One-factor Markov-functional interest rate models 
and pricing of Bermudan swaptions

Sönke Blunck
March 1, 2012

Premia 14

1 Preliminaries and notation

Most of what is presented here is taken from [HKP]. Let \( P(t, T) \) denote the value at time \( t \) of a zero-coupon bond which matures and pays unity at time \( T \). We denote by \( \mathcal{F}_t \) the information available at time \( t \) from observing the values of these assets, i.e. \( \mathcal{F}_t := \sigma(P(t, T); t \in \mathbb{R}_+) \). Let \((N, \mathbb{N})\) be a numeraire pair, i.e. a numeraire \((N_t)\) and a measure \( \mathbb{N} \) equivalent to the original measure such that the \( \tilde{P}(t, T) := \frac{P(t, T)}{N_t} \) are \( \{\mathcal{F}_t\} \)-martingales.

Given payment dates \( S = (S_1, \ldots, S_M) \) and daycount fractions \( \tau = (\tau_1, \ldots, \tau_M) \), we define
\[
A^{S,\tau}_t := \sum_{j=1}^M \tau_j P(t, S_j) \quad \text{principal value of basis point (PVBP).}
\]

Given, in addition, a (swap starting) date \( T \), we define
\[
R^{S,\tau,T}_t := \frac{P(t, T) - P(t, S_M)}{A^{S,\tau}_t} \quad \text{swap rate}.
\]

The corresponding (payer) swaption with maturity \( T \) and strike \( K \) is defined by the following payoff (at \( T \)):
\[
A^{S,\tau}_T(R^{S,\tau,T}_T - K)_+ \quad \text{(payoff of swaption)}.
\]

The corresponding digital (payer) swaption with maturity \( T \) and strike \( K \) is defined by the following payoff (at \( T \)):
\[
A^{S,\tau}_T 1_{K^{S,\tau,T}_T > K} \quad \text{(payoff of digital swaption)}.
\]

Note that, in the particular case \( M = 1 \), the quantity \( R^{S,\tau,T}_t \) is nothing but the (simply compounded) forward rate as seen at time \( t \) for the period \([T, S]\).
2 The general model

For \( i = 0, \ldots, m - 1 \), we fix payment dates \( S^i = (S^i_1, \ldots, S^i_M) \), daycount fractions \( \tau^i = (\tau^i_1, \ldots, \tau^i_M) \) and a swap starting date \( T_i \). Now we denote

\[
A^i_t := A^i_t S^i, \tau^i \quad \text{and} \quad R^i_t := R^i_t S^i, \tau^i, T_i.
\]

We make the following hypotheses:

(i) \( (x_t) \) is a one-dimensional Markov process under \( N \) with a known law.
(ii) For all \( i = 0, \ldots, m - 2 \), we have \( R^i_{T_i} = R_i(x_{T_i}) \) for some strictly increasing (but apriori unknown !) function \( R_i \). [Here we use the fact that \( (x_t) \) is one-dimensional.]
(iii) We have \( N_{T_{m-1}} = N_{m-1}(x_{T_{m-1}}) \) for some (known) function \( N_{m-1} \).
(iv) For all \( i = 0, \ldots, m - 2 \) and \( j = 1, \ldots, M_i \), we have: if \( S^i_j \notin \{T_{i+1}, \ldots, T_{m-1}\} \), then \( S^i_j > T_{m-1} \) and \( P(T_{m-1}, S^i_j) = P_{i,j}(x_{T_{m-1}}) \) for some (known) function \( P_{i,j} \).

In order to price e.g. Bermudan swaptions with our model by using a tree for the process \( (x_t) \), it is crucial to find the functional forms \( N_{T_i} = N_i(x_{T_i}) \) for \( i = 0, \ldots, m - 2 \); see Section 6 for details. A first step towards these functional forms is the following lemma. We employ the usual evolution family of operators \( (U_{t,s})_{t\geq s\geq 0} \) associated to the process \( (x_t) \):

\[
U_{t,s}f(y) := E^N(f(x_t) \mid x_s = y).
\]

Recall that we have the following property:

\[
E^N(f(x_t) \mid F_s) = U_{t,s}f(x_s).
\]

Lemma 2.1. Let \( i \in \{0, \ldots, m - 2\} \). Suppose that, for all \( k = i + 1, \ldots, m - 1 \), we have \( N_{T_k} = N_k(x_{T_k}) \) for some (known) function \( N_k \).

(a) For all \( j = 1, \ldots, M_i \), we have

\[
\tilde{P}(T_i, S^i_j) = \tilde{P}_{i,j}(x_{T_i}), \text{ where } \tilde{P}_{i,j} := \begin{cases} U_{T_i, T_j} \frac{1}{N_k} & S^i_j = T_k \text{ with } k \in \{i + 1, \ldots, m - 1\} \\ U_{T_{m-1}, T_j} \frac{P_{i,j}}{N_{m-1}} & \text{otherwise} \end{cases}.
\]

(b) We have

\[
\tilde{A}_{T_i} = \bar{A}_{T_i}, \text{ where } \bar{A}_{T_i} := \sum_{j=1}^{M_i} \tau^i_j \tilde{P}_{i,j}.
\]

Proof. (a) In the first case, the assertion follows from our hypothesis on the \( N_{T_k} \):

\[
\tilde{P}(t, S^i_j) = E^N(\tilde{P}(T_k, T_k) \mid F_t) = E^N(\frac{1}{N_k(x_{T_k})} \mid F_t) = \left(U_{T_k-t, T_k} \frac{1}{N_k}\right)(x_t).
\]

In the second case, the assertion is seen as follows:

\[
\tilde{P}(t, S^i_j) = E^N(\tilde{P}(T_{m-1}, S^i_j) \mid F_t) = E^N(\frac{P_{i,j}(x_{T_{m-1}})}{N_{m-1}(x_{T_{m-1}})} \mid F_t) = \left(U_{T_{m-1}, t} \frac{P_{i,j}}{N_{m-1}}\right)(x_t),
\]
where we used the hypotheses (iii) and (iv) in the second step.
(b) follows directly from (a) and the definition of \( \tilde{A}_{T_i} \):
\[
\tilde{A}_{T_i} = \sum_{j=1}^{M_i} \tau_j^i \tilde{P}(T_i, S_j^i) = \sum_{j=1}^{M_i} \tau_j^i \tilde{P}_i,j(x_{T_i}) . \quad \Box
\]

By now, we know how to compute \( \tilde{A}_i \) if we have the \( N_i+1, \ldots, N_{m-1} \). But how to compute \( N_i \) in order to pass to the next iteration step? At first, we compute \( R_i \) by calibrating our model to the digital \( R_{T_i} \)-swaption. Obviously, its value at time 0 given by our model is
\[
V_{i,N}^i(K) := E^N\left( \frac{N_i}{N_{T_i}} \tilde{A}_i^i, 1_{R_{T_i}>K} \right) = N_0 E^N\left( \tilde{A}_i^i, 1_{R_{T_i}>K} \right) .
\]

In order to represent its market value at time 0, we consider strictly decreasing functions \( V_{i,mkt}^i : \mathbb{R}_+ \to \mathbb{R}_+ \).

**Proposition 2.2.** Let \( i \in \{0, \ldots, m-2\} \). Suppose that, for all \( k = i+1, \ldots, m-1 \), we have \( N_{T_k} = N_k(x_{T_k}) \) for some (known) function \( N_k \). Suppose furthermore that we calibrate our model to the digital \( R_{T_i} \)-swaption, i.e.
\[
V_{i,mkt}^i(K) = V_{i,N}^i(K) \quad \text{for all strikes } K.
\]

(a) We have
\[
R_i = \left( V_{i,mkt}^i \right)^{-1} \circ J_i , \quad \text{where } J_i(y) := N_0 U_{T_i,0}(\tilde{A}_i^i 1_{(y,\infty)})(x_0) .
\]
(b) We have \( N_{T_i} = N_i(x_{T_i}) \), where the function \( N_i \) is given by
\[
\frac{1}{N_i} = \tilde{P}_{i,M_i} + \tilde{A}_i R_i .
\]

**Proof.** (a) is obvious in view of
\[
V_{i,mkt}^i(K) = V_{i,N}^i(K) = N_0 E^N\left( \tilde{A}_i^i, 1_{R_{T_i}>K} \right)
\]
\[
= N_0 E^N\left( \tilde{A}_i(x_{T_i}) 1_{R_{T_i}>K} \right) = N_0 E^N\left( \tilde{A}_i^i(x_{T_i}) 1_{(R_{i}^{-1}(K),\infty)}(x_{T_i}) \right)
\]
\[
= N_0 U_{T_i,0}(\tilde{A}_i^i 1_{(R_{i}^{-1}(K),\infty)})(x_0) = J_i(R_{i}^{-1}(K)) ,
\]
where we used hypothesis (ii) in the (third and) fourth step. (b) follows directly from
\[
\frac{1}{N_{T_i}} = \tilde{P}(T_i, S_{M_i}^i) + \tilde{A}_{T_i} R_{T_i}^i
\]
which is just a reformulation of the definition of \( R_{T_i}^i \). \( \Box \)
Remark 2.3. Recall that if the swap rate \((R_i^t)\) is of the type
\[
dR_i^t = \tilde{\sigma}_i^t R_i^t dW_i^A_t
\]
then the value at time \(0\) of the digital \(R_{T_i}^{i,s}\)-swaption is given by Black’s formula:
\[
V_{0,i,A}^i = A_i^0 E^A_i(1_{R_{T_i}^{i,s}>K}) = A_i^0 \Phi \left( \frac{\log \left( \frac{R_0^i}{K} \right) - \tilde{\sigma}_i^t T_i^i}{\sigma_i^t \sqrt{T_i^i}} \right),
\]
where \(\Phi\) denotes the cumulative normal distribution function. If we suppose \(V_{0,i,mkt}^i\) to be of this type, then one easily checks that
\[
\left( V_{0,i,mkt}^i \right)^{-1}(x) = R_0^i \exp \left( -\tilde{\sigma}_i^t T_i^i - \tilde{\sigma}_i^t \sqrt{T_i^i} \Phi^{-1}(\frac{x}{A_i^0}) \right).
\]

3 A LIBOR model

Here we consider the particular case of our general model where \(M_i = 1\) and \(S_i^1 = T_{i+1}\) for \(i = 0, \ldots, m-1\) and \(T_m\) is some final payment date. In particular, hypothesis (iv) is empty. We denote
\[
\tilde{P}_i := \tilde{P}_{i,1} \quad \text{and} \quad \tau_i := \tau_{i,1} = \tau(T_i, T_{i+1}).
\]
We have \(A_i^t = \tau_i P(t, T_{i+1})\) and \(R_i^t = R(t, T_i, T_{i+1})\), the forward rate, hence
\[
\tilde{P}_i = U_{T_{i+1}, T_i} \frac{1}{\lambda_{i+1}} \quad \text{and} \quad \tilde{A}_i = \tau_i \tilde{P}_i
\]
in the notation of Lemma 2.1. Suppose
\[
dR_{m^{-1}}^{m^{-1}} = \sigma_t^{m^{-1}} R_t^{m^{-1}} dW_t^N, \quad \text{where} \quad \sigma_t^{m^{-1}} = \sigma e^{at}
\]
for some \(\sigma > 0\) and some mean reversion parameter \(a\). We choose
\[
N_t := P(t, T_m) \quad \text{and} \quad x_t := \int_0^t \sigma_s^{m^{-1}} dW_s^N.
\]
Then the functional form of \(R_{T_{m^{-1}}}^{m^{-1}}\) is evident:
\[
R_{T_{m^{-1}}}^{m^{-1}} = R_0^{m^{-1}} \exp \left( -\frac{1}{2} \int_0^{T_{m^{-1}}} (\sigma_s^{m^{-1}})^2 ds + x_{T_{m^{-1}}} \right) = \mathcal{R}_{m^{-1}}(x_{T_{m^{-1}}}),
\]
where the function \(\mathcal{R}_{m^{-1}}\) is obviously given by
\[
\mathcal{R}_{m^{-1}}(x) := R_0^{m^{-1}} \exp \left( -\frac{1}{2} \int_0^{T_{m^{-1}}} (\sigma_s^{m^{-1}})^2 ds + x \right)
\]
\[
= \tau_{m^{-1}1} \left( \frac{P(0, T_{m^{-1}})}{P(0, T_m)} - 1 \right) \exp \left( -\frac{1}{2} \Sigma_{T_{m^{-1}}, 0} + x \right), \quad \Sigma_{t,s} := \sigma^2 e^{2at-2as}.
\]
Hence, since \( N_{T_m-1} = P(T_{m-1}, T_m) = (1 + \tau_{m-1} R_{T_{m-1}}^m)^{-1} \), the functional form of \( N_{T_m-1} \) required in hypothesis (iii) is easy to deduce: \( N_{T_m-1}(x_{T_{m-1}}) = (1 + C_2 e^x)^{-1} \), where

\[
C_2 := \left( \frac{P(0, T_{m-1})}{P(0, T_m)} - 1 \right) \exp\left( -\frac{1}{2} \Sigma^2_{T_{m-1}, 0} \right).
\]

Obviously, \( x_t \) given \( x_s \) is normally distributed with mean \( x_s \) and variance \( \Sigma^2_{t,s} \). In other words:

\[
U_{t,s} f(y) = \frac{1}{\sqrt{2\pi\Sigma_{t,s}}} \int_{\mathbb{R}} f(x) \exp\left( -\frac{(y-x)^2}{2\Sigma^2_{t,s}} \right) \, dx.
\]

For the iteration step (to deduce \( N_i \) from \( N_{i+1} \)), it suffices to represent \( 1/N_i \) in terms of \( \tilde{P}_i \) since

\[
\tilde{P}_i = U_{T_{i+1}, T_i} (1/N_{i+1}).
\]

This representation is obtained from Proposition 2.2:

\[
\frac{1}{N_i} = \tilde{P}_i \left( 1 + \tau_i (V^{i,mkt}_0)^{-1} \circ J_i \right),
\]

where the function \( J_i \) is given by

\[
J_i(y) := P(0, T_m) \tau_i U_{T_i, 0} (\tilde{P}_i 1_{(y,\infty)})(0).
\]

We can summarize the algorithm for the computation of the functional forms \( N_{m-1}, \ldots, N_0 \) as follows:

1. Initialization (at time \( T_{m-1} \)): Choose \( N_{m-1} \) as in (2).
2. For \( i = m-2, \ldots, 0 \): Define \( \tilde{P}_i \) as in (4) and then \( J_i \) as in (6). Now obtain \( N_i \) via (5).

Observe that the calibration instruments corresponding to the \( V^{i,mkt}_0 \) are the digital \((T_i, T_{i+1})\)-caplets defined by the following payoff at \( T_i \):

\[
\tau_i P(T_i, T_{i+1}) 1_{R(T_i, T_i, T_{i+1}) > K}.
\]

For \( i = m-1 \), it can be evaluated explicitly due to the dynamics in (1). This could be used for the choice of the parameter \( \sigma \) in (1).

**Proposition 3.1.** The current value of the digital \((T_{m-1}, T_m)\)-caplet in our LIBOR model is

\[
V^{m-1,N}_0(K) := \tau_{m-1} P(0, T_m) \Phi\left( \sigma Q^{-1}\left[ \log\left( \frac{R(0, T_{m-1}, T_m)}{K} \right) - \frac{\sigma^2}{2} \right] \right),
\]
where the parameter $\sigma_Q$ is given by

$$\sigma_Q := \sigma \sqrt{\frac{e^{2\alpha T_m - 1} - 1}{2\alpha}}.$$  

Moreover, we have for all $x \in (0, \tau_m - 1 P(0, T_m))$ that $V_0^{m-1,N}(K) = x$ if and only if

$$\sigma = \sqrt{\frac{e^{2\alpha T_m - 1} - 1}{2\alpha}} \cdot \left(2\Phi^{-1}\left(\frac{x}{\tau_m - 1 P(0, T_m)}\right), \frac{\alpha \log K - \log N}{K}\right).$$

The proof is straightforward and therefore omitted.

4 A (cancellable) swap model

Here we consider briefly the particular case of our general model where $M_i = m - i$ and $S_{i}^j = T_{i+j}$ for $i = 0, \ldots, m - 1, j = 1, \ldots, M_i$ and $T_m$ is some final payment date.

Since $S^i = (T_{i+1}, \ldots, T_m)$, we only have to give the functional form of $P(T_{m-1}, T_m)$ in order to check hypothesis (iv). But if we take the numeraire $N_t = P(t, T_m)$ as in the LIBOR model in Section 3, then $P(T_{m-1}, T_m) = N_{T_{m-1}} = N_{m-1}(x_{T_{m-1}})$, hence hypothesis (iv) is implied by hypothesis (iii). Moreover, we have

$$A_t^i = \sum_{j=1}^{m-i} \tau_j^i P(t, T_{i+j}) . \quad (7)$$

As in the LIBOR model, we suppose

$$dR_t^{m-1} = \sigma_t^{m-1} R_t^{m-1} dW_t^N , \text{ where } \sigma_t^{m-1} = \sigma e^{at}$$

for some $\sigma > 0$ and some mean reversion parameter $a$ and choose as before

$$x_t := \int_0^t \sigma_s^{m-1} dW_s^N .$$

Now we can again compute the desired functional forms but, due to $(7)$, they are more complicated than in the LIBOR model in Section 3 where we had $A_t^i = \tau_i^i P(t, T_{i+1})$.

Observe that here the natural calibration instruments are the digital (European) $(T_i, \ldots, T_{m-1})$-swaptions.
5 Numerical results: Bermudan swaption pricing in the LIBOR model

In this section, we will apply the (standard) tree method from Section 6 in order to price Bermudan swaptions in the LIBOR model of Section 3. Recall that, in this case, the calibrating instruments used in Proposition 2.2 are the digital \((T_i, T_{i+1})\)-caplets with the following payoff at \(T_i\):

\[\tau_i \bar{P}(T_i, T_{i+1}) 1_{R(T_i, T_{i+1}) > K} .\]

Since we do not have real data for their market prices \(V_{i,mkt}^0(K)\), we assume them to be given by a standard Hull-White model for the short rate \((r_t)\):

\[dr_t = \left[ \bar{\theta}_t - \bar{\sigma} r_t \right] dt + \bar{\sigma} dW_t . \tag{8}\]

The proof of the following result on the current price of digital caplets in the Hull-White model is straight-forward and therefore omitted.

**Proposition 5.1.** Consider the digital \((T, S)\)-caplet defined by the payoff at \(T\) of

\[\tau P(T, S) 1_{R(T, S) > K},\]

where \(\tau\) denotes the year fraction from \(T\) to \(S\). Its current value in the Hull-White model (8) is

\[V_{\text{HW}}^0(K) := \tau P(0, S) \Phi\left(\sigma_p^{-1}\left[\log\left(\frac{R(0, T, S)}{K + \tau^{-1}}\right) - \frac{\sigma_p^2}{2}\right]\right),\]

where the parameter \(\sigma_p\) is given by

\[\sigma_p := \bar{\sigma} \frac{e^{-\bar{a}T} - e^{-\bar{a}S}}{\bar{a}} \sqrt{\frac{e^{2\bar{a}T} - 1}{2\bar{a}}} .\]

Moreover, we have for all \(x \in (0, \tau P(0, S))\):

\[(V_{\text{HW}}^0)^{-1}(x) = \tau^{-1} \frac{P(0, T)}{P(0, S)} \exp\left(-\frac{\sigma_p^2}{2} - \sigma_p \Phi^{-1}\left(\frac{x}{\tau P(0, S)}\right)\right) - \tau^{-1} .\]

In the following, we denote

\[V_{i,\text{HW}}^0(K) := V_{\text{HW}}^0(K) \quad \text{for} \quad T = T_i, S = T_{i+1}, \tau = \tau_i .\]

We proceed as follows. We fix the Hull-White parameters \(\bar{a}\) and \(\bar{\sigma}\) and assume that the market prices \(V_{i,mkt}^0(K)\) are given by the corresponding Hull-White prices:

\[V_{i,mkt}^0(K) = V_{i,\text{HW}}^0(K) \quad \text{for} \quad i = 0, \ldots, m - 2 \text{ and all } K .\]
Now we choose our LIBOR model parameters $a$ and $\sigma$ in (1). Then iterative calibration to the digital ($T_i, T_{i+1}$)-caplets for $i = m - 2, \ldots, 0$ is used as in Proposition 2.2 [see (5) and (6)] to obtain the functional forms $N_{m-2}, \ldots, N_0$. In other words, we suppose that

$$V_0^{i,N}(K) = V_0^{i,HW}(K) \quad \text{for } i = 0, \ldots, m - 2 \text{ and all } K.$$  

Note that the iterations $i = m - 2, \ldots, 0$ involve (iterated) numerical integration.

Finally, will price the Bermudan (payer) swaption explained in Section 6.3: with strike $K_0$, with $n$ exercise times $T_0, \ldots, T_{n-1}$ and $m$ swap payment dates $T_1, \ldots, T_m$. The Bermudan swaption is priced on the one hand in our LIBOR model via a tree for the process $(x_t)$ with $N_x$ time steps as explained in Section 6, on the other hand in our Hull-White model via a tree for the short rate $(r_t)$ with $N_r$ time steps. We denote by $N_{\text{disc}}$ the number of discretizations steps for the functional forms $N_{m-2}, \ldots, N_0$. Our parameter values are:

$$\bar{a} = 0.1, \quad \bar{\sigma} = 0.01$$

$$a = \bar{a}, \quad \sigma = 0.09$$

**ITM:** $K_0 = 0.0589092$, **ATM:** $K_0 = 0.0687274$, **OTM:** $K_0 = 0.0785456$

$$n = 1, 3, 5, \quad m = 5, \quad T_i = 2 + \frac{i}{2}$$

Moreover, we use the standard (non-flat) PREMIA data for the initial yield curve. One obtains the following prices (given in BP); the third column of prices can be seen as Hull-White benchmarks.

<table>
<thead>
<tr>
<th>$n$</th>
<th>Strike $K_0$</th>
<th>$N_x = 50$, $N_{\text{disc}} = 5000$</th>
<th>$N_r = 150$</th>
<th>$N_r = 1500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ITM</td>
<td>231.33</td>
<td>231.77</td>
<td>231.75</td>
</tr>
<tr>
<td>1</td>
<td>ATM</td>
<td>97.73</td>
<td>97.70</td>
<td>97.76</td>
</tr>
<tr>
<td>1</td>
<td>OTM</td>
<td>28.59</td>
<td>27.96</td>
<td>27.92</td>
</tr>
<tr>
<td>3</td>
<td>ITM</td>
<td>249.38</td>
<td>249.85</td>
<td>249.93</td>
</tr>
<tr>
<td>3</td>
<td>ATM</td>
<td>122.60</td>
<td>123.16</td>
<td>122.98</td>
</tr>
<tr>
<td>3</td>
<td>OTM</td>
<td>48.83</td>
<td>47.89</td>
<td>47.87</td>
</tr>
<tr>
<td>5</td>
<td>ITM</td>
<td>252.15</td>
<td>253.35</td>
<td>253.36</td>
</tr>
<tr>
<td>5</td>
<td>ATM</td>
<td>127.68</td>
<td>129.01</td>
<td>128.94</td>
</tr>
<tr>
<td>5</td>
<td>OTM</td>
<td>54.51</td>
<td>54.41</td>
<td>54.30</td>
</tr>
</tbody>
</table>

With only one fixed value for the LIBOR model parameters $a$ and $\sigma$ it might be hopeless to reobtain all the Hull-White prices of the rather different swaptions we consider: European ($n = 1$) and Bermudan ($n = m$) swaptions which ITM, ATM or OTM.
6 Pricing of Markov-functional Bermudan options via trees and Monte Carlo (Appendix)

Consider the Bermudan option given by the payoffs \( h_0, \ldots, h_{n-1} \) at the exercise times \( 0 < T_0 < \ldots < T_{n-1} \). Its discounted value \( \tilde{V}_{T_0} \) at time \( T_0 \) is given by

\[
\tilde{V}_{T_0} = \sup_{\tau \in \mathcal{T}_{(0,\ldots,n-1)}} \mathbb{E}(\tilde{h}_\tau | \mathcal{F}_{T_0}) ,
\]

where \( \tilde{h}_i := \frac{h_i}{N_{T_i}} \), \( N_t \) is the numeraire and \( \mathcal{T}_{(0,\ldots,n-1)} \) denotes the set of stopping times with values in \( \{0,\ldots,n-1\} \). The discounted value \( \tilde{V}_0 \) at time 0 can be computed as follows via dynamic programmation:

\[
\begin{align*}
\tilde{V}_{T_{n-1}} &= \tilde{h}_{n-1} \\
\tilde{V}_{T_i} &= \mathbb{E}(\tilde{V}_{T_{i+1}} | \mathcal{F}_{T_i}) \vee \tilde{h}_i \quad \text{for } i = n-2, \ldots, 0 \\
\tilde{V}_0 &= \mathbb{E}(\tilde{V}_{T_0})
\end{align*}
\]

Now suppose that the \( \tilde{h}_i \) have the following Markov-functional form:

\[
\tilde{h}_i = f_i(x_{T_i}) \quad \text{for } i = 0, \ldots, n-1 . \tag{9}
\]

Here \( (x_t) \) is a Markov process with values in \( \mathbb{R}^D \). Then simulating \( (x_t) \) by trinomial trees or Monte Carlo yields standard methods to approximate \( \tilde{V}_0 \).

6.1 Trinomial trees

Suppose \( (D = 1 \text{ and}) \) that, for our Markov process \( (x_t) \), we are given a trinomial tree built for the time instants

\[
0 = t_0 < t_1 < \ldots < t_N = T_{n-1} .
\]

For \( i = 0, \ldots, n-1 \), let \( t_{d(i)} = T_i \), in particular \( d(n-1) = N \). Suppose that, at time \( t_i \), the tree has \( S_i \) nodes and that, from the \( j \)-th node at time \( t_i \), one can move to the \( (k_{l,j} + 1) \)-th, the \( k_{l,j} \)-th and the \( (k_{l,j} - 1) \)-th node at time \( t_{l+1} \). In order to approximate the discounted present value \( \tilde{V}_0 \) of the Bermudan option using our given trinomial tree, we only need (apart from the payoff functions \( f_0, \ldots, f_{n-1} \)) its following quantities:

- For \( l = 0, \ldots, N-1 \) and \( j = 0, \ldots, S_l - 1 \), let \( p_{l,j}^u, p_{l,j}^m \) and \( p_{l,j}^d \) be the up-, middle- and down-probability to move from the \( j \)-th node at time \( t_l \) to the \( (k_{l,j} + 1) \)-th, the \( k_{l,j} \)-th and the \( (k_{l,j} - 1) \)-th node at time \( t_{l+1} \).
• For \( i = 0, \ldots, n - 1 \) and \( j = 0, \ldots, S_{d(i)} - 1 \), let \( x_{d(i),j} \) be the value of \( x \) at the \( j \)-th node at time \( t_{d(i)} = T_i \) (in other words, the \( x_{d(i),j} \) are the values of \( x_{T_i} \) in the tree).

Then the following tree algorithm yields the approximation \( \tilde{v}_{0,0} \) of \( \tilde{V}_0 \). The \( \tilde{v}_{l,j} \) represent the discounted value of the Bermudan option at time \( t_l \).

1. Initialization (at time \( T_{n-1} = t_{d(n-1)} = T_N \)):
   \[
   \tilde{v}_{N,j} := f_{n-1}(x_{N,j}) \quad \text{for} \quad j = 0, \ldots, S_N - 1 .
   \]

2. For \( i = n - 1, \ldots, 1 \):
   (a) For \( l = d(i) - 1, \ldots, d(i) - 1 \), we set
   \[
   \tilde{v}_{l,j} := p^{u}_{l,j} \tilde{v}_{l+1,k_{l,j}+1} + p^{m}_{l,j} \tilde{v}_{l+1,k_{l,j}} + p^{d}_{l,j} \tilde{v}_{l+1,k_{l,j}-1} \quad \text{for} \quad j = 0, \ldots, S_l - 1 .
   \]
   (b) Early exercise at \( T_{l-1} = t_{d(i-1)} \):
   \[
   \tilde{v}_{d(i-1),j} := \tilde{v}_{d(i-1),j} \lor f_{d(i-1)}(x_{d(i-1),j}) \quad \text{for} \quad j = 0, \ldots, S_{d(i-1)} - 1 .
   \]

3. For \( l = d(0) - 1, \ldots, 0 \), we set
   \[
   \tilde{v}_{l,j} := p^{u}_{l,j} \tilde{v}_{l+1,k_{l,j}+1} + p^{m}_{l,j} \tilde{v}_{l+1,k_{l,j}} + p^{d}_{l,j} \tilde{v}_{l+1,k_{l,j}-1} \quad \text{for} \quad j = 0, \ldots, S_l - 1 .
   \]

### 6.2 Monte Carlo (Longstaff-Schwartz algorithm)

Suppose that, for our Markov process \((x_t)\), we are given \( M \) Monte Carlo samples \((x^m_{T_0}, \ldots, x^m_{T_{n-1}})\), where \( m = 0, \ldots, M - 1 \). Suppose furthermore that, for \( i = 0, \ldots, n - 2 \), we have suitably chosen functions \( g^1_0, \ldots, g^1_{d(i)-1} \) representing a basis of a \( d(i) \)-dimensional subspace of \( L_2(\mathbb{R}^D, \mu_i) \), where \( \mu_i \) denotes the law of \( x_{T_i} \). For \( \alpha \in \mathbb{R}^{d(i)} \) and \( x \in \mathbb{R}^D \), we denote \((\alpha.g)(x) = \sum_{j=0}^{d(i)-1} \alpha_j g^j_1(x)\).

Then, the following Longstaff-Schwartz algorithm approximates the current discounted value \( \tilde{V}_0 \) of our Bermudan option. Here, at the \( i \)-th iteration step, \( \tilde{v} \) represents \( \tilde{V}_{T_i} \), the discounted value of the Bermudan option at \( T_i \).

1. Initialization (at time \( T_{n-1} \)):
   \[
   \tilde{v}_m := f_{n-1}(x^m_{T_{n-1}}) \quad \text{for} \quad m = 0, \ldots, M - 1 .
   \]

2. For \( i = n - 2, \ldots, 0 \):
   (a) Let \( \alpha \in \mathbb{R}^{d(i)} \) be the unique solution of the least square problem
   \[
   \min_{\alpha \in \mathbb{R}^{d(i)}} \sum_{m=0}^{M-1} \left( (\alpha.g)(x^m_{T_i}) - \tilde{v}_m \right)^2 .
   \]
   (b) For \( m = 0, \ldots, M - 1 \): if \( f_i(x^m_{T_i}) > (\alpha.g)(x^m_{T_i}) \) then \( \tilde{v}_m := f_i(x^m_{T_i}) \).
3. Return the estimate $\frac{1}{M} \sum_{m=0}^{M-1} \tilde{v}_m$ of the current discounted value $\tilde{V}_0$.

6.2.1 Modification for large dimensions (explanatory process)

If the dimension $D$ of our driving process $(x_t)$ is too large ($D > 10$), a reasonable basis $g^i$ of functions on $\mathbb{R}^D$ would need too many functions. Hence the parameter $d(i)$ would be too large for a sufficiently fast solution of the least square problem. This difficulty arises for example in LIBOR Market models where $(x_t)$ represents a vector of $D$ different LIBOR rates.

In this situation, one modifies the approach from above by considering - besides the driving process $(x_t)$ - an “explanatory process” $(y_t)$ with values in $\mathbb{R}^d$ and $d << D$. It should be chosen such that simulating $(x_t)$ in order to obtain our Monte Carlo samples $(x^{mT}_0, \ldots, x^{mT}_{n-1})$ yields also Monte Carlo samples $(y^{mT}_0, \ldots, y^{mT}_{n-1})$ without additional computational costs. Natural choices of $(y_t)$ could be $y_t = W_t$ [if $(x_t)$ is a diffusion with Brownian motion $(W_t)$] or $y_t = F(t, x_t)$. The latter choice is made e.g. in [PPR] where, in the LIBOR Market model situation we just mentioned, the authors consider the case $y =$ swap-rate.

Suppose that, for $i = 0, \ldots, n - 2$, we have suitably chosen functions $g^0, \ldots, g^{d(i)-1}$ representing a basis of a $d(i)$-dimensional subspace of $L_2(\mathbb{R}^d, \nu_i)$, where $\nu_i$ denotes the law of $y_{T_i}$.

Now, in the modified Longstaff-Schwartz algorithm, one only has to replace all occurrences of $(\alpha.g^i)(x^{mT}_{T_i})$ by $(\alpha.g^i)(y^{mT}_{T_i})$.

6.3 Example: Bermudan swaptions in the Markov-functional LIBOR model

Consider an interest rate swap first resetting in $T_0$ and paying at $T_1, \ldots, T_m$, with fixed rate $K_0$ and year fractions $\tau_0, \ldots, \tau_{m-1}$. Assume that one has the right to enter the swap at the times $T_0, \ldots, T_{n-1}$, where $n \leq m$.

Then the corresponding Bermudan (payer) swaption fits in our general setting from above as the following particular case:

$$h_i = \left( \text{value of the interest rate swap at } T_i \right)_+$$

$$= \left( 1 - P(T_i, T_m) - K_0 \sum_{k=i+1}^{m} \tau_{k-1} P(T_i, T_k) \right)_+ .$$

(10)

In the notation of our Markov-functional LIBOR model in Section 3, we can rewrite line (10) as follows:

$$\tilde{h}_i = \left( \frac{1}{\tau_{T_i}} - \tilde{P}(T_i, T_m) - K_0 \sum_{k=i+1}^{m} \tau_{k-1} \tilde{P}(T_i, T_k) \right)_+ .$$
Since $N_t = P(t, T_m)$, we have $\tilde{P}(T_i, T_m) = 1$. Moreover, for $k = i + 1, \ldots, m - 1$,
\[
\tilde{P}(T_i, T_k) = E^{\tilde{N}}(\tilde{P}(T_k, T_k) | \mathcal{F}_{T_k}) = E^{\tilde{N}}(\frac{1}{N_k(x_{T_k})} | \mathcal{F}_{T_k}) = (U_{T_k,T_i} \cdot \frac{1}{N_k})(x_{T_k}).
\]
Hence, we obtain the desired Markov-functional forms in (9) as follows:
\[
\tilde{h}_i = f_i(x_{T_i}),
\]
where the function $f_i$ is obviously given by
\[
f_i(x) := \left( \frac{1}{N_i(x)} - (1 + K \tau_{m-1}) - K_0 \sum_{k=i+1}^{m-1} \tau_{k-1} (U_{T_k,T_i} \cdot \frac{1}{N_k})(x) \right)_+.
\]

6.4 Example: (European) digital caplets in the Markov-functional LIBOR model

In order to test the calibration of our Markov-functional LIBOR model to a Hull-White model as in Section 5, one might wish to price the calibrating instruments which are the digital $(T_i, T_{i+1})$-caplets. This does not involve the functional forms $N_0, \ldots, N_{i-1}$, hence by replacing $m$ by $m - i$ if necessary, we can assume $i = 0$.

The digital $(T_0, T_1)$-caplet fits into our general setting from above as the following particular case: $n = 1$ (European) and
\[
h_0 = \tau_0 P(T_0, T_1) 1_{R(T_0, T_0, T_1) > K}.
\]
Since $\tau_0 R(T_0, T_0, T_1) = P(T_0, T_1)^{-1} - 1$, we can rewrite this as follows, denoting $K_1 := \tau_0 K + 1$:
\[
\tilde{h}_0 = \tau_0 \tilde{P}(T_0, T_1) 1_{P(T_0, T_1)^{-1} > K_1}.
\]
Notice that $\tilde{P}(T_0, T_1) = (U_{T_0,T_1} \cdot \frac{1}{N_1})(x_{T_0}) =: \mathcal{L}(x_{T_0})$ as before and
\[
P(T_0, T_1)^{-1} = P(T_0, T_m)^{-1} \tilde{P}(T_0, T_1)^{-1} = \frac{1}{N_0 \mathcal{L}(x_{T_0})} =: \mathcal{M}(x_{T_0}).
\]
Hence, we obtain the desired Markov-functional form in (9) as follows:
\[
\tilde{h}_0 = f_0(x_{T_0}),
\]
where the function $f_0$ is obviously given by
\[
f_0(x) := \tau_0 \mathcal{L}(x) 1_{\mathcal{M}(x) > K_1}.
\]
7 An explicit formula for $\mathcal{N}_{m-2}$ in the LIBOR model (Appendix)

The following lemma is helpful for a (more or less) explicit formula for the functional form $\mathcal{N}_{m-2}$ in the LIBOR model. It can be used to avoid the first numerical integration in the iterations. On the other hand, one needs an approximation of the cumulative normal distribution function $\Phi$.

**Lemma 7.1.** We have for all $x, y \in \mathbb{R}$:

$$U_{t,s}(\exp 1_{(y,\infty)})(x) = e^{\frac{1}{2} \Sigma^2_{t,s} x} \Phi\left(\frac{\sqrt{\Sigma}_{t,s} x + \Sigma_{t,s}}{\Sigma_{t,s}}\right)$$

$$U_{t,s}(1_{(y,\infty)})(x) = \Phi\left(\frac{\sqrt{\Sigma}_{t,s} x}{\Sigma_{t,s}}\right)$$

The proof of Lemma 7.1 is elementary and therefore omitted.

**Corollary 7.2.** We have for all $x, y \in \mathbb{R}$:

$$\tilde{P}_{m-2}(x) = 1 + C_0 e^x$$

$$J_{m-2}(y) = P(0, T_m) \tau_{m-2}\left(\Phi\left(-\frac{y}{\sqrt{T_{m-2,0}} + \Sigma_{T_{m-2,0}}}\right) + C_1 \Phi\left(-\frac{y}{\sqrt{T_{m-2,0}} + \Sigma_{T_{m-2,0}}}\right)\right)$$

Here we denote, using the constant $C_2$ from (3):

$$C_0 := C_2 \exp\left(\frac{1}{2} \Sigma^2_{T_{m-1, T_{m-2}}}\right) \quad \text{and} \quad C_1 := C_0 \exp\left(\frac{1}{2} \Sigma^2_{T_{m-2,0}}\right).$$

**Proof.** We have $\mathcal{N}_{m-1} = (1 + C_2 \exp)^{-1}$, hence Lemma 7.1 (for $y = -\infty$) yields the first assertion:

$$\tilde{P}_{m-2}(x) = (U_{T_{m-1, T_{m-2}}, \mathcal{N}_{m-1}})(x) = (U_{T_{m-1, T_{m-2}}, (1 + C_2 \exp)})(x)$$

$$= 1 + C_2 e^{\frac{1}{2} \Sigma^2_{T_{m-1, T_{m-2}}} x} = 1 + C_0 e^x.$$

Now the second assertion can be deduced from the first and again Lemma 7.1:

$$N_{0}^{-1} J_{m-2}(y) = U_{T_{m-2,0}, (\hat{A}_{m-1}(y,\infty))}(x_0) = \tau_{m-2} U_{T_{m-2,0}, (1 + C_0 \exp)1_{(y,\infty)}}(0)$$

$$= \tau_{m-2} \left(\Phi\left(-\frac{y}{\sqrt{T_{m-2,0}} + \Sigma_{T_{m-2,0}}}\right) + C_0 e^{\frac{1}{2} \Sigma^2_{T_{m-2,0}}} \Phi\left(-\frac{y}{\sqrt{T_{m-2,0}} + \Sigma_{T_{m-2,0}}}\right)\right).$$

References
