OPTIMAL CONTROL APPLICATIONS AND METHODS
Optim. Control Appl. Meth. 2012; **33**:352–362
Published online 27 April 2011in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/oca.1000

10991514, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

Efficient calibration of the Hull White model

Sebastian Schlenkrich*,†

d-fine GmbH, Frankfurt/Main, Germany

SUMMARY

In this paper we study the calibration of the Hull White interest rate model for the pricing of Bermudan swaptions and Bermudan bond options. A common modelling approach for the calibration of the Hull White model is to choose the model parameters such that market prices of corresponding European derivatives are replicated by the model. This requires that a multidimensional non-linear optimization problem has to be solved. The focus of this paper lies in the efficient formulation and numerical solution of this optimization problem. We investigate the numerical properties of the iterative solution of the optimization problem by means of a Gauss Newton and an Adjoint Broyden quasi-Newton method. Required derivatives of the objective function are evaluated by techniques of Automatic Differentiation. The algorithms are benchmarked to Minpack's general purpose solver HYBRD. Numerical results for the calibration of long-term Bermudan swaptions are reported. Copyright © 2011 John Wiley & Sons, Ltd.

Received 24 September 2010; Revised 16 January 2011; Accepted 23 February 2011

KEY WORDS: Automatic Differentiation; interest rate modelling; nonlinear equations; quasi-Newton methods

1. INTRODUCTION

The Hull White interest rate model is one of the classical interest rate models in finance. It was proposed in [1] as an extension of the Vasicek model. The model yields analytical formulas for bonds and European bond options. With time inhomogeneous model parameters it can be fitted to an observed term structure of interest rates and a term structure of volatilities. The resulting calibrated model can then be used to price more exotic interest rate derivatives. The Hull White model is particularly applied to price Bermudan bond options and Bermudan swaptions.

In this paper we analyse the problem of calibrating the Hull White model to a term structure of interest rates and to a term structure of short rate volatilities. We proceed specifying the notation for our financial products required. In Section 1.1 we give the market formulas to derive prices from implicit Black'76 volatilities. We incorporate the volatility smile into these prices for our calibration instruments. The required analytical formulas of the Hull White model are cited in Section 1.2.

Section 2 describes the calibration problem. This includes the formulation as a non-linear least squares problem and its iterative solution. In Section 3 we report some numerical results comparing the performance of the iterative solvers analysed. Finally, we give some concluding remarks in Section 4.

Notation

We consider the pricing of Bermudan swaptions. The underlying swap is assumed to exchange a fixed simple compounding rate R against a floating rate. A typical floating rate index is the

*Correspondence to: Sebastian Schlenkrich, d-fine GmbH, Opernplatz 2, D-60313 Frankfurt, Germany.
†E-mail: sebastian.schlenkrich@d-fine.de

Copyright © 2011 John Wiley & Sons, Ltd.

10991514, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

EURIBOR index. The fixed leg coupon payment dates are denoted by $S_1, \dots, S_M$ and S_0 is the start date of the first coupon period. The year fractions associated with the fixed leg coupon periods are $\tau_1, \dots, \tau_M$. Moreover we assume that the fixed leg payment dates are a subset of the floating rate payment dates.

The price of a risk-free zero coupon bond at observation time t with maturity T is given by $P(t,T)$. The price of the underlying swap at time $t \leqslant S_0$ is determined by discounting the fixed and floating leg cash flows. For a (fixed) receiver swap it becomes

$$\mathrm{Swap}(t) = \underbrace{R\sum_{i=1}^{M} \tau_i P(t,S_i)}_{\mathrm{FixedLeg}(t)} - \underbrace{[P(t,S_0) - P(t,S_M)]}_{\mathrm{FloatLeg}(t)}.$$

Rearranging terms yields that the swap can also be interpreted as the time t price of a risk-free (forward) bond contract with unit bond price paid at S_0, fixed coupons $R\tau_i$ paid at S_i for $i = 1, \dots, M$ and unit notional payment at S_M. That is

$$\mathrm{Bond}(t) = \underbrace{-P(t,S_0)}_{\text{bond price}} + \underbrace{\sum_{i=1}^{M} R\tau_i P(t,S_i) + P(t,S_M)}_{\text{coupons and notional}}$$

and

$$\mathrm{Swap}(t) = \mathrm{Bond}(t).$$

A swaption gives its holder the right to enter a swap at a given strike rate R. The swaption is considered to be of European style if the right may be exercised at a single predefined date. Bermudan swaptions bare the right to enter a fixed maturity swap at several predefined exercise dates. At an exercise date T with $S_{i_0} = \min_{i=1,\dots,M}\{S_i | S_i \geqslant T\}$ the payoff of the swaption is given by:

$$\mathrm{Swaption}(T) = [\omega \mathrm{Swap}(T)]^+ = \left[\omega\left\{R\sum_{i=i_0+1}^{M} \tau_i P(T,S_i) - [P(T,S_{i_0}) - P(T,S_M)]\right\}\right]^+.$$

Here $\omega \in \{-1,+1\}$ distinguishes between a fixed rate payer (-1) and a fixed rate receiver $(+1)$ swaption. The notation $[\cdot]^+$ abbreviates $\max\{\cdot, 0\}$.

1.1. Black'76 formula for European swaptions

European swaptions may be valued in the Black'76 framework. For a derivation of the formulas see, for example, [2]. The payoff of the swaption is rewritten as:

$$\mathrm{Swaption}(T) = \left(\sum_{i=i_0+1}^{M} \tau_i P(T,S_i)\right)\left[\omega\left(R - \frac{P(T,S_{i_0}) - P(T,S_M)}{\sum_{i=i_0+1}^{M} \tau_i P(T,S_i)}\right)\right]^+.$$

In this representation the annuity is given by

$$\mathrm{Annuity}(T) = \sum_{i=i_0+1}^{M} \tau_i P(T,S_i)$$

and the (forward) par swap rate is denoted by:

$$Y(T) = \frac{P(T,S_{i_0}) - P(T,S_M)}{\sum_{i=i_0+1}^{M} \tau_i P(T,S_i)}.$$

Copyright © 2011 John Wiley & Sons, Ltd.

DOI: 10.1002/oca

With this notation the swaption payoff becomes

$$\text{Swaption}(T) = \text{Annuity}(T)[\omega(R - Y(T))]^{+} = \text{Annuity}(T)[-\omega(Y(T) - R)]^{+}$$

Thus, a European receiver (payer) swaption is equivalent to a European put (call) on the forward par swap rate $Y(T)$ with strike R.

The Black'76 model assumes that the forward par swap rate $Y(T)$ is stochastic and in particular log-normally distributed. Its mean is the time-t $(t<T)$ observable forward par swap rate $Y(t)$ and its variance is $\sigma_{\mathrm{B76}}^2 \delta(t, T)$. The function $\delta(\cdot, \cdot)$ denotes the year fraction function applied to scale the Black'76 volatility σ_{B76}. The resulting price of the European swaption at observation time t becomes

$$\text{Swaption}(t) = \text{Annuity}(t) \cdot \text{Black76}(Y(t), R, \sigma_{\mathrm{B76}}, \delta(t, T), -\omega). \tag{1}$$

The Black'76 formula for European puts $(w = -1)$ and calls $(w = +1)$ with forward price F, strike K, volatility σ, and time to maturity τ is

$$\text{Black76}(F, K, \sigma, \tau, w) = w[F\Phi(w\, d_1) - K\Phi(w\, d_2)],$$

$$d_{1,2} = \frac{\log(F/K)}{\sigma\sqrt{\tau}} \pm \frac{\sigma\sqrt{\tau}}{2}.$$

In the market prices of European swaptions are quoted in terms of implicit Black'76 volatilities. These quotes are given for several times to maturity $\delta(t, T)$ and remaining swap tenors $\delta(S_{i_0}, S_M)$. Moreover, the prices and implicit volatilities depend on the moneyness of the swaption strike rate R. The moneyness for swaptions is measured by the absolute difference $R - Y(t)$. As a result a swaption volatility cube is spanned by the time to exercise, the remaining swap tenor, and the moneyness. To evaluate prices of specific swaptions for which no direct volatility quote is available interpolation schemes are used.

Modelling the smile. For our calibration problem, we require swaption volatilities for several exercise dates and swap tenors. However, the fixed rates of the swaptions in question are equal to the corresponding fixed rate of the underlying Bermudan swaption. Since the forward par swap rates $Y(T_i)$ vary for the European swaptions also the moneynesses $R - Y(T_i)$ of the European swaptions differ. Therefore, we have to interpolate the smiles of the swaption volatilities.

As reported, for example, in [2] it is market practice to interpolate volatility smiles by the approximated formulas of the SABR model [3]. We calibrate the SABR model separately to available market quotes of swaption volatilities for a given pair of exercise date and swap tenor. Using the resulting grid of SABR parameters, we build up a volatility surface for the required strike R. Then the resulting volatility surface is interpolated linearly and extrapolated constantly. A graph of the SABR swaption volatility surface used for the numerical tests is given in Figure 3.

1.2. Analytical formulas for the Hull White model

A key parameter for pricing financial derivatives is the yield curve or its equivalent counterpart the discount factor curve. At observation time t we can in principle observe the curve of discount factors $T \mapsto P(t, T)$ for all $T \geqslant t$. In practice values of discrete discount factors $P(t, T_k)$ $(k = 1, 2, \ldots)$ are derived from market quotes of, for example, deposit and swap rates. Typical observation points T_k correspond to maturities $\delta(t, T_k)$ of one month to one year for deposit rates and two years to 30 years for swap rates. The discrete grid T_k of discount factors $P(t, T_k)$ is usually interpolated to get discount factors for values of T in between the grid points.

Instead of modelling the discount factor curve directly, one may consider certain interest rates. We assume that the discount factor curve is positive and continuously differentiable. Then we may define the continuous forward rate curve $T \mapsto f(t, T)$ as:

$$f(t, T) = -\frac{\partial \log(P(t, T))}{\partial T}.$$

Copyright © 2011 John Wiley & Sons, Ltd.

DOI: 10.1002/oca

10991514, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

10991514, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

The definition of the forward rate curve implies immediately that a discount factor $P(t,T)$ is given by:

$$P(t,T) = \exp\left\{ -\int_t^T f(t,s)\,\mathrm{d}s \right\}.$$

The forward rate curve may be modelled in a general form within the HJM framework [4]. A particular modelling approach is to consider the instantaneous short rate

$$r(t) = f(t,t).$$

The Hull White model [1] specifies a stochastic process for the short rate $r(t)$. The model is given by:

$$\mathrm{d}r(t) = [\theta(t) - ar(t)]\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W(t).$$

Here $\theta(t)$ denotes the risk neutral drift, a the constant mean reversion parameter, and $\sigma(t)$ the volatility of the short rate. It is common to assume that the volatility is piecewise constant between two exercise dates of a Bermudan swaption. Let $T_0 = t$ and denote the Bermudan exercise dates with $T_1, \ldots, T_N$ then we have that

$$\sigma(t) = \sigma_j \quad \text{for } t \in (T_{j-1}, T_j],\ \ j = 1, \ldots, N.$$

Given that the Hull White model is calibrated to data at time t the price of a zero coupon bond at an exercise date T_j, with maturity S, and realized short rate r becomes

$$\mathrm{ZCB}(t; T_j, S, r) = A(t; T_j, S)\ \mathrm{e}^{-B(T_j,S)r}$$

with

$$A(t; T_j, S) = \frac{P(t,S)}{P(t,T_j)} \exp\left\{ B(T_j,S) f(t,T_j) - \frac{B(T_j,S)^2}{2} \mathrm{e}^{-2a\,\delta(t,T_j)} C(t;T_j) \right\},$$

$$B(T_j, S) = \frac{1}{a}(1 - \mathrm{e}^{-a\delta(T_j,S)}),$$

$$C(t; T_j) = \sum_{k=1}^{j} \frac{\sigma_k^2}{2a}(\mathrm{e}^{2a\delta(t,T_k)} - \mathrm{e}^{2a\delta(t,T_{k-1})}),$$

$$f(t, T_j) = -\frac{\partial \log(P(t,T_j))}{\partial T_j}.$$

A coupon bond with cash flows c_i at coupon payment dates S_i is determined by the sum of the scaled zero coupon bond prices, i.e.

$$\mathrm{CB}(t; T_j, S_1, \ldots, S_M, r) = \sum_{S_i \geqslant T_j} c_i\ \mathrm{ZCB}(t; T_j, S_i, r).$$

The time t price of an option on a zero coupon bond with exercise date T_j, bond maturity S, and strike price K paid at T_j is given by

$$\mathrm{ZCO}(t; T_j, S, K, \omega) = P(t,T_j)\ \mathrm{Black76}(P(t,S)/P(t,T_j), K, \sigma_P, 1, \omega).$$

with

$$\sigma_P = \frac{1}{a}(\mathrm{e}^{-a\delta(t,T_j)} - \mathrm{e}^{-a\delta(t,S)})\sqrt{C(t;T_j)}.$$

Note that the notation $\tau = 1$ in the Black'76 formula above is no loss of generality. The temporal scaling of the volatility σ_P is already incorporated in the terms $(\mathrm{e}^{-a\delta(t,T_j)} - \mathrm{e}^{-a\delta(t,S)})$ and $C(t;T_j)$.

Copyright © 2011 John Wiley & Sons, Ltd.

DOI: 10.1002/oca

An option on a coupon bond with cash flows c_i at coupon payment dates S_i, exercise date T_j, and strike price K may be valued using Jamshidian's decomposition [5]. This approach requires to solve the following equation:

$$\mathrm{CB}(t; T_j, S_1, \ldots, S_M, r^\star) = K$$

for the short rate $r^\star$. Using the resulting short rate $r^\star$ we can evaluate corresponding individual strikes K_i by

$$K_i = \mathrm{ZCB}(t; T_j, S_i, r^\star).$$

With these individual strikes the coupon bond option can be prices as a sum of zero coupon bond options. We have that

$$\mathrm{CBO}(t, T_j, S_1, \ldots, S_M, \omega) = \sum_{S_i \geqslant T_j} c_i \ \mathrm{ZCO}(t, T_j, S_i, K_i, \omega). \tag{2}$$

For the calibration procedure, it is important to note that the price of a coupon bond option with exercise date T_j depends only on short rate volatilities σ_1 to σ_j. The price is independent from volatilities corresponding to times larger than T_j.

2. THE NON-LINEAR CALIBRATION PROBLEM

In Section 1 we illustrate that a forward receiver swap contract is equivalent to a contract to buy a coupon bond. Consequently, a European receiver swaption with strike rate R and fixed leg year fractions $\tau_1, \ldots, \tau_M$ is equivalent to a call option on a coupon bond with coupons $R\tau_i$ and unit strike price.

Calibrating the Hull White model means choosing the model parameters such that the model prices for European coupon bond options given by Equation (2) coincide in a well-defined way with market prices of European swaptions determined from quoted Black'76 swaption volatilities and Equation (1). In our setting the model parameters are the piecewise constant Hull White volatility values σ_j and the mean reversion parameter a.

In a test case, we analyse the calibration of the mean reversion a and a constant short rate volatility $\bar{\sigma}$ to 10 EUR denominated at-the-money European swaptions with exercises ranging from 10 to 19 years and fixed maturity in 20 years. We use market data as of September 2009. The resulting objective function aimed to be minimized and the convergence history of an optimization run with a constrained Newton's method are illustrated in Figure 1.

The numerical example demonstrates that there is only very limited progress w.r.t. the objective function in the last iterations. Moreover, there is a steep descent towards a convergence valley. This observation shows that calibrating the mean reversion to European swaptions is a rather ill-posed problem. As a result we assume the mean reversion parameter a to be predefined. However, Bermudan prices do depend on the choice of the mean reversion. Hence, if market prices for

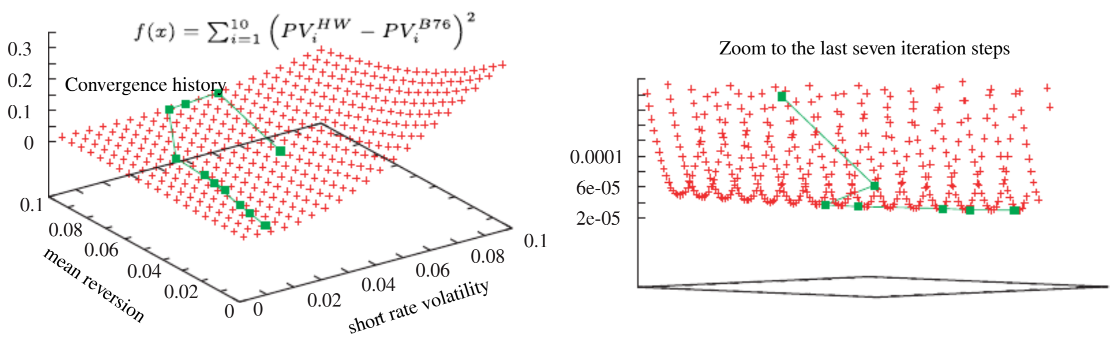

Figure 1. Convergence history for the simultaneous calibration of volatility and mean reversion.

Copyright © 2011 John Wiley & Sons, Ltd.

DOI: 10.1002/oca

10991514, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

10991514, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

Bermudan swaptions are available then the mean reversion parameter can be chosen to replicate these prices. This coincides with procedures proposed, for example, in [6].

2.1. *Formulation of the optimization problem*

For the calibration of the Hull White model, we specify a mean reversion parameter a. Then we consider the difference between the Hull White model price and the market price of a set of European swaptions. For each exercise date $j=1,\dots,N$ of the underlying Bermudan swaption, a European swaption with swap maturity equal to the Bermudan swap maturity is chosen. That is, in terms of Equations (2) and (1) we consider the differences

$$\mathrm{CBO}(t, T_j, S_1, \dots, S_M, \omega) - \mathrm{Swaption}(t) \quad \text{for } j=1,\dots,N.$$

Since we use a piecewise constant short rate volatility we may write the Hull White model price in terms of the relevant short rate volatility values. Thus, we have

$$\mathrm{CBO}(\sigma_1, \dots, \sigma_j; t, \dots) - \mathrm{Swaption}(t) \quad \text{for } j=1,\dots,N.$$

The resulting objective function is formulated by reordering independent and dependent variables. We define

$$x = (x_1, \dots, x_N)^\top = (\sigma_N, \dots, \sigma_1)^\top,$$

$$f_i(x) = \mathrm{CBO}(x; T_{N-i}, S_1, \dots, S_M, \omega) - \mathrm{Swaption}(t) \quad \text{for } i=0,\dots,N-1,$$

and

$$F : \mathbb{R}^N \to \mathbb{R}^N, \quad F(x) = (f_1(x), \dots, f_N(x))^\top.$$

The reordering of the volatilities and the reference instruments has no influence on the modelling of the calibration problem. However, it yields that each individual function $f_i(x)$ only depends non trivially on elements x_i to x_N of x. This gives the advantageous property that the Jacobian of F is an upper triangular matrix.

In principal one can aim at solving the non-linear system of equations $F(x)=0$. Since the Jacobian of F has upper triangular form this problem could be split into a sequence of one-dimensional non-linear problems. One could start to solve for f_N, then f_{N-1} up to f_1. Unfortunately, this approach fails if at least one of the market prices could not be replicated by the model exactly. However, we want our calibration be robust enough to cope with such market situations.

Since *a priori* we cannot assure that the set of non-linear equations $F(x)=0$ can be solved exactly we consider the non-linear least squares formulation

$$\min_x \frac{1}{2} F(x)^\top F(x).$$

Moreover, we want to require the volatilities to lie within some reasonable bounds. In particular, we want to ensure that the volatilities are positive. Therefore, we consider some scalar boundaries l and u and the component wise box constraint $l \leqslant x \leqslant u$. Hence, we aim at solving

$$\min_{l \leqslant x \leqslant u} \frac{1}{2} F(x)^\top F(x). \tag{3}$$

The formulation as non-linear least squares problem also allows us to choose a coarser time discretization of the short rate volatilities. In market situations with few liquid volatility quotes this yields more stable calibration results.

2.2. *Iterative solution of the non-linear problem*

The non-linear least squares problem in (3) may be solved iteratively by a Gauss–Newton method. For further reference see, for example, [7]. In each Gauss–Newton iteration step with given iterate x a step direction s is evaluated by solving the linear system

$$F'(x)^\top F'(x)s = -F'(x)^\top F(x).$$

Copyright © 2011 John Wiley & Sons, Ltd.

DOI: 10.1002/oca

In general the linear system is solved by computing a QR factorization of the Jacobian $F'(x)$. This requires $O(N^3)$ operations. Since the Jacobian $F'(x)$ of the Hull White calibration problem is already an upper triangular matrix its QR factorization is trivial. Thus, the linear system may be solved within $O(N^2)$ operations.

The Gauss–Newton iteration is stabilized by a line search in the range of the objective function F. For given function values $F(x)$ and $F(x+s)$ a step multiplier $\lambda \in \mathbb{R}$ is determined by:

$$\lambda = \frac{F(x)^\top [F(x) - F(x+s)]}{[F(x) - F(x+s)]^\top [F(x) - F(x+s)]}.$$

The state $x+s$ is accepted as the next iterate if $\lambda \geqslant \frac{1}{2} + \varepsilon$ for a suitable choice of $\varepsilon \in (0, \frac{1}{6})$. If $x+s$ is not accepted the Gauss–Newton step s is replaced by $\bar{\lambda} \cdot s$ with

$$\bar{\lambda} = \begin{cases} \lambda & \text{for } \varepsilon < |\lambda| < \frac{1}{2} + \varepsilon \\ \varepsilon & \text{for } 0 \leqslant \lambda \leqslant \varepsilon \\ -\varepsilon & \text{for } -\varepsilon \leqslant \lambda < 0 \\ -(\frac{1}{2} + \varepsilon) & \text{for } \lambda < -(\frac{1}{2} + \varepsilon) \end{cases}$$

The line search procedure is repeated with the function values $F(x)$ and $F(x + \bar{\lambda} \cdot s)$ until acceptance. Further references and convergence results concerning this line search approach can be found in [8].

If we find that a Gauss–Newton step $\bar{x} = x + s$ would result in an iterate violating the box constraints $l \leqslant \bar{x} \leqslant u$ then the components of $\bar{x}$ are projected to the bounds. That is, for $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_N)^\top$ we set $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_N)^\top$ with

$$\hat{x}_i = \begin{cases} l & \text{for } \bar{x}_i < l \\ \bar{x}_i & \text{for } l \leqslant \bar{x}_i \leqslant u \\ u & \text{for } \bar{x}_i > u \end{cases}, \quad i = 1, \ldots, N.$$

In such a case we get a new step direction $\hat{s} = \hat{x} - x$. The line search is then performed along the altered step direction $\hat{s}$.

The Gauss–Newton iteration may be summarized as follows: Starting with an initial guess $x^{(0)}$ we evaluate for $i = 0, 1, 2, \ldots$

1. solve $F'(x^{(i)})^\top F'(x^{(i)}) s^{(i)} = -F'(x^{(i)})^\top F(x^{(i)})$,
2. determine a suitable step multiplier $\lambda^{(i)} \in \mathbb{R}$ (by considering the box constraints),
3. evaluate the next iterate $x^{(i+1)} = x^{(i)} + \lambda^{(i)} \cdot s^{(i)}$.

Derivative Evaluation by Automatic Differentiation. Each iteration step of the Gauss–Newton method requires the evaluation of the Jacobian $F'(x^{(i)})$. For the Hull White calibration problem, this means differentiating the coupon bond option formula with respect to the individual volatilities. Since the coupon bond option formula is given only implicitly there is no closed-form formula for the Jacobian of F. A common approach is to approximate the Jacobian by finite differences. However, this approach suffers from approximation and round off errors depending on the choice of the finite difference increment. Furthermore, finite difference methods used within general purpose solvers usually do not allow one to exploit the triangular structure of the Jacobian of F. That is for two-sided finite differences applied to the function F each of the $N^2 - N$ zeros in the lower left triangle of the Jacobian $F'(x)$ are essentially evaluated by two evaluations of the CBO$(\cdots)$ formula with each identical arguments.

An alternative approach for the evaluation of the Jacobian is given by the methods of Automatic Differentiation. For further details see [9]. The key idea is applying the chain rule of differentiation to the computer program for the objective function F. This allows the evaluation of tangents $F'(x)s$ and gradients $F'(x)^\top v$ within machine precision. Moreover, the computational effort for the evaluation of these objects is bounded by a small multiple of the evaluation of the objective

Copyright © 2011 John Wiley & Sons, Ltd.

DOI: 10.1002/oca

10991514, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

function itself. Methods of Automatic Differentiation are applied to financial models, for example, in the studies [10, 11].

We apply Automatic Differentiation tool ADOL-C [12] to differentiate the function for the evaluation of the Hull White coupon bond option. This allows us to compute the Jacobian $F'(x)$. Our numerical results indicate that the application of Automatic Differentiation to the Hull White calibration problem is much more efficient than finite difference approximations.

Adjoint Broyden quasi-Newton updates. Automatic Differentiation allows to compute tangents and gradients in a computationally efficient way. However, in general a Jacobian ($N \times N$ matrix) evaluation still requires the effort of about N function evaluations. For large numbers of N and complex function evaluations, this effort can be crucial for the optimization procedure. In particular, for the Hull White calibration problem, the function evaluation of F requires the evaluation of N coupon bond options. Each coupon bond option itself contains the iterative procedure to solve for the short rate $r^\star$ (Jamshidian's trick). If we have many exercise dates the Hull White calibration can be computationally rather expensive.

The computational effort for the Jacobian evaluation in each Gauss–Newton iteration step can be reduced by approximating the Jacobians $F'(x^{(i)})$ successively by low rank updates. Such methods are known as quasi-Newton methods. Well-known representatives of quasi-Newton methods are Broyden's update for the approximation of Jacobians and the BFGS update for the approximation of Hessians. These updates only use function values of F. For further details see, for example, [7] and references therein.

For this study we apply Adjoint Broyden updates. The Adjoint Broyden updates make explicit use of gradient and tangent information to improve the approximation to the Jacobian [13]. Instead of evaluating the Jacobian $F'(x^{(i)})$ in each Gauss–Newton step, we choose to approximate it by the following rank-1 update:

$$A_i = A_{i-1} = \frac{vv^\top (F'(x^{(i)}) - A_{i-1})}{v^\top v} \quad \text{with } v = (F'(x^{(i)}) - A_{i-1}) s_{i-1}.$$

In the update formula the matrix A_{i-1} is the Jacobian approximation (or the Jacobian) of the preceding iteration and s_{i-1} is the current iteration step direction. The tangent $F'(x^{(i)})s_i$ and gradient $v^\top F'(x^{(i)})$ are evaluated by the forward and reverse mode of Automatic Differentiation, respectively.

In our implementation we realize the Adjoint Broyden updates by updating a QR factorization. The QR update can be evaluated by $O(N^2)$ operations. Moreover, our numerical tests indicate that some (usually three to five) initial Gauss–Newton steps with Jacobian evaluation speed up the global convergence properties of the iteration substantially.

Benchmark algorithm HYBRD. We compare the Gauss–Newton least squares algorithm (with and without Adjoint Broyden updates) with MINPACK's [14] general purpose solver HYBRD [15]. The HYBRD algorithm approximates required Jacobians by finite differences. Iterations are speeded up by Broyden updates. To achieve global convergence the iteration is stabilized by a (dog leg) trust region method. With these features the HYBRD algorithm differs significantly from our Gauss–Newton method with line search in the range of F.

3. NUMERICAL RESULTS

We implement the Hull White calibration problem including the pricing formulas in C (and some C++). The yield curve is interpolated log-linearly between given discount factors. The swaption volatility surface is interpolated bilinear with constant extrapolation. For our test cases we applied market data as of February 2010. The applied yield curve in terms of continuous compounded zero rates is given in Figure 2.

The implicit swaption volatilities for a fixed strike are determined by interpolating given quotes using the SABR model [3]. The resulting volatility surface for a strike of 4% is illustrated in Figure 3.

Copyright © 2011 John Wiley & Sons, Ltd.

10991514, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

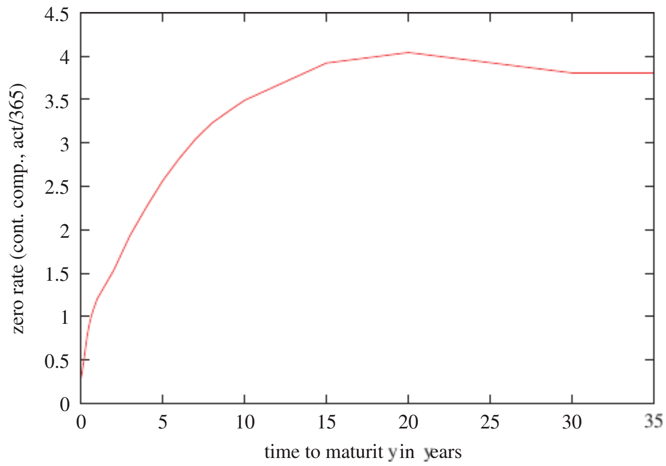

Figure 2. EUR yield curve.

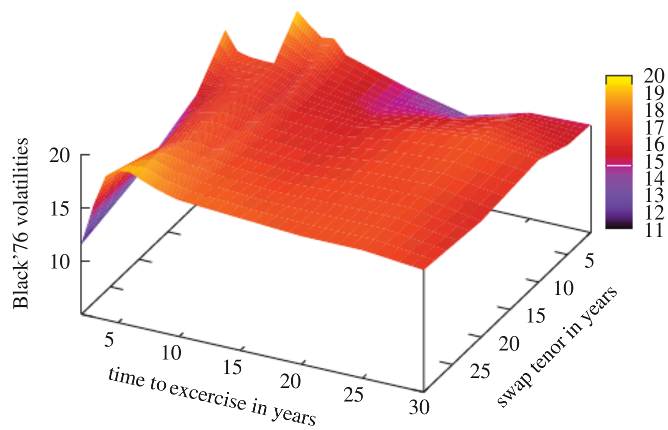

Figure 3. Implicit EUR swaption volatility surface for a strike of 4%.

As a first test case we choose to calibrate the Hull White short rate volatility for an EUR annually callable 30Y Bermudan swaption. The underlying Swap pays 4% (annual, 30/360) against 6M EURIBOR (semi annual, act/360). The first exercise is 1Y ahead of the valuation date. With this specification the dimension of the optimization problem is $N = 30$.

For a second test case, we enlarged the swap maturity to 100Y. Such a swaption is of course not very likely to be traded in the market. Nevertheless it could be considered as an (simple) approximation for the optional part of a perpetual Bermudan callable bond. This test case is of interest for us since the problem dimension is increased substantially, i.e. $N = 100$.

The Hull White model is parameterized with a mean reversion of 5%. The upper bound of the short rate volatility is given by 50%. For the lower bound of the short rate volatility we choose 0.1%. The iterations are initialized with a constant volatility of 10% for the first test case and 5% for the second test case.

We compare the required runtimes for the Gauss Newton, Adjoint Broyden, and HYBRD method to reach a given accuracy. As a common measure for the calibration accuracy we choose the residual $\|F(x^{(i)})\|_2$ in the Euclidean norm in each iteration i. Figure 4 illustrates the convergence of the iterations for the 30Y swaption calibration.

The results for the calibration of the 30Y swaption in Figure 4 show a rapid and continuous decay of the residual for the Gauss–Newton method. One iteration of the Gauss–Newton method without additional line search trials requires approximately 0.11 s. For the Adjoint Broyden quasi-Newton method, we used five initial steps with Jacobian evaluation. It turns out that fewer initial

Copyright © 2011 John Wiley & Sons, Ltd.

DOI: 10.1002/oca

10991514, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

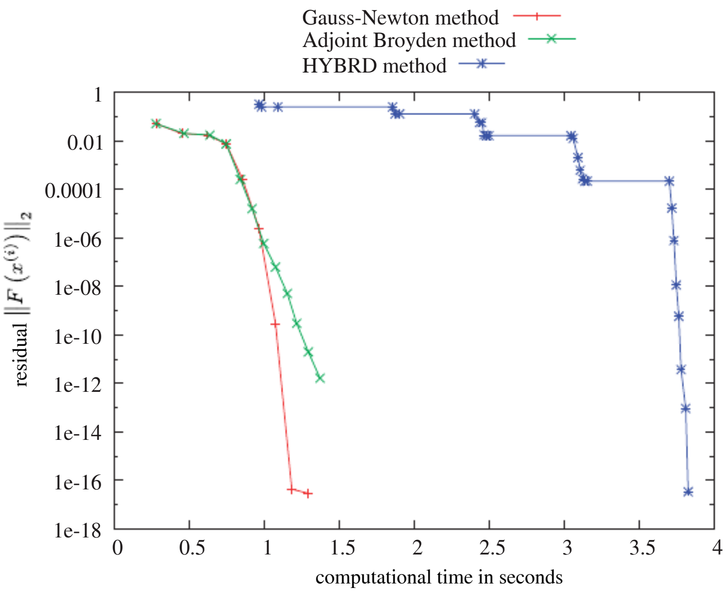

Figure 4. Convergence history for the 30Y swaption calibration.

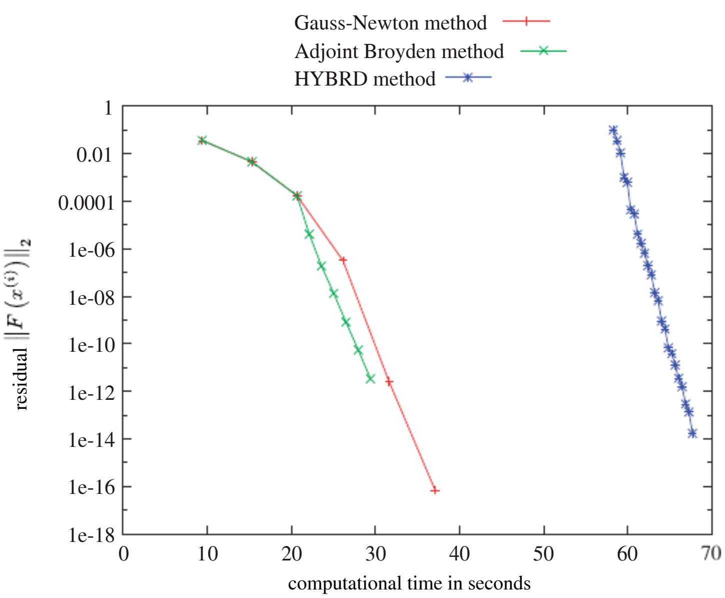

Figure 5. Convergence history for the 100Y swaption calibration.

Gauss–Newton steps result in substantially more Adjoint Broyden steps for this test case. This would reduce the overall performance of the method. Iterates of the Adjoint Broyden method are evaluated much faster than iterates of the Gauss–Newton method, as soon as Jacobian evaluations are replaced by rank-1 updates. However, the local quadratic convergence order of the Gauss–Newton method outperforms the Adjoint Broyden method in this test case.

For the benchmark algorithm HYBRD we see some kind of step wise convergence. The steps with no decay correspond to iterates where a new Jacobian approximation is evaluated by finite differences. Such steps require approximately 0.55 s. In this test case, Automatic Differentiation allows an evaluation of the Jacobian in only 20% of the runtime required for finite differences.

The steps with Broyden updates are much faster and show a rapid convergence. However, due to the time-consuming Jacobian evaluations by finite differences, the HYBRD algorithm requires substantially more runtime than the Gauss–Newton and Adjoint Broyden method.

Copyright © 2011 John Wiley & Sons, Ltd.
DOI: 10.1002/oca

1099151 4, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles

Figure 5 illustrates the results for the calibration of the 100Y swaption. The problem dimension is substantially larger than for the first test case. Hence, more computational time is required by the iterative solvers. Again the Gauss–Newton method shows a rapid and continuous decay. One iteration requires approximately 5.5 s. The Adjoint Broyden update method is slightly faster than the Gauss Newton method since iterates are evaluated much more quickly.

For the HYBRD algorithm, we also see a fast convergence. However, substantial initial computational effort is required to evaluate the initial Jacobian approximation by finite differences. This initial step requires about 58 s. Again we see that Automatic Differentiation speeds up the Jacobian evaluation significantly. In this test case finite differences is about 10 times slower than Automatic Differentiation.

4. CONCLUSIONS

In this paper we studied the problem of calibrating a piecewise constant short rate volatility in the Hull White model. This calibration requires the iterative solution of a non-linear least squares problem. We describe a Gauss Newton and an Adjoint Broyden quasi-Newton method. These methods make explicit use of derivative evaluations by methods of Automatic Differentiation. As a benchmark method we use the general purpose solver HYBRD which approximates derivatives by finite differences.

For two calibration problems, we compare the numerical convergence properties of the algorithms. We find that the Gauss–Newton method and the Adjoint Broyden method converge significantly faster than the HYBRD method for the test cases considered. The improvement in terms of computational costs by the Gauss–Newton and Adjoint Broyden method is due to the derivative evaluation by Automatic Differentiation. In particular, we find that for our Hull White calibration test cases a Jacobian evaluation by Automatic Differentiation requires only 10 or 20% of the runtime compared with a finite difference approximation.

Our results indicate that for financial calibration problems, the additional implementation effort for the incorporation of exact derivatives can gain a substantial improvement of the numerical methods. Further investigation could focus on the combination derivative evaluation and existing general purpose solvers. For example, Minpack's solver HYBRJ could be combined with Jacobian evaluations by Automatic Differentiation.

REFERENCES

1. Hull J, White A. Pricing interest-rate-derivative securities. *The Review of Financial Studies* 1990; **3**:573–592.
2. Brigo D, Mercurio F. *Interest Rate Models—Theory and Practice*. Springer: Berlin, 2007.
3. Hagan P, Kumar D, Lesniewski A, Woodward D. Managing smile risk. *Wilmott magazine*, September 2002.
4. Heath D, Jarrow R, Morton A. Bond pricing and the term structure of interest rates: a new methodology. *Econometrica* 1992; **60**:77–105.
5. Jamshidian F. An exact bond option pricing formula. *The Journal of Finance* 1989; **44**:205–209.
6. Hagan PS. Evaluating and hedging exotic swap instruments via LGM. *Bloomberg Technical Report*.
7. Nocedal J, Wright SJ. *Numerical Optimization*. Springer: Berlin, 2006.
8. Griewank A. The 'global' convergence of Broyden-like methods with a suitable line search. *Journal of the Australian Mathematical Society*, *Series B* 1986; **28**:75–92.
9. Griewank A, Walther A. *Evaluating Derivatives*: *Principles and Techniques of Algorithmic Differentiation* (2nd edn). SIAM: Philadelphia, PA, 2008.
10. Giles M, Glasserman P. Smoking adjoints: fast monte carlo greeks. *Risk*, January 2006.
11. Kaebe C, Maruhn J, Sachs E. Adjoint based monte carlo calibration of financial market models. *Journal of Finance and Stochastics* 2009; **13**:351–379.
12. Griewank A, Juedes D, Utke J. ADOL-C: A package for automatic differentiation of algorithms written in C/C++. *TOMS* 1996; **22**:131–167.
13. Griewank A, Schlenkrich S, Walther A. A quasi-Newton method with optimal R-order without independence assumption. *Journal of Optimization Methods & Software* 2008; **23**(2):215–225.
14. MINPACK, 2010. Available from: http://www.netlib.org/minpack/.
15. Powell MJD. A hybrid method for nonlinear equations. In *Numerical Methods for Nonlinear Algebraic Equations*, Rabinowitz P (ed.). Gordon and Breach: London, 1970; 87–114.

Copyright © 2011 John Wiley & Sons, Ltd.

DOI: 10.1002/oca

1099151x, 2011, 33, Downloaded from https://onlinelibrary-wiley-com.emedien.ub.uni-muenchen.de By Ludwig-Maximilians-Universität- on [26/09/2021]. Re-use and distribution is strictly not permitted, except for Open Access articles