inside bid-ask spreads, which are at least 10 bps. These results also confirm the conclusions in Hull & White (1993), namely that the quadratic interpolation seems to improve the method when K is low. The table also shows computation times and it is obvious that the method is quite efficient compared to this particular MC implementation. The interesting thing here is not whether the numbers are exactly equal, as we know that in the limit both methods will give us the correct values. The basic point is that the PDE is fully able to handle this path dependency; there is no need to simulate in this case.

4.2 Collateralized Mortgage Obligations

CMO's are constructed by allocating the payments from the underlying collateral (usually MBS) into different new securities (called tranches). Depending on the redistribution of the payments, these tranches can have characteristics that are indeed very different from those of the collateral. In practice all kinds of CMO's are created to fit investor preferences.

McConnell & Singh (1994) propose a two-step procedure to value CMO's under rational prepayments. The rational exercise of the prepayment option precludes MC as a feasible solution procedure regarding the prepayment decisions, so they find the optimal exercise boundary by finite difference. In the second step they use MC to work forward in time distributing the cash flows using the optimal exercise boundary found in step one as a prepayment function. In relation to the augmented state space approach, McConnell & Singh (1994) claim that it is necessary to include a state variable for each tranche making the approach technically unfeasible. However, if the allocation of the cash flow is based on the remaining debt alone, we do not need more than one state variable per sub pool of borrowers. We do not need a state variable for each tranche.

We now show that some of the most widely used CMO structures can be priced using

Table 2: Convergence of the PDE solution for the Mortgage Backed Security

Linear No. Aug	Comp. Time	PDE Short Rate 2%	4.8%	8%	12%	PDE-MC (bps) 2%	4.8%	8%	12%
3	1	101.484	101.023	96.570	88.949	2	44	46	35
5	1	101.476	100.810	96.317	88.767	1	23	21	16
7	2	101.471	100.728	96.231	88.707	0	15	12	10
9	3	101.469	100.690	96.191	88.679	0	11	8	8
11	3	101.468	100.666	96.169	88.664	0	9	6	6
13	3	101.467	100.651	96.155	88.654	0	7	5	5
15	4	101.466	100.642	96.146	88.648	0	6	4	4
17	4	101.466	100.636	96.140	88.643	0	5	3	4
19	5	101.465	100.630	96.135	88.640	0	5	3	4
21	5	101.465	100.627	96.131	88.638	0	5	2	3
31	8	101.464	100.616	96.122	88.631	0	4	1	3
41	9	101.464	100.612	96.119	88.629	0	3	1	3
61	14	101.464	100.609	96.116	88.627	0	3	1	2
81	18	101.464	100.608	96.115	88.626	0	3	1	2
MC	598	101.47	100.58	96.11	88.60				
Std.Dev		0.00	0.01	0.02	0.02	0.2	1.2	2.1	2.0

Quadratic No. Aug.	Comp. Time	PDE Short Rate 2%	4.8%	8%	12%	PDE-MC (bps) 2%	4.8%	8%	12%
3	1	101.475	100.695	96.169	88.664	1	11	6	6
5	2	101.460	100.512	96.023	88.563	-1	-7	-9	-4
7	2	101.458	100.544	96.064	88.592	-1	-4	-5	-1
9	3	101.460	100.572	96.088	88.609	-1	-1	-2	1
11	4	101.462	100.587	96.101	88.618	-1	1	-1	1
13	5	101.462	100.596	96.108	88.623	0	2	0	2
15	6	101.463	100.603	96.113	88.626	0	2	0	2
17	6	101.463	100.606	96.115	88.627	0	2	1	2
19	7	101.463	100.608	96.117	88.628	0	3	1	3
21	7	101.463	100.609	96.118	88.629	0	3	1	3
31	11	101.463	100.610	96.118	88.629	0	3	1	3
41	14	101.464	100.610	96.117	88.628	0	3	1	2
61	21	101.464	100.608	96.116	88.627	0	3	1	2
81	27	101.464	100.608	96.115	88.626	0	3	1	2
MC	598	101.47	100.58	96.11	88.60				
Std.dev		0.00	0.01	0.02	0.02	0.2	1.2	2.1	2.0

This table illustrate the convergence of the PDE approach using a Linear and Qudratic interpolation scheme. No. Aug denotes the number of spatial grid points in the augmented state-space and Comp. Time the calculation time in seconds. MC denotes the Monte-Carlo estimates for various levels of initial short rate and Std.Dev. is the standard deviation. PDE-MC is the difference between PDE and the MC estimates measured in basis points. In the PDE implementation a Crank-Nicolson scheme was used with 80 spatial grid points and 24 steps per year.

the augmented state variable approach. There are only a few limitations. For example, we will not be able to calculate measures such as weighted average life (WAL), as the state price distribution in the augmented state space is unknown.

4.2.1 Mortgage Strips

One of the most natural ways to split the total cash flow received from the collateral, is into principal and interest payments. These mortgage strips are also known as Interest Only (IO) and Principal Only (IO). The holder of an IO receives all interest payments from the collateral, while the PO holders receive the scheduled as well as unscheduled repayment on the principal. It is clear that the value of the IO and the PO together should equal that of the collateral, i.e.

$$V^C = V^{IO} + V^{PO}.$$

If we use this fact to rewrite equation (6), it follows that

$$V^C(t^-, r_t, B_{j-1}) = \theta_j \cdot 1 + (1-\theta_j) \cdot V^C\left(t^+, r_t, U\left(t, r_t, B_{j-1}\right)\right)$$

which implies

$$\begin{aligned} V^{IO}(t^-, r_t, B_{j-1}) + V^{PO}(t^-, r_t, B_{j-1}) &= \theta_j \cdot 1 + (1-\theta_j) \cdot V^{IO}\left(t^+, r_t, U\left(t, r_t, B_{j-1}\right)\right) \\ &\quad + (1-\theta_j) \cdot V^{PO}\left(t^+, r_t, U\left(t, r_t, B_{j-1}\right)\right) \end{aligned}$$

The PO receives all repayments, and the IO looses the future interest corresponding to the prepaid principal. Hence, the jump equations due to prepayment will look like

$$\begin{aligned} V^{IO}(t^-, r_t, B_{j-1}) &= (1-\theta_j) \cdot V^{IO}\left(t^+, r_t, U\left(t, r_t, B_{j-1}\right)\right), \\ V^{PO}(t^-, r_t, B_{j-1}) &= \theta_j \cdot 1 + (1-\theta_j) \cdot V^{PO}\left(t^+, r_t, U\left(t, r_t, B_{j-1}\right)\right). \end{aligned}$$

From these it is clear that both these tranches are mildly path-dependent if the prepayment function is path-independent. Hence, the mortgage strips can be evaluated in exactly the same way as the collateral.

4.2.2 Sequential Pay Tranches

As mentioned in section (4.1) there are several reasonable measures for the historical interest rate and prepayment behavior, but the pool factor definition chosen above has the additional advantage that it can also be used to value CMO structures, where we can not apply the similarity reduction. The sequential pay tranches are examples of such structures, as the value of the tranches are *not* linear in remaining debt.

An example Consider two tranches T_1 and T_2 on a collateral of 100 units of the MBS from before. Tranche 1 receives the first W_1 percent of the collateral, and when all

CMO	Nominal	Coupon
Collateral	100	C
Tranche A	$W_1 \cdot 100$	C_1
Tranche B	$W_2 \cdot 100$	C_2

Table 3: Example of Sequential Pay CMO

principal in tranche T_1 has been redeemed, tranche T_2 starts receiving principal. Both tranches receive interest on the remaining principal. Notice that if $C_1 > C_2$ there will be an interest deficit after the first installment on the principal, and an interest excess if $C_1 < C_2$. In these cases issuers often add a residual class - a so called Z-bond, but we will not go into these details. The number of CMO constructions is almost infinite and only the inventiveness seems to set the limit.

To keep things simple let us assume that tranche T_1 receives the first 60% of the principal and that tranche T_2 gets the rest, but that they both pay the same interest as the collateral, i.e. $W_1 = 60\%$, $W_2 = 40\%$, and $C = C_1 = C_2 = 8\%$.

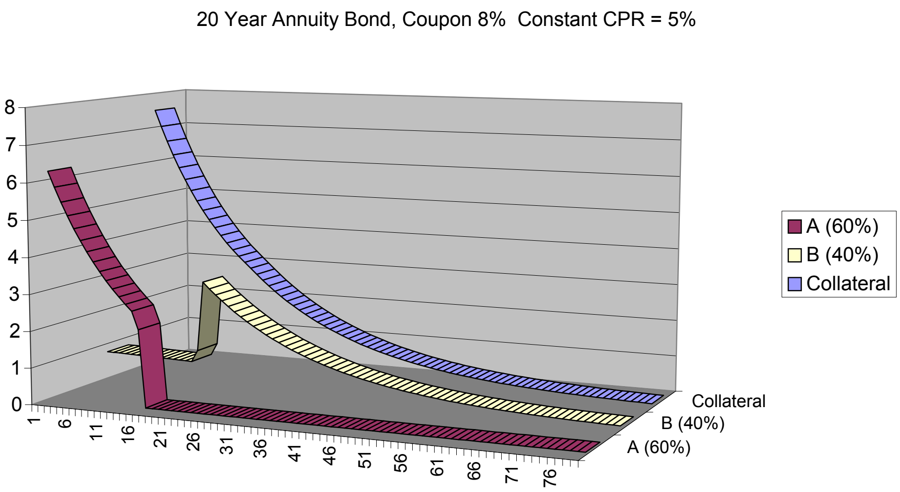

Figure 3: Illustration of cash flows for the tranches, given the cashflow from the collateral.

Notice that this construction has no similarity reduction as the amortization is *not* linear in the principal. Here the use of the augmented state variable is crucial, even if the prepayment function is path-independent. By using the pool factor defined above as state variable, we will be able to decide how much the individual tranches should receive at a given time for a given spot rate. To find the nominal value of the remaining debt we multiply the pool factor with the scheduled remaining principal in case of no prepayments. Denote the scheduled remaining principal in case of no prepayments by $\hat{P}_j$ and the actual remaining principal after the $j'th$ payment P_j, both measured in percent of initial principal. Then by definition of the pool factor,

$$P_j = B_j \cdot \hat{P}_j.$$

Given the nominal value we can allocate the cash flow to the tranches in accordance with the definition as we do when we go forward during the MC simulation. At time j we let Z_j denote the total repayment, TZ_j the total repayment since time 0, I_j^i and Z_j^i denote the interest and repayment for the i'th tranche, and $\bar{I}$ is the number of trances.

$$\begin{aligned} Z_j &= P_{j-1} - P_j, \\ TZ_j &= 1 - P_j, \\ P_j^i &= \left(\left(\sum_{m=1}^{i} W_m - TZ_j \right) \wedge W_i \right)^+, \quad i = 1, .., \bar{I} \\ Z_j^i &= \left(Z_j - \sum_{m=1}^{i-1} Z_j^m \right) \wedge P_{j-1}^i, \quad i, ..., \bar{I} \\ I_j^i &= P_{j-1}^i \cdot C_i, \quad i = 1, ..., \bar{I}. \end{aligned}$$

Given these expressions we can now state the following jump condition for the value of tranche i at time j

$$V^i\left(t^-, r_t, B_{j-1}\right) = V^i\left(t^-, r_t, B_j\right) + Z_j^i + I_j^i, \quad i = 1, ..., \bar{I}.$$

Numerical Results for Sequential Tranches As for the numerical results for the sequential tranches reported in Tables 4 and 5, there are at least three things worth mentioning. First, the differences between the MC results and the PDE approach are small when the number of state levels K is high enough. Secondly, however, as opposed to the conclusions in Hull & White (1993) and the results for the collateral, the tranches are quite sensitive with regard to the number of state levels. We need much more than 6 levels in order to obtain reasonable results. Thirdly, we also see that for small values of K the quadratic interpolation scheme performs worse than the linear scheme. The two latter points are not that surprising though, as the value function is not smooth in the state variable.

4.3 Average Rate Capped Amortizing ARM

We examine a security traded in the Danish mortgage market named BoligX. The construction of the security is quite non-standard for several reasons. The BoligX loan is a 5-year adjustable rate mortgage ARM that can be issued with or without an embedded 5-year cap. Usually a cap on an ARM is paid for separately, but the BoligX loan is a genuine pass-through in the sense that payments from the borrowers are paid directly to the mortgage holders, and the cap with strike κ is paid for through a premium rate ρ.

There are quarterly payments which are settled in pairs twice a year. The size of the payments are based on the borrower having an adjustable rate annuity mortgage with m payments, typically 80 or 120 corresponding to 20- or 30-year. The coupon on the underlying mortgage is reset twice a year as a day arithmetic average of the 6-month Cibor rate over a prespecified 10 days fixing period.

This means that at the $n'th$ fixing time, the next two payments are equal to the payment received from an annuity with $m - 2n$ periods and a coupon rate that is C_n. On top of the average will be a coupon premium to pay for the cap. Due to this construction of the security there will be repayment on the principal, and this repayment increases/decreases as interest rates decrease/increase.

Let N denote the number of fixing periods and $\left\{s_n^1, .., s_n^{10}\right\}$ the set of dates in the n'th fixing period. Furthermore, let t_n^1 and t_n^2 denote payment times for the payments settled at time s_n^{10}. Hence, the n'th coupon rate will be given as

$$C_n = \min\left(A_n + \rho, \kappa\right),$$

where $A_n = \frac{1}{10}\sum_{i=1}^{10} r\left(s_n^i\right)$. The size of the payments settled in period n can then be found from the standard annuity formula

$$Y_n = P_n \frac{C_n}{1 - \left(1 + C_n\right)^{-(m-2(n-1))}}, \quad n = 1, ..., \frac{m}{2}$$

where $m - 2\left(n - 1\right)$ is the number of remaining payments and P_n the remaining principal outstanding at fixing time n.

In order to model the settlement of the coupon rate as an average of previous interest rates, we let the state variable A be the discretely sampled average of the short rate . The update rule in the case of a discretely sampled average as a state variable, can be written as

$$A\left(s_n^i\right) = U\left(s_n^i, r\left(s_n^i\right), A\left(s_n^{i-1}\right)\right) = \frac{1}{i} r\left(s_n^i\right) + \frac{i-1}{i} A\left(s_n^{i-1}\right).$$

Another non-standard feature of the BoligX loan is that the sampling takes place before the actual accruement period. But as the payments are known at the fixing time

Table 4: Convergence of the PDE solution for Tranche A

Linear		PDE				PDE-MC			
No.	Comp.	Short Rate				(bps)			
Aug.	Time	2%	4.8%	8%	12%	2%	4.8%	8%	12%
3	3	61.326	69.465	66.295	58.456	47	924	848	494
5	3	60.931	63.081	60.962	55.451	7	285	315	193
7	4	60.885	61.382	59.248	54.469	2	115	143	95
9	5	60.872	60.744	58.501	54.032	1	51	68	51
11	6	60.862	60.346	58.062	53.776	0	12	25	26
13	6	60.866	60.424	57.981	53.662	1	19	16	14
15	8	60.868	60.479	58.019	53.649	1	25	20	13
17	8	60.867	60.453	58.016	53.649	1	22	20	13
19	10	60.865	60.376	57.965	53.631	0	15	15	11
21	10	60.862	60.285	57.897	53.603	0	5	8	8
31	15	60.861	60.267	57.853	53.565	0	4	4	5
41	19	60.861	60.257	57.835	53.553	0	3	2	3
61	29	60.861	60.251	57.821	53.543	0	2	0	2
81	38	60.861	60.248	57.815	53.540	0	2	0	2
MC	574	60.86	60.23	57.82	53.52				
Std.Dev		0.00	0.01	0.01	0.01	0.0	0.5	1.2	1.2

Quadratic		PDE				PDE-MC			
No.	Comp.	Short Rate				(bps)			
Aug.	Time	2%	4.8%	8%	12%	2%	4.8%	8%	12%
3	2	59.862	31.239	21.175	23.957	-100	-2899	-3664	-2956
5	4	59.415	43.296	48.234	49.676	-145	-1693	-958	-384
7	5	60.885	61.382	59.248	54.469	2	115	143	95
9	7	60.872	60.744	58.501	54.032	1	51	68	51
11	8	60.862	60.346	58.062	53.776	0	12	25	26
13	10	60.866	60.424	57.981	53.662	1	19	16	14
15	11	60.868	60.479	58.019	53.649	1	25	20	13
17	12	60.867	60.453	58.016	53.649	1	22	20	13
19	13	60.865	60.376	57.965	53.631	0	15	15	11
21	15	60.862	60.285	57.897	53.603	0	5	8	8
31	21	60.861	60.267	57.853	53.565	0	4	4	5
41	28	60.861	60.257	57.835	53.553	0	3	2	3
61	42	60.861	60.251	57.821	53.543	0	2	0	2
81	54	60.861	60.248	57.815	53.540	0	2	0	2
MC	574	60.86	60.23	57.82	53.52				
Std.dev		0.00	0.01	0.01	0.01	0.0	0.5	1.2	1.2

This table illustrate the convergence of the PDE approach using a Linear and Qudratic interpolation scheme. No. Aug denotes the number of spatial grid points in the augmented state-space and Comp. Time the calculation time in seconds. MC denotes the Monte-Carlo estimates for various levels of initial short rate and Std.Dev. is the standard deviation. PDE-MC is the difference between PDE and the MC estimates measured in basis points. In the PDE implementation a Crank-Nicolson scheme was used with 80 spatial grid points and 24 steps per year.

Table 5: Convergence of the PDE solution for Tranche B

Linear No. Aug.	Comp. Time	PDE Short Rate 2%	4.8%	8%	12%	PDE-MC (bps) 2%	4.8%	8%	12%
3	3	40.115	30.847	29.542	29.967	-49	-950	-875	-512
5	3	40.517	37.326	34.971	33.045	-9	-302	-332	-204
7	4	40.566	39.083	36.740	34.067	-4	-127	-155	-102
9	5	40.582	39.755	37.519	34.528	-2	-60	-77	-56
11	6	40.594	40.177	37.980	34.799	-1	-17	-31	-29
13	6	40.591	40.115	38.075	34.924	-2	-24	-22	-16
15	8	40.590	40.071	38.047	34.943	-2	-28	-25	-14
17	8	40.592	40.105	38.057	34.949	-1	-25	-24	-13
19	10	40.595	40.190	38.115	34.971	-1	-16	-18	-11
21	10	40.598	40.286	38.187	35.002	-1	-6	-11	-8
31	15	40.600	40.319	38.244	35.050	-1	-3	-5	-3
41	19	40.601	40.337	38.267	35.065	-1	-1	-3	-2
61	29	40.601	40.348	38.287	35.079	0	0	-1	-1
81	38	40.602	40.354	38.295	35.083	0	0	0	0
MC	574	40.61	40.35	38.29	35.08				
Std.Dev		0.00	0.01	0.01	0.01	0.1	0.8	1.0	0.9

Quadratic No. Aug.	Comp. Time	PDE Short Rate 2%	4.8%	8%	12%	PDE-MC (bps) 2%	4.8%	8%	12%
3	2	41.588	69.304	74.925	64.667	98	2895	3663	2958
5	4	42.185	59.347	49.621	40.072	158	1900	1133	499
7	5	40.613	39.634	37.205	34.385	1	-72	-109	-70
9	7	40.613	40.111	37.815	34.728	1	-24	-48	-36
11	8	40.616	40.422	38.181	34.934	1	7	-11	-15
13	10	40.607	40.292	38.218	35.019	0	-6	-7	-6
15	11	40.603	40.205	38.153	35.013	0	-15	-14	-7
17	12	40.602	40.208	38.138	35.002	0	-14	-16	-8
19	13	40.603	40.271	38.177	35.012	0	-8	-12	-7
21	15	40.605	40.351	38.237	35.034	0	0	-6	-5
31	21	40.604	40.348	38.263	35.061	0	0	-3	-2
41	28	40.603	40.350	38.276	35.070	0	0	-2	-1
61	42	40.602	40.352	38.289	35.080	0	0	0	0
81	54	40.602	40.355	38.296	35.084	0	0	0	0
MC	574	40.61	40.35	38.29	35.08				
Std.Dev		0.00	0.01	0.01	0.01	0.1	0.8	1.0	0.9

This table illustrate the convergence of the PDE approach using a Linear and Qudratic interpolation scheme. No. Aug denotes the number of spatial grid points in the augmented state-space and Comp. Time the calculation time in seconds. MC denotes the Monte-Carlo estimates for various levels of initial short rate and Std.Dev. is the standard deviation. PDE-MC is the difference between PDE and the MC estimates measured in basis points. In the PDE implementation a Crank-Nicolson scheme was used with 80 spatial grid points and 24 steps per year.

s_n^{10}, the time s_n^{10} present value of the payments settled is just $Y_n \left(B\left(s_n^{10}, t_n^1\right) + B\left(s_{n,}^{10} t_n^2\right)\right)$, where $B(t,T)$ denotes the time t value of a discount bond maturing at time T.

A part of these two payments is amortized principal, and hence we will need to incorporate this into a jump condition. It can easily be shown that the amortization rate θ due to the first two payments of an annuity bond with an initial nominal of P_n, $m-2(n-1)$ payments and coupon C_n is

$$\theta_n = \frac{Y_n \cdot (1+C_n)^{-(m-2(n-1))} \cdot (2+C_n)}{P_n}.$$

We are now ready to state the jump conditions for the sampling dates in period n

$$V(s_n^{j-}, r(s_n^j), A(s_n^{j-1})) = V(s_n^{j+}, r(s_n^j), A(s_n^j)), \quad j = 1, ..., 9$$

At the last sampling date in period n we also add the present value of the two payments and apply the jump condition due to the armortized principal.

$$\begin{aligned} V(s_n^{10-}, r\left(s_n^{10}\right), A\left(s_n^9\right)) &= Y_n \cdot \left(B(s_n^{10}, t_n^1) + B(s_n^{10}, t_n^2)\right) \\ &\quad + (1-\theta_n) \cdot V\left(s_n^{10+}, r\left(s_n^{10}\right), A\left(s_n^{10}\right)\right), \end{aligned}$$

At the very last payment date the investor also receives the remaining principal, while the intermediary issues a new BoligX loan on behalf of the borrower.

Numerical Results BoligX The premium ρ in the example is 20 bps and the cap rate κ is 7.7%. In table 6 we see that only the out the money value is more than two standard deviations away from the MC value. As there are no differences in the performance of the linear or quadratic interpolation scheme when $K \geq 5$, there is no reason to use anything other than linear interpolation.

5 Conclusions

In this paper we have analyzed a numerical method that efficiently allows valuation of a class of path-dependent interest rate derivatives in a finite difference setup. We have focused on mortgage backed security valuation in particular and we show that this method is able to handle both the American feature but also path-dependencies present in MBS's. Furthermore, the method is at least as efficient as standard Monte Carlo techniques for similar precision, even when we consider 20- or 30-year products.

There are of course limitations to the application of this method due to the curse of dimensionality. If the dimension of the augmented state vector is high, we will not only have to make use of a high dimensional interpolation scheme, but the number of points in the discretized augmented state space will increase exponentially with the dimension. For example, suppose we have a mortgage pool that consists of say 4 sub pools or more with different prepayment behavior. The valuation of a sequential pay CMO, would require us to use a 4-dimensional state vector to summarize all possible combinations of remaining debt or equivalently burnout in the sub pools.

At last we mention that this method can also be used to model and access the value of the delivery options embedded in for example Danish mortgage backed bonds. A delivery option gives the borrower the right to buy back her own loan from the mortgage pool at market value. The presence of this option means that we almost never see prepayments below par. In order to model this option we will need to know the market value of the mortgage at each future time and state. That is, we need not only know the values of the loans in individual sub pools but also their relative share of the total mortgage pool.

Crank-Nicolson Scheme, S=80, SPY=96

Linear		PDE				PDE-MC			
No.	Comp.	Short Rate				(bps)			
Aug.	Time	2%	4.8%	8%	12%	2%	4.8%	8%	12%
3	3	98.637	97.438	94.827	89.093	-23	-37	-37	-17
5	3	98.675	97.501	94.871	89.116	-19	-31	-33	-14
7	4	98.722	97.570	94.917	89.139	-15	-24	-28	-12
9	4	98.758	97.627	94.958	89.161	-11	-18	-24	-10
11	4	98.783	97.627	94.993	89.179	-9	-14	-21	-8
13	6	98.801	97.698	95.022	89.193	-7	-11	-18	-7
15	6	98.815	97.719	95.046	89.205	-5	-9	-15	-5
21	8	98.841	97.761	95.098	89.232	-3	-5	-10	-3
31	11	98.861	97.794	95.147	89.256	-1	-2	-5	0
41	14	98.873	97.813	95.179	89.272	0	0	-2	1
61	20	98.882	97.827	95.204	89.284	1	2	1	3
81	25	98.882	97.828	95.205	89.285	1	2	1	3
101	32	98.882	97.828	95.205	89.285	1	2	1	3
MC	475	98.87	97.81	95.20	89.26				
std.dev		0.008	0.011	0.013	0.011	0.8	1.1	1.3	1.1

Quadratic		PDE				PDE-MC			
No.	Comp.	Short Rate				(bps)			
Aug.	Time	2%	4.8%	8%	12%	2%	4.8%	8%	12%
3	3	99.356	98.948	98.111	96.586	49	114	291	733
5	4	98.675	97.501	94.871	89.116	-19	-31	-33	-14
7	4	98.722	97.570	94.917	89.139	-15	-24	-28	-12
9	5	98.758	97.627	94.958	89.161	-11	-18	-24	-10
11	6	98.783	97.668	94.993	89.179	-9	-14	-21	-8
13	7	98.801	97.698	95.022	89.193	-7	-11	-18	-7
15	8	98.815	97.719	95.046	89.205	-5	-9	-15	-5
21	10	98.841	97.761	95.098	89.232	-3	-5	-10	-3
31	14	98.861	97.794	95.147	89.256	-1	-2	-5	0
41	19	98.873	97.813	95.179	89.272	0	0	-2	1
61	27	98.882	97.827	95.204	89.284	1	2	1	3
81	36	98.882	97.828	95.205	89.285	1	2	1	3
101	44	98.882	97.828	95.205	89.285	1	2	1	3
MC	475	98.87	97.81	95.20	89.26				
std.dev		0.008	0.011	0.013	0.011	0.8	1.1	1.3	1.1

Table 6: BoligX: Convergence of the PDE approach as we increase density in the augmented state space. Monte Carlo estimates and standard deviations are below.

References

Andersen, L. B. G. (1996), Five Essays on the Pricing of Contingent Claims, PhD thesis, The Aarhus School of Business.

Dewynne, J. & Wilmott, P. (n.d.), Untitled.

Duffie, D. (1996), *Dynamic Asset Pricing*, Princeton University Press.

Forsyth, P., Vetzal, K. & Zvan, R. (2000), 'Convergence of lattice and PDE methods for pricing asian options', *Working Paper, University of Waterloo* .

Hilliard, J., Kau, J., Keenan, D. & Muller III, W. (1995), 'Pricing a class of american and european path dependent securities', *Management Science* **41**(12), 1892–1899.

Hull, J. & White, A. (1993), 'Efficient procedures for valuing european and american path-dependent options', *The Journal of Derivatives* pp. 21–31.

Jakobsen, S. (1992), Prepayment and the Valuation of Mortgage-Backed Bonds, PhD thesis, The Aarhus School of Business.

James, J. & Webber, N. (2000), *Interest Rate Modelling*, John Wiley and Sons, Ltd.

McConnell, J. & Singh, M. (1994), 'Rational prepayments and the valuation of collateralized mortgage obligations', *The Journal of Finance* **XLIX**(3), 891–921.

Press, W., Flannery, B., Teukolsky, S. & Vetterling, W. (1989), *Numerical Recipes in Pascal*, Cambridge University Press.

Richard, S. & Roll, R. (1989), 'Prepayments on fixed-rate mortgage backed securities', *Journal of Portfolio Management* pp. 73–82.

Schwartz, E. & Torous, W. (1989), 'Prepayment and the valuation of mortgage-backed securities', *The Journal of Finance* **XLIV**(2), 375–392.

Stanton, R. (1995), 'Rational prepayment and the valuation of mortgage-backed securities', *Review of Financial Studies* **8**, 677–708.

Stanton, R. & Wallace, N. (1999), 'Anatomy of and ARM: The interest-rate risk of adjustable-rate mortgages', *Journal of Real Estate Finance and Economics* **19**(1), 49–67.

Tavella, D. & Randall, C. (2000), *Pricing Financial Instruments - The Finite Difference Method*, John Wiley and Sons, Inc.

Vetzal, K. (1998), 'An improved finite difference approach to fitting the initial term structure', *The Journal of Fixed Income* pp. 62–81.

Wilmott, P., Dewynne, J. & Howison, S. (1993), *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press.

A Appendix

A.1 The Derivation of the Fundamental PDE

If we assume that the value function $V(t, r_t, A_t)$ satisfies the usual regularities we can apply Itô's lemma to find the dynamics for the value of the claim

$$\begin{aligned} dV(t, r_t, A_t) &= \frac{\partial V}{\partial r_t} dr_t + \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial A_t} dA_t + \tfrac{1}{2} \frac{\partial^2 V}{\partial r_t^2} d\langle r_t \rangle \\ &= \left(\mu(r_t, t) \frac{\partial V}{\partial r_t} + \frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} + f(r_t, t) \frac{\partial V}{\partial A_t} \right) dt \\ &\quad + \sigma(r_t, t) \frac{\partial V}{\partial r_t} dW_t. \end{aligned}$$

Let now $X_t = F(r_t, t)$ denote the price of another security depending on the short rate e.g. a zero coupon bond of maturity T, governed by the following SDE

$$\begin{aligned} dX_t &= \frac{\partial F}{\partial r_t} dr_t + \frac{\partial F}{\partial t} dt + \tfrac{1}{2} \frac{\partial^2 F}{\partial r_t^2} d\langle r_t \rangle \\ &= \frac{\partial F}{\partial r_t} dr_t + \frac{\partial F}{\partial t} dt + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 F}{\partial r_t^2} dt \end{aligned}$$

If we sell α_t units of X_t and for each unit of V, we see that the value of the portfolio changes as

$$\begin{aligned} d(V_t - \alpha_t X_t) &= dV_t - \alpha_t dX_t \\ &= \left(\frac{\partial V}{\partial r_t} dr_t + \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial A_t} dA_t + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} dt \right) \\ &\quad - \alpha_t \left(\frac{\partial F}{\partial r_t} dr_t + \frac{\partial F}{\partial t} dt + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 F}{\partial r_t^2} dt \right) \end{aligned}$$

By choosing α_t such that

$$\left(\frac{\partial V}{\partial r_t} - \alpha_t \frac{\partial F}{\partial r_t} \right) dr_t = 0 \Leftrightarrow \alpha_t = \frac{\partial V / \partial r_t}{\partial F / \partial r_t},$$

the change in value of the portfolio is deterministic. Hence the drift should be equal to the short rate,

$$\begin{aligned} d(V_t - \alpha_t X_t) &= d\left(V_t - \frac{\partial V/\partial r_t}{\partial F/\partial r_t} X_t \right) = \left(V_t - \frac{\partial V/\partial r_t}{\partial F/\partial r_t} X_t \right) r_t dt \\ &\Leftrightarrow \\ &\quad \left(\frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial A_t} dA_t + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} dt \right) \\ &\quad - \frac{\partial V/\partial r_t}{\partial F/\partial r_t} \left(\frac{\partial F}{\partial t} + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 F}{\partial r_t^2} \right) dt = \left(V_t - \frac{\partial V/\partial r_t}{\partial F/\partial r_t} X_t \right) r_t dt \\ &\Leftrightarrow \\ &\quad \frac{\partial V}{\partial t} dt + \frac{\partial V}{\partial A_t} dA_t + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} dt - V_t r_t dt \\ &= \frac{\partial V/\partial r_t}{\partial F/\partial r_t} \left(\frac{\partial F}{\partial t} + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 F}{\partial r_t^2} - X_t r_t \right) dt \\ &\Leftrightarrow \\ &\quad \frac{\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} - V_t r + \frac{\partial V}{\partial A_t} f(r_t, t) \right)}{\partial V / \partial r_t} dt \\ &= \frac{\left(\frac{\partial F}{\partial t} + \frac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 F}{\partial r_t^2} - X_t r_t \right)}{\partial F / \partial r_t} dt \end{aligned}$$

As the securities were arbitrarily chosen, the equation cannot depend on them, hence leaving the right hand side equal to some function $g(r_t, t)$ depending only on time and the short rate. A standard trick is to write this as a function of the drift and the volatility for some function $\lambda(r_t, t)$ which is denoted market price of risk. Define $\lambda(r_t, t)$ such that $g(r_t, t) = -(\mu(r_t, t) - \lambda(r_t, t)\sigma(r_t.t))$, leading to

$$\begin{aligned}
g(r_t, t) &= \frac{\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} - V_t r + \frac{\partial V}{\partial A_t} f(r_t, t)\right)}{\partial V / \partial r_t} \\
&\Leftrightarrow \\
0 &= \frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} - V_t r_t + \frac{\partial V}{\partial A_t} f(r_t, t) - g(r_t, t)\frac{\partial V}{\partial r_t} \\
&\Leftrightarrow \\
r_t V_t &= \frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma(r_t, t)^2 \frac{\partial^2 V}{\partial r_t^2} + \frac{\partial V}{\partial A_t} f(r_t, t) + (\mu(r_t, t) - \lambda(r_t, t)\sigma(r_t, t)) \frac{\partial V}{\partial r_t}.
\end{aligned}$$

A.2 The Finite Difference Schemes

We will use the "delta" method in order facilitate shifts between various finite difference schemes, thus letting

$$\begin{aligned}
\frac{\partial}{\partial t} u(x_s, t_n) &= \frac{u_s^{n+1} - u_s^n}{\delta_t^n} + O(\delta_t^n), \\
\frac{\partial}{\partial x} u(x_s, t_n) &= \omega_1 \frac{u_{s+1}^n - u_{s-1}^n}{2\delta_x} + (1 - \omega_1) \frac{u_{s+1}^{n+1} - u_{s-1}^{n+1}}{2\delta_x} + O\left((\delta_x)^2\right), \\
\frac{\partial^2}{\partial^2 x} u(x_s, t_n) &= \omega_1 \frac{u_{s+1}^n - 2u_s^n + u_{s-1}^n}{(\delta_x)^2} + (1 - \omega_1) \frac{u_{s+1}^{n+1} - 2u_s^{n+1} + u_{s-1}^{n+1}}{(\delta_x)^2} + O\left((\delta_x)^2\right).
\end{aligned}$$

Notice, that setting ω_1 equal to 1 corresponds to a pure implicit scheme, 0 to a pure explicit scheme and ω_1 equal to $\frac{1}{2}$ is the Crank-Nicolson scheme. To simplify the notation let

$$\begin{aligned}
\beta(x, t) &= \tfrac{1}{2}\sigma^2 \pi^2 x^4, \\
\alpha(x, t) &= \pi x^2 \left(\sigma^2 \pi x - \tilde{\mu}\right), \\
\gamma(t) &= \frac{\delta_t^n}{(\delta_x)^2}, \\
\omega_2 &= 1 - \omega_1.
\end{aligned}$$

Substituting the finite difference approximations into the PDE (5) and simplifying, we get the following equation for an inner point (n, s) of the grid

$$\begin{aligned}
& -\omega_1 l_s^n u_{s-1}^n + (1 + \omega_1 \delta_t^n r(x_s) - \omega_1 m_s^n) u_s^n - \omega_1 h_s^n u_{s+1}^n \\
= \; & \omega_2 l_s^n u_{s-1}^{n+1} + (1 - \omega_2 \delta_t^n r(x_s) + \omega_2 m_s^n) u_s^{n+1} + \omega_2 h_s^n u_{s+1}^{n+1},
\end{aligned}$$

where

$$\begin{aligned}
m_s^n &= -2\gamma\beta, \\
l_s^n &= \gamma\left(\beta - \tfrac{1}{2}\alpha\delta_x\right), \\
h_s^n &= \gamma\left(\beta + \tfrac{1}{2}\alpha\delta_x\right).
\end{aligned}$$

For the two boundary equations we will use implied boundary conditions, but we will not be able to use a central derivative to approximate u_{xx} and u_x. Furthermore, as we do not wish to spoil the second order of the Crank-Nicholson scheme, by using a simple one sided difference approximation, which is only accurate to order $O(\delta_x)$. Instead we will use the following one sided approximations when x is at the boundaries, as they are accurate of order $O\left((\delta_x)^2\right)$. On the upper boundary x_{S+1}

$$
\begin{aligned}
\frac{\partial}{\partial x}u\left(x_{S+1},t_n\right) &= \omega_1\frac{u^n_{S-1}-4u^n_S+3u^n_{S+1}}{2\delta_x}+\omega_2\frac{u^{n+1}_{S-1}-4u^{n+1}_S+3u^{n+1}_{S+1}}{2\delta_x}, \\
\frac{\partial}{\partial x\partial x}u\left(x_{S+1},t_n\right) &= \omega_1\frac{u^n_{S-1}-2u^n_S+u^n_{S+1}}{(\delta_x)^2}+\omega_2\frac{u^{n+1}_{S-1}-2u^{n+1}_S+u^{n+1}_{S+1}}{(\delta_x)^2}.
\end{aligned}
$$

which lead to

$$
\begin{aligned}
&-\omega_1\gamma\left(\beta+\tfrac{1}{2}\alpha\delta_x\right)u^n_{S-1}+\omega_1 2\gamma\left(\alpha\delta_x+\beta\right)u^n_S \\
&+\left(1+\omega_1 r(x_{S+1})\delta^n_t-\omega_1\alpha\tfrac{3}{2}\gamma\delta_x-\omega_1\beta\gamma\right)u^n_{S+1} \\
=\;&\omega_2\gamma\left(\beta+\tfrac{1}{2}\alpha\delta_x\right)u^{n+1}_{S-1}-\omega_2 2\gamma\left(\alpha\delta_x+\beta\right)u^{n+1}_S \\
&+\left(1-\omega_2 r(x_{S+1})\delta^n_t+\omega_2\alpha\tfrac{3}{2}\gamma\delta_x+\omega_2\beta\gamma\right)u^{n+1}_{S+1}.
\end{aligned}
$$

Similar approximations on the lower boundary x_0 lead to

$$
\begin{aligned}
&\left(1+\omega_1 r(x_0)\delta^n_t+\omega_1\gamma\left(\alpha\tfrac{3}{2}\delta_x-\beta\right)\right)u^n_0 \\
&+\omega_1 2\gamma\left(\beta-\alpha\delta_x\right)u^n_1-\omega_1\gamma\left(\beta-\alpha\tfrac{1}{2}\delta_x\right)u^n_2 \\
=\;&\left(1-\omega_2 r(x_0)\delta^n_t-\omega_2\gamma\left(\alpha\tfrac{3}{2}\delta_x-\beta\right)\right)u^{n+1}_0 \\
&-\omega_2 2\gamma\left(\beta-\alpha\delta_x\right)u^{n+1}_1+\omega_2\gamma\left(\beta-\alpha\tfrac{1}{2}\delta_x\right)u^{n+1}_2.
\end{aligned}
$$

These equations can be expressed as follows. Notice, that the one sided, but second order, approximations come with a (very) small price tag, namely that the system of equations that we end up with, is not a truly tri-diagonal system.

$$
\begin{bmatrix}
B_0 & C_0 & D_0 & 0 & & & \\
A_1 & B_1 & C_1 & 0 & & & \\
0 & A_2 & B_2 & C_2 & & & \\
 & & & \cdots & & & \\
 & & A_{S-1} & B_{S-1} & C_{S-1} & 0 \\
 & & 0 & A_S & B_S & C_S \\
 & & 0 & E_{S+1} & A_{S+1} & B_{S_{+1}}
\end{bmatrix}
\begin{bmatrix}
u^n_0 \\ u^n_1 \\ u^n_2 \\ \vdots \\ u^n_{S-1} \\ u^n_S \\ u^n_{S+1}
\end{bmatrix}
=\mathbf{rh}\left(\mathbf{u}^{n+1}\right),
$$

$$
\mathbf{rh}\left(\mathbf{u}^{n+1}\right)=
\begin{bmatrix}
b_0 & c_0 & d_0 & 0 & & & \\
a_1 & b_1 & c_1 & 0 & & & \\
0 & a_2 & b_2 & c_2 & & & \\
 & & & \cdots & & & \\
 & & a_{S-1} & b_{S-1} & c_{S-1} & 0 \\
 & & 0 & a_S & b_S & c_S \\
 & & 0 & e_{S+1} & a_{S+1} & b_{S_{+1}}
\end{bmatrix}
\begin{bmatrix}
u^{n+1}_0 \\ u^{n+1}_1 \\ u^{n+1}_2 \\ \vdots \\ u^{n+1}_{S-1} \\ u^{n+1}_S \\ u^{n+1}_{S+1}
\end{bmatrix}.
$$

However, we will only need two additional row operations in order to obtain a true tri-diagonal system. Then we can use standard routines to solve the system.

A.3 The Monte Carlo Setup

The simulation setup used in this paper is based on the excellent paper on efficient simulation in non-linear one-factor interest rate models by Andersen (1996). We apply the extended version of the second order Milstein discretization scheme and the antithetic variate technique for variance reduction. Admittedly, there are several techniques that could possibly reduce the variance of the Monte Carlo estimates further.