

Accelerated Simulation for Pricing Asian Options

Felisa J. Vázquez-Abad

Department of Computer Science and Operations Research
University of Montreal, Montréal, Québec H3C 3J7
e-mail: vazquez@iro.umontreal.ca

and

Daniel Dufresne

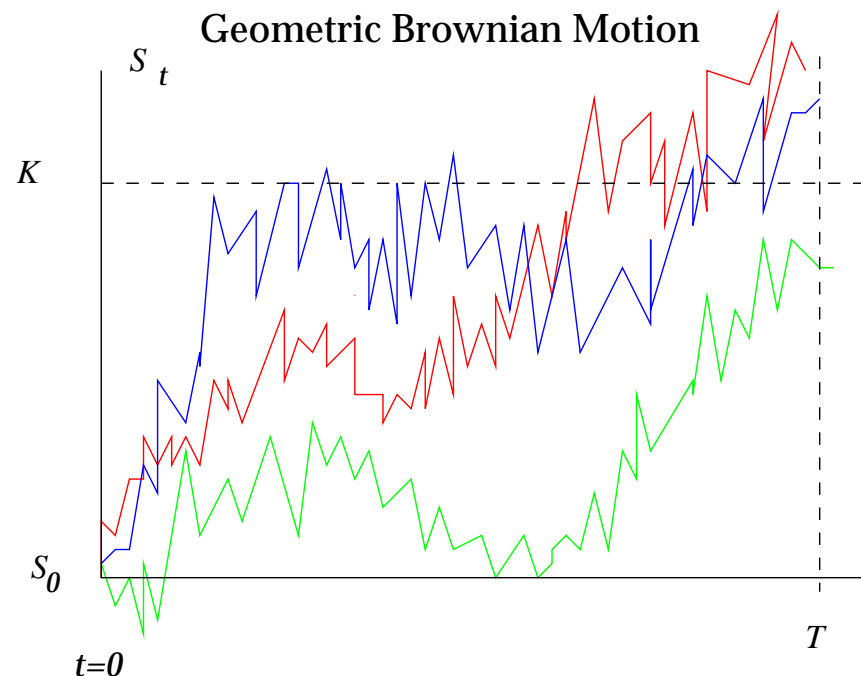
Department of Economics
University of Melbourne
e-mail: dufresne@myriad.its.unimelb.edu.au

Invited paper to the *Winter Simulation Conference*, Washington, D.C., December 1998.

INTRODUCTION

Two assets: a risk-free asset yielding interest r , and a risky asset, with price $S_t = S_0 e^{\mu t + \sigma W_t}$.

- **European Call Option:** Option to buy one unit of stock at strike price K at the exercise time T . Only if $S_T > K$ the holder of the option will exercise it. Profit: $(S_T - K)_+$.
- **European Average Call Option:** Instead of the terminal value S_T , the average is compared to K .



Asset Pricing: Under the *risk-neutral* measure \mathbf{P} , discounted asset prices $\{e^{-rt}S_t\}$ form a martingale. The price π is the expectation w.r.t. \mathbf{P} of the discounted gain. Discrete model:

$$S_{i+1} = S_i \exp \left[\left(r + \frac{\sigma^2}{2} \right) h + \sigma \sqrt{h} Z_i \right], \quad Z_i \text{ i.i.d. } \sim \mathcal{N}(0, 1), \quad h = \frac{T}{N}$$

$$\pi = \mathbf{E} \left[e^{-rT} \left(\frac{1}{N} \sum_{i=1}^N S_i - K \right)_+ \right] \quad \text{If } K \gg S_0 \text{ out of the money } \Rightarrow \textit{rare event estimation.}$$

SIMULATION METHODS I

$$\begin{aligned}
Z_i &\sim \mathcal{N}(0, 1), \quad i = 1, \dots, N, \text{ are independent standard normal variables.} \\
X_i^u &= u_\sigma h + \sigma \sqrt{h} Z_i, \quad 1 \leq i \leq N, \text{ where } u_\sigma = u - \frac{\sigma^2}{2} \Rightarrow X_i^u \sim \mathcal{N}(u_\sigma h, \sigma^2 h), \\
S_i^u &= S_{i-1}^u \exp(X_i^u), \quad 1 \leq i \leq N, \quad \text{are the asset prices,} \\
A_u &= \frac{1}{N} \sum_{i=1}^N S_i^u \text{ is the arithmetic average of asset prices,}
\end{aligned}$$

Naïve Estimation: Take $u = r$ above, and use

$$D_0 = Y_1^r = e^{-rT} (A_r - K)_+.$$

Confidence Intervals: Approximate Confidence Interval: apply the Central Limit Theorem to M replications of the simulation. Use a confidence level $\alpha = 0.05$.

$$\hat{\pi} \pm 1.96 \frac{\sqrt{\widehat{\text{Var}}[D_0]}}{M}$$

The *precision* of the estimation can be reduced if we use other estimators with smaller variance than the naïve estimator. We shall discuss the methods:

- CONTROL VARIABLE
- CHANGE OF MEASURE
- HYBRID ESTIMATION

SIMULATION METHODS II

$$\begin{aligned}
Z_i &\sim \mathcal{N}(0, 1), \quad i = 1, \dots, N, \quad X_i^u = u_\sigma h + \sigma \sqrt{h} Z_i, \\
S_i^u &= S_{i-1}^u \exp(X_i^u), \quad 1 \leq i \leq N, \\
A_u &= \frac{1}{N} \sum_{i=1}^N S_i^u \text{ is the arithmetic average of asset prices,} \\
G_u &= \left(\prod_{i=1}^N S_i^u \right)^{\frac{1}{N}} \text{ is the geometric average of asset prices.}
\end{aligned}$$

The Control Variable Method: Take $u = r$, and let Y_2^r be some other estimator Y_2^r whose expected value is known. Use:

$$D_1 = Y_1^r + \alpha(\mathbf{E} Y_2^r - Y_2^r), \quad \text{which implies that } \mathbf{E} D_1 = \pi$$

It can be shown that $\alpha = \mathbf{Cov}(Y_1^r, Y_2^r) / \mathbf{Var} Y_2^r$ minimizes the variance. Use as a control variable the option price when the *geometric mean* is used in the average, that is:

$$Y_2^r = e^{-rT} (G_r - K)_+.$$

Mean and variance of Y_2^r are known. Variance reduction:

$$\mathbf{Var}[D_1] = \mathbf{Var}[D_0] \left(1 - \frac{\mathbf{Cov}(Y_1^r, Y_2^r)}{\sqrt{\mathbf{Var}(Y_1^r) \mathbf{Var}(Y_2^r)}} \right) \leq \mathbf{Var}[D_1]$$

SIMULATION METHODS III

$$Z_i \sim \mathcal{N}(0, 1), \quad i = 1, \dots, N, \quad B_i = Z_1 + \dots + Z_i, \quad 1 \leq i \leq N, \quad B_0 = 0$$

$$X_i^u = u_\sigma h + \sigma \sqrt{h} Z_i, \quad S_i^u = S_{i-1}^u \exp(X_i^u), \quad 1 \leq i \leq N, \quad A_u = \frac{1}{N} \sum_{i=1}^N S_i^u.$$

The Change of Measure Approach: Girsanov's Theorem for one-dimensional Brownian Motion implies for the discrete model:

$$\forall v \in \mathbb{R}, \quad \mathbf{E} f(Z_1, \dots, Z_N) = \mathbf{E} e^{-\frac{Nv^2}{2} - vB_N} f(Z_1 + v, \dots, Z_N + v).$$

Let $v = (u - r)\sqrt{h}/\sigma$, then for

$$f(Z_1, \dots, Z_N) = A_r = \frac{1}{N} \sum_{i=1}^N S_{i-1}^r e^{X_i^r}, \quad f(Z_1 + v, \dots, Z_N + v) = \frac{1}{N} \sum_{i=1}^N S_{i-1}^u e^{X_i^u} = A_u$$

Call $L_u = e^{-\frac{Nv^2}{2} - vB_N} = \exp \left\{ -\frac{N}{2} \left[\frac{(u - r)\sqrt{h}}{\sigma} \right]^2 - \frac{(u - r)\sqrt{h}}{\sigma} B_N \right\}$, then it follows that:

$$D_2(u) = L_u Y_1^u = L_u (A_u - K)_+ \quad \text{is unbiased for } \pi, \forall u \in \mathbb{R}.$$

Now we simulate at drift u_σ instead of r_σ . The **optimal** value of u is that which minimizes $\text{Var} D_2(u)$. We shall write:

$$D_2 = D_2(u^*)$$

Likelihood Ratio Estimators, rare events: Importance Sampling.

SIMULATION METHODS IV

The Hybrid Estimators: Adding a control variable to D_2 , we simulate at u and thus we can calculate G_u , yielding another unbiased estimator:

$$D_3(u) = L_u Y_1^u + \alpha(\mathbf{E}Y_2^u - Y_2^u), \quad D_3 = D_3(u^*)$$

where u^* minimizes the variance.

Finally, consider changing the measure of the controlled estimator D_1 :

$$D_4(u) = L_u Y_1^u + \alpha(\mathbf{E}Y_2^r - L_u Y_2^u), \quad D_4 = D_4(u^*)$$

where u^* minimizes the variance.

Summarizing:

- D_0 : Naïve Estimator
- D_1 : Controlled Estimator
- D_2 : Naïve Estimator under Importance Sampling
- D_3 : Controlled Likelihood Ratio Estimator
- D_4 : Likelihood Ratio Estimator under Importance Sampling

EFFICIENT IMPORTANCE SAMPLING

The change of measure approach yields estimators that can improve efficiency. When rare events are involved, Importance Sampling will simulate more often the rare situations and L_u weighs appropriately the estimate to yield unbiasedness.

Problem: For a likelihood estimator $D(u)$, the optimal value of u^* is problem-dependent.

$$\text{Minimize: } \mathbf{Var}[D(u)], \quad \Rightarrow \quad \text{Find } u^* : \left. \frac{\partial}{\partial u} \mathbf{Var}[D(u)] \right|_{u=u^*} = 0$$

Solutions:

- The first solution and most commonly used is to perform *pilot* simulations. We used Functional Estimation to save computational time.

Use the same random numbers to evaluate in parallel the Likelihood Ratio Estimators for different values of u and each value of K .

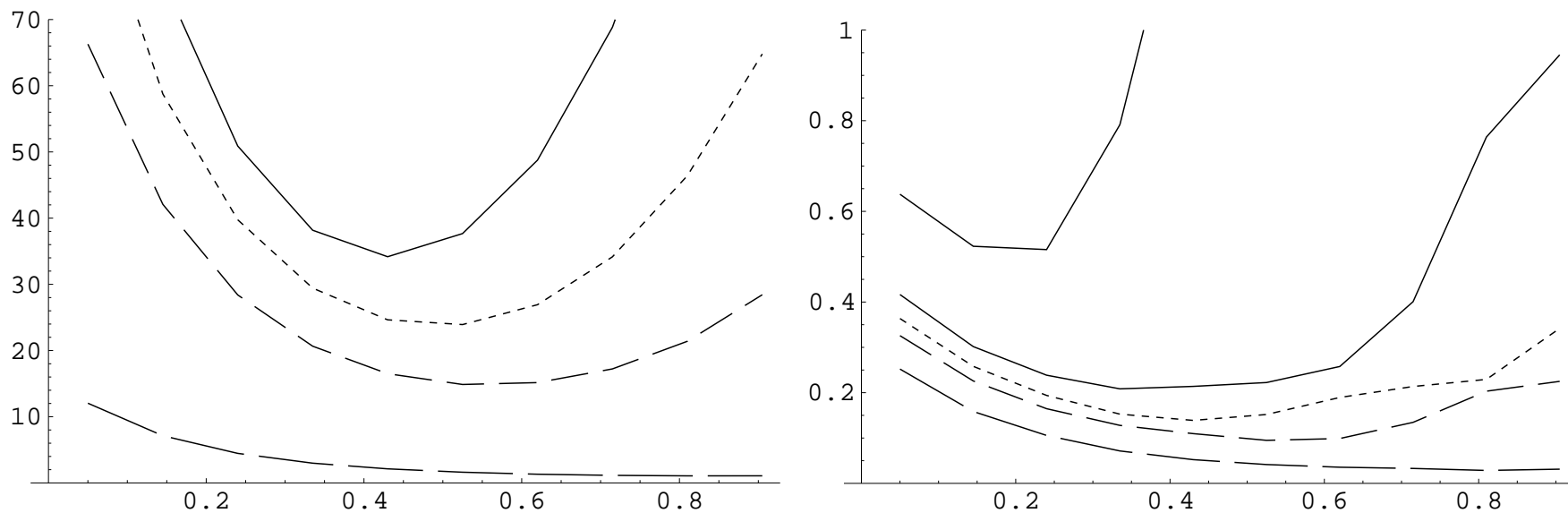
$$Z_i \sim \mathcal{N}(0, 1), \quad i = 1, \dots, N, \quad B_i = Z_1 + \dots + Z_i, \quad 1 \leq i \leq N, \quad B_0 = 0$$

$$A_u = \frac{S_0}{N} \sum_{i=1}^N e^{i(u+\frac{\sigma}{2})h + \sigma\sqrt{h}B_i}, \quad L_u = e^{\left\{ -\frac{(u-r)^2 T}{2\sigma^2} - \frac{(u-r)\sqrt{h}B_N}{\sigma} \right\}}$$

- New Approach: optimize the parameter u at the same time $D(u)$ is simulated.

FUNCTIONAL ESTIMATION

Functional Estimation: We performed functional estimation for 10 values of $u \in [0, 1]$, with $r = 0.05$, $\sigma^2 = 0.2$, $S_0 = 50$, $T = 1.0$ and $M = 10,000$. It took 20-28 seconds for each K , both for D_2 and D_4 .



Estimated Variance of D_2 and D_4 as a function of u .

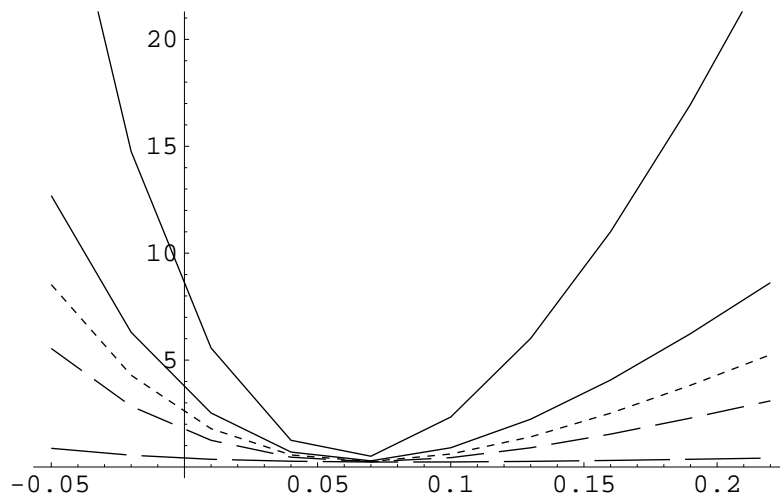
Notice the difference in scale of $\widehat{\text{Var}}(D_2)$ compared with $\widehat{\text{Var}}(D_4)$.

Conjecture: The variance is locally convex and has a unique minimum, which depends on K/S_0 . The optimal values of u for D_2 and D_4 seem to be the same (or very close).

OPTIMAL VALUES OF u^*

Optimal values of u^* found by pilot simulations with Functional Estimation (≈ 24 secs CPU).

Optimal Values for Change of Measure: u^*					
Method	K= 30	K= 45	K=50	K= 55	K = 75
D_2	0.25	0.40	0.50	0.60	0.80
D_3	0.07	0.07	0.07	0.07	0.07
D_4	0.25	0.40	0.50	0.60	0.80



Estimated Variance of D_3

Summarizing: The behaviour of the variances of the estimators:

- $\text{Var}D_2(u)$ and $\text{Var}D_4(u)$ are similar, but $D_4(u)$ has variances several orders of magnitude smaller.
- For D_2 and D_4 the optimal u^* seems to be the same and it depends on K/S_0 .
- D_3 does not reduce the variance considerably, but u^* seems to be independent of K/S_0 .

COMPARISON OF THE METHODS

$$\begin{aligned}
 D_1 &= Y_1^r + \hat{\alpha}^*(\mathbf{E}Y_2^r - Y_2^r), & D_3 &= L_u Y_1^u + \hat{\alpha}^*(\mathbf{E}Y_2^r - Y_2^r), & Y_1^u &= (A_u - K)_+ \\
 D_0 &= Y_1^r, & D_2 &= L_u Y_1^u, & D_4 &= L_u Y_1^u + \hat{\alpha}^*(\mathbf{E}Y_2^r - L_u Y_2^u), & Y_2^u &= (G_u - K)_+
 \end{aligned}$$

D_1, D_3, D_4 : Control Variable methods, estimated parameter $\hat{\alpha}^*$.

D_2, D_3 and D_4 : Likelihood Ratio estimators, estimated optimal value u^* .

$r = 0.05, \sigma^2 = 0.2, S_0 = 50, T = 1.0$ and $M = 10,000$					
Estimators					
Method	K= 30	K= 45	K=50	K= 55	K = 75
D_0	20.46 ± 0.26	8.45 ± 0.216	5.80 ± 0.189	3.83 ± 0.160	0.630 ± 0.068
D_2	20.34 ± 0.137	8.32 ± 0.115	5.66 ± 0.096	3.74 ± 0.075	0.583 ± 0.020
D_1	20.31 ± 0.016	8.28 ± 0.013	5.64 ± 0.012	3.72 ± 0.011	0.585 ± 0.010
D_3	20.31 ± 0.014	8.28 ± 0.011	5.64 ± 0.010	3.71 ± 0.010	0.583 ± 0.009
D_4	20.31 ± 0.014	8.27 ± 0.009	5.62 ± 0.008	3.70 ± 0.006	0.573 ± 0.003
Variance					
Method	K= 30	K= 45	K=50	K= 55	K = 75
D_0	176.09	121.70	92.58	66.28	12.04
D_2	49.04	34.59	23.76	14.95	1.07
D_1	0.64	0.42	0.36	0.33	0.25
D_3	0.48	0.28	0.25	0.24	0.23
D_4	0.53	0.207	0.150	0.095	0.028
$\hat{\alpha}^*$	0.998	1.05	1.07	1.10	1.20

ACCELERATING THE SIMULATION

Of all the estimators, D_4 seems to be a better Likelihood Ratio estimator, especially in the cases of interest, when the option is out of the money. But *efficiency* is a measure of precision and speed, and pilot tests take long time.

Problem: For a likelihood estimator $D(u)$, the optimal value of u^* is problem-dependent.

$$\text{Minimize: } \mathbf{Var}[D(u)], \quad \Rightarrow \quad \text{Find } u^* : \left. \frac{\partial}{\partial u} \mathbf{Var}[D(u)] \right|_{u=u^*} = 0$$

Solutions:

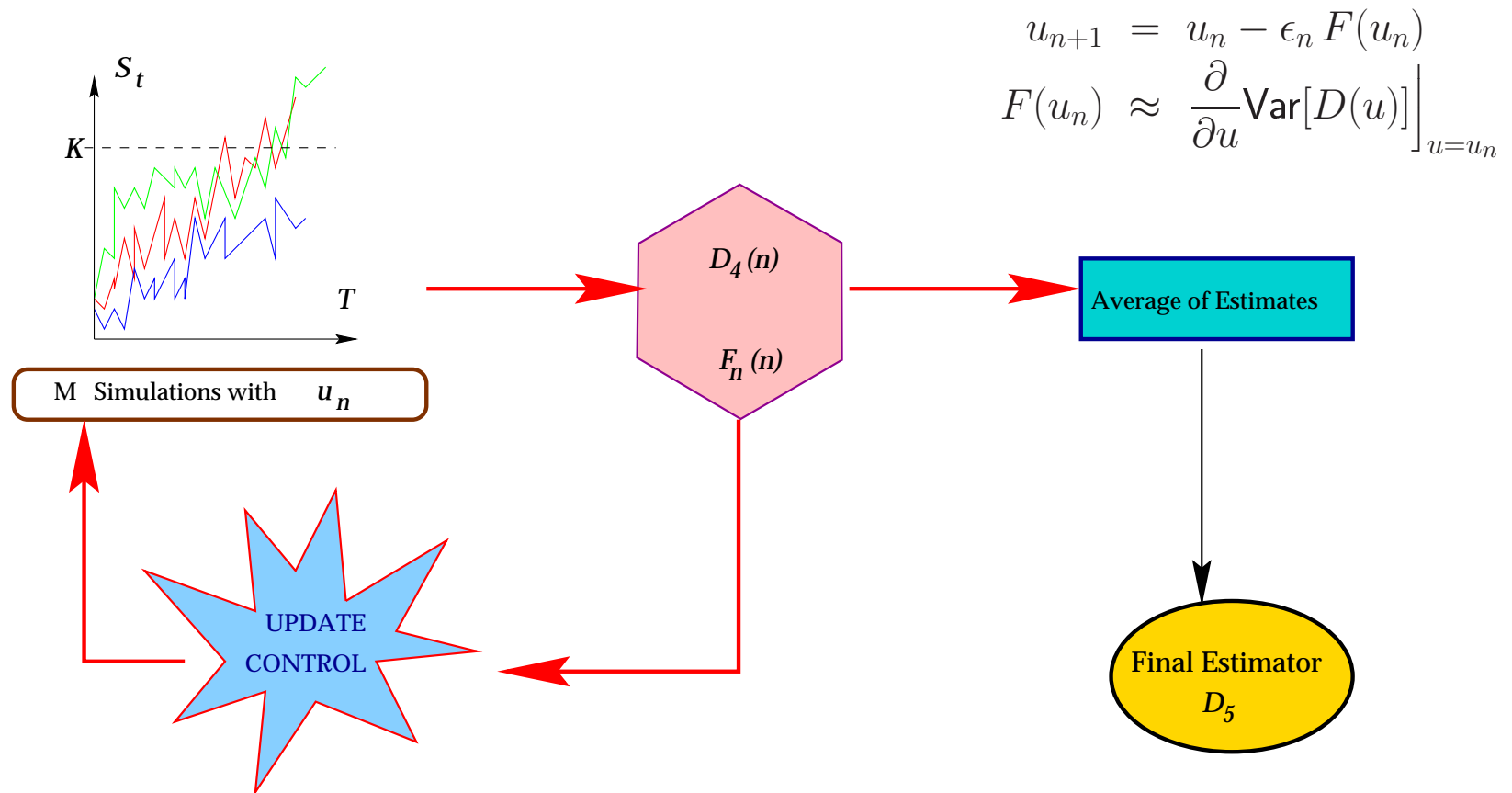
- Pilot tests with Functional Estimation. It took about five times as long to simulate as a single run.
- Proposed approach: apply Robbins-Monroe procedure to search for the optimal u^* at the same time that $D(u)$ is simulated.

$$\begin{aligned} u_{n+1} &= u_n - \epsilon_n F(u_n) \\ F(u_n) &\approx \left. \frac{\partial}{\partial u} \mathbf{Var}[D(u)] \right|_{u=u_n} \end{aligned}$$

Adjust the value of the parameter u in the direction of variance reduction.

\Rightarrow **Derivative Estimation ...**

STOCHASTIC APPROXIMATION



Parallel Estimation and Optimization

- It has the potential of variance improvement compared to the control variable estimator,
- It does not need preliminary simulations, but rather finds the most variance reduction as it estimates the option value.

DERIVATIVE ESTIMATION VIA IPA (INFINITESIMAL PERTURBATION ANALYSIS)

IPA uses the *stochastic derivative* to estimate the derivative of an expectation. Consider the estimator $D_2(u) = L_u(A_u - K)_+$. Both L_u and A_u are differentiable functions of u , given a fixed trajectory Z_1, \dots, Z_N , that is:

$$\begin{aligned} A_u &= \frac{S_0}{N} \sum_{i=1}^N e^{i(u+\frac{\sigma}{2})h + \sigma\sqrt{h}B_i}, & L_u &= \exp \left\{ -\frac{(u-r)^2 T}{2\sigma^2} - \frac{(u-r)\sqrt{h}B_N}{\sigma} \right\} \\ A'_u &= \frac{1}{N} \sum_{i=0}^N ihS_{ih}^u, & l'_u &= - \left(\frac{(u-r)T + \sqrt{\sigma^2 h}B_N}{\sigma^2} \right) \end{aligned}$$

Use the fact that $\mathbf{E}[L_u Y_i^u]$ is independent of u and show that we can interchange the derivative and the expectation to get:

$$\frac{\partial}{\partial u} \mathbf{Var}[D_2(u)] = \frac{\partial}{\partial u} \mathbf{E}\{L_u^2[(A_u - K)_+]^2\} = \mathbf{E}[2L_u^2 l'_u ((Y_1^u)^2) + 2L_u^2 Y_1^u A'_u].$$

Theorem 1 *The IPA estimators*

$$\begin{aligned} F_2(u) &= 2D_2(u)(l'_u D_2(u) + L_u A'_u) \\ F_4(u) &= 2D_4(u) \left(l'_u D_4(u) + L_u \left[A'_u \mathbf{1}_{\{Y_1^u > 0\}} - \alpha \frac{(T+h)}{2} G_u \mathbf{1}_{\{Y_2^u > 0\}} \right] \right) \end{aligned}$$

are unbiased, that is:

$$\mathbf{E}[F_i] = \frac{\partial}{\partial u} \mathbf{Var}_u(D_i), \quad i = 2, 4$$

IPA ESTIMATION

The table shows the results for $r = 0.05$, $\sigma^2 = 0.2$, $S_0 = 50$, $K = 50$, $T = 1.0$ and $M = 50,000$.

Derivative Estimation via IPA				
Value of u	$\text{Var}(D_2)$	F_2	$\text{Var}(D_4)$	F_4
0.2	45.32	-175.5 ± 15.7	0.21	-2.08 ± 0.29
0.3	32.01	-93.4 ± 9.2	0.16	-1.13 ± 0.17
0.4	25.44	-38.7 ± 7.3	0.15	-0.28 ± 0.35
0.5	23.69	3.89 ± 8.3	0.17	0.25 ± 0.77
0.6	26.05	45.44 ± 12.0	0.20	0.34 ± 0.53
0.7	32.94	94.88 ± 20.2	0.22	0.36 ± 0.78
0.8	45.80	168.82 ± 41.6	0.49	6.29 ± 21.68

Conjecture: The optimal values of u for D_2 and D_4 are close, but by construction, $\text{Var}[D_4] < \text{Var}[D_2]$ and it's easier to estimate $\text{Var}[D_2]$.

Idea: Drive the procedure *faster* towards the optimum, with a *correction* for asymptotic optimality towards the solution of

$$\frac{\partial}{\partial u} \text{Var}[D(u)] = 0$$

Use first the (larger) IPA estimate F_2 and F_4 later.

STOCHASTIC APPROXIMATION FOR ACCELERATED SIMULATION

To speed up the simulation, we let the algorithm *change* the value of u towards the optimum in an adaptive way using derivative information. We implement a stochastic approximation algorithm as:

$$u_{n+1} = u_n - \epsilon_n \bar{F}_n(u_n) \quad (1)$$

where:

$$\bar{F}_n(u) = \rho_n F_2(u) + (1 - \rho_n) F_4(u), \quad \rho_n = \rho_0^n, \quad 0 < \rho_0 < 1,$$

then $\mathbf{E}[\bar{F}_n(u)] \rightarrow \mathbf{E}[F_4(u)]$ as $n \rightarrow \infty$. Let u be constrained to some compact interval U . It can be shown that

$$u_n \rightarrow u^* \quad \text{with probability 1,}$$

provided that:

$$\begin{aligned} \forall u \in U, \quad \mathbf{E}[\bar{F}_n(u)] &= \frac{\partial}{\partial u} \mathbf{Var}(D_4) + \beta_n, \quad \sup_{u \in U, n} \mathbf{Var}[\bar{F}_n(u)] < \infty \\ \sum_{n=1}^{\infty} \epsilon_n &= +\infty, \quad \sum_{n=1}^{\infty} \epsilon_n^2 < \infty, \quad \sum_{n=1}^{\infty} \epsilon_n \beta_n < \infty \end{aligned}$$

Our proposed estimator is:

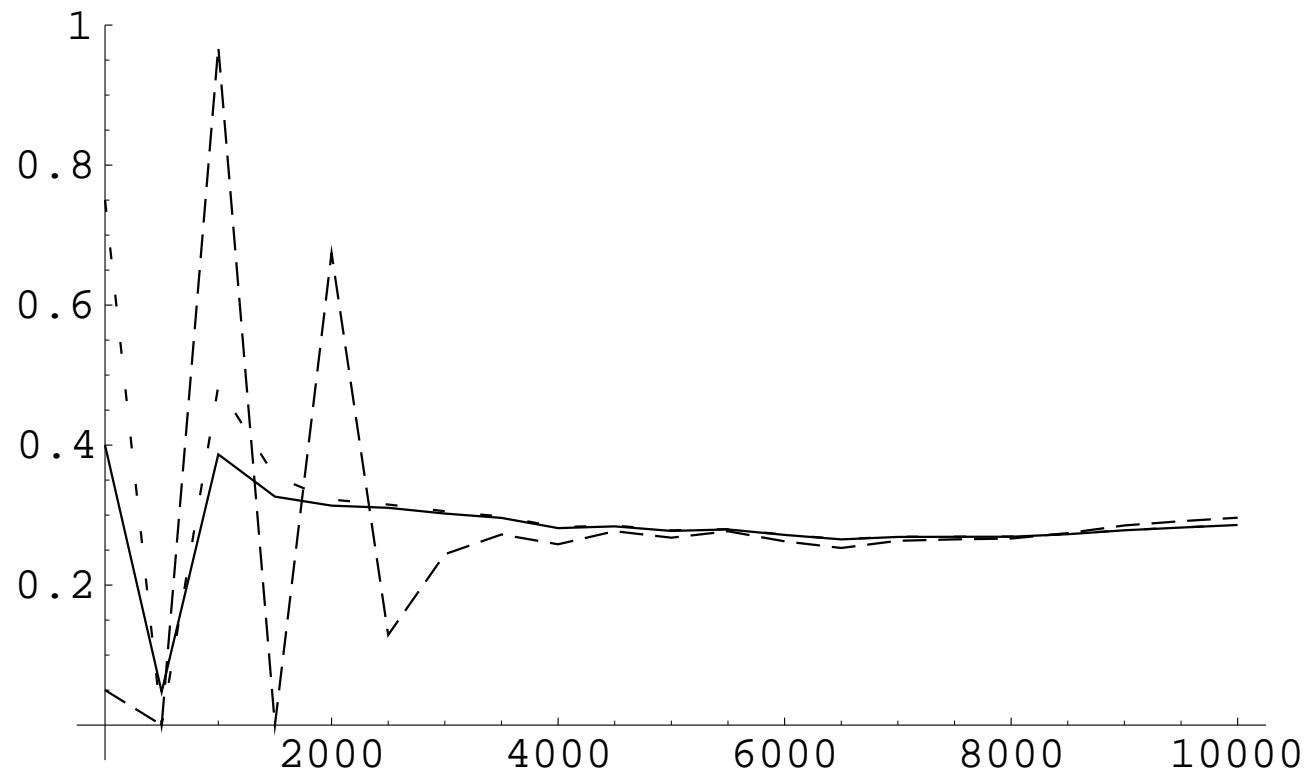
$$D_5 = \frac{1}{B} \sum_{n=1}^B D_4(u_n)$$

It uses u_n for M simulations to obtain the sample mean of $D_4(u_n)$ while estimating the *sensitivity* $\bar{F}_n(u_n)$. Then the control value is changed in the direction of variance reduction, with $\epsilon_n = \epsilon_0/n$.

RESULTS OF THE ACCELERATED SIMULATION METHOD.

Control value u_n against the number of simulations performed. $r = 0.05, \sigma^2 = 0.2, S_0 = 50, T = 1.0, K = 30, M = 500, B = 20, \epsilon_0 = 5 \times 10^{-4}, U = [0, 1]$. Total number of simulations: $MB = 10,000$.

The optimal value found by inspection was $u^* \approx 0.25$. In the plot the simulation starts at three different initial values for u_0 . Convergence can be achieved within the first iterations.



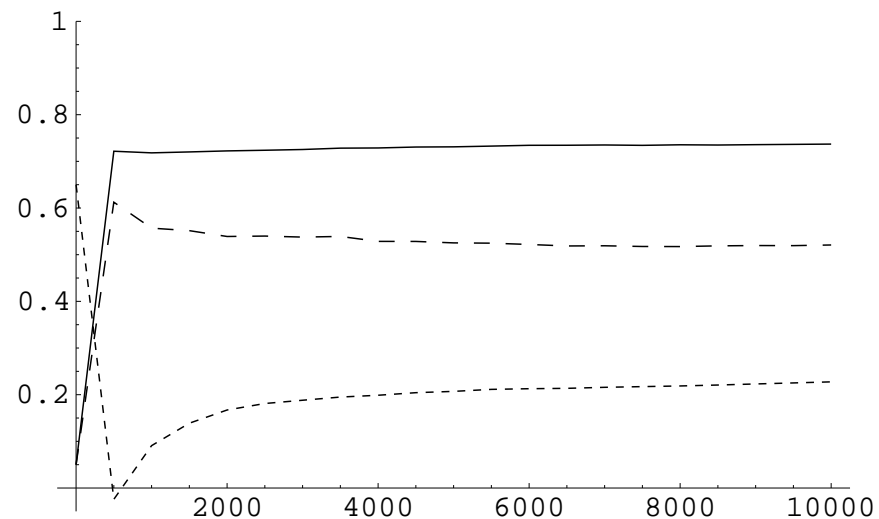
Values of u for $K = 30$ and different μ_0 .

PLOT OF THE ACCELERATED SIMULATION.

$r = 0.05, \sigma^2 = 0.2, S_0 = 50, T = 1.0, K = 30, M = 500, B = 20, U = [-0.05, 1]$. Total number of simulations: $MB = 10,000$.

- Solid line: $K = 75$,
 $\epsilon_0 = 0.008, u^* \approx 0.75$.
- Long dashes: $K = 50$,
 $\epsilon_0 = 0.001, u^* \approx 0.5$.
- Short dashes: $K = 30$,
 $\epsilon_0 = 0.0001, u^* \approx 0.26$.

We chose initial values of u which we knew to be far from the optimum.



It should be clear from these plots why the variance of our estimator is very close to the optimal one (which was used for D_4), since in all cases convergence was achieved within the first three or four iterations of the stochastic approximation.

STATISTICAL ANALYSIS THE SELF-OPTIMIZED ESTIMATOR

Table comparing D_1 with D_4 and D_5 . The CPU time reported in the table considers that a pilot test with Functional Estimation was made for D_4 .

$r = 0.05, \sigma^2 = 0.2, S_0 = 50, T = 1.0$ and $M = 500, B = 20$				
Estimators				
Method	K= 30	K=50	K = 75	
D_1	20.31 ± 0.016	5.64 ± 0.012	0.585 ± 0.010	
D_4	20.31 ± 0.014	5.62 ± 0.008	0.573 ± 0.003	
D_5	20.31 ± 0.015	5.62 ± 0.008	0.578 ± 0.004	
Estimated Variance				CPU Time
Method	K= 30	K=50	K = 75	in seconds
D_1	0.64	0.36	0.25	5
D_4	0.53	0.15	0.03	26 + 6
D_5	0.54	0.18	0.04	6

Notice that for $K = 75, S_0 = 50$, D_5 is six times more efficient than D_1 .

- The Self-Optimized Estimator D_5 seems to show no larger variances than the currently used Controlled Estimator D_1 .
- It can considerably improve the precision when the option is out of the money.
- The added computational effort for IPA derivatives is negligible.

A SECOND MODEL

Overall behaviour is consistent for all the examples that we simulated: while D_4 presents a lower variance, it takes extra simulation time to determine its optimal parameter for the change of measure. CPU times: about 5 seconds for each run of length 10,000 simulations, plus 26 seconds of pilot tests to determine u^* for D_4 .

$r = 0.09, \sigma^2 = 0.04, S_0 = 100, T = 0.4$ and $M = 500, B = 20$						
Estimators						
Method	K= 90	K=95	K = 100	K = 105	K = 110	K=130
D_0	11.6194	7.3281	3.9317	1.7225	0.6249	0.001366
D_1	11.5486	7.2401	3.8398	1.6732	0.5930	0.001443
D_4	11.5473	7.2386	3.8378	1.6701	0.5913	0.001380
D_5	11.5473	7.2380	3.8375	1.6700	0.5914	0.001394
Variance						
Method	K= 90	K=95	K = 100	K = 105	K = 110	K=130
D_0	50.960	41.400	26.200	12.430	4.400	0.006233
D_1	0.0129	0.0103	0.0083	0.0069	0.0061	0.000245
D_4	0.0101	0.0060	0.0040	0.0023	0.0014	0.000004
D_5	0.0143	0.0077	0.0057	0.0032	0.0018	0.000005

Comment: Example from Fu, Madan and Wang (1997) “Pricing Asian Options: a Comparison of Analytical Results and Monte Carlo Methods”, Working paper.

As expected, our estimator is noticeably more efficient for larger values of K .

CONCLUSIONS AND OPEN QUESTIONS

Comparison between our estimator and the control variable estimator:

- Accelerated Simulation can be shown to achieve always greater precision in the limit, as the learning algorithm (the stochastic approximation) gathers more samples.
- We use the variance of D_2 to guide our search for u^* and achieve faster convergence.
- The extra calculations compared to the control variable estimator take negligible computational time.
- The gain in precision can be very large when out of the money. Call/Put parity of Options allows indirect estimation for options that are in the money.

Open Research Topics:

- Convexity of the variance as a function of u for different options.
- Relationship between the variance of D_2 and D_4 and their corresponding optima.
- Extension to pricing of general financial derivatives.
- Choice of ϵ_0 , ρ_0 , M and B . The choice of parameters for adaptive learning is usually a hard problem. Given a computational budget in CPU time, for example, we would like to be able to determine the parameters of the Accelerated Simulation program.
- State space: for now we have used $U = [0, 1]$, which we decided by inspection of our preliminary simulations. The question of how to project stochastic approximations to constrain the control values u to a compact set U is not always obvious.