Numerical methods for optimal hedging portfolios

M. R. Grasselli Dept. of Mathematics and Statistics
McMaster University
Sharcnet AGM - University of Waterloo

June 24, 2004

1. Introduction

- Market Model: We consider an n−factor Markovian market with state variables $(S^1, \ldots, S^d, Y^1, \ldots, Y^{n-d})$ where S_t is the $\mathbb{R}^d$−valued process which describes the discounted prices of traded assets and Y_t is the $\mathbb{R}^{n-d}$−valued process corresponding to the values of nontraded quantities such as stochastic volatilities which may or may not be observed directly.

  For example, we treat a two factor stochastic volatility model

$$\begin{aligned} dS_t &= S_t[(\mu(t, Y_t) - r)dt + \sigma(t, Y_t)dW_t^1] \\ dY_t &= a(t, Y_t)dt + b(t, Y_t)[\rho dW_t^1 + \sqrt{1 - \rho^2}dW_t^2] \end{aligned} \qquad (1)$$

  with initial values $S_0, Y_0 \geq 0$, for deterministic functions μ, a, b and independent one dimensional P−Brownian motions W_t^1 and W_t^2 with constant correlation $|\rho| \leq 1$.

- Optimal hedging portfolio: the strategy followed by an investor who, when faced with a (discounted) financial liability B maturing at a future time T, tries to solve the stochastic control problem

$$u(x) = \sup_{H \in \mathcal{A}} E\left[U\left(X_T - B\right)\right], \tag{2}$$

where $X_t = x + (H \cdot S)_t$ is the discounted terminal wealth obtained when investing according to the self financing portfolio $H_t = (H_t^1, \ldots, H_t^d)$ and the (discounted) liability B is assumed to be a random variable of the form $B = B(S_T, Y_T)$.

- Utility function: $U(x) = -\frac{e^{-\gamma x}}{\gamma}$, where $\gamma > 0$ is the risk aversion parameter.

2. Utility based pricing

For such Markovian markets we can embed the optimal hedging problem (2) into the larger class of optimization problems defined by

$$u(t, x, s, y) = \sup_{H \in \mathcal{A}_t} E_{t,s,y}[U(X_T - B(S_T, Y_T))|X_t = x], \qquad (3)$$

for $t \in (0, T)$, where $x \in \mathbb{R}$ denotes some arbitrary level of wealth, $\mathcal{A}_t$ denotes admissible portfolios starting at time t and $E_{t,s,y}[\cdot]$ denotes expectation with respect to the joint probability law at time t of the processes S_u, Y_u, for $u \geq t$, with initial condition $S_t = s$ and $Y_t = y$.

Suppose that (3) has an optimizer H_t^B, that is, assume that

$$u(t,x,s,y) = E_{t,s,y}[U(x + (H^B \cdot S)_t^T - B(S_T, Y_T))],$$

Define the certainty equivalent for the claim B at time t as the process $c_t^B = c^B(t,x,s,y)$ satisfying the equation

$$U(x - c_t^B)) = E_{t,s,y}[U(x + (H^B \cdot S)_t^T - B(S_T, Y_T))]. \qquad (4)$$

If we set $B = 0$, then the optimal hedging problem becomes the Merton optimal investment problem and we denote the certainty equivalent by $c_t^0 = c^0(t,x,s,y)$.

The indifference price for the claim B is defined to be solution $\pi^B = \pi^B(t, x, s, y)$ to the equation

$$\sup_{H \in \mathcal{A}_t} E_{t,s,y}[U(x+(H \cdot S)_t^T] = \sup_{H \in \mathcal{A}_t} E_{t,s,y}[U(x+\pi^B+(H \cdot S)_t^T - B(S_T, Y_T)]. \tag{5}$$

From the definition of the certainty equivalent, we see that this equation is equivalent to

$$U(x - c_t^0) = U(x + \pi^B - c_t^B), \tag{6}$$

so that the indifference price is given by

$$\pi^B(t, x, s, y) = c^B(t, x, s, y) - c^0(t, x, s, y). \tag{7}$$

3. Discrete time hedging

We now consider portfolio processes of the form

$$H_t = \sum_{k=1}^{K} H_k \, \mathbf{1}_{(t_{k-1}, t_k]}(t) \tag{8}$$

where each H_k is an $\mathbb{R}^d$–valued $\mathcal{F}_{k-1}$ random variable. We take the discrete time partition of the interval $[0, T]$ to be of the form

$$t_0 = 0 < t_1 = \frac{T}{K} < \ldots < t_k = \frac{kT}{K} \ldots < t_K = T$$

and use the notation $S_j := S_{t_j}$ for discrete time stochastic processes.

The discounted wealth for self–financing portfolios is

$$X_j = x + (H \cdot S)_j, \tag{9}$$

with the notation $(H \cdot S)_k^j := (H \cdot S)_j - (H \cdot S)_k$, where

$$(H \cdot S)_j := \sum_{k=1}^{j} H_k \Delta S_k \tag{10}$$

and $\Delta S_k := S_k - S_{k-1}$.

Now the dynamic programming problem for the optimal hedge falls into K subproblems

$$u_{k-1}(x) = \sup_{H_k \in \mathcal{F}_{k-1}} E_{k-1}[u_k(x + H_k \Delta S_k)], \tag{11}$$

for $k = K, K-1, \ldots, 1$, with $u_K(x) = U(x - B)$. Similarly, the certainty equivalent value process $c_k^B(x)$ is defined iteratively by

$$U(x - c_{k-1}^B(x)) = \sup_{H_k \in \mathcal{F}_{k-1}} E_{k-1}[U(x + H_k \Delta S_k - c_k^B(x + H_k \Delta S_k)] \tag{12}$$

with $c_K^B(x)$ taken equal to the terminal discounted claim B.

In our Markovian setting and with an exponential utility, the solution of (11) and (12) as well as the optimal allocation H^B have the form wealth independent form

$$u_k = g_k(S_k, Y_k) \tag{13}$$
$$c_k^B = c_k(S_k, Y_k) \tag{14}$$
$$H_{k+1}^B = h_{k+1}(S_k, Y_k) \tag{15}$$

for (deterministic) Borel scalar functions $\{g_k, c_k\}_{k=0}^{K-1}$ and $\mathbb{R}^d$–valued functions $\{h_{k+1}\}_{k=0}^{K-1}$ on the state space $\mathcal{S}$.

4. The exponential utility allocation algorithm

We want an algorithm which will generate an approximate trading rule, based on a data set

$$\{(S^i_k, Y^i_k)\}_{i=1,\dots,N;k=0,\dots,K}$$

where $(S^i_k, Y^i_k) \in \mathbb{R}^n$ denotes the state of the ith sample path at time $t_k = kT/K$ for the processes (S_t, Y_t). In the special case of an exponential utility, the theoretical optimal rule

$$H^B_{k+1} = h_k(S^i_k, Y^i_k)$$

in (15) depends only on the directly observed data $\{S^i_k, Y^i_k\}$ and is independent of the wealth X^i_k. For this reason our algorithm is at this point restricted to exponential utility functions, and we take $\gamma = 1$ for simplicity.

1. Step $k = K$: The final optimal allocation is the $\mathcal{F}_{K-1}$–random variable H_K^B which solves

$$\min_{H_K \in \mathcal{F}_{K-1}} E[\exp(-H \cdot \Delta S_K + B)]. \tag{16}$$

Since the solution is known to be given by $H_K^B = h_K(S_{K-1}, Y_{K-1})$ for some deterministic function $h_K \in \mathcal{B}(\mathcal{S})$ (the set of Borel functions on $\mathcal{S}$), we write this as

$$\min_{h \in \mathcal{B}(\mathcal{S})} E[\exp(-h(S_{K-1}, Y_{K-1}) \cdot \Delta S_K + B)]. \tag{17}$$

On a finite set of data, we can pick an R–dimensional subspace $\mathcal{R}(\mathcal{S}) \subset \mathcal{B}(\mathcal{S})$ of functions on $\mathcal{S}$ and attempt to "learn" a suboptimal solution

$$\arg\min_{h \in \mathcal{R}(\mathcal{S})} E[\exp(-h(S_{K-1}, Y_{K-1}) \cdot \Delta S_K + B)].$$

By the central limit theorem, the expectation above can be approximated by the finite sample estimate

$$\Psi_K(h) = \frac{1}{N}\sum_{i=1}^{N} \exp\Big(-h(S^i_{K-1}, Y^i_{K-1}) \cdot \Delta S^i_K + B(S^i_K, Y^i_K)\Big) \tag{18}$$

This leads to the estimator $h^{\mathcal{R}}_K$ based on $\{S^i_k, Y^i_k\}$ and the choice of subspace $\mathcal{R}$ defined by

$$h^{\mathcal{R}}_K = \operatorname*{arg\,min}_{h \in \mathcal{R}(\mathcal{S})} \Psi_K(h) \tag{19}$$

2. Inductive step for $k = K-1, \ldots, 2$: The estimate $h_k^{\mathcal{R}}$ of the optimal rule h_k, for the intermediate time steps $2 \leq k < K-1$ is determined inductively given the estimates $h_{k+1}^{\mathcal{R}}, \ldots, h_K^{\mathcal{R}}$. It is defined to be

$$h_k^{\mathcal{R}} = \operatorname*{arg\,min}_{h \in \mathcal{R}(\mathcal{S})} \Psi_k(h; h_{k+1}^{\mathcal{R}}, \ldots, h_K^{\mathcal{R}}) \tag{20}$$

where

$$\Psi_k(h) = \frac{1}{N} \sum_{i=1}^{N} \exp\left(-h(S_k^i, Y_k^i) \cdot \Delta S_{k+1}^i + c_k^i(h_{k+1}^{\mathcal{R}}, \ldots, h_K^{\mathcal{R}}, S_K^i, Y_K^i)\right), \tag{21}$$

with

$$c_k^i(h_{k+1}^{\mathcal{R}}, \ldots, h_K^{\mathcal{R}}, S_K^i, Y_K^i) = B(S_K^i, Y_K^i) - \sum_{j=k+1}^{K} h_j^{\mathcal{R}}(S_{j-1}^i, Y_{j-1}^i) \cdot \Delta S_j^i \tag{22}$$

3. Final step $k = 1$: This step is degenerate since the initial values (S_0, Y_0) are constant over the sample. Therefore we determine the optimal constant vector $h_1 \in \mathbb{R}^d$ by solving

$$h_1 = \underset{h \in \mathbb{R}^d}{\arg\min}\, \Psi_1(h; h_2^{\mathcal{R}}, \ldots, h_K^{\mathcal{R}}) \tag{23}$$

Finally, the optimal value

$$\Psi_1 = \frac{1}{N} \sum_{i=1}^{N} \exp\left(-h_1(S_0, Y_0) - \sum_{j=2}^{K} h_j^{\mathcal{R}}(S_{j-1}^i, Y_{j-1}^i) \cdot \Delta S_j^i + B(S_K^i, Y_K^i) \right),$$

is an estimate of the quantity $\exp(c_0^B)$, where c_0^B is the certainty equivalent value of the claim B at time $t = 0$

5. Numerical results

Geometric Brownian motion

We start with a one dimensional complete market in order to test the algorithm against well known exact solutions. Consider a market where the stock price process, discounted by the constant interest rate r, satisfy

$$\frac{dS_t}{S_t} = (\mu - r)dt + \sigma dW, \tag{24}$$

where μ and $\sigma > 0$ are constants and W is a one–dimensional P–Brownian motion.

As it is well known, the unique equivalent martingale measure Q has desity dQ/dP given by the stochastic exponential of the constant market price of risk $\lambda = (\mu - r)/\sigma$ and the Merton portfolio for this market is given by

$$\widehat{H}_t = \frac{\mu - r}{\gamma\sigma^2}\frac{1}{S_t}. \tag{25}$$

We can now compare the hedging portfolio "learned" by our algorithm with the "true" optimal hedging portfolio given

$$H_t^B = \widehat{H}_t + \mathcal{H}_t^B, \tag{26}$$

where $\mathcal{H}_t^B$ is the Black–Scholes *delta hedging* portfolio replicating B. Similarly, the indifference prices calculated by the algorithm can be compared with the Black–Scholes price for the same claim.

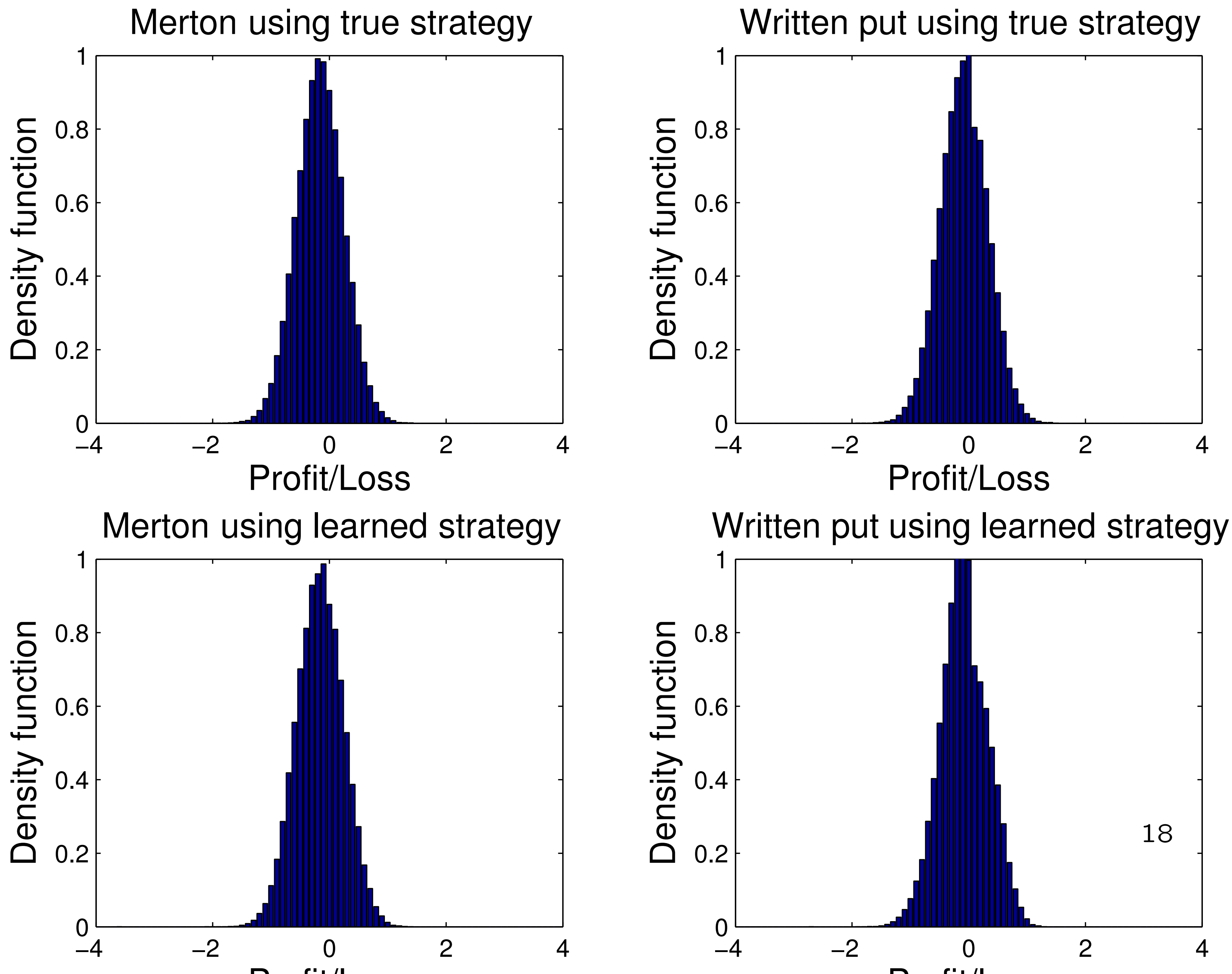
We fix the parameters of the model at

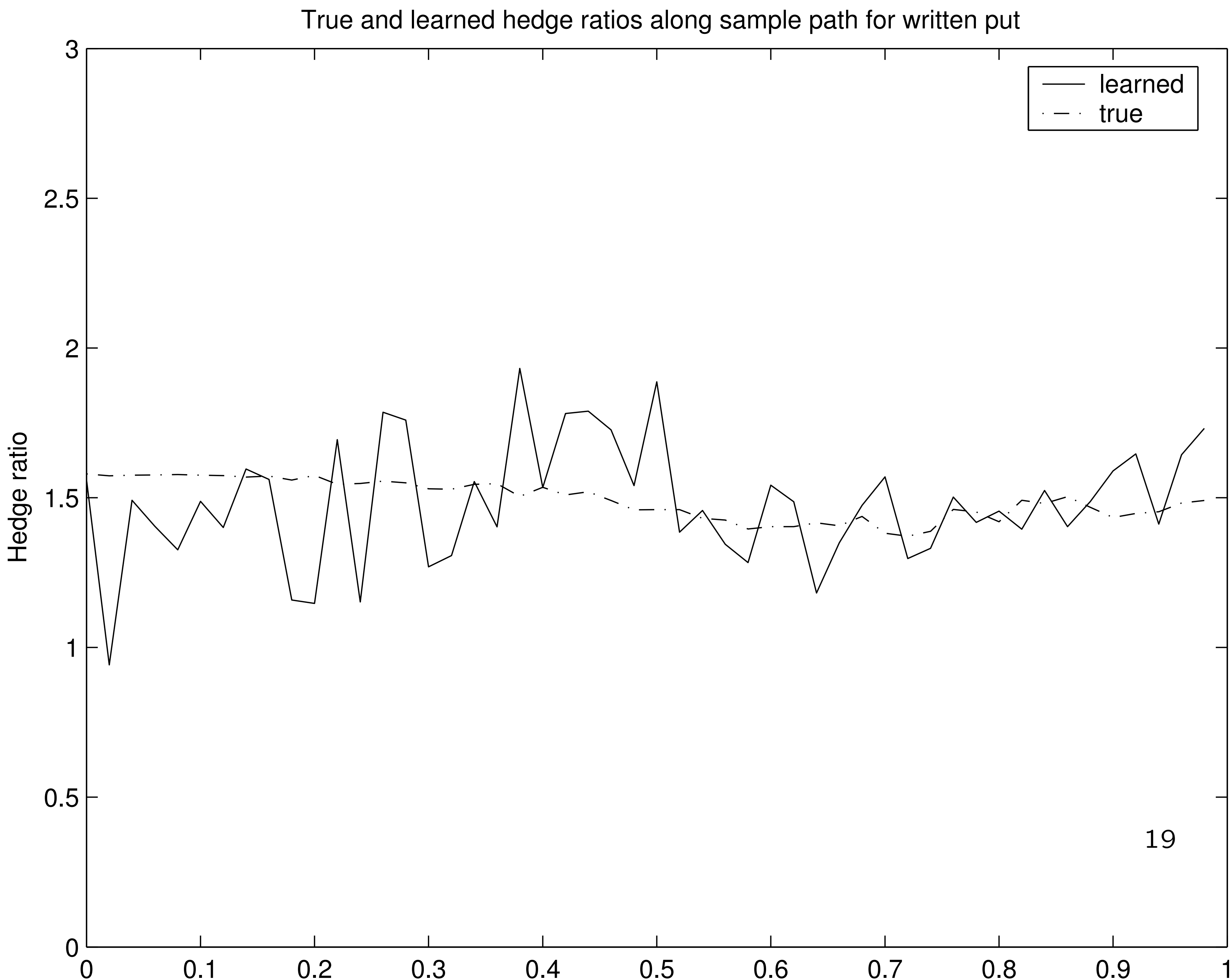
$$S_0 = 1, \quad \mu = 0.1, \quad \sigma = 0.2 \quad \text{and } r = 0.02$$

over the period of one year $T = 1$ and discrete time intervals of $1/50$.

We apply the allocation algorithm with $N = 100000$ to two scenarios: (i) the Merton investment problem; and (ii) the hedging problem for the *writer* of a single written at–the–money European put. Then, for comparison to theory, we use the same Monte Carlo simulations, but rehedged weekly according to the theoretical formula (26).

As for the subspace $\mathcal{R}(\mathcal{S})$, we use the three dimensional space spanned by the functions $\{1, s, s^2\}$.


Merton using true strategy
Written put using true strategy
Merton using learned strategy
Written put using learned strategy
Density function
Profit/Loss
1
0.8
0.6
0.4
0.2
0
–4
–2
0
2
4


True and learned hedge ratios along sample path for written put
Hedge ratio
learned
true
0
0.5
1
1.5
2
2.5
3
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

Reciprocal affine stochastic volatility models

We now take μ and r to be constants and $\sigma(t, Y_t) = \sqrt{Y_t}$ in (1). Define

$$R_t = R(t, Y_t) = \frac{(1-\rho^2)(\mu - r)^2}{2Y_t}, \tag{27}$$

which we postulate to be a CIR process, that is

$$dR_t = \alpha(\kappa - R_t)dt + \beta\sqrt{R_t}\left[\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2\right], \tag{28}$$

for constants $\alpha, \kappa, \beta > 0$ with $4\alpha\kappa > \beta^2$.

We then obtain from the Itô formula that

$$a(t, Y_t) = \alpha Y_t + \frac{2(\beta^2 - \alpha\kappa)}{(1-\rho^2)(\mu - r)^2} Y_t^2, \tag{29}$$

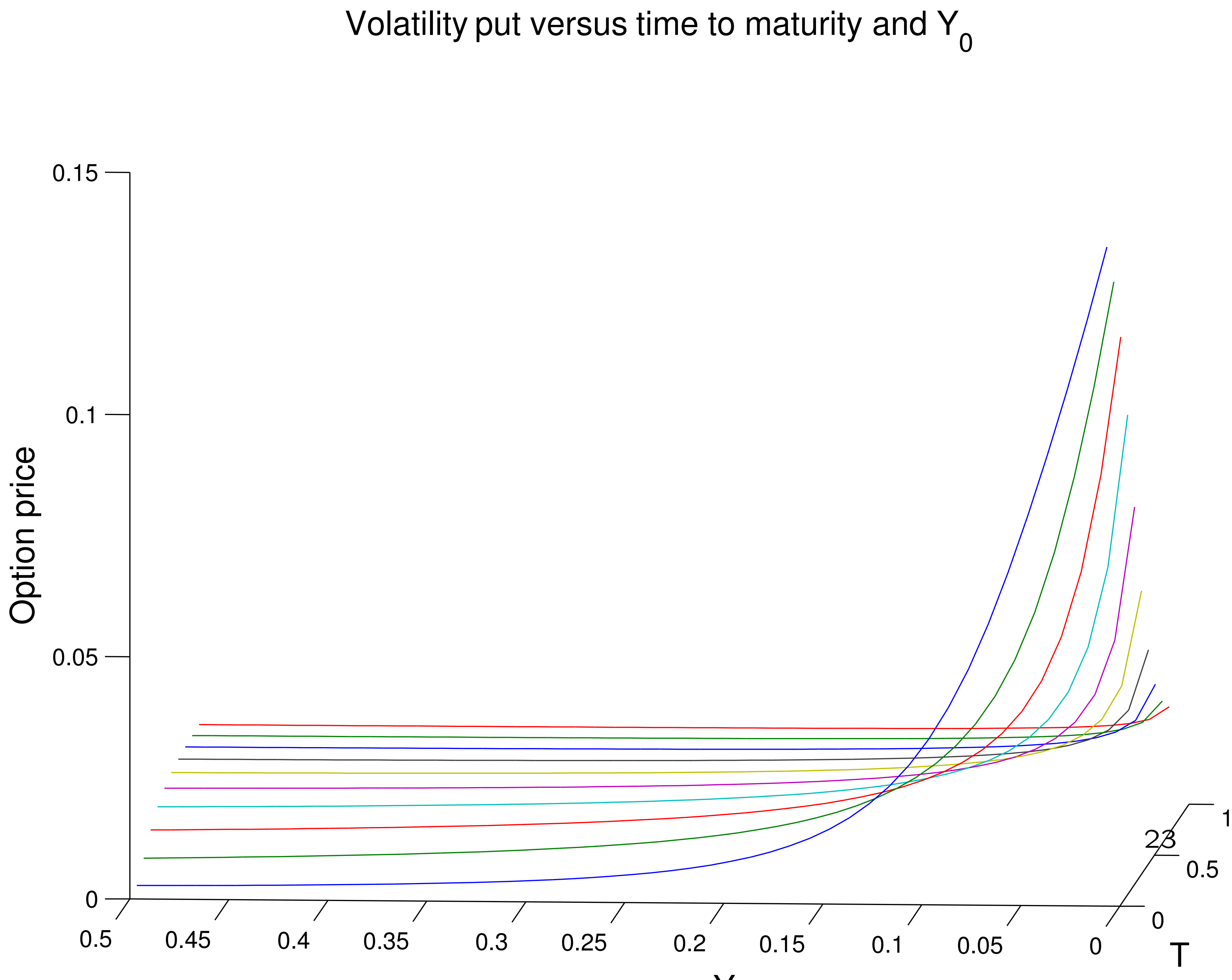
$$b(t, Y_t) = -\left(\frac{2}{1-\rho^2}\right)^{1/2} \frac{\beta}{(\mu - r)} Y_t^{3/2}. \tag{30}$$

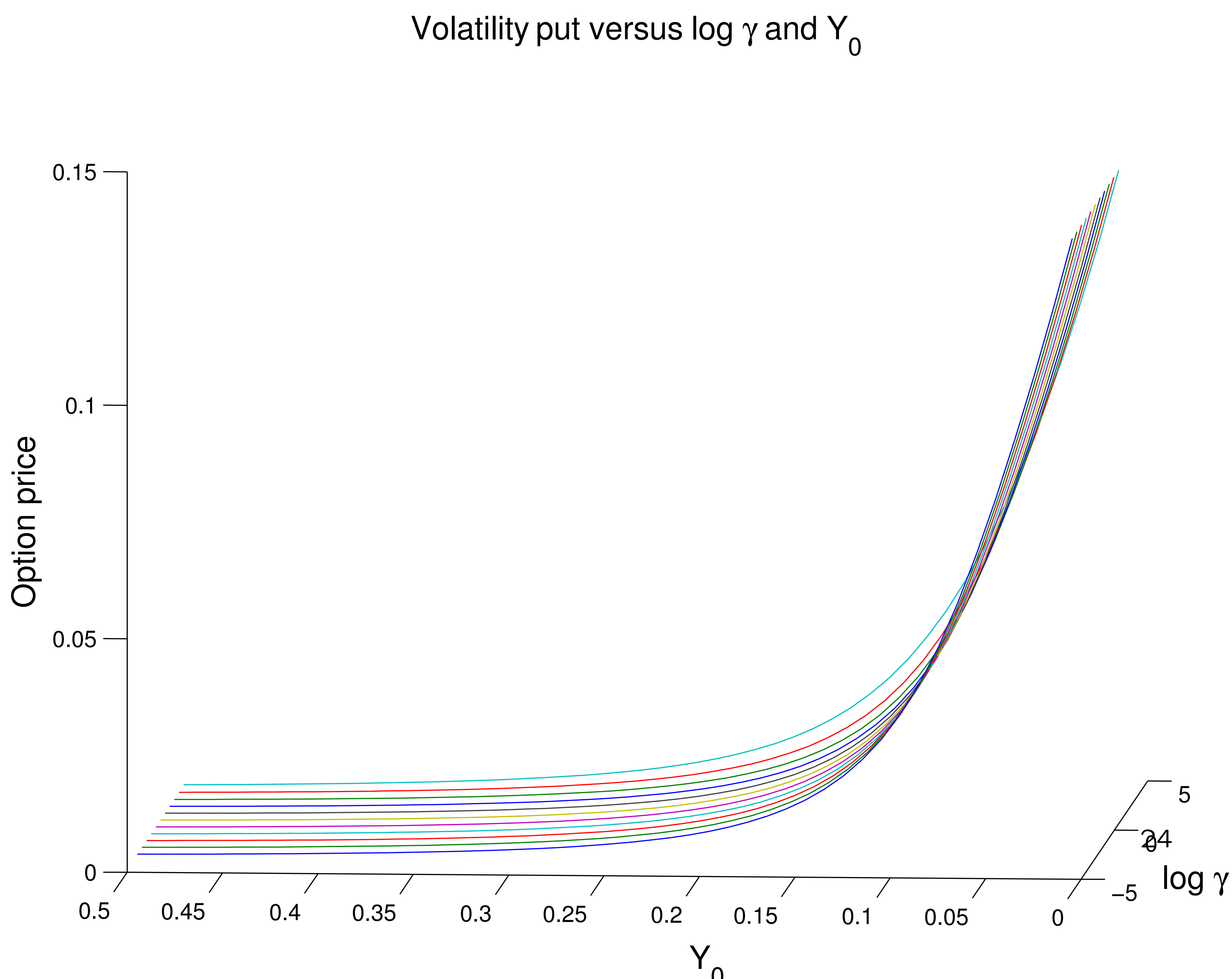
It follows from the *Hamilton–Jacobi–Bellman* equations associated with (3), that indifference prices and optimal hedging portfolios for pure volatility claims of the form $B = B(Y_T)$ can be explicitly computed using a Fourier trasnform technique. We use these as benchmarks for our more general Monte Carlo algorithm.

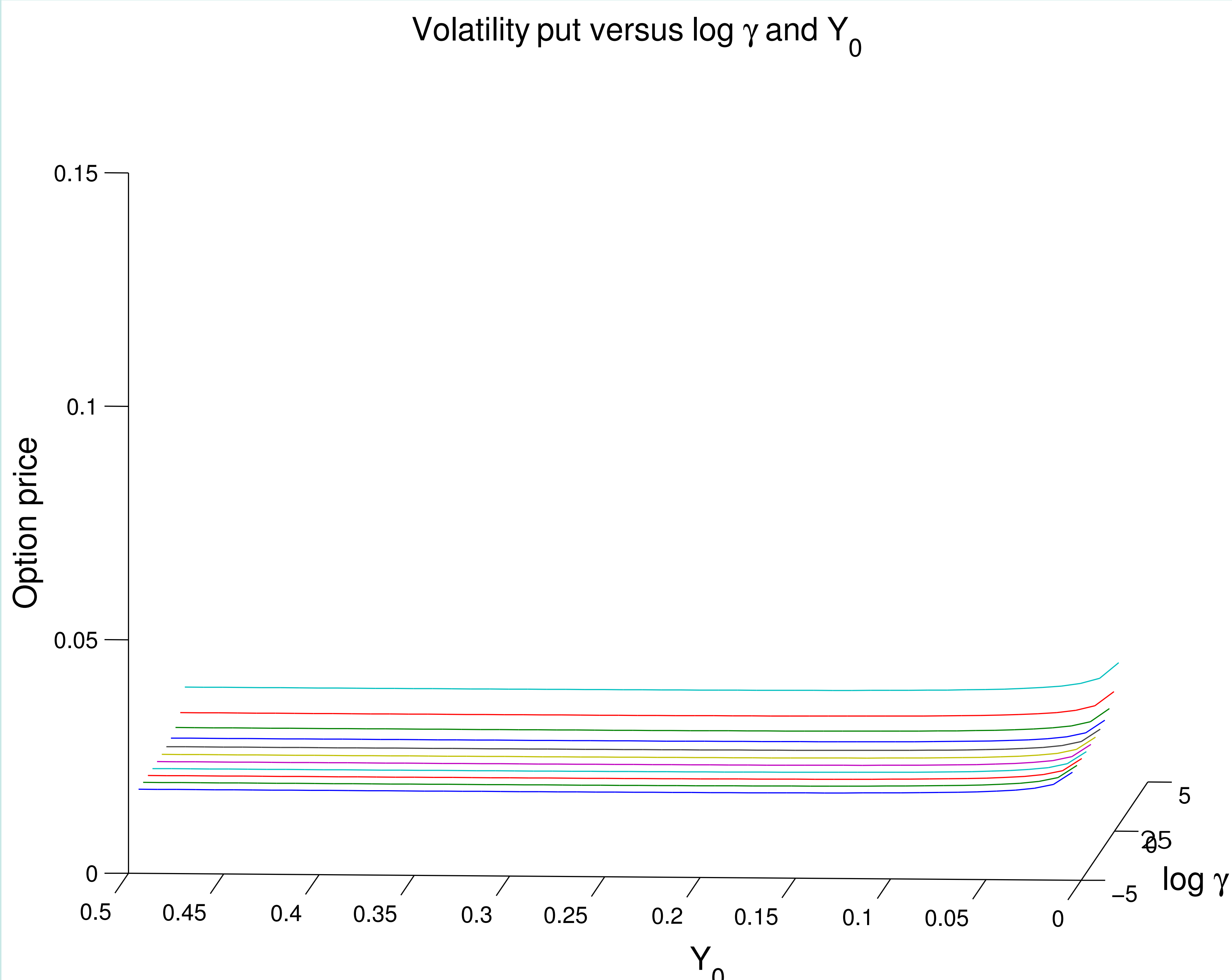
We fix the model parameters at reasonable values:

$$\begin{aligned} \alpha &= 5, \quad \beta = 0.04, \quad \kappa = 0.001, \\ \mu &= 0.04, \quad r = 0.02, \quad \rho = 0.5 \end{aligned}$$

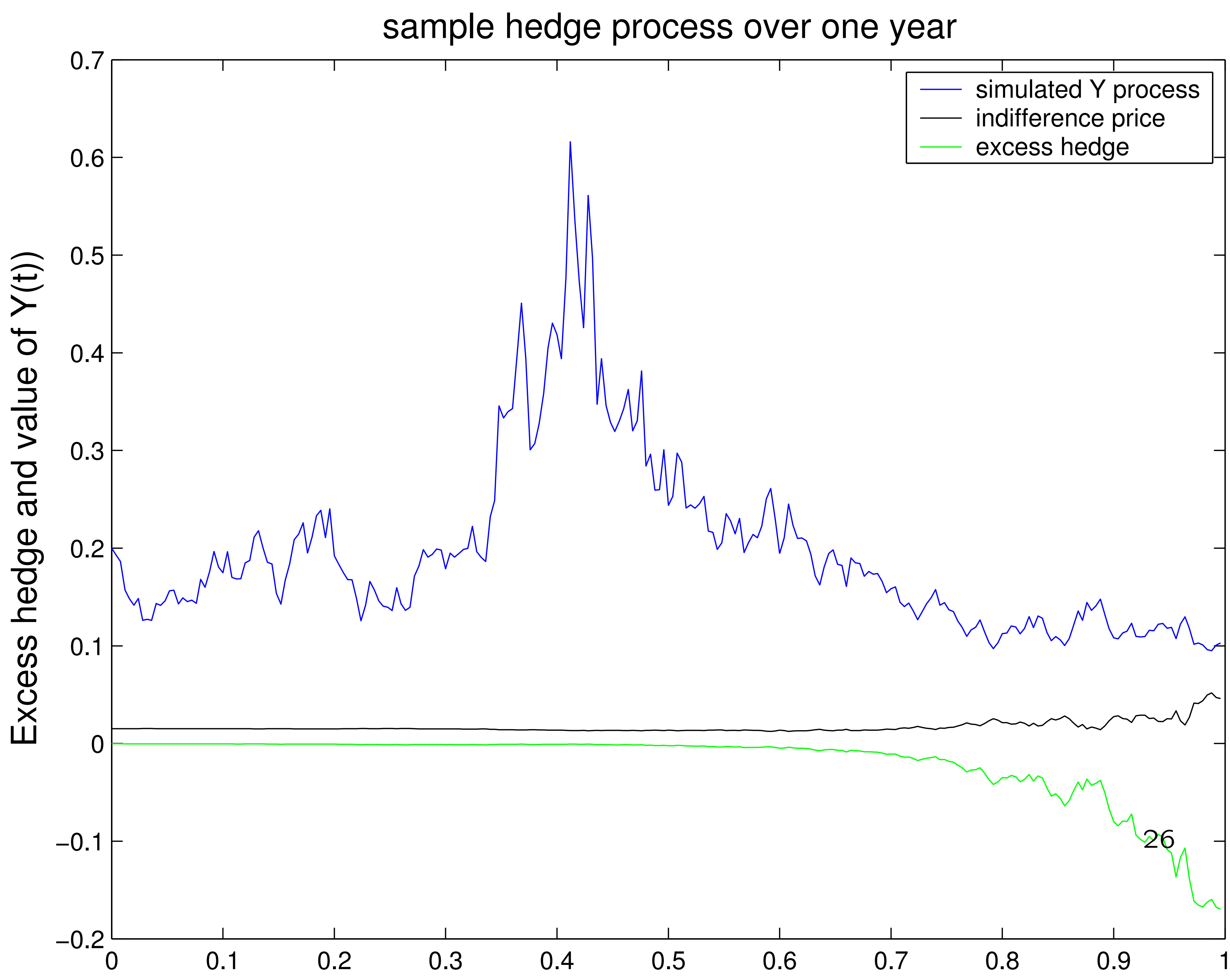
and initial squared volatility ranging in the interval $[0, 0.5]$. With these parameters the squared volatility process has a mean reversion time of approximately two months and an equilibrium distribution with expected value approximately 40%. We calculate the price of a put option on volatility with payoff $(0.15 - \sigma_T^2)^+$. When not mentioned the risk aversion parameter is set to $\gamma = 1$.


Volatility put versus time to maturity and Y_0
Option price
0.15
0.1
0.05
0
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
T
1
0.5
0
23


Volatility put versus log γ and Y0
Option price
0.15
0.1
0.05
0
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Y0
log γ
5
−5



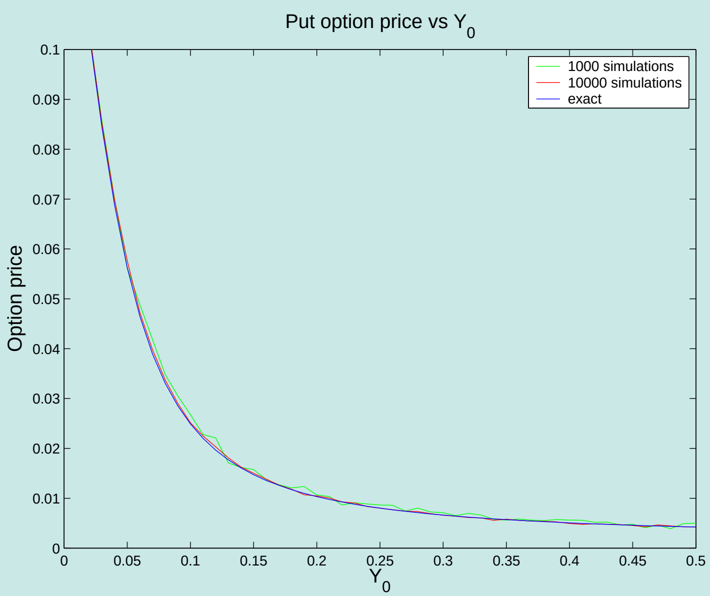
Volatility put versus log γ and Y_0
Option price
0.15
0.1
0.05
0
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Y_0
5
2.5
0
−5
log γ



sample hedge process over one year
Excess hedge and value of Y(t))
simulated Y process
indifference price
excess hedge
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

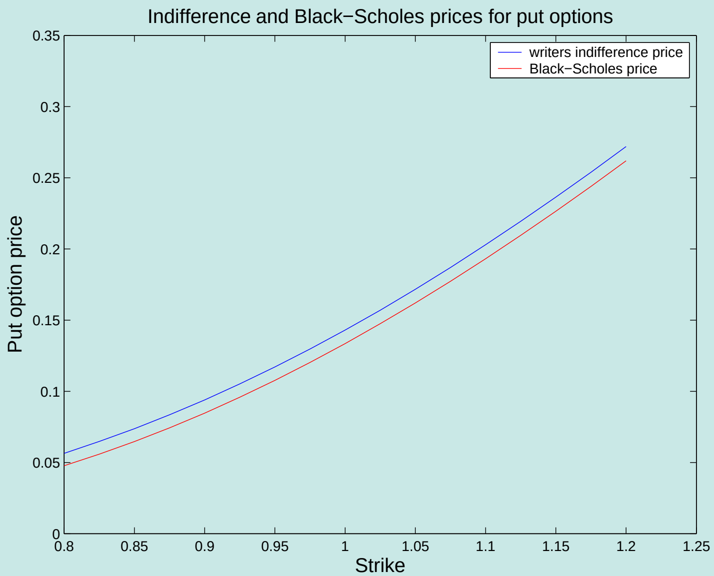
We now run the algorithm with the same model parameters as before (in particular $\gamma = 1$). To account for the portfolio dependence in both S_t and Y_t we took $\mathcal{R}(\mathcal{S})$ to be the six–dimensional space spanned by the functions $\{1, y, y^2, s, sy, s^2\}$.

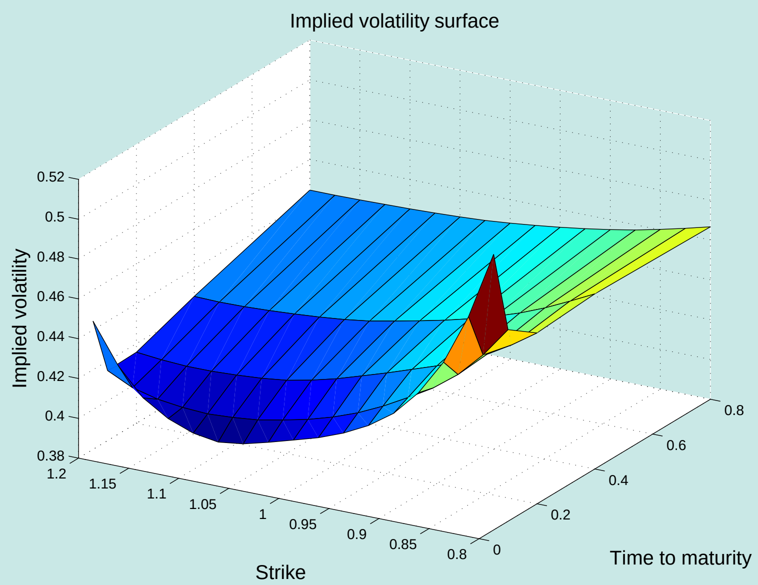
We first applied the allocation algorithm to a volatility put option with payoff $(0.15 - \sigma_T^2)^+$ and time to maturity at $T = 0.2$ and computed the indifference prices with Y_0 varying in the interval $[0, 0.5]$.



Put option price vs Y_0
Option price
Y_0
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
1000 simulations
10000 simulations
exact

Next we consider a put option on the stock, that is, with payoff $(K - S_T)^+$. The following pictures show the indifference prices and implied volatility surface with $N = 10000$ simulations.


Indifference and Black−Scholes prices for put options
writers indifference price
Black−Scholes price
Put option price
Strike
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.8
0.85
0.9
0.95
1
1.05
1.1
1.15
1.2
1.25


Implied volatility surface
Implied volatility
0.52
0.5
0.48
0.46
0.44
0.42
0.4
0.38
1.2
1.15
1.1
1.05
1
0.95
0.9
0.85
0.8
Strike
0
0.2
0.4
0.6
0.8
Time to maturity