# Weighted Monte Carlo: A New Technique for Calibrating Asset-Pricing Models

Marco Avellaneda, Robert Buff, Craig Friedman,
Nicolas Grandchamp, Lukasz Kruk, and Joshua Newman *

**Abstract**

A general approach for calibrating Monte Carlo models to the market prices of benchmark securities is presented. Starting from a given model for market dynamics (price diffusion, rate diffusion, etc.), the algorithm corrects for price-misspecifications and finite-sample effects in the simulation by assigning "probability weights" to the simulated paths. The choice of weights is done by minimizing the Kullback-Leibler relative entropy of the posterior measure to the empirical measure. The resulting ensemble prices the given set of benchmark instruments exactly or in the sense of penalized least squares. We discuss pricing and hedging in the context of these weighted Monte Carlo models. Significant reduction of variance due to the model calibration is demonstrated theoretically as well as numerically. Concrete applications to the calibration of stochastic volatility models and term-structure models with up to forty benchmark instruments are presented. Implied volatilities, forward-rate curves and exotic option pricing are investigated with several examples.

## 1 Introduction

According to Asset-Pricing Theory, security prices should be equal to the expectations of their discounted cash-flows under a suitable probability measure. This "risk-neutral" measure represents the economic value of consuming one unit of account on a given future date and state of the economy. A risk-neutral probability implemented in the context of a specific market is often called a *pricing model*. It is self-evident that a pricing model should reproduce correctly the prices of liquid instruments which are

*Address: Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY, 10012. We are grateful to Graciela Chichilnisky, Freddy Delbaen, Raphael Douady, Darrell Duffie, Nicole El Karoui, David Faucon, Olivier Floris, Helyette Geman, Jonathan Goodman, Robert Kohn, Jean-Pierre Laurent, Jean-Michel Lasry, Jerome Lebuchoux, Marek Musiela, Jens Nonnemacher, and Frank Zhang for their enlightening comments and suggestions. We acknowledge the hospitality and generous support of Banque Paribas, Ecole Polytechnique, Ecole Normale Superieure de Cachan, ETH-Zurich and the Center for Financial Studies (Frankfurt). This work was partially supported by the U.S. National Science Foundation (DMS-9973226).

actively traded in the market. This ensures that "off-market", less liquid, instruments are realistically priced by the model.[1]

Here, we consider pricing models based on Monte Carlo (MC) simulations of future market scenarios ("paths"). [2] Prices are computed by averaging discounted cashflows over the different paths. We shall be concerned with the *calibration* of such models, *i.e.* with specifying the statistics of the sample paths in such a way that the model matches the prices of benchmark instruments traded in the market.

Most calibration procedures rely on the existence of explicit formulas for the prices of the benchmark instruments. The unknown parameters of the underlying stochastic process are found by inverting such pricing formulas, either exactly or in the sense of least-squares. Unfortunately, in Monte Carlo simulations, this method may not be sufficiently accurate enough due to sampling errors (the finite sample effect). Furthermore, closed-form solutions for prices may not always be available or easy to code. In the latter case, fitting the model to market prices implies searching the parameter space through direct simulation, a computationally expensive proposition.

This paper consisders a general, non-parametric, approach for calibrating Monte Carlo models and applies it to several practical situations. The main idea behind our method is to put the emphasis on determining directly the risk-neutral *probabilities* of the future states of the market, as opposed to finding the *parameters* of the differential equations used to generate the paths for the state-variables.

One way to motivate our algorithm is to observe that Monte Carlo simulations can be divided (somewhat arbitrarily) into two categories: those that are *uniformly weighted* and those that are *non-uniformly weighted*. To wit, consider a set of sample paths, denoted by $\omega_1, ...\omega_\nu$, generated according to some simulation procedure. By definition, a uniformy weighted simulation is such that all sample paths are assigned the same probability. Thus, a contingent claim that pays the holder $h_i$ dollars if the path $\omega_i$ ocurrs, has model value

$$\Pi_h \;=\; \frac{1}{\nu}\sum_{i=1}^{\nu} h_i \;. \tag{1}$$

A non-uniformly weighted simulation is one in which the probabilities are not necessarily equal. Suppose that we assign, respectively, probabilities $p_1, \ldots p_\nu$ to the different paths. The value of the contingent claim according to the corresponding "non-uniformly weighted" simulation is

$$\Pi_h \;=\; \sum_{i=1}^{\nu} h_i\, p_i \;. \tag{2}$$

Our approach is based on non-uniformly weighted simulations. First, we simulate a large number of paths of a stochastic process followed by the state-variables (prices, rates, etc. ) under a *prior* distribution. Second – and here we depart from the

[1] Throughout this paper, a pricing model refers to a model for pricing less liquid instruments relatively to more liquid ones (the benchmarks) in the context of a particular market. This type of financial model is used by most large investment banks to manage their market positions.

[2] See Dupire (1998) for an up-to-date collection of papers by academics and practitioners on Monte Carlo methods in quantitative finance.

conventional Monte Carlo method – we assign a different probability to each path. Probabilities are determined in such a way that (i) the expected values of the discounted cash-flows of benchmark instruments coincide exactly or within tolerance with the market prices of these securities and (ii) they are as close as possible to uniform probabilities ($p_i = 1/\nu$) coresponding to the simulated prior.

This method allows us to incorporate market information in two stages. The first step gives a prior probability measure that corresponds to our best guess for the risk-neutral measure given the information available. This guess may involve real statistics, such as estimates of rates of return, historical volatilities, correlations. It may also use parameters which are implied from market prices (implied volatilities, cost-of-carry, etc). In other words, the path simulation is used to construct a "backbone" or "prior" for the model which incorporates econometric or market-implied data. The second step has two purposes: it reconciles the econometric/prior information with the prices observed at any given time and also corrects finite-sample errors on the prices of the benchmark instruments which arise from Monte Carlo simulation.

We shall denote the mid-market prices of $N$ benchmark instruments by $C_1, \ldots, C_N$ and represent the present value of the cashflows of the $j^{th}$ benchmark instrument by

$$g_{1j},\ g_{2j}\ ,\ \ldots\ g_{\nu j}\quad j = 1, \ldots N\ . \tag{3}$$

The price relations for the benchmark instruments can then be written in the form

$$\sum_{i=1}^{\nu} p_i\, g_{ij}\ =\ C_j\ ,\qquad j = 1, \ldots N\ , \tag{4}$$

where $(p_1, \ldots p_\nu)$ are the probabilitites that we need to determine. Generically, this (linear) system of equations admits infinitely many solutions because the number of paths $\nu$ is greater than the number of constraints. [3] The criterion that we propose for finding the calibrated probability measure is to minimize the *Kullback-Leibler relative entropy* of the non-uniformly sampled simulation with respect to the prior. Recall that if $p_1, \ldots p_\nu$ and $q_1\ \ldots\ q_\nu$ are probability vectors on a probability space with $\nu$ states, the relative entropy of $p$ with respect to $q$ is defined as

$$D(p|q)\ =\ \sum_{i=1}^{\nu}\ p_i\ \log\left(\frac{p_i}{q_i}\right)\ . \tag{5}$$

In the case of Monte Carlo simulation with $q_i = 1/\nu \equiv u_i$ we have[4]

$$D(p|u)\ =\ \log\nu + \sum_{i=1}^{\nu}\ p_i\ \log p_i\ . \tag{6}$$

We minimize this function under the linear constraints implied by (4). To this effect, we implement a dual, or Lagrangian, formulation which transforms the problem into

[3] It is also possible that the system of equations admits no solutions if the prior is inadequate of if the prices give rise to an arbitrage opportunity. We shall not dwell on this here.

[4] We shall denote the uniform probability vector by $u$, i.e. $u = (1/\nu, \ldots 1/\nu)$.

an unconsntrained minimization over $N$ variables. Optimization of the dual objective function is made with L-BFGS (Byrd *et al* (1994)), a gradient-based quasi-Newton optimization routine.

The use of minimization of relative entropy as a tool for computing Arrow-Debreu probabilities was introduced by Buchen and Kelly (1996) and Gulko (1995, 1996) for single-period models. Other calibration methods based on minimizing a least-squares penalization function of the Arrow-Debreu probabilities were proposed earlier by Rubinstein (1994) and Jackwerth and Rubinstein (1995). Samperi (1997), Avellaneda *et al* (1997) and Avellaneda (1998) generalized the minimum-entropy method to intertemporal lattice models and diffusions. More recently Laurent and Leisen (1999) considered the case of Markov chains. These studies suggest this is a computationally feasible approach that works in several classical context, such as generalizations of the Black-Scholes model with volatility skew or for one-factor interest rate models.

The use of minimum relative entropy for selecting Arrow-Debreu probabilitites can be justified on economic grounds. Samperi (1997) shows that there exists a one-to-one correspondence between the calibration of a model starting with a prior probability measure and using a "penalization function" on the space of probabilities and the calculation of state-prices via utility maximization. More precisely, the Arrow-Debreu prices generated by minimizing relative entropy coincide with the marginal utilities for consumption obtained by maximizing the expectation of the utility function $U(x) = -exp(-\alpha\, x)$ by investing in a portfolio of benchmark instruments. This correspondence is quite general. It implies, most notably, that other "distances" or "penalization functions" for Arrow-Debreu probabilities of the form

$$\tilde{D}(p|q) \;=\; \sum_{i=1}^{\nu} \psi\left(\frac{p_i}{q_i}\right) q_i \quad , \qquad \psi(x) \;\; \text{convex} \tag{7}$$

can be used instead of relative entropy (which corresponds to the special case $\psi(x) = x \log x$). For each such penalization function, there exists a corresponding concave utility $U(x)$, obtained via a Legendre transformation, such that the Arrow-Debreu probabilities are consistent with an agent maximizing his/her expected utility for terminal wealth by investing in a portfolio of benchmarks. [5]

The use of relative entropy also has consequences in terms of price-sensitivity analysis and hedging. Avellaneda (1998) shows that the sensitivities of model values with respect to changes in the benchmark prices are equal to the linear regression coefficients of the contingent claim on the linear span of the cashflows of the benchmark instruments. In particular, the price-sensitivities can be computed directly using a single Monte Carlo simulation, *i.e.* without having to perturb the $N$ input prices

[5] While the particular choice of the mathematical distances $\tilde{D}(p|q)$ remains to be justified, the different distances between probabilites which result are, roughly speaking, economically equivalent – except possibly for the particular choice of smooth, increasing, convex utility that might represent the agent's preferences. The Kullback-Leibler distance is convenient because it leads to particularly simple mathematical computations, as we shall see hereafter. Another important feature of relative entropy is that it is invariant under changes of variables and therefore independent of the choice parameterization used to describe the system (Cover and Thomas ( ). We refer the reader to Samperi (1999) for an in-depth discussion of this correspondence principle.

prices and to repeat the calibration procedure each time. Thus, we hope that this method may provide an alternative approach to computing hedge-ratios as well.

Practical considerations in terms of model implementation are studied in the last 4 sections of the paper. First, we observe that calibration of Monte Carlo models to the prices of benchmark instruments results in a strong reduction of variance, or simulation noise. This is due to the fact that the model effectively averages the residual cash-flows (modulo the linear space spanned by the benchmarks). Therefore, instruments which are well-approximated by benchmarks have very small Monte Carlo variance. In particular, the interpolation of implied volatilities and prices of options between strikes and maturities is numerically efficient.

In practice, the success of any model-calibration method depends on the characteristics of the market where it is applied. To evaluate the perfomance and the output of the algorithm, we provide several concrete examples in which the algorithm is put to work. In particular, we study option-pricing models in the foreign-exchange and equity markets, using forwards and liquidly traded options as benchmarks. The models that we use incorporate stochastic volatility and are calibrated to the observed volatility skew. We also discuss the calibration of fixed-income models, and apply the algorithm to the construction of forward-rate curves based on the prices of on-the-run US Treasury securities.

# 2 Relative entropy distance and the support of the risk-neutral measure

Relative entropy measures the deviation of the calibrated model from the prior. Intiutively, if the relative entropy is small the model is "close" to the prior and thus is "more desirable" than a model that has large distance from the prior. Let us make this statement more precise in the context of Monte Carlo simulations. The relative entropy distance,

$$D(p|u) \;=\; \log\nu \;+\; \sum_{i=1}^{\nu} p_i \log p_i \;, \tag{8}$$

takes values in the interval $[0, \log\nu]$. The value zero corresponds to $p_i \;=\; 1/\nu$ (the prior) whereas a value of $\log\nu$ is obtained when all the probability is concentrated on a single path. More generally, consider a probability distribution which is supported on a subset of paths of size $\mu$ and is uniformly distributed on these paths. If we take $\mu = \;\nu^\alpha$, with $0 < \alpha < 1$, and substitute the corresponding probabilities in (8), we find that

$$D(p|u) \;=\; \log\nu + \;\log\left(\frac{1}{\nu^\alpha}\right) \;= (1-\alpha)\log\nu \;. \tag{9}$$

Within this class of measures, the relative entropy distance counts the number of paths in the support on a logarithmic scale. If $\frac{D(p|u)}{\log\nu} \ll 1$ the support of the calibrated measure is of size $\nu$, whereas $\frac{D(p|u)}{\log\nu} \approx 1$ corresponds to a measure with a "thin

support". Thin supports are inefficient from a computational viewpoint. They imply that the calibration algorithm "discards" a large number of simulated paths. In this case, the *a priori* support of the distribution constructed by simulation will be very different from the the *a posteriori* support. This confirms the intuition whereby calibrations with small relative entropy are desirable.

Figure 1: Schematic graph of the relative entropy function. A probability with $D(p|u) = (1-\alpha)\log\nu$ is supported essentially on a subset of paths of cardinality $\nu^\alpha$. Probabilities with small $D(p|u)$ have large support whereas probabilities supported on a single path have the highest Kullback-Leibler distance, $\log\nu$.

This analysis can be applied to more general probability distributions. Let us write

$$p_i \;=\; \frac{1}{\nu^{\alpha_i}}\;,\quad i = 1,\,2,\,\ldots\nu\;. \tag{10}$$

Let $N_\alpha$ represent the number of paths with $\alpha_i = \alpha$, so that we have

$$\sum_\alpha N_\alpha \;=\; \nu\;,\qquad \sum_\alpha \frac{N_\alpha}{\nu^\alpha} \;=\; 1\;. \tag{11}$$

Substituting (10) into (8), we find that

$$
\begin{aligned}
D(p|u) &= \log \nu \left( 1 + \sum_{\alpha} \frac{N_\alpha}{\nu^\alpha} \log \left( \frac{1}{\nu^\alpha} \right) \right) \\
&= \log \nu \left( 1 - \sum_{\alpha} \frac{N_\alpha}{\nu^\alpha} \alpha \right) \\
&= \log \nu \ ( 1 - \mathbf{E}^p (\alpha)) , \qquad (12)
\end{aligned}
$$

which shows that the relative entropy increases if the expected value of $\alpha$ is large. Due to the constraints implied by (11), this is possible only if there is a wide range of exponents $\alpha_i$. Since probabilities are measured on a logarithmic scale, the measure will be concentrated on those paths which correspond to small values of $\alpha$. A wide mismatching of probabilities between the calibrated measure and the prior is undesirable because this means that certain state-contingent claims will have very different values under the prior and the posterior measures.

# 3 Calibration algorithm

We describe the algorithm for calibrating Monte Carlo simulations under market price constraints. It is a simple adaptation of the classical dual program used for entropy optimization (see Cover and Thomas (1991)). The new idea proposed here is to apply the algorithm to the state-space which consists of a collection of sample paths generated by Monte Carlo simulation of the prior.

To fix ideas, we shall consider a model in which paths are generated as solutions of the stochastic difference equations

$$
X_{n+1} = X_n + \sigma(X_n, n) \cdot \xi_{n+1} \sqrt{\Delta T} + \mu(X_n, n) \, \Delta T, \qquad n = 1, 2, \dots M \qquad (13)
$$

where $M \Delta T = T_{\max}$ is the horizon time. Here $X_n \in \mathbf{R}^d$ is a vector of state variables and as a multidimensional process with values and $\xi_n \in \mathbf{R}^{d'}$ is as a vector of independent Gaussian shocks ($d, d'$ are positive integers). The variance-covariance structure is represented by the $\nu \times \nu'$ matrix $\sigma(X, t)$ and the drift is the $\nu$-vector $\mu(X, t)$. [6]

Using a pseudo-random number generator, we construct a set of sample paths of (41) of size $\nu$, which we denote by

$$
\omega^{(i)} = \left( X_1(\omega^{(i)}), \ \dots X_M(\omega^{(i)}) \right) \qquad i = 1, 2, \dots \nu \, . \qquad (14)
$$

We assume throughout this paper that the benchmark instruments are such that their *cashflows along each path $\omega$ are completely determined by the path itself.* In the case of equities, where the components of the state-vector $X$ generally represent stock prices, instruments satisfying this assumption include forwards, futures and standard

[6] This formulation extends trivially to the case of jump-diffusions or more general Markov processes and the MRE algorithm applies to these more general stochastic processes.

European options. It is also possible to use barrier options or average-rate options. American-style derivatives do not satisfy this assumption because the early-exercise premium depends on the value of the option (and hence on the full pricing measure defined on the paths) as well as on the current state of the world. For fixed-income securites, benchmark instruments can include interest rate forwards, futures contracts, bonds, swaps, caps and European swaptions.[7] Under these circumnstances, the price relations can be written in the form (4) where $g_{ij}$ is the present value of the cash-flows of the $j^{th}$ instrument along the $i^{th}$ path. The mathematical problem is to minimize the convex function of $p$

$$D(p|u) = \log \nu + \sum_{i=1}^{\nu} p_i \log p_i \tag{15}$$

under linear constraints. This problem has been well-studied (Cover and Thomas, 1991). Introducing Lagrange multipliers $\lambda_1, \ldots \lambda_N$), we can reformulate it as a minimax program (the "dual" formulation of the constrained problem)

$$\min_{\lambda} \max_{p} \left\{ - D(p|u) + \sum_{j=1}^{N} \lambda_j \left( \sum_{i=1}^{\nu} p_i \, g_{ij} \; - \; C_j \right) \right\} \tag{16}$$

A straighforward argument shows that probability vector that realizes the supremum for each $\lambda$ has the Boltzmann-Gibbs form

$$p_i = p(\omega^{(i)}) = \frac{1}{Z(\lambda)} \exp \left( \sum_{j=1}^{N} g_{ij} \, \lambda_j \right) \; . \tag{17}$$

To determine the Lagrange multipliers, define the "partition function", or normalization factor,

$$Z(\lambda) = \frac{1}{\nu} \sum_{i=1}^{\nu} \exp \left( \sum_{j=1}^{N} g_{ij} \, \lambda_j \right) \; . \tag{18}$$

and consider the function

$$\begin{aligned} W(\lambda) &= \log \left( Z(\lambda) \right) \; - \; \sum_{j=1}^{N} \lambda_j \, C_j \\ &= \log \left\{ \frac{1}{\nu} \sum_{i=1}^{\nu} \exp \left( \sum_{j=1}^{N} g_{ij} \, \lambda_j \right) \right\} - \sum_{j=1}^{N} \lambda_j \, C_j \; . \end{aligned} \tag{19}$$

We shall denote by $g_j(\omega)$ the present value of the caashflows of the $j^{th}$ instrument along the path $\omega$. (Thus, $g_j(\omega_i) = g_{ij}$). At a critical point of $W(\lambda)$, we have

[7] American-style securities, such as Bermudan swaptions or callable bonds do not satisfy this assumption.

$$
\begin{aligned}
0 &= \frac{1}{Z}\frac{\partial Z}{\partial \lambda_k} - C_k \\
&= \frac{1}{Z(\lambda)} \sum_{i=1}^{\nu} g_{ik} \exp\left( \sum_{j=1}^{N} g_{ij}\, \lambda_j \right) - C_k \\
&= \mathbf{E}^p \{ g_k(\omega) \} - C_k \ . \qquad (20)
\end{aligned}
$$

Hence, if $\lambda$ is a critical point of $W(\lambda)$, the probability vector defined by equation (28) is calibrated to the benchmark instruments.

Notice that the function $W(\lambda)$ is convex: differentiating both sides of equation (19) with respect to $\lambda$ yields

$$
\frac{\partial^2 W(\lambda)}{\partial \lambda_j \, \partial \lambda_k} = \mathbf{Cov}^p \{ g_j(\omega)\, g_k(\omega) \} \ , \qquad (21)
$$

which is a non-negative definite matrix. In particular, the critical point, if it exists, should correspond to a *minimum* of $W(\lambda)$.

Based on this, we have the following algorithm:

(i) Construct a set of sample paths using the difference equations (1) and a pseudo-random-number generator.

(ii) Compute the cashflow matrix $\{g_{ij}, i = 1, ...\nu, \quad j = 1, 2, ..., N, \}$.

(iii) Using a gradient-based optimization routine, minimize the function $W(\lambda)$ in (19)).[8]

(iv) Compute the risk-neutral probabilities $p_i \ , \ i = 1, 2, ...\nu$ for each path using equation (28) and the optimal values of $\lambda_1 \ \ldots \lambda_N$.

# 4 Implementation using weighted least-squares

It may not always be desirable to match model values to the price data exactly due to bid-ask spreads and liquidity considerations. Alternatively, we can minimize the sum of the weighted least-squares residuals and the relative entropy. We define the sum of the weighted least-squares residuals

$$
\chi^2_w = \frac{1}{2} \sum_{j=1}^{N} \frac{1}{w_j} \left( \mathbf{E}^p \{ g_j(\omega) \} - C_j \right)^2 \ , \qquad (22)
$$

where the $w = (w_1 \ \ldots \ w_N)$ is a vector of positive weights.

The proposal is to minimize the quantity

$$
\chi^2_w + D(p|u) \qquad (23)
$$

[8] In our implementation, we use the L-BFGS algorithm.

over all probability vectors $p = (p_1 \ldots p_\nu)$. Notice that the limit $w_i \ll 1$ corresponds to exact matching of the constraints. The discrepancy between the model value and market price with a weight $w_i$ is typically of order $\frac{1}{\sqrt{w_i}}$.

We indicate how to modify the previous algorithm to compute the probabilities $(p_1 \ldots p_\nu)$ that minimize $\chi^2_w + D(p|u)$. Using the inequality

$$a\,b \leq \frac{1}{2}a^2 + \frac{1}{2}b^2 \,, \tag{24}$$

we find that, for all $p$,

$$\chi^2_w \geq -\sum_{j=1}^{N} \lambda_j \left(\mathbf{E}^p \{g_j(\omega)\} - C_j\right) - \frac{1}{2}\sum_{j=1}^{N} w_j \lambda_j^2 \,. \tag{25}$$

It follows that

$$\inf_p \left[D(p|u) + \chi^2_w\right] \geq \tag{26}$$

$$\begin{aligned} & \sup_\lambda \left\{ \inf_p \left[ D(p|u) - \sum_{j=1}^{N} \lambda_j \left(\mathbf{E}^p \{g_j(\omega)\} - C_j\right)\right] - \frac{1}{2}\sum_{j=1}^{N} w_j \lambda_j^2 \right\} \\ = \; & -\inf_\lambda \left[ \log\left(Z(\lambda)\right) - \sum_{j=1}^{N} \lambda_j C_j + \frac{1}{2}\sum_{j=1}^{N} w_j \lambda_j^2 \right] \\ = \; & -\inf_\lambda \left[ W(\lambda) + \frac{1}{2}\sum_{j=1}^{N} w_j \lambda_j^2 \right] \,. \end{aligned} \tag{27}$$

Here, $W(\lambda) = \log(Z(\lambda) - \sum_j \lambda_j C_j$ is the function used in the case of exact fitting.

The inequality expressed in (27) is in fact an equality. To see this, observe that the function $D(p|u) + \chi^2_w$ is convex in $p$ and grows quadratically for $p \gg 1$. Therefore, a probability vector realizing the infimum exists and is characterized by the vanishing of the first variation in $p$. A straightforward calculation shows that if $p$ is a minimum of this function, we have

$$p_i^* = \frac{1}{Z(\lambda^*)} \exp\left(\sum_{j=1}^{N} \lambda_j^* g_{ij}\right) \tag{28}$$

with

$$\lambda_j^* = -\frac{1}{w_j}\left(\mathbf{E}^{p^*}\{g_j(\omega)\} - C_j\right) \,. \tag{29}$$

In particular, notice that this value of $\lambda$ is such that (25) is an equality. Furthermore, the probability (28) is of exponential type, so we have

$$\begin{aligned} D(p^*|u) + \chi^2_w &= D(p^*|u) - \sum_{j=1}^N \lambda_j^* \left( \mathbf{E}^{p^*} \{ g_j(\omega) \} - C_j \right) - \frac{1}{2} \sum_{j=1}^N w_j (\lambda_j^*)^2 \\ &= \inf_p \left[ D(p|u) - \sum_{j=1}^N \lambda_j^* \left( \mathbf{E}^p \{ g_j(\omega) \} - C_j \right) \right] - \frac{1}{2} \sum_{j=1}^N w_j \lambda_j^2 \\ &= -\log(Z(\lambda^*)) + \sum_{j=1}^N \lambda_j^* C_j - \frac{1}{2} \sum_{j=1}^N w_j \lambda_j^2 \\ &\leq -\inf_\lambda \left( W(\lambda) + \frac{1}{2} \sum_{j=1}^N w_j \lambda_j^2 \right), \end{aligned} \tag{30}$$

so equality must hold. This calculation shows that the pair $(\lambda^*, p^*)$ is a saddlepoint of the min-max problem and that there is no "duality gap" in (27) and (30).

The algorithm for finding the probabilities that mimimise $\chi^2_w$ under the entropy penalization consists in minimizing

$$\log(Z(\lambda)) - \sum_{j=1}^N \lambda_j \left( \mathbf{E}^{p^*} \{ g_j(\omega) \} - C_j \right) + \frac{1}{2} \sum_{j=1}^N w_j \lambda_j^2 \, . \tag{31}$$

among candidate vectors $\lambda$. This algorithm represents a small modification of the one corresponding to the exact fitting of prices and can be implemented in the same way, using L-BFGS. [9]

# 5 Price sensitivities and hedge-ratios

The MRE setting provides a simple method for computing portfolio price-sensitivities, under the additional assumption that *the prior measure remains fixed* as we perturb the benchmark prices and recalibrate. [10] We show, under this assumption, that sensitivities can be related to regression coefficients of the target contingent claim on the cashflows of the benchmarks. For simplicity, we discuss only the case of exact fitting, but the analysis carries over to the case of least-squares residuals with minor modifications.

---

[9] More generally, we can consider the minimization of $W(\lambda) + \sum_{j=1}^N w_j \psi(\lambda_j)$, where $\psi$ is a convex function. An argument entirely similar to the one presented above shows that this program corresponds to minimizing the quantity $\sum \frac{1}{w_j} \psi^*(\mathbf{E}^{p^*} \{ \Gamma_j(\omega) \} - C_j)$, where $\psi^*$ is the Legendre dual of $\psi$. The case $\psi(x) = |x|$ can be used to model proportional bid-ask spreads in the prices of the benchmark instruments, for example.

[10] This assumes, implicitly, that the prior probability represents information that "varies slowly" with respect to the observed market prices. For example, the assumption is consistent with interpreting the prior as a historical probability.

Let $F\left(\omega^{(i)}\right)$, $i = 1, \ldots \nu$ represent the discounted cash-flows of a portfolio or contingent claim. To compute the price-sensitivities of the model value of the portfolio we utilize the "chain rule", differentiating first with respect to the Lagrange multipliers. More precisely,

$$\frac{\partial \mathbf{E}^p\left(F(\omega)\right)}{\partial C_k} = \sum_{j=1}^{N_{sec}} \frac{\partial \mathbf{E}^p\left(F(\omega)\right)}{\partial \lambda_j} \frac{\partial \lambda_j}{\partial C_k} \tag{32}$$

We note, using equation (**??**) for the probability $p_i$ , that

$$\frac{\partial \mathbf{E}^p\left(F(\omega)\right)}{\partial \lambda_j} = \mathbf{Cov}^p\left\{F(\omega)\,,\; g_j(\omega)\right\}\;. \tag{33}$$

Moreover, we have, on account of equation (21),

$$\begin{aligned} \frac{\partial}{\partial \lambda_j}\left(\mathbf{E}^p\left(g_k(\omega)\right)\right) &= \frac{\partial}{\partial \lambda_j}\left(\frac{\partial \log(Z(\lambda))}{\partial \lambda_k}\right) \\ &= \mathbf{Cov}^p\left\{g_j(\omega),\, g_k(\omega)\right\} \end{aligned} \tag{34}$$

In particular,

$$\frac{\partial C_k}{\partial \lambda_j} = \mathbf{Cov}^p\left\{g_j(\omega)\, g_k(\omega)\right\}\;. \tag{35}$$

Substitution of the expressions in (34) and (35) into equation (32) gives

$$\nabla_C\, \mathbf{E}^p\left\{F(\omega)\right\} = \mathbf{Cov}^p\left\{F(\omega)\; g_\cdot(\omega)\right\} \cdot \left[\, \mathbf{Cov}^p\left\{g_\cdot(\omega)\, g_\cdot(\omega)\right\}\,\right]^{-1}\;, \tag{36}$$

with the obvious matrix notation. [11] This implies, in turn, that the sensitivities of the portfolio value with respect to the input prices are the linear regression coefficients of $F(\omega)$ with respect to $g_j(\omega)$. Namely, if we solve

$$\min_{\beta}\left\{\mathbf{Var}^p\left[\; F(\omega) \;-\; \beta_0 \;-\; \sum_{j=1}^{N} \beta_j\; g_j(\omega)\;\right]\right\}\;, \tag{37}$$

we obtain, from (36),

$$\beta_k = \frac{\partial \mathbf{E}^p\left(F(\omega)\right)}{\partial C_k} \qquad k = 1,\; \ldots\; N \tag{38}$$

and[12]

[11] The invertibility of the covariance matrix presupposes that the cashflow vectors of the benchmark instruments, $g_j(\omega)\; j = 1, \ldots N$, are linearly independent. This assumption is discussed, for example, in Avellaneda(1998).

[12] This result can be interpreted as follows. Assume that an agent hedges the initial portfolio by shorting $\beta_j$ units of the $j^{th}$ benchmark instrument for $j = 1, \ldots, N$. In this case, the model value of the net holdings (initial portfolio + hedge) is $\beta_0$. It represents the expected cost of dynamic replication of the residual.

$$\beta_0 \;=\; \mathbf{E}^p\left(F(\omega)\right) - \sum_{j=1}^{N} \beta_j \; \mathbf{E}^p\left(g_j(\omega)\right) \;. \tag{39}$$

We conclude that a sensitivity analysis with respect to variations of the input prices can be done without the need to perform additional Monte Carlo runs and to perturb the input prices one by one. Instead, the MRE framework allows us to compute prices and hedges with a single Monte Carlo simulation, which is much less costly.[13]

Notice that the characterization of the hedge-ratios as regression coefficients shows that they are "stable" in the sense that they vary continuously with input prices. In practice, the significance of this hedging technique depends on details of the implementation procedure, such as the number of paths used in the simulation. The main issue is whether the support of the probability measure induced by the prior – the basic scenarios of the simulation – is sufficiently "rich in scenarios", for example.

# 6 Variance reduction

The calibration of Monte Carlo simulations significantly reduces noise in the pricing of many cash-flow structures. Claims that are "well-replicated" by the benchmarks – in the sense that the variance in (37) is small compared to the variance of $F(\omega)$[14] – will benefit from a significant noise reduction in comparison with standard MC evaluation.

In fact, given any vector $\zeta \;=\; (\zeta_1, \ldots, \zeta_N)$, we have

$$\mathbf{E}^p(F(\omega)) \;=\; \mathbf{E}^p\left\{F(\omega) - \sum_{j=1}^{N} \zeta_j \, g_j(\omega)\right\} \;+\; \sum_{j=1}^{N} \zeta_j \, C_j \;. \tag{40}$$

Since the second term on the right-hand is constant, the variance of the Monte Carlo method for pricing the cash-flow $F$ is the same as the one associated with $F - \zeta \cdot g$. This statement is true for any value of the vector $\zeta$ and so, in particular, for the regression coefficients $(\beta_1 \ldots \beta_N)$. Since, by definition, $F - \beta \cdot g$ has the least possible true variance among all choices of $\zeta$, the cash-flow $\beta \cdot \Gamma$ is an "optimal control variate" for the simulation. Our method implicitly uses such control variates.

To measure this variance reduction experimentally in a simple framework, we considered the problem of calibrating a Monte Carlo simulation to the prices of European stock options, assuming a lognormal price with constant volatility.

We considered European options on a stock with a spot price of 100 with no dividends. The interest rate was taken to be zero. Taking a 'maximum horizon" for the model of 120 days, we used as benchmarks all European options with maturities of 30, 60 and 90 days and strikes of 90, 100 and 110, as well as forward contracts with maturities of 30, 60 and 90 days. We assumed that the prices of the benchmarks

[13] In contrast, a perturbation analysis that uses centered differences to approximate the partial derivatives with respect to input instruments requires $2\,N \;+\; 1$ Monte Carlo simulations.

[14] The ratio of the variances is the statistic $1 - R^2$ in the risk-neutral measure.

were given by the Black-Scholes formula with a volatility of 25 percent.The prior was taken to be a geometric Brownian motion with drift zero and volatility 25%.[15]

The test consisted of pricing various options (target options) with strike/maturity distributed along a regular grid (maturities from 20 days to 120 days with 1-day intervals; all integer strikes lying between two standard deviations from the mean of the distribution). For each option, we compared the variances resulting from pricing with the simulated lognormal process with and without calibrating to the "benchmarks". As a matter of general principle, when pricing an option contract, we also include the forward contract corresponding to the option's expiration date in the set of calibration instruments.[16]

All Monte Carlo runs were made with 2000 paths. Each run took roughly half a second of CPU time on a SunOS 5.6. This includes the time required to search for the optimal lambdas. We verified the correctness of the scheme by checking that all model prices fell within 3 (theoretical) standard deviations of the true price, both with and without the min-entropy adjustment.

Figure 2: Variance Improvement Ratio

| Maturity | Strike | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (Days) | 80 | 85 | **90** | 95 | **100** | 105 | **110** | 115 | 120 |
| 20 | N/A | N/A | 1.03 | 2.22 | 8.24 | 2.78 | 1.54 | N/A | N/A |
| **30** | N/A | N/A | INF | 13.66 | INF | 19.11 | INF | 3.09 | N/A |
| 45 | N/A | 1.25 | 2.38 | 5.21 | 14.83 | 6.63 | 3.75 | 2.25 | N/A |
| **60** | N/A | 5.63 | INF | 49.83 | INF | 73.98 | INF | 10.83 | 3.03 |
| 75 | 1.54 | 3.04 | 6.25 | 10.61 | 25.71 | 14.08 | 9.12 | 5.45 | 3.05 |
| **90** | 2.47 | 8.79 | INF | 92.40 | INF | 150.36 | INF | 22.96 | 5.77 |
| 120 | 1.96 | 2.77 | 4.09 | 6.08 | 13.57 | 8.34 | 5.77 | 4.15 | 3.03 |

We found that there was significant variance reduction for all cases, with the exception of options having strikes far from the money and maturities which did not match the benchmark maturities. As expected, the best results were observed for those options with strikes and maturities near to those of the benchmark options. In particular, options with the "benchmark maturities" (30, 60 and 90 days) yielded some of the best results for most strikes which were not too far away from the money. We also obtained some of the best results of options with strikes at or close to the forward values. The following table gives the factor by which the entropy method improved the variance for selected strikes and maturity dates. Note that the table below includes the benchmark instruments which yield an infinite improvement since the entropy method always prices them correctly (indicated by INF on Figure 2). Benchmark strikes and maturities are shown in boldface.

The variance reduction from the entropy method translated into some excellent

[15]We assumed that all benchmark options were correctly priced with the prior to separate the issues of calibration and variance reduction, focussing on the latter.

[16]Doing so guarantees that the mean of the distribution of the asset price at the expiration date is fitted exactly.

data for the computed standard errors. The figure below contains this information. The data is given in terms of Black-Scholes implied volatility and is obtained by taking the standard error of the option price and dividing by the Black-Scholes value of vega.

Figure 3: Standard Errors from the Entropy Method (In percentage of implied volatility)

| Maturity | Strike | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (Days) | 80 | 85 | **90** | 95 | **100** | 105 | **110** | 115 | 120 |
| 20 | N/A | N/A | 0.56 | 0.37 | 0.32 | 0.37 | 0.51 | N/A | N/A |
| **30** | N/A | N/A | 0 | 0.16 | 0 | 0.15 | 0 | 0.38 | N/A |
| 45 | N/A | 0.47 | 0.35 | 0.28 | 0.26 | 0.28 | 0.34 | 0.47 | N/A |
| **60** | N/A | 0.25 | 0 | 0.10 | 0 | 0.09 | 0 | 0.21 | 0.45 |
| 75 | 0.50 | 0.32 | 0.23 | 0.21 | 0.21 | 0.22 | 0.24 | 0.30 | 0.42 |
| **90** | 0.38 | 0.19 | 0 | 0.07 | 0 | 0.07 | 0 | 0.15 | 0.32 |
| 120 | 0.40 | 0.33 | 0.30 | 0.29 | 0.29 | 0.29 | 0.31 | 0.35 | 0.40 |

Finally, we examined the $R^2$ statistic given by the entropy method. We found that $R^2$ was greatest for values with benchmark maturity dates and strikes whose values are close to that of the forward. The results are given in the table below. Our interpretation is that the options with benchmark dates or near-the-money strikes have only a small component of their cashflows which is orthogonal to the benchmark instruments, and conversely. One would expect both greater variance reduction and dependence on the values of the benchmark instruments in these cases. Note that the variance reduction data given in figure 2 confirms this interpretation.

Figure 4: R Squared Statistic from the Entropy Method

| Maturity | Strike | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| (Days) | 80 | 85 | **90** | 95 | **100** | 105 | **110** | 115 | 120 |
| 20 | N/A | N/A | 0.25 | 0.57 | 0.85 | 0.65 | 0.44 | N/A | N/A |
| **30** | N/A | N/A | 1 | 0.93 | 1 | 0.95 | 1 | 0.75 | N/A |
| 45 | N/A | 0.40 | 0.63 | 0.80 | 0.91 | 0.83 | 0.72 | 0.55 | N/A |
| **60** | N/A | 0.83 | 1 | 0.97 | 1 | 0.98 | 1 | 0.91 | 0.68 |
| 75 | 0.41 | 0.68 | 0.83 | 0.89 | 0.95 | 0.92 | 0.88 | 0.81 | 0.67 |
| **90** | 0.62 | 0.89 | 1 | 0.99 | 1 | 0.99 | 1 | 0.95 | 0.80 |
| 120 | 0.51 | 0.63 | 0.73 | 0.79 | 0.90 | 0.85 | 0.80 | 0.73 | 0.64 |

# 7 Example: fitting a volatility skew

We apply the algorithm to calibrate a model using forwards and the prices of European options with different strikes and maturities. This example is take from

the interbank foreign exchange market. It is well-known that options with different strikes/maturities trade with different implied volatilities. The goal is to construct a pricing model that incorporates this effect. Notice that such problem has been addressed by many authors in the context of the so-called "volatility surface"( Dupire (1994), Derman and Kani (1994), Rubinstein (1994), Chriss (1996); see Avellaneda *et al*(1997) for references on this problem up to 1997). The method presented here is completely different since we do not iterpolate option prices or use a parameterization of the local volatility function $\sigma(S,\, t)$.

Figure 5: Data used for fitting the implied volatilities of options. The implied volatilities, displayed on the left-hand side of the graph, range from 13% to 14.5%. From Avellaneda and Paras (1996).

We considered a dataset consisting of 25 contemporaneous USD/DEM option prices obtained from a major dealer in the interbank market on August 25, 1995. The maturities are 30, 60, 90, 180 and 270 days. Strikes (quoted in DEM) correspond to 50-, 20- and 25-delta puts and calls. Aside from these options, we introduced 5 additional "zero-strike options" which correspond to the present value of a dollar in

DEM for delivery at the different expiration dates (see Table 5) – the forward prices implied by the interest rates and the spot price. Including these forward prices in the set of benchmark instruments ensures that the model is calibrated to the forward rates and hence that there is no net bias in the forward prices.

As a prior, we considered the system of stochastic differential equations:

$$\begin{aligned} \frac{dS_t}{S_t} &= \sigma_t \, dZ_t \; + \; \mu \, dt \\ \frac{d\sigma_t}{\sigma_t} &= \kappa dW_t \; + \; \nu_t \, dt \; , \end{aligned} \tag{41}$$

where $Z_t$ and $W_t$ are Brownian motions such that $\mathbf{E}(dZ_t \; dW_t) \; = \; \rho \; dt$ . In equation (41), $S_t$ represents the value of one US Dollar in DM. The instantaneous volatility is denoted by $\sigma_t$. The additional parameters are: $\mu$, the cost-of-carry (interest rate differential), $\kappa$, the volatility of volatility, and $\nu_t$ is the drift of the volatility. Therefore, we are calibrating a two-factor stochastic volatility model. We assume the following numerical values for the parameters that define the prior dynamics:

1. $S_0$= midmarket USD/DEM spot exchange rate = 1.4887
2. US rate=5.91%
3. DM rate=4.27%
4. $\mu = -1.64\%$ (For convenience, we take $\mu$ =DM rate - US rate in the prior, *i.e.* we adjust the model to the standard risk-neutral drift).
5. $\sigma_0$ = Initial value of the prior volatility of USD/DEM = 14%. (This is essentially the average of the observed implied volatilities).
6. $\kappa = 50\%$
7. $\rho$ = - 50%

We simulated 5000 paths of equations (41), consisting of 2500 paths and their antithetics. The gradient tolerance in the BFGS routine was set to $10^{-7}$ and we used equal weights $w_i = 10^{-5}$ for the least-squares approximation. We found that the difference between model prices and market prices was typically on the order of $10^{-4} \; - \; 10^{-5}$ DM, representing relative errors of 1% in the deep-out-of-the money short-term options and much less 0.1% for at the money options (See Figure 6).

The algorithm initiated with $\lambda_i \; = \; 0, \; i = 1, \ldots 30$ converges after approximately 20 iterations of the BGFS routine. The entire calibration procedure takes about 4 seconds on a desktop PC with a Pentium II 330 Mhz processor. In practice, computation times are much faster because the values of lambdas from the previous runs can be stored and used as better initial guesses.

Figure 6: Fitting errors and values of lambda for the 30 instruments used in the calibration.

The relative entropy of the calibrated risk-neutral measure was found to be $D(p|u) = 7.39 \times 10^{-3}$. We can interpret this result in terms of the parameter $\alpha$ of Section 2. We find a value of $\alpha = 1 - \frac{D(p|u)}{\log \nu} = 0.99913$, which would correspond to an "effective number of paths" $\nu^{\alpha} \approx 4963$ according to the heuristics of Section 2. This represents an excellent fit in terms of the support of the calibrated measure. In Figures 7, we present descriptive statistics for the calibrated probabilities, in Figure 8, we plot the probabilities, which appear to be randomly distributed with a small mean about the uniform value $1/5000 = 0.0002$. Figure 9 displays a histogram of the logarithms of the probabilities. These results indicate a relatively small scatter about the mean and a large set of paths which support the posterior measure.

Figure 7: Descriptive statistics for the vector of calibrated probabilities $(p_1, \ldots, p_{5000})$.

Figure 8: Snapshot of the 5000 probabilities obtained by the method. The values oscillate about $1/5000 = 0.0002$.

The values of the lambdas were between $-0.84$ and $0.79$, which correspond to moderate variations of the probabilities about their mean. In Figure 9, we present

a histogram of the 5000 probabilities obtained The distibution of probabilites – or, equivalently, the distribution of state-price deflators – is unimodal and strongly peaked about its mean $1/5000 = 0.0002$. This confirms that the risk-neutral measure is supported on the full set of paths generated using the prior.

Figure 9: Histogram of the logarithms of the calibrated probabilities multiplied by 5000. The distribution is unimodal and strongly peaked about $log(1/5000) = -8.517$. Observe that there are some outliers corresponding to small probabilities.

To gain insight into the calibrated model, we generated an *implied volatility surface*, by repricing a set of options on a fine grid in strike/maturity space. We determined the highest and lowest strikes in the input option, which are, respectively, 1.67 and 1.34 DEM, corresponding to the 20-delta options with the longest maturity (270 days). We then considered strike increments going from the maximum to the minimum value according to the rule

$$k_{max} = 1.67 = k_0 \quad , \quad k_i \; = \; \frac{k_{i-1}}{1.01} \; , \quad i = 1, 2, ...20 \tag{42}$$

and maturities

$$t_{min} = 30 = t_0 \quad , \quad t_i \; = \; t_{i-1} \; + \; 10 \; , \quad i = 1, 2, ...24. \tag{43}$$

We repriced these 480 options with the weighted Monte Carlo and generated a surface by interpolating the implied volatilities linearly between strikes and dates. The

interpolation was done graphically using the Excel 5.0 graphics package. The results are exhibited in Figures 10 and 11.

Finally we tested the model by pricing a barrier option with the following characteristics:

1. Option type: USD put, DM call
2. Notional amount: 1 USD
3. Maturity: 180 days
4. Strike: 1.48 DM
5. Knockout barrier: 1.38 DM

and comparing the results with the price and delta hedge obtained using the Black-Scholes lognormal model.

Figure 10: Implied volatility surface for the USD/DEM options market based on the data in Figure 5. The prior volatility is 14%.

Closed-form solutions for the price of barrier options in the lognormal setting were computed by Reiner and Rubinstein (1991). To compare with our model, we used a Black-Scholes constant volatility of 14% – which corresponds to the implied

volatilities of short-maturity options – as well as with a constant volatility of 13%. The latter corresponds to the implied volatility of the ATM forward call expiring in 180 days. The differences in the Black Scholes prices and deltas by changing the volatilities from 13 to 14% are quite small. In Figure 12, we compare the results of applying the closed-form solution with 14% volatility are compared with the Monte Carlo simulation model described above. The values obtained with both models agree to 3 significant digits.

Figure 11: Call prices associated with the implied volatility surface of Figure 10.

To evaluate the quality of the hedges, we computed the "net delta" of the exotic option given by the model and compared it with the Black-Scholes delta. For this, we assumed that the input options have the same deltas under BS than under our model. Of course, this a simplification, since the deltas of the individual options may be affected globally by the differences in implied volatilities. However, experience shows that such approximation is reasonable, in the sense that we do not expect the volatility skew in this market to distort significantly the deltas of the plain-vanilla options. The total amount of forward dollars needed to hedge the knockout is obtained by converting each option hedge into an equivalent forward position and summing over all contracts with the same maturity. To this amount we also add the corresponding sensitivity to the forward contract (modeled here as a call with strike 0).

Figure 12: Price and hedge report for the reverse-knockout barrier option. Notice that the calibrated model computes exposures to all the options and forwards entered as reference instruments. Instruments that have an exposure (in notional terms) of more than 10% of the notional amount of the exotic option are labeled with asterisks. We observe significant exposure (1) at 60 and 90 days near the knockout barrier and (2) at the expiration date in ATM and low-strike options. Hedges at th barrier involve a spread in contracts with neighboring strikes, as espected.

Adding the "forward deltas" associated with the different maturities, we find a total of 1.57% which is near the Black-Scholes value of 1.80 The results obtained in this example appear reasonable, despite the fact that the computation was done using Monte Carlo simulation with only 5000 paths. We believe that this is due in part to the reduction of variance which results from calibration process.

# 8 Example 2: Fitting the smile in AmericaOnline options on May 1999

We calibrated another stochastic volatility model to the mid-market prices of 35 America Online call prices recorded on May 10, 1999 at the close.

To ensure sufficient liquidity of the benchmarks, we took only options with traded volume above 100 for shorter maturities and above 50 for the maturity longer than

| Maturity | Strike | Price | IVOL | Maturity | Strike | Price | IVOL |
|---|---|---|---|---|---|---|---|
| 12 | 120 | 12.125 | 78.33 | 68 | 150 | 12.25 | 87.99 |
| 12 | 130 | 7.125 | 83.73 | 68 | 170 | 7.625 | 87.60 |
| 12 | 135 | 5.125 | 83.29 | 68 | 175 | 6.625 | 86.80 |
| 12 | 140 | 3.625 | 83.40 | 68 | 180 | 6 | 87.54 |
| 12 | 145 | 2.625 | 85.16 | 68 | 200 | 3.75 | 87.84 |
| 40 | 115 | 21.75 | 85.34 | 159 | 120 | 32.625 | 83.90 |
| 40 | 120 | 18.875 | 85.27 | 159 | 125 | 30.25 | 83.19 |
| 40 | 125 | 16.25 | 84.96 | 159 | 150 | 21.5 | 83.43 |
| 40 | 135 | 12.25 | 86.86 | 159 | 160 | 19 | 84.19 |
| 40 | 140 | 10.625 | 87.80 | 257 | 100 | 48.375 | 81.21 |
| 40 | 160 | 5.5 | 87.65 | 257 | 110 | 43.375 | 80.80 |
| 68 | 100 | 35.625 | 88.52 | 257 | 120 | 38.75 | 80.07 |
| 68 | 110 | 29 | 86.98 | 257 | 130 | 35.125 | 80.76 |
| 68 | 115 | 25.875 | 85.55 | 257 | 150 | 28.125 | 79.73 |
| 68 | 120 | 23.25 | 85.61 | 257 | 160 | 25.375 | 79.77 |
| 68 | 125 | 20.875 | 85.78 | 257 | 170 | 23 | 79.98 |
| 68 | 135 | 17.125 | 87.80 | 257 | 200 | 16.625 | 78.93 |
| 68 | 145 | 13.625 | 87.54 | | | | |

Figure 13: America Online call (mid-market) prices on May 10, 1999

half a year. It can be seen from the table that the implied volatilities of the benchmark calls vary in the range 78.33-88.52%. These extreme volatilities correspond to deeply in- or out-of-the money short term options. We also included forward prices for the stock at the different delivery dates, namely 12, 40, 68, 159 and 257 days. Thus, we calibrated the simulation to 40 benchmark prices.

We simulated $\nu = 10,000$ Monte Carlo paths from this distribution on the time-horizon of 258 days with 1 time-step per day. We used the following parameters: $S_0 = 128.375$ (the America Online closing price on May 10, 1999), $\sigma_0 = 0.86$, $\rho = -0.5$, $r = 5\%$ and $\kappa = 0.5$. Then, we applied the MRE method described in section 3.2 to calibrate the uniform distribution on the obtained sample paths to the set of 35 European call prices on America Online given in Figure 13.

The program matched all the given 40 benchmark prices with the predefined accuracy to 4 decimal places using 181 iterations starting from $\lambda = 0$. The obtained entropy of the calibrated measure on the path space was 0.66 with the maximum possible entropy being $\log 10000 = 9.21$, indicating that the prior distribution its not far in the entropy distance from the calibrated one.

Figure 14: Implied volatility surface for AOL option closing prices, calibrated to the prices of Figure 13. The additional parameters are spot price=128.312, $\sigma_0 = 86\%, r = 5\%, \kappa = 50\%$ and $\rho = -50\%$. The relative entropy distance to the prior is $D = 0.66$.

Figure 14 displays the implied volatility surface associated with the calibrated model. This surface was obtained by pricing a dense grid of plain-vanilla options with the calibrated Monte Carlo. Figure 15 represents the surface of corresponding call option prices. Notice that the shapes of the implied volatility surfaces in both examples are quite different Of course, the call price surfaces appear to be more similar: from well-know no-arbitrage relations, they are both convex and decreasing in the strike direction and monote-increasing with expiration.

Figure 15: Surface of call prices corresponding to the implied volatility surface of Fig. 14.

# 9 Example 3: constructing a US Treasury yield curve

We considered the following calibration problem: given the current prices of on-the-run treasury securities, construct a smooth forward rate curve consistent with these pricesand perform sensitivity analysis around them (as would be done for

the hedging of a hypothetical fixed-income portfolio).

Table 16 shows the on-the-run instruments and the prices observed in the morning of Thursday, April 15 1999.

A stochastic short-rate model was used to discount future cash flows. As a prior, we considered the modified Vasicek model

$$dr = \alpha\,(m(t) - r)\,dt + \sigma\,dW \tag{44}$$

Here, $m(t)$ is the—possibly time-dependent—level of mean reversion, and the constant $\alpha$ controls the rate of mean reversion. We experimented with two types of mean-reversion levels: constant levels and time-dependent levels, where the latter were

| Maturity | Coupon | Price |
|---|---|---|
| 07/15/99 | – | 98.955 |
| 10/14/99 | – | 97.823 |
| 03/30/00 | – | 95.725 |
| 03/31/01 | 4.875 | 99.875 |
| 02/15/04 | 4.75 | 98.812 |
| 11/15/08 | 4.75 | 97.219 |
| 02/15/29 | 5.25 | 96.250 |

Figure 16: Seven benchmark US-treasury bills and bonds. Prices are as quoted on 04/15/99 and not adjusted for accrued interest. The alignment of prices reflects the direction of the sensitivities listed below in Figure 1: left alignment indicates positive sensitivity, right alignment negative sensitivity

| Scenario | $\alpha$ | $m(t)$ | $\sigma$ |
|---|---|---|---|
| I | 0.25 | 0.0426035 | **0.01** |
| II | 0.25 | 0.0426035 | **0.05** |
| III | 0.25 | bootstrap | **0.01** |
| IV | 0.25 | bootstrap | **0.05** |

Figure 17: Various prior instantiations of the coefficients of (44). 0.0426035 is the rate of the first leg of the piecewise constant bootstrapped forward rate curve

taken to be $m(t) =$ the piecewise-constant (bootstrapped) instantaneous forward-rate curve.[17]

Figure 17 shows several prior instantiations of the coefficients of (44).

We calibrated the modified Vasicek model with 15000 Monte-Carlo paths and 24 time steps per year. Figure 18 shows the calibrated forward-rate curve and zero-coupon-bond yield curve for scenarios I and II, with constant level of mean reversion. Figure 19 shows the calibrated forward-rate curve and zero-coupon-bond yield curve for scenarios III and IV. In these scenarios, the level of mean reversion $m(t)$ is set to the piecewise constant bootstrapped forward rate curve consistent with the data in Figure 16.

In accordance with the work of Samperi(1997) and others, the optimal Lagrange multipliers $\lambda_j^*$ can be interpreted in terms of an optimal investment portfolio. Specifically, consider an expected CARA utility function defined on the space of static

[17]These are arbitrary modeling choices. For example, we could start with an econometrically calibrated forward rate curve or with a level of mean-reversion that corresponds to an estimate of forward rates for long maturities. Bootstrapping is standard method for building a forward rate curve: it works by assembling forward rates instrument by instrument, earlier maturities first. Rates are constant between maturities and jump at maturities.

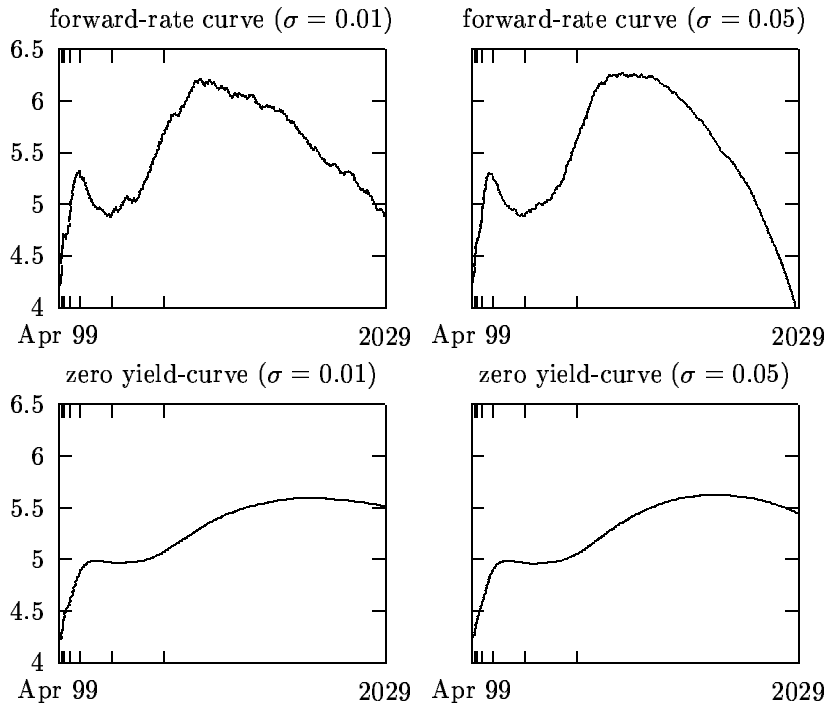


Figure 18: Forward-rate curve and zero-coupon-bond yield curve for scenario I ($\sigma = 0.01$) on the left side, and scenario II ($\sigma = 0.05$) on the right side

portfolios $(\theta_1, ...\theta_N)$ as follows:

$$U(\theta) \;=\; -\,\frac{1}{\nu}\sum_{i=1}^{\nu} e^{-\sum_{j=1}^{N}\theta_j\,(g_{ij}-C_j)}\;. \tag{45}$$

Since $\lambda^*$ minimizes $log(Z(\lambda)) - \lambda \cdot C$, it follows from the analysis of §3 that the vector of Lagrange multipliers and the optimal portfolio are in simple correspondence: we have

$$\lambda_j^* \;=\; -\,\theta_j\;,\; j = 1, \ldots N\;. \tag{46}$$

Sensitivities are the opposites of the optimal portfolio weights for (45). A negative lambda corresponds to a "cheap" instrument (hence $\theta > 0$) and a positive lambda to a "rich" instrument (hence $\theta < 0$). The Lagrange multipliers, or sensitivities, $\lambda_1^*, \ldots, \lambda_7^*$ for all four scenarios are summarized in Table 1. Making $m(t$ time-dependend does not change the sensitivities significantly (scenario I versus III and II versus IV). The Lagrange multipliers corresponding to short-term instruments, however, are very high if the prior volatility is low ($\sigma = 0.01$ in scenarios I and III).

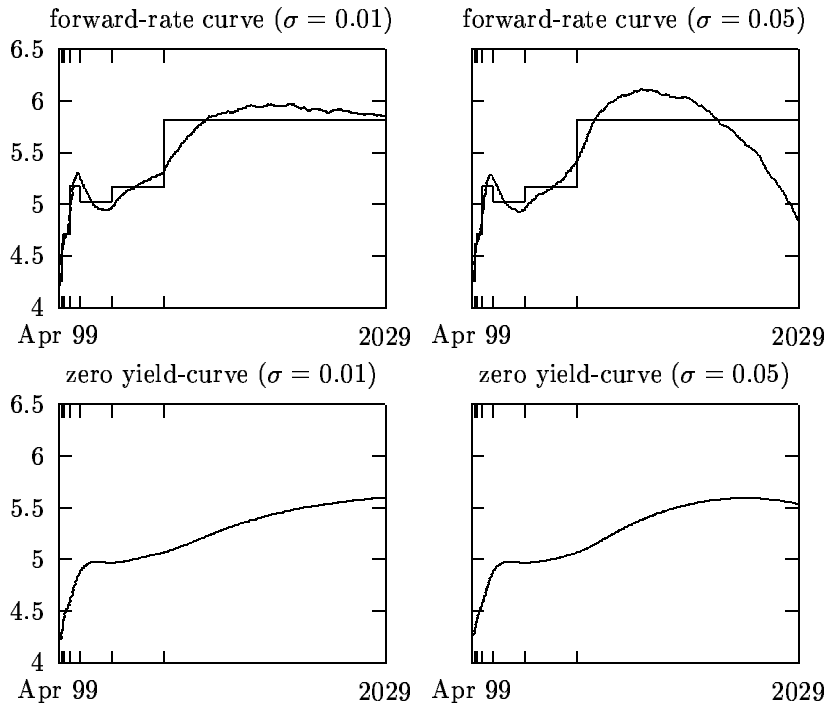


Figure 19: Forward-rate curve and zero-coupon-bond yield curve for scenario III ($\sigma = 0.01$) on the left side, and scenario IV ($\sigma = 0.05$) on the right side. Both scenarios revert to the piecewise constant bootstrapped prior $m(t)$ superimposed in the top row

This last application of the MRE algorithm has also been implemented by one of the authors (R. Buff) as a prototype software operating remotely via the Internet. The software, which uses periodically updated Treasury securities prices and/or bond prices entered by the users, is accessible in the Courant Finance Server (http://www.courantfinance.cims.nyu.edu).

# 10 Conclusions

We have presented a very simple approach for calibrating Monte Carlo simulations to the price of benchmark instruments. This approach is based on minimizing the Kullback-Leibler relative entropy between the posterior measure and a prior measure. In this context, the prior corresponds to the uniform measure over simulated paths (hence to the "classical" Monte Carlo simulation). This approach is known to be equivalent to finding the Arrow-Debreu prices which are consistent with an investor which maximizes an expected utility of exponential type. The advantage of the minimum-etropy algorithm is that (i) it is non-parametric (and thus not market

| | Sensitivities for scenario... | | | |
|---|---|---|---|---|
| Maturity | I | II | III | IV |
| 07/15/99 | **32.507** | 0.582 | **29.421** | 0.569 |
| 10/14/99 | **−17.638** | −0.242 | **−15.307** | −0.298 |
| 03/30/00 | **5.359** | 0.114 | **4.607** | 0.123 |
| 03/31/01 | **−1.667** | −0.056 | **−1.692** | −0.055 |
| 02/15/04 | 0.085 | 0.009 | 0.211 | 0.010 |
| 11/15/08 | 0.340 | 0.013 | 0.013 | 0.005 |
| 02/15/29 | −0.356 | −0.021 | −0.050 | −0.011 |

Table 1: The sensitivities for the seven benchmark instruments in Table 16, in each of the four prior scenarios. Sensitivities with absolute value greater than 1 are typeset in boldface

or model specific) and (ii) it allows the modeler to incorporate econometric information and *a-priori* information on the market dynamics, effectively separating the specification of the dynamics from the issue of price-fitting.

We showed that the algorithm can be implemented as an exact fit to prices or in the sense of least-squares. The notion of entropy distance can be interpreted as a measure of the logarithm of *effective number of paths* which are active in the posterior measure. Large entropy distances correspond therefore to "thin" supports and thus to an ortogonality (in the measure-theoretic sense) between the prior and posterior measures.

The sensitivities produced by the model can be computed via regression, without the need to simulate the market dynamics and to recalibrate each time that we perturb the price of a benchmark instrument. Another interesting feature of the weighted Monte Carlo algorithm is the reduction of variance which results from the exact pricing of benchmark instruments. In fact, the simulation effectively evaluates the "residual risk" obtained after projecting the payoff of interest onto the space of portfolios spanned by the benchmark instruments. Numerical experiments indicate that the reduction of variance can be significant.

We discussed concrete implementations of the algorithm for the case of foreign-exchange and equity options, calibrated the underlying dynamics to two-factor stochastic volatility models. We have exhibited numerical evidence that shows that such an algorithm can be implemented in practice on desktop computers.

# 11 References

Avellaneda M. (1998) Minimum-entropy Calibration of Asset-Pricing Models, *International Journal of Theoretical and Applied Finance* 1(4) 447

– (1997) C. Friedman, R. Holmes, D. Samperi Calibrating Volatility Surfaces via

Relative-Entropy Minimization, *Applied Mathematical Finance*, March, 4 (1) 37-64

Avellaneda, M. and A. Paras (1996) Managing the volatility risk of portfolios of derivative securities: The Lagrangian Uncertain Volatility Model. *Applied Mathematical Finance* 3, 21-52

Buchen P. W., M. Kelly The Maximum Entropy Distribution of an Asset Inferred from Option Prices (1996) *Journal of Financial and Quantitative Analysis*, vol. 31, n. 1, March, p. 143 159.

Chriss, N. (1996), Transatlantic Trees, *RISK*, 9, 7.

Derman, E. and Kani, I.,(1994) Riding on a Smile, *RISK*, 7, 2.

Dupire, B.(1994), Pricing with a Smile, *RISK*, 7, 1.

Dupire, B.(1998), Monte Carlo Methodlogies and Applications for Pricing and Risk Management , RISK Publications. London

Cover T. M., J. A. Thomas (1991)*Elements of Information Theory* , Wiley, New York

Gulko L. (1995) The Entropy Theory of Option Pricing, Yale University Working Paper.

– (1996) The Entropy Theory of Bond Pricing, Yale University Working Paper.

Jackwerth, J.C. and Rubinstein, M., (1995) Recovering Probability Distributions from Contemporaneous Security Prices, Berkley University Hass School of Business Working Paper.

Laurent, J.P. and D. Leisen (1999), Building a Consistent Pricing Model from Observed Option Prices, Working Paper, the Hoover Institute, Stanford University

Samperi, D. (1997) *Inverse Problems, Model Selection and Entropy in Derivative Security Pricing*, Ph. D. Thesis, New York University

Rubinstein, M. (1994) , Implied Binomial Trees, *The Journal of Finance*, July, v 69, n 3,771-818

Rubinstein, M. and E. Reiner (1991) , Breakind Down the Barriers, *RISK*, July, 4 (8), 28-35

Zhu, C., Boyd, R.H., Lu, P. and Nocedal, J. (1994),*L-BFGS-B: FORTRAN Subroutines for Large-Scale Bound-Constrained Optimization*, Northwestern University, Department of Electrical Engineering

Zhu, Y. and M. Avellaneda (1998), A risk-neutral stochastic volatility model *International Journal of Theoretical and Applied Finance*, V. 1 (2), 289