

WORKING PAPER

Adaptive Milstein Methods for Diffusions with Discontinuous Drift

Strong error bounds and an application to threshold rebalancing

Supervised by Dr. T. Zamrik

25-OCT-2026

Gradient flow of the Peaks function: $\dot{\mathbf{x}} = \nabla f(\mathbf{x})$, $f(x, y) = |3(1-x)^2 e^{-x^2-(y+1)^2} - 10(x/5 - x^3 - y^5) e^{-x^2-y^2} - \frac{1}{3} e^{-(x+1)^2-y^2}|$



WP-2026-16039117
zamrik.com



Zamrik Quant Research

The research arm of Zamrik Quant Lab

Copyright

© 2026 T. Zamrik. All rights reserved. Published by Zamrik Quant Research. No part of this document may be reproduced, stored, distributed or adapted, in whole or in part, by any means, without the prior written permission of the publisher, except for brief quotation with full citation. Requests for permission: info@zamrik.com.

Suggested Citation

Stochastic Analysis Desk, Zamrik Quant Research (2026). ‘Adaptive Milstein Methods for Diffusions with Discontinuous Drift: Strong error bounds and an application to threshold rebalancing’. *Zamrik Quant Research Working Paper WP-2026-16039117*. Available at <https://zamrik.com/research-items/wp-2026-16039117/>.

About Zamrik Quant Research

Zamrik Quant Research is the research arm of Zamrik Quant Lab. It works through five research desks, from market microstructure and derivatives to optimal control, stochastic analysis and applied mathematics, and publishes everything it finds in the open: research papers, notebooks that run on Kaggle, free tools that answer one practical question, and essays.

The five desks are Market Microstructure & Trading; Derivatives & Volatility; Optimal Control & Decision; Stochastic Analysis; and Applied Mathematics. Each paper, notebook, tool and essay belongs to exactly one desk. Zamrik Quant Research is run by a small, independent team, headed by Dr. T. Zamrik. We are self-funded: no sponsors, no advertisers, no outside capital.

Every publication passes through the same five functions before anyone outside sees it: research, data, software, verification and production. We never retrofit the mathematics to a conclusion chosen in advance. Every number is checked two independent ways. Every figure is re-run from its code. When we find an error, or replace a piece with a better one, we record it in the public corrections log at <https://zamrik.com/corrections/>.

Legal Notice

This paper is research and education. Nothing in it is investment, financial, legal or engineering advice; nothing is a recommendation to buy or sell shares of stock, derivative contracts, bonds, or any other instrument or property; and nothing is a forecast. You are responsible for your own decisions. Figures and code are provided as they are. Care and independent checks reduce errors but cannot rule them out, so no warranty is given. The views expressed are those of the Stochastic Analysis Desk.



WORKING PAPER

Adaptive Milstein Methods for Diffusions with Discontinuous Drift

Strong error bounds and an application to threshold rebalancing

Supervised by Dr. T. Zamrik¹

25-OCT-2026

WP-2026-16039117

Abstract. We apply an adaptive interface method to a two-dimensional diffusion with finite-speed rebalancing outside an elliptical no-trade region. An explicit change of variables removes the drift jump, preserves commutativity of the noise, and admits a quantitative global inverse certificate. The resulting approximation has terminal root-mean-square error $O(\delta)$ at expected Brownian observation cost $O(\delta^{-1})$. Terminal quadratic tracking risk has L^1 error $O(\delta)$; a randomized within-step estimator of gross turnover has L^1 error $O(\delta^{2/3})$. A synthetic experiment illustrates the risk–turnover trade-off but finds no runtime advantage over ordinary Euler at the tested tolerances. The complete underlying interface proof is included in the proof sections below. Independent proof review and a wider originality assessment remain outstanding.

WP-2026-16039117 · Stochastic Analysis Desk · Working Paper · Published
25-OCT-2026 · Reading time: about 21 min

1. Introduction

No-trade regions balance deviations from a target allocation against the cost of restoring it. They occur in rebalancing research [7]; gradual rebalancing is also studied in practice [8]. We use this motivation to construct an explicit curved-interface diffusion, rather than derive an optimal trading rule or calibrate a market model.

Main findings

- An anisotropic, finite-speed rebalancing model satisfies the interface theorem after an explicit transformation with a certified Lipschitz inverse.
- The approximation bounds extend to terminal quadratic risk and a randomized estimator of discontinuous turnover, at respective L^1 rates $O(\delta)$ and $O(\delta^{2/3})$.
- Across three synthetic thresholds, estimated mean turnover falls from 0.2330 to 0.0730 while mean terminal risk rises from 0.1565 to 0.2569.
- At the first tested levels with estimated endpoint RMS below 10^{-3} , Euler takes 0.212 ms per path and adaptive transformed Milstein takes 2.833 ms per path; the experiment provides no runtime-speedup claim.

¹Correspondence: Stochastic Analysis Desk, Zamrik Quant Research, info@zamrik.com.



The analytic contribution proposed here is a modest multidimensional extension of existing adaptive and transformation arguments [1], [2], [3], [4], [5]. The application supplies explicit assumptions and output functionals; it does not itself settle whether the underlying extension is publishably new.

2. Finite-speed rebalancing

Let $e_t \in \mathbb{R}^2$ denote errors in two risky exposures relative to a target. We study

$$de_t = u_r(e_t) dt + \Sigma dW_t, \quad e_0 = 0, \quad u_r(e) = -k \frac{Qe}{\|Qe\|} \mathbf{1}_{\{e^\top Qe > r^2\}}. \quad (2.1)$$

The value of u_r on the ellipse is set to zero. Here Q is positive definite, Σ is nonsingular, and $k, r > 0$ are fixed. The control has Euclidean speed k outside the ellipse and vanishes inside it. These are unconstrained exposure errors, not weights required to sum to one.

Our two outputs are terminal tracking risk and gross risky-exposure turnover:

$$R(e) = e^\top Qe, \quad A = \int_0^1 \ell(e_t) dt, \quad \ell(e) = k \frac{\|Qe\|_1}{\|Qe\|} \mathbf{1}_{\{e^\top Qe > r^2\}}. \quad (2.2)$$

Turnover counts absolute trading in the two risky exposures, without a one-half convention or a cash leg. It is not total variation of the noisy exposure process. It satisfies $0 \leq A \leq k\sqrt{2}$; a proportional trading cost would be a fixed multiple of A .

The experiment fixes

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 0.35 & 0 \\ 0.10 & 0.25 \end{pmatrix}, \quad k = 0.4, \quad r \in \{0.25, 0.40, 0.55\}, \quad T = 1.$$

The units are synthetic. No prices, returns, dividends, transaction-cost data, or market calibration enter this experiment. As an analytic scale check, with control switched off, $\mathbb{E}R(e_1) = \text{tr}(Q\Sigma\Sigma^\top) = 0.4125$. More generally, Itô's formula gives

$$\mathbb{E}R(e_1) = 0.4125 - 2k\mathbb{E} \int_0^1 \|Qe_t\| \mathbf{1}_{\{e_t^\top Qe_t > r^2\}} dt.$$

This identity explains the risk-reducing contribution of the control, but does not prove monotonicity in the threshold r .

3. An explicit transformation and its certificate

Set $x = Q^{1/2}e/r$, $S = Q^{1/2}\Sigma/r$, $C = SS^\top$, and $K = k/r$. The interface becomes the unit circle and the drift becomes



$$b_-(x) = 0, \quad b_+(x) = -K \frac{Qx}{\sqrt{x^\top Qx}} \quad (\|x\| > 1).$$

Write $d = \|x\| - 1$, $n = x/\|x\|$, and choose $w = 0.1$. In the annulus $|d| < w$, define

$$\begin{aligned} f(d) &= d|d| (1 - d^2/w^2)^6, \\ \alpha(x) &= \frac{KQx\|x\|^2}{2\sqrt{x^\top Qx}(x^\top Cx)}, \quad G(x) = x + \alpha(x)f(d). \end{aligned} \quad (3.1)$$

Outside the annulus set $G(x) = x$. The cutoff is C^5 at its outer edges, which is sufficient here; G is globally $C^{1,1}$ and is C^4 on each open side of the circle, with bounded derivatives of orders one through four. On the circle $G(x) = x$ and $DG(x) = I$.

Let q_-, q_+ and c_-, c_+ be the extreme eigenvalues of Q and C . Put

$$A_0 = \frac{Kq_+}{2\sqrt{q_-}c_-}, \quad B_0 = 2 + q_+/q_- + 2c_+/c_-.$$

Direct differentiation on $\|x\| \geq 1/2$ gives $\|\alpha\| \leq A_0$ and $\|D\alpha\| \leq 2A_0(1 + B_0)$. The elementary polynomial bounds $\sup |f| \leq w^2/16$ and $\sup |f'| \leq w/4$ imply

$$\|DG - I\|_\infty \leq \varepsilon_G := \frac{2A_0(1 + B_0)w^2}{16} + \frac{A_0w}{4} < 1. \quad (3.2)$$

For completeness, the maximum of $z^2(1 - z^2)^6$ is $6^6/7^7 < 1/16$; the interior extrema of $2z(1 - z^2)^5(1 - 7z^2)$ occur at $z^2 = (16 \pm \sqrt{165})/91$ and have absolute value below $1/4$.

Writing $G = I + g$, the equation $x = z - g(x)$ is a contraction. Consequently G is a global bijection and $\text{Lip}(G^{-1}) \leq (1 - \varepsilon_G)^{-1}$. For the three thresholds, respectively,

$$\varepsilon_G \leq 0.08625, \quad 0.13799, \quad 0.18974, \quad \text{Lip}(G^{-1}) \leq 1.09439, \quad 1.16009, \quad 1.23417.$$

These are analytic upper bounds, not maxima inferred from a sampled grid.

For $z = G(x)$ off the circle the transformed coefficients are

$$a(z) = DG(x)b(x) + \frac{1}{2}D^2G(x) : C, \quad B(z) = DG(x)S.$$

The contraction is componentwise. The term $\frac{1}{2}D^2G : C$ has interface traces $\pm\alpha(n^\top Cn) = \mp b_+/2$. Thus both traces of a equal $b_+/2$, and we extend a by this common value. The transformed coefficients are bounded, globally Lipschitz, piecewise C^2 with bounded derivatives, and globally uniformly elliptic. Moreover

$$DB_j(z)B_k(z) = D^2G(x)[S_k, S_j] = DB_k(z)B_j(z).$$



The generalized Itô formula for $C^{1,1}$ maps applies because the nondegenerate diffusion spends zero Lebesgue time on the circle. This also gives strong existence and uniqueness for (2.1) by transformation and inversion. All assumptions of Theorem 1 are therefore satisfied.

3.1 Derivative bounds behind the inverse certificate

The numerical bounds on the inverse have an analytic basis. To make that certificate reproducible, write $R = \|x\|$, $u = x^\top x$, $s = x^\top Qx$, and $t = x^\top Cx$ on the annulus supporting the transformation. Introduce the positive scalar amplitude $a_0(x)$ by

$$a_0(x) = \frac{Ku}{2\sqrt{s}t}, \quad \alpha(x) = a_0(x)Qx.$$

The symbol a_0 here denotes this scalar amplitude, not the transformed drift a . Logarithmic differentiation avoids expanding a vector quotient into many individual terms:

$$\nabla \log a_0(x) = \frac{2x}{u} - \frac{Qx}{s} - \frac{2Cx}{t}.$$

The eigenvalue bounds $s \geq q_- R^2$ and $t \geq c_- R^2$ imply

$$a_0(x)\|Q\| \leq \frac{A_0}{R}, \quad \|\alpha(x)\| \leq A_0, \quad \|\nabla \log a_0(x)\| \leq \frac{B_0}{R}.$$

Differentiating $\alpha = a_0 Qx$ now yields the matrix bound directly:

$$D\alpha = a_0 Q + \alpha(\nabla \log a_0)^\top, \quad \|D\alpha\| \leq \frac{A_0(1+B_0)}{R} \leq 2A_0(1+B_0).$$

The last inequality uses $R \geq 1 - w > 1/2$. Since the signed distance has unit gradient on this annulus, differentiating $G(x) = x + \alpha(x)f(d(x))$ gives

$$DG - I = f(d)D\alpha + \alpha f'(d)n^\top.$$

Combining these expressions with the polynomial bounds on f and f' proves the global perturbation estimate stated above. Outside the annulus the perturbation vanishes. Its derivative also vanishes at the outer edges, so the estimate extends across the boundary of its support without adding an interface term.

The singular-value implication is quantitative: if $z = G(x)$, then

$$\sigma_{\min}(B(z)) \geq (1 - \varepsilon_G)\sigma_{\min}(S) > 0.$$

Thus the same certificate proves both invertibility of the transformation and uniform ellipticity of the transformed noise. This is stronger than checking nonsingularity at sampled points. It also gives a deterministic interpretation of a computed inverse residual: for any approximate inverse v ,



$$\|v - G^{-1}(z)\| \leq (1 - \varepsilon_G)^{-1} \|G(v) - z\|.$$

This last inequality controls a single inverse evaluation. It does not, by itself, bound the accumulation of inverse errors through the full stochastic algorithm; the implementation limitation stated later therefore remains in force.

4. Algorithm and output guarantees

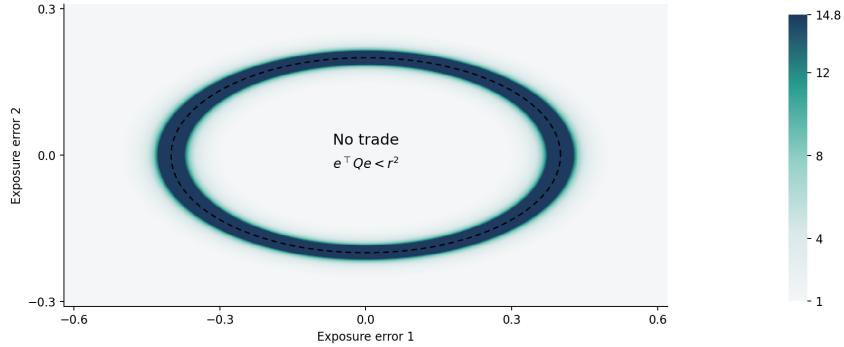
For $\psi = \beta \log^4(1/\delta)$, with fixed $\beta > 0$, take

$$h(y) = \min\{\delta, \max\{\delta^2\psi, d_\Theta(y)^2/\psi\}\}, \quad \Theta = \{\|y\| = 1\}.$$

The underlying theorem uses $\beta = 1$. Its proof works for every fixed positive β : replace L^4 by ψ and L^2 by $\sqrt{\psi}$; the exceptional probability is still superpolynomial and the cost identities are unchanged. Constants and the admissible asymptotic range may depend on β . The experiment uses $\beta = 1/64$ to avoid excessive refinement at its finite set of levels. We do not assert that all tested levels lie inside the unspecified asymptotic range of the theorem.

Adaptive refinement near the trading boundary

$r = 0.40$ | $\delta = 1/128$ | Refinement = maximum step / local step



Dashed ellipse: trading boundary. Refinement is evaluated after transformation.

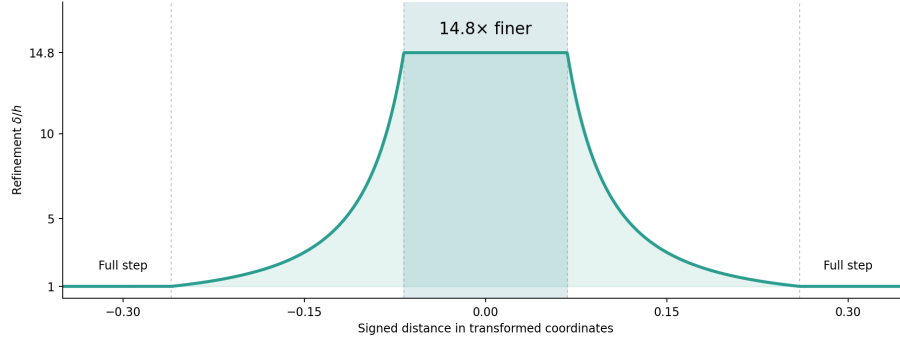
Near the boundary, local steps are 14.8 times smaller than the maximum step.

Figure 1: Spatial refinement of the adaptive mesh at $r = 0.40$ and $\delta = 1/128$. The colour gives maximum step divided by local step, evaluated at $y = G(Q^{1/2}e/r)$; the dashed ellipse is the trading boundary. This is an evaluation of the mesh formula, not a sample-path simulation.



How the adaptive step shrinks near the interface

$r = 0.40$ | $\delta = 1/128$ | Shaded band: minimum-step region



Maximum step: 0.007812 Minimum step: 0.000529

The distance is measured from the unit circle after the drift-removing transformation.

Figure 2: Refinement factor as a function of signed transformed distance $\|y\| - 1$. Near the interface steps are about 14.8 times smaller. The central band uses the minimum step; distant states use the full step. Parameters are the same as in Figure 1.

On a step starting at $y = G(x)$, with $\xi = S\Delta W$, the transformed Milstein update simplifies to

$$y' = G(x) + DG(x)(b(x)h + \xi) + \frac{1}{2}D^2G(x)[\xi, \xi]. \quad (4.1)$$

The Hessian drift term cancels the diagonal Milstein correction. At a starting point on the circle use the common transformed drift trace and set the Milstein derivative coefficient to zero, as in the underlying theorem. The interpolant is obtained by replacing $h, \Delta W$ by elapsed time and the corresponding Brownian increment. Set $\hat{e}_t = rQ^{-1/2}G^{-1}(Y_t)$.

Corollary 1. For the fixed parameters above, with exact coefficient and inverse evaluation, the adaptive approximation satisfies

$$\sup_{t \leq 1} \mathbb{E}\|e_t - \hat{e}_t\|^2 \leq C\delta^2, \quad \mathbb{E}N \leq C\delta^{-1}, \quad \mathbb{E}|R(e_1) - R(\hat{e}_1)| \leq C\delta.$$

Proof. The first two claims follow from Theorem 1 and the Lipschitz inverse. Bounded coefficients of the transformed SDE and interpolant give uniform second moments. Since $|R(e) - R(v)| \leq \|Q\|\|e - v\|(\|e\| + \|v\|)$, Cauchy–Schwarz proves the risk claim. $\square$

To estimate turnover, take independent uniform random variables U_i , independent also of the Brownian motion. With the last step clipped to $T = 1$, put $s_i = \tau_i + U_i h_i$ and

$$\hat{A} = \sum_{i < N} h_i \ell(\hat{e}_{s_i}). \quad (4.2)$$

Each s_i costs one extra Brownian vector observation and one inverse evaluation. The state update and subsequent grid selection use only step endpoints, never the interior mark. The



total observation count is $2N$.

Proposition 1. The estimator (4.2) satisfies

$$\mathbb{E} |A - \widehat{A}| \leq C\delta^{2/3}.$$

The isolated randomized-quadrature error has root-mean-square order δ .

Proof. The turnover rate is bounded and Lipschitz on each side of its interface; absolute values in the ℓ^1 norm do not require differentiability here. For $v, e \in \mathbb{R}^2$ its difference is at most $C\|e - v\| + C$ times the indicator that the segment from e to v intersects the ellipse. Such an intersection implies $d(e, \Theta_e) \leq \|e - v\|$, where Θ_e is the physical ellipse. A signed-distance local-time argument, with bounded drift and constant nonsingular diffusion, gives $\mathbb{E} \int_0^1 \mathbf{1}_{\{d(e_t, \Theta_e) \leq \epsilon\}} dt \leq C\epsilon$. Indeed, saturate signed distance smoothly outside a fixed tube, apply Tanaka's formula to bound expected local time uniformly, and use the positive lower bound on normal quadratic variation inside the tube. The state bound and Markov's inequality now yield

$$\mathbb{E} \int_0^1 |\ell(e_t) - \ell(\hat{e}_t)| dt \leq C\delta + C\epsilon + C\delta^2/\epsilon^2.$$

Choosing $\epsilon = \delta^{2/3}$ proves order $\delta^{2/3}$ for the continuous-integral error. For the quadrature part set $F(y) = \ell(rQ^{-1/2}G^{-1}(y))$. This is bounded and piecewise Lipschitz across the unit circle. Conditionally on the entire Brownian path, the grid and interpolant are fixed, and the independent uniforms make (4.2) unbiased for $\int_0^1 F(Y_t) dt$. Conditional variances add and are bounded by

$$\sum_{i < N} h_i \int_{\tau_i}^{\tau_i + h_i} |F(Y_t) - F(Y_{\tau_i})|^2 dt. \quad (4.3)$$

The squared difference is bounded by $C\|Y_t - Y_{\tau_i}\|^2 + C\mathbf{1}_{S(Y_{\tau_i}, Y_t)}$, where S denotes a segment meeting the circle. The expected contribution of the first term to (4.3) is at most $C\delta^2$, by (8.3) and $h_i \leq \delta$. On the good event F_δ used in the underlying proof, a crossing can occur only when $d_\Theta(Y_{\tau_i}) \leq r_2 = \delta\psi$, hence $h_i \leq h_* = \delta^2\psi$. Throughout such a step $d_\Theta(Y_t) \leq 2r_2$. Lemma 1 bounds the expected time in this tube by $Cr_2 + C_q\delta^q$. Thus the crossing contribution is at most $Ch_*r_2 + C_q\delta^q = C\delta^3\psi^2 + C_q\delta^q$. On the exceptional event, boundedness of F bounds the whole sum by $C\delta$, and its probability is superpolynomial. Since $\delta\psi^2$ is bounded as $\delta \downarrow 0$, the expected conditional variance is $O(\delta^2)$. Cauchy-Schwarz and the continuous-integral bound finish the proof. $\square$

The exponent $2/3$ is a conservative upper guarantee obtained from a tube bound and an L^2 state estimate. It is neither an optimality claim nor an empirically fitted convergence order.

Algorithm 1 — Adaptive simulation and randomized turnover

Inputs are the model parameters, horizon one, step parameter δ , initial exposure and independent uniform and Gaussian draws. The procedure implements the approximation in



Theorem 1 and Corollary 1.

1. Normalize the initial exposure, evaluate the transform, and set the clock and turnover sum to zero.
2. Compute the distance of the transformed state to the unit circle. Evaluate the three-region mesh and clip the proposed step at the remaining horizon.
3. Draw an independent uniform fraction. Draw independent Brownian increments for the two subintervals; their sum supplies the endpoint increment.
4. Evaluate the transformed Milstein polynomial at the interior observation and endpoint using the same left state and coefficients.
5. Invert both transformed values. Add the step length times the interior turnover rate to the turnover sum. Use only the endpoint state to choose subsequent steps.
6. Advance the clock and repeat until the horizon. Return physical exposure, quadratic risk, turnover and the total Brownian observation count.

In the experiment, damped Newton iteration stops at residual $2 \times 10^{-13}(1 + \|y\|)$; a failure to converge is reported, never silently accepted. The theorem treats inverse evaluation as exact. Each state step costs two Brownian vector observations for the two requested functionals. The extra quadrature point does not feed back into the mesh, which is essential to the conditional-unbiasedness proof of Proposition 1.

5. Numerical experiment

We compare Euler in the original coordinates with fixed-step and adaptive transformed Milstein. Because the original noise is additive, original-coordinate Milstein would equal Euler and is not a separate baseline. Every method uses the same randomized turnover rule and is charged for both endpoint and interior Brownian observations.

There are 1,024 independent paths. All methods share each path's Brownian skeleton on a 65,536-step grid. The reference is original-coordinate Euler on that grid; a nested 32,768-step reference measures sensitivity. Adaptive and interior query values are sampled by exact conditional Brownian bridges, retaining earlier observations in the current skeleton cell. Different methods have independent conditional bridges given the common skeleton; we do not claim their off-grid observations all belong to one identical realized continuous path. Each method-reference pair nevertheless has the required Brownian coupling.

Errors below are relative to the fine numerical reference, not the exact solution. At $r = 0.40$, the fine-coarse endpoint RMS difference is 2.894×10^{-5} (Monte Carlo SE 6.82×10^{-7}). It is 15.4% of the finest adaptive error, so that point is flagged and is not treated as evidence of an asymptotic rate. The reference difference is a sensitivity diagnostic, not a bound on the reference's true error.



Table 1: Endpoint accuracy and production cost at $r = 0.40$. Level is $1/h$ for fixed methods and $1/\delta$ for adaptive. Observation count is twice mean steps. Observed order is the adjacent log error ratio divided by the log mean-step ratio; it is tabulated rather than represented by a log-log line. Reference ratio is the nested reference difference divided by reported RMS. $\dagger$ marks a ratio above 0.10. CPU timings exclude compilation and reference construction.

Method / level	Mean steps	RMS $\times 10^3$	Order	Ref. ratio	ms/path
Euler / 64	64	4.619	–	0.006	0.029
Euler / 128	128	2.580	0.84	0.011	0.056
Euler / 256	256	1.437	0.84	0.020	0.108
Euler / 512	512	0.849	0.76	0.034	0.212
Euler / 1024	1024	0.511	0.73	0.057	0.423
Euler / 2048	2048	0.295	0.79	0.098	0.838
Fixed transformed / 64	64	8.087	–	0.004	0.169
Fixed transformed / 128	128	4.880	0.73	0.006	0.329
Fixed transformed / 256	256	3.083	0.66	0.009	0.683
Fixed transformed / 512	512	2.106	0.55	0.014	1.340
Fixed transformed / 1024	1024	1.353	0.64	0.021	2.538
Fixed transformed / 2048	2048	0.887	0.61	0.033	5.255
Adaptive transformed / 32	77	3.391	–	0.009	0.608
Adaptive transformed / 64	172	1.851	0.76	0.016	1.295
Adaptive transformed / 128	369	0.980	0.83	0.030	2.833
Adaptive transformed / 256	775	0.565	0.74	0.051	6.716
Adaptive transformed / 512	1619	0.330	0.73	0.088	14.612
Adaptive transformed / 1024	3363	0.188 †	0.77	0.154	31.237

Table 1 shows adaptive observed orders between 0.73 and 0.83 over the tested refinements. These finite-level measurements do not establish the theoretical order-one asymptote. The adaptive method improves on fixed transformed Milstein in Brownian observation cost; its advantage over the much cheaper original Euler is small and tolerance dependent in this experiment.

For example, the first tested configurations with estimated RMS below 10^{-3} are Euler with 512 steps, fixed transformed Milstein with 2,048 steps, and adaptive transformed Milstein with $\delta = 1/128$. Their RMS estimates (Monte Carlo SE) are respectively 0.000849 (0.000019), 0.000887 (0.000021), and 0.000980 (0.000022). Their CPU times are 0.212, 5.255, and 2.833 ms per path. This is a comparison using point-estimate selec-



tion from a finite grid, not a confidence-certified attainment of the tolerance. The adaptive configuration uses about 737 observations versus Euler’s 1,024, but takes about 13.4 times as much CPU time.

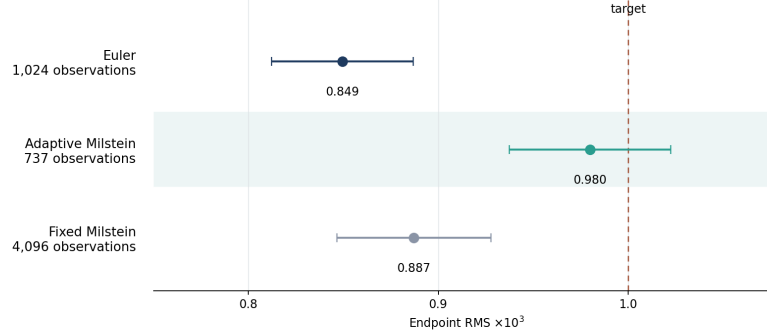
Timings use production Gaussian increments, without reference-grid lookups, on one CPU process after compilation. Four distinct blocks of 64 paths are timed per configuration. Uniform draws, the additional interior Gaussian observations, geometry, inversions, and turnover evaluation are included. They are implementation-specific timings, not an arithmetic lower bound.

Table 2: Fine-reference policy estimates and Monte Carlo standard errors. The same 1,024 Brownian skeletons are used across thresholds.

r	Mean risk	Risk SE	Mean turnover	Turnover SE
0.25	0.15653	0.00622	0.23304	0.00339
0.40	0.20469	0.00699	0.13288	0.00335
0.55	0.25689	0.00794	0.07295	0.00281

Endpoint accuracy at the selected observation budgets

$r = 0.40$ | Fastest tested levels with RMS point estimate below 10^{-3}



Error bars: approximate 95% Monte Carlo intervals from 1,024 paths.

The adaptive interval crosses the target; numerical-reference bias is not included.

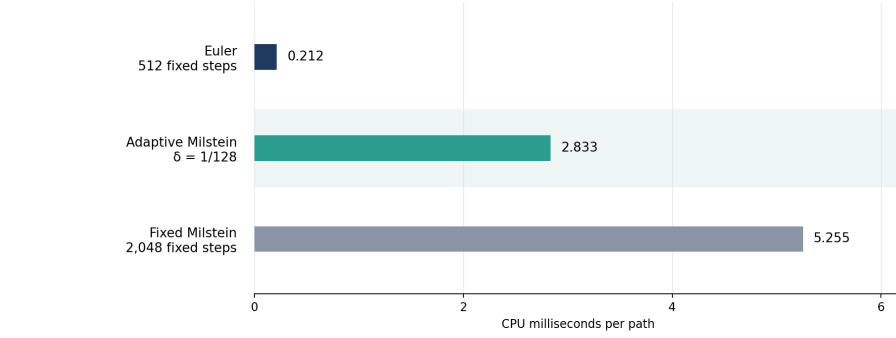
Observations include the turnover query. Both Milstein methods use transformed coefficients.

Figure 3: Endpoint accuracy and mean Brownian observation counts at $r = 0.40$. Each row selects the fastest tested configuration with RMS point estimate below 10^{-3} . Error bars are approximate pointwise 95% Monte Carlo intervals; the adaptive interval crosses the target, and numerical-reference bias is not included. Both Milstein methods use transformed coefficients.



CPU time at the selected accuracy levels

Adaptive Milstein takes 13.4× Euler CPU time at these tested levels.

 $r = 0.40$. Same configurations as the endpoint-accuracy figure.

256 fresh paths per configuration; compilation and reference construction excluded.

Turnover queries and inversions included. Both Milstein methods use transformed coefficients.

Figure 4: Production CPU time for the same selected configurations as Figure 3. Adaptive Milstein takes about 13.4 times Euler’s CPU time at these levels. Timings include turnover queries and inversions, but exclude compilation and numerical-reference construction.

Table 2 gives the application result: enlarging the no-trade region reduces estimated turnover and increases estimated terminal tracking risk in this experiment. Fine–coarse mean risk differences range in magnitude from 1.64×10^{-6} to 4.92×10^{-6} ; mean turnover differences range from 0.645×10^{-6} to 2.10×10^{-6} . These are much smaller than the Monte Carlo standard errors of the policy means.

6. Validation and scope

Analytic Jacobians and Hessians were checked against independent automatic differentiation at 72 off-interface points, with maximum absolute differences 5.56×10^{-16} and 8.89×10^{-16} . The inverse round-trip error was below 2.73×10^{-13} . Two-sided transformed-drift traces, the polynomial cutoff bounds, the global inverse certificates, and conditional bridge covariance were also checked. With control set to zero, all three schemes reproduce ΣW_1 to 3.56×10^{-16} and have zero turnover. These checks validate local formulas and implementation features; they do not certify the convergence proof or originality.

The inverse is computed by damped Newton iteration with residual tolerance $2 \times 10^{-13}(1 + \|y\|)$ and an explicit failure condition. Across the reported simulations the maximum residual is below 4.20×10^{-13} . The analytic result uses exact inverses; accumulated finite-precision and inexact-inversion error have not been given a separate global bound.

This model omits self-financing constraints, wealth dynamics, estimation of the target, stochastic volatility, and an optimization objective. Its value here is an explicit, genuinely two-dimensional application of the interface theorem with a discontinuous path functional. The proposed paper should be assessed on that mathematical result and its proof, not on a claim of faster practical portfolio simulation.



7. Assumptions and algorithm

Let W be a d -dimensional Brownian motion, $d \geq 2$, on a filtered probability space satisfying the usual conditions. Consider

$$dZ_t = a(Z_t) dt + B(Z_t) dW_t, \quad Z_0 = z_0 \in \mathbb{R}^d, \quad 0 \leq t \leq 1.$$

Assume the following.

1. $a : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $B : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ are bounded and globally Lipschitz. There is $\kappa > 0$ such that the smallest singular value of $B(x)$ is at least κ for every x .
2. Θ is a compact separating smooth hypersurface. For some $\rho > 0$ there is $g \in C_b^2(\mathbb{R}^d)$ such that

$$\{|g| < \rho\} = \{d_\Theta < \rho\}, \quad |g(x)| = d_\Theta(x), \quad \|Dg(x)\| = 1 \quad (d_\Theta(x) < \rho),$$

where $d_\Theta(x) = \inf_{y \in \Theta} \|x - y\|$. A circle has this property by smoothly saturating its signed distance outside a tubular neighbourhood.

3. On $\mathbb{R}^d \setminus \Theta$, a and B are C^2 , and their first and second derivatives are bounded uniformly over the complementary regions. Their derivative traces may jump on Θ .
4. Writing B_j for column j , we have $DB_j B_k = DB_k B_j$ off Θ .

The SDE has a unique strong solution by global Lipschitz continuity. Set derivatives of a and B equal to zero on Θ when defining the numerical method, and put

$$H_{jk}(x) = DB_j(x) B_k(x), \quad H_{jk}|_\Theta = 0.$$

These functions are bounded and satisfy $H_{jk} = H_{kj}$ everywhere.

For small $\delta > 0$, write

$$L = \log(1/\delta), \quad r_1 = \sqrt{\delta} L^2, \quad r_2 = \delta L^4, \quad h_* = \delta^2 L^4.$$

Fix δ_0 sufficiently small that $L \geq 1$ and $r_2 \leq r_1 < \rho/4$ for $0 < \delta \leq \delta_0$. Use the step function

$$h(x) = \begin{cases} \delta, & d_\Theta(x) \geq r_1, \\ d_\Theta(x)^2/L^4, & r_2 < d_\Theta(x) < r_1, \\ h_*, & d_\Theta(x) \leq r_2. \end{cases} \quad (7.1)$$

This is the three-region mesh from the scalar adaptive method [1], now applied to distance from a hypersurface. In particular, $h_* \leq h \leq \delta$.



Starting with $\tau_0 = 0$ and $Y_0 = z_0$, let $h_i = h(Y_{\tau_i})$ and $\tau_{i+1} = \tau_i + h_i$. For $s \in [\tau_i, \tau_{i+1}]$, define

$$\begin{aligned} Y_s &= Y_{\tau_i} + a(Y_{\tau_i})(s - \tau_i) + B(Y_{\tau_i})\Delta W_s \\ &\quad + \frac{1}{2} \sum_{j,k=1}^d H_{jk}(Y_{\tau_i})(\Delta W_s^j \Delta W_s^k - \mathbf{1}_{j=k}(s - \tau_i)), \end{aligned} \quad (7.2)$$

where $\Delta W_s = W_s - W_{\tau_i}$. The symmetry of H makes this the commutative Milstein interpolant.

Let $N = \min\{n \geq 1 : \tau_n \geq 1\}$. The algorithm observes W at $\tau_1, \dots, \tau_{N-1}, 1$ and evaluates the last formula at time 1; it does not require an observation beyond 1. Full-step extensions beyond 1 are used only in estimates. Each τ_i is a stopping time, h_i is $\mathcal{F}_{\tau_i}$ -measurable, and

$$N \leq \lceil h_*^{-1} \rceil, \quad \sum_{i < N} h_i = \tau_N \leq 1 + \delta. \quad (7.3)$$

Write $\eta(s) = \tau_i$ for $\tau_i \leq s < \tau_{i+1}$, and $U_s = Y_s - Y_{\eta(s)}$.

Theorem 1. Under the stated assumptions there are constants $C, \delta_0 > 0$, independent of δ , such that

$$\sup_{0 \leq t \leq 1} \mathbb{E} \|Z_t - Y_t\|^2 \leq C\delta^2, \quad \mathbb{E} N \leq C\delta^{-1} \quad (0 < \delta \leq \delta_0).$$

The first estimate is a supremum of expectations, not an expectation of a path supremum. Only terminal-time output is asserted computable from the counted observations.

Proof. The complete proof is given in the next four sections: increment estimates, weighted occupation and cost, interface-crossing control, and the final Grönwall estimate. The proof concludes immediately before the auxiliary constant-drift example. $\square$

8. Increment estimates and ellipticity of the interpolant

Throughout, C may change from line to line and depend on the fixed coefficients, dimension and surface, but not on δ . Let

$$M_i = \sup_{0 \leq u \leq h_i} \|W_{\tau_i+u} - W_{\tau_i}\|.$$

Bounded coefficients in (7.2) give

$$\sup_{0 \leq u \leq h_i} \|Y_{\tau_i+u} - Y_{\tau_i}\| \leq C(h_i + M_i + M_i^2). \quad (8.1)$$

Conditional on $\mathcal{F}_{\tau_i}$, Brownian strong Markov and scaling give, for every fixed $p \geq 2$,



$$\mathbb{E} \left[\int_0^{h_i} \|Y_{\tau_i+u} - Y_{\tau_i}\|^p du \mid \mathcal{F}_{\tau_i} \right] \leq C_p h_i^{1+p/2}. \quad (8.2)$$

Summing over $i < N$, using that $\{\tau_i < 1\}$ is known at τ_i and using (7.3), proves

$$\mathbb{E} \int_0^1 \|U_s\|^p ds \leq C_p \delta^{p/2}. \quad (8.3)$$

All conditioning here is at actual stopping times, not at the random last grid point before a deterministic time.

The interpolant is an Itô process with drift $a(Y_{\eta(s)})$ and diffusion columns

$$\Sigma_{j,s} = B_j(Y_{\eta(s)}) + \sum_k H_{jk}(Y_{\eta(s)}) \Delta W_s^k.$$

Choose a fixed $m_0 > 0$ small enough that $\|\Sigma_s - B(Y_{\eta(s)})\|_{\text{op}} \leq \kappa/2$ whenever $\|\Delta W_s\| \leq m_0$. Let E_δ be the event that $M_i \leq m_0$ for every $i < N$. Conditional Brownian maximal tails and the deterministic count yield

$$\mathbb{P}(E_\delta^c) \leq C[h_*^{-1}]e^{-c/\delta} \leq C_q \delta^q \quad \text{for every fixed } q > 0. \quad (8.4)$$

On E_δ the smallest singular value of Σ_s is at least $\kappa/2$ for all $s \leq 1$. This argument uses global uniform ellipticity and bounded H ; normal nondegeneracy alone is not substituted for these assumptions.

9. Weighted occupation and expected cost

Lemma 1. For every $q > 0$, $0 < \varepsilon < \rho$, and bounded nonnegative measurable $f : [0, \varepsilon] \rightarrow \mathbb{R}$,

$$\mathbb{E} \int_0^1 f(d_\Theta(Y_s)) \mathbf{1}_{\{d_\Theta(Y_s) < \varepsilon\}} ds \leq C \int_0^\varepsilon f(r) dr + C_q \delta^q \|f\|_\infty. \quad (9.1)$$

Proof. Put $V = g(Y)$. Itô's formula gives drift

$$A_s = Dg(Y_s)a(Y_{\eta(s)}) + \frac{1}{2} \text{tr}(D^2g(Y_s)\Sigma_s\Sigma_s^\top)$$

and quadratic variation $d\langle V \rangle_s = \|Dg(Y_s)\Sigma_s\|^2 ds$. Bounded g, Dg, D^2g, a, B, H and Brownian maximal moments imply

$$\mathbb{E} \int_0^1 (|A_s| + \|\Sigma_s\|^2) ds \leq C.$$

For example $\|\Delta W_s\| \leq 2 \sup_{u \leq 1} \|W_u\|$ gives a bound independent of N . Thus the argument does not presuppose the expected-cost estimate. Tanaka's formula gives $\sup_{v \in \mathbb{R}} \mathbb{E} L_1^v(V) \leq C$: the absolute-value difference is bounded by $2\|g\|_\infty$, the finite-variation term has bounded



expectation, and the martingale has expectation zero. On E_δ and the tube in the lemma,

$$d\langle V \rangle_s \geq \kappa^2 ds/4.$$

Bound the contribution of E_δ^c by $\|f\|_\infty \mathbb{P}(E_\delta^c)$. For the remaining contribution, apply the occupation formula to V and integrate the uniform local-time bound over $(-\varepsilon, \varepsilon)$. Equation (8.4) proves (9.1). $\square$

Define F_δ by requiring, for each $i < N$,

$$\sup_{u \leq h_i} \|Y_{\tau_i+u} - Y_{\tau_i}\| \leq \frac{1}{2} \sqrt{h_i} L^2.$$

For sufficiently small δ_0 , (8.1) shows that this holds whenever $M_i \leq c_0 \sqrt{h_i} L^2$, for a fixed sufficiently small $c_0 > 0$. Indeed $\sqrt{h_i} L^2 \leq r_1 \rightarrow 0$. Therefore

$$\mathbb{P}(F_\delta^c) \leq C[h_*^{-1}]e^{-cL^4} \leq C_q \delta^q \quad (q > 0). \quad (9.2)$$

On an inner step this increment threshold is $r_2/2$, on a middle step it is $d_\Theta(Y_{\tau_i})/2$, and on an outer step it is $r_1/2 \leq d_\Theta(Y_{\tau_i})/2$.

The pathwise identity for the number of completed steps gives

$$N \leq 1 + \int_0^1 h(Y_{\eta(s)})^{-1} ds.$$

Outer steps contribute at most δ^{-1} . On F_δ , inner grid positions imply $d_\Theta(Y_s) < 2r_2$, while middle grid positions give

$$\frac{1}{2} d_\Theta(Y_{\eta(s)}) \leq d_\Theta(Y_s) \leq \frac{3}{2} d_\Theta(Y_{\eta(s)}).$$

Consequently, with q chosen large enough,

$$\begin{aligned} \mathbb{E}N &\leq 1 + \delta^{-1} + h_*^{-1}(Cr_2 + C_q \delta^q) \\ &\quad + CL^4 \left(\int_{r_2/2}^{3r_1/2} r^{-2} dr + C_q \delta^q r_2^{-2} \right) + h_*^{-1} \mathbb{P}(F_\delta^c). \end{aligned}$$

Here Lemma 1 is applied to the constant function and to r^{-2} truncated below at $r_2/2$. Since $r_2/h_* = L^4/r_2 = \delta^{-1}$, taking, for example, $q = 4$ proves $\mathbb{E}N \leq C\delta^{-1}$.

10. Interface-crossing remainder

Let $S(x, y)$ denote the event that the straight segment joining x to y meets Θ . This includes a starting point on Θ and includes segments with endpoints on the same side that pass through the surface.



Lemma 2. The numerical process satisfies

$$\mathbb{E} \int_0^1 \|U_s\|^2 \mathbf{1}_{S(Y_{\eta(s)}, Y_s)} ds \leq C\delta^2. \quad (10.1)$$

Proof. On F_δ , a segment can meet Θ only on an inner step, since its length on any other step is less than the distance of its initial point to Θ . Conditional estimate (8.2) with $p = 2$ and the already proved cost estimate give

$$\mathbb{E} \sum_{i < N} \mathbf{1}_{\{d_\Theta(Y_{\tau_i}) \leq r_2\}} \int_0^{h_i} \|Y_{\tau_i+u} - Y_{\tau_i}\|^2 du \leq Ch_*^2 \mathbb{E} N \leq C\delta^3 L^8.$$

The exceptional event is bounded, by Cauchy–Schwarz, Jensen and (8.3), by

$$\mathbb{E} \left[\mathbf{1}_{F_\delta^c} \int_0^1 \|U_s\|^2 ds \right] \leq \left(\mathbb{E} \int_0^1 \|U_s\|^4 ds \right)^{1/2} \mathbb{P}(F_\delta^c)^{1/2} \leq C\delta \mathbb{P}(F_\delta^c)^{1/2}.$$

Use (9.2) with $q = 4$ and the boundedness of δL^8 on $(0, \delta_0]$ to obtain (10.1). $\square$

For $F = a$ or a column of B , bounded derivatives off the surface and global Lipschitz continuity imply

$$\|F(y) - F(x) - DF(x)(y - x)\| \leq C\|y - x\|^2 + C\|y - x\| \mathbf{1}_{S(x,y)}. \quad (10.2)$$

If the segment avoids the surface, integrate the derivative along that segment and use the bounded second derivative. Otherwise use Lipschitz continuity and bounded first derivative. The prescribed zero derivative on the surface is compatible with the second bound.

11. Strong-error closure on the random mesh

The increment after subtracting its leading Brownian part is

$$K_s = U_s - B(Y_{\eta(s)}) \Delta W_s.$$

By the explicit polynomial formula and conditional Brownian moments, the same stopped-step summation as in (8.2) yields

$$\mathbb{E} \int_0^1 \|K_s\|^2 ds \leq C\delta^2.$$

Combining this with (10.2), (8.3) for $p = 4$ and Lemma 2 gives

$$\mathbb{E} \int_0^1 (\|R_s^a\|^2 + \|R_s^B\|_F^2) ds \leq C\delta^2, \quad (11.1)$$



where

$$\begin{aligned} R_s^a &= a(Y_s) - a(Y_{\eta(s)}) - Da(Y_{\eta(s)})B(Y_{\eta(s)})\Delta W_s, \\ R_{j,s}^B &= B_j(Y_s) - B_j(Y_{\eta(s)}) - \sum_k H_{jk}(Y_{\eta(s)})\Delta W_s^k. \end{aligned}$$

It remains to control the leading drift term. Put $A_i = Da(Y_{\tau_i})B(Y_{\tau_i})$. For every fixed deterministic $t \leq 1$, stochastic Fubini, step by step, gives

$$\begin{aligned} J_t &= \int_0^t Da(Y_{\eta(s)})B(Y_{\eta(s)})\Delta W_s ds \\ &= \sum_i \int_{\tau_i}^{\tau_{i+1} \wedge t} (\tau_{i+1} \wedge t - u) A_i dW_u, \end{aligned} \tag{11.2}$$

with terms having $\tau_i \geq t$ omitted. This integrand is adapted: $\tau_{i+1} = \tau_i + h_i$ is already known at τ_i . Its norm is at most $C\delta$, and thus

$$\mathbb{E}\|J_t\|^2 \leq C\delta^2. \tag{11.3}$$

We do not condition on $\eta(t)$ as though it were a stopping time, and we do not use (11.2) to assert a path-supremum estimate.

Subtract the interpolant equation from the exact SDE. For $e_t = Z_t - Y_t$,

$$\begin{aligned} e_t &= \int_0^t (a(Z_s) - a(Y_s))ds + \int_0^t (B(Z_s) - B(Y_s))dW_s \\ &\quad + J_t + \int_0^t R_s^a ds + \int_0^t R_s^B dW_s. \end{aligned}$$

Global Lipschitz bounds, Cauchy–Schwarz, Itô isometry, (11.1) and (11.3) imply

$$\mathbb{E}\|e_t\|^2 \leq C \int_0^t \mathbb{E}\|e_s\|^2 ds + C\delta^2.$$

Grönwall proves the first assertion of Theorem 1; the expected-cost assertion was proved earlier. This completes its proof.

12. An auxiliary constant-drift example

Let $u, v \in \mathbb{R}^2$ be fixed distinct vectors and let

$$dX_t = \mu(X_t)dt + dW_t, \quad \mu(x) = u \text{ } (|x| < 1), \quad \mu(x) = v \text{ } (|x| > 1).$$

Values on the circle do not affect the solution. The choice $u = (1, 0)$ and $v = (0, 1)$ is not



radially symmetric.

Let $r(x) = |x| - 1$, $\alpha = (u - v)/2$, and choose $\chi \in C_c^\infty((-1, 1))$ equal to one near zero. For a fixed small $w < 1/2$, define

$$G(x) = x + \alpha r(x)|r(x)|\chi(r(x)/w)$$

in the annulus $|r(x)| < w$, and $G(x) = x$ elsewhere. The perturbation has derivative bounded by $C|\alpha|w$. Choose w so that this bound is below $1/2$.

Then G is globally $C^{1,1}$, is smooth on each side of the circle, and is a global bi-Lipschitz bijection: writing $G = \text{id} + k$, the equation $G(x) = z$ is the contraction equation $x = z - k(x)$. Its Jacobian has singular values in $[1/2, 3/2]$, and G fixes the circle pointwise, with $DG = I$ there. Injectivity implies that no point off the circle maps onto it.

For $z = G(x)$ off the circle define

$$a(z) = DG(x)\mu(x) + \frac{1}{2}\Delta G(x), \quad B(z) = DG(x).$$

As $r \rightarrow 0^\pm$, the Laplacian limits are $\Delta G \rightarrow \pm 2\alpha$. Hence the two traces of the transformed drift coincide:

$$v + \alpha = u - \alpha = (u + v)/2.$$

Extend a by this common value on the circle. The coefficients a, B are bounded and globally Lipschitz, and are C^2 with uniformly bounded first and second derivatives on the two open sides. These facts follow from the explicit smooth formulas on the compact annulus and the constant formulas outside it. Bounded derivatives along line segments, split at their finitely many intersections with the circle, give the global Lipschitz bound.

The diffusion is uniformly elliptic. Moreover, off the circle,

$$DB_j(z)B_k(z) = D^2G(x)[e_k, e_j] = D^2G(x)[e_j, e_k] = DB_k(z)B_j(z).$$

The generalized Itô formula for $C^{1,1}$ transformations therefore gives $Z = G(X)$ solving the transformed SDE. It can also be obtained by mollification; the second derivatives are bounded, and the diffusion spends zero Lebesgue time on the circle. The inverse transformation gives the corresponding existence and uniqueness statement. This is the standard drift-removing construction, not a proposed new transformation [3], [5].

Corollary 2. For the explicit circle model, $\widehat{X}_1 = G^{-1}(Y_1)$ satisfies

$$(\mathbb{E}|X_1 - \widehat{X}_1|^2)^{1/2} \leq C\delta, \quad \mathbb{E}N \leq C\delta^{-1}.$$

Proof. Apply Theorem 1 and the Lipschitz bound for G^{-1} . The inverse evaluation requires no extra Brownian observations. $\square$



There is a concrete reason why a theorem assuming globally smooth diffusion cannot simply be used after transforming. At a point on the circle with unit normal n , the two Hessian limits are

$$D^2G_{\pm}[p, q] = \pm 2\alpha(n \cdot p)(n \cdot q).$$

Thus DB_j has jump $4\alpha n_j n^\top$. For $u \neq v$, at least one diffusion-column derivative jumps at every point of the circle. The transformation removes the drift jump but does not remove the derivative interface of the diffusion.

13. What the closest theorems do and do not establish

Scalar adaptive approximation [1]

The cited result numbered 2 proves order one and expected cost $O(\delta^{-1})$ for a scalar equation with finitely many drift discontinuities. Its three-region mesh is adopted here. The state space in that theorem is $\mathbb{R}$; its statement does not apply to the coupled planar circle model. This does not diminish the substantial methodological debt to its proof.

Multidimensional adaptive Euler [2]

The paper supplies hypersurface occupation theory and an adaptive Euler method of order one half up to logarithmic terms in average cost. The cited result numbered 2.8 concerns a general Itô process with suitable boundedness and normal nondegeneracy. Those are familiar ingredients, but the stated approximation result does not provide the order-one Milstein error-and-cost pair derived above.

Multidimensional Milstein [3]

The cited result numbered 2 proves fixed-grid order $3/4 - \epsilon$ under broader coefficient assumptions, with Brownian evaluations alone in the commutative case. The cited result numbered 1 has a faster $1 - \epsilon$ case when the diffusion derivative has no exceptional hypersurface. The explicit derivative jump just calculated prevents use of that special case here. The paper's general theorem concerns deterministic meshes, so it also cannot be applied by substituting a random step count.

Adaptive jump-diffusion extension [4]

The cited result numbered 4.3 supplies order one and expected-cost bounds, but the setting explicitly has scalar state and scalar coefficient functions. Setting the jump term to zero does not extend its state dimension.

General adaptive convergence [6]

This work treats adaptive meshes that may depend on future noise and proves pathwise convergence for sufficiently smooth vector fields. Its abstract and stated scope do not supply the present nonsmooth-interface order-one error-and-cost result. Only the abstract and overview were used for this exclusion; no stronger claim about its entire theorem family is made.

Assessment



None of the four directly compared approximation theorems states the result of this note or yields it by direct substitution. Combining and adapting their techniques is precisely what the proof does. The contribution, if not already recorded elsewhere, is a modest multidimensional extension under stronger boundedness and ellipticity assumptions, with an explicit genuine curved-interface example. It should not be marketed as a new general theory of adaptive approximation.

14. Limits, checks and remaining work

- This note contains a full analytic proof under its stated assumptions and a direct verification of an explicit discontinuous-drift application. It does not prove the result for arbitrary unbounded piecewise-smooth drift, merely normal ellipticity, noncommutative noise, or all L^p norms.
- No lower bound has been proved; “order one” is an achieved upper rate, not a claim of optimality.
- Expected Brownian observation cost is not total runtime. The geometric transform, its derivatives, and its inverse are treated as exact evaluation oracles. Numerical inversion errors and their accumulation require separate analysis for an implementation.
- The logarithmic mesh thresholds are asymptotic and can force very small admissible δ . No practical speedup is inferred from the theorem.
- Proof review must distinguish $\sup_t \mathbb{E}|e_t|^2$ from $\mathbb{E} \sup_t |e_t|^2$, and must not replace conditioning at τ_i by conditioning at $\eta(t)$.
- The closest theorem statements and relevant proof sections were checked. No matching theorem was found in the targeted search, but the search does not prove global absence. No author contact or independent referee review has occurred.
- The main text reports a separate rebalancing experiment. Those numerical checks do not certify this proof or its novelty.



References

- [1] Yaroslavtseva, L. (2022). An adaptive strong order 1 method for SDEs with discontinuous drift coefficient. *Journal of Mathematical Analysis and Applications*. arXiv:2102.13155
- [2] Neuenkirch, A., Szölgényi, M., & Szpruch, L. (2019). An adaptive Euler-Maruyama scheme for stochastic differential equations with discontinuous drift and its convergence analysis. *SIAM Journal on Numerical Analysis*. arXiv:1802.04521
- [3] Rauhögger, C. (2025). Milstein-type methods for strong approximation of systems of SDEs with a discontinuous drift coefficient. arXiv:2505.15509
- [4] Schwarz, V. (2026). Strong order 1 adaptive approximation of jump-diffusion SDEs with discontinuous drift. arXiv:2504.09452
- [5] Leobacher, G., & Szölgényi, M. (2017). A strong order 1/2 method for multidimensional SDEs with discontinuous drift. *Annals of Applied Probability*. arXiv:1512.02807
- [6] Foster, J., & Jelinčíč, A. (2023). On the convergence of adaptive approximations for stochastic differential equations. arXiv:2311.14201
- [7] El Bernoussi, R., & Rockinger, M. (2023). Rebalancing with transaction costs: theory, simulations, and actual data. *Financial Markets and Portfolio Management*. <https://doi.org/10.1007/s11408-022-00419-6>
- [8] Dimensional Fund Advisors (2026). *Portfolio Rebalancing: All at Once or Little by Little?*.



Zamrik Quant Research
The research arm of Zamrik Quant Lab

About the Desk

Stochastic Analysis Desk

Zamrik Quant Research

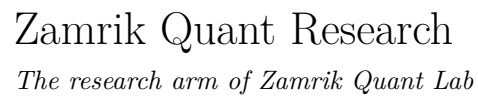
How random processes behave — Itô calculus, diffusions, local time, Fokker–Planck — in theory and in application.

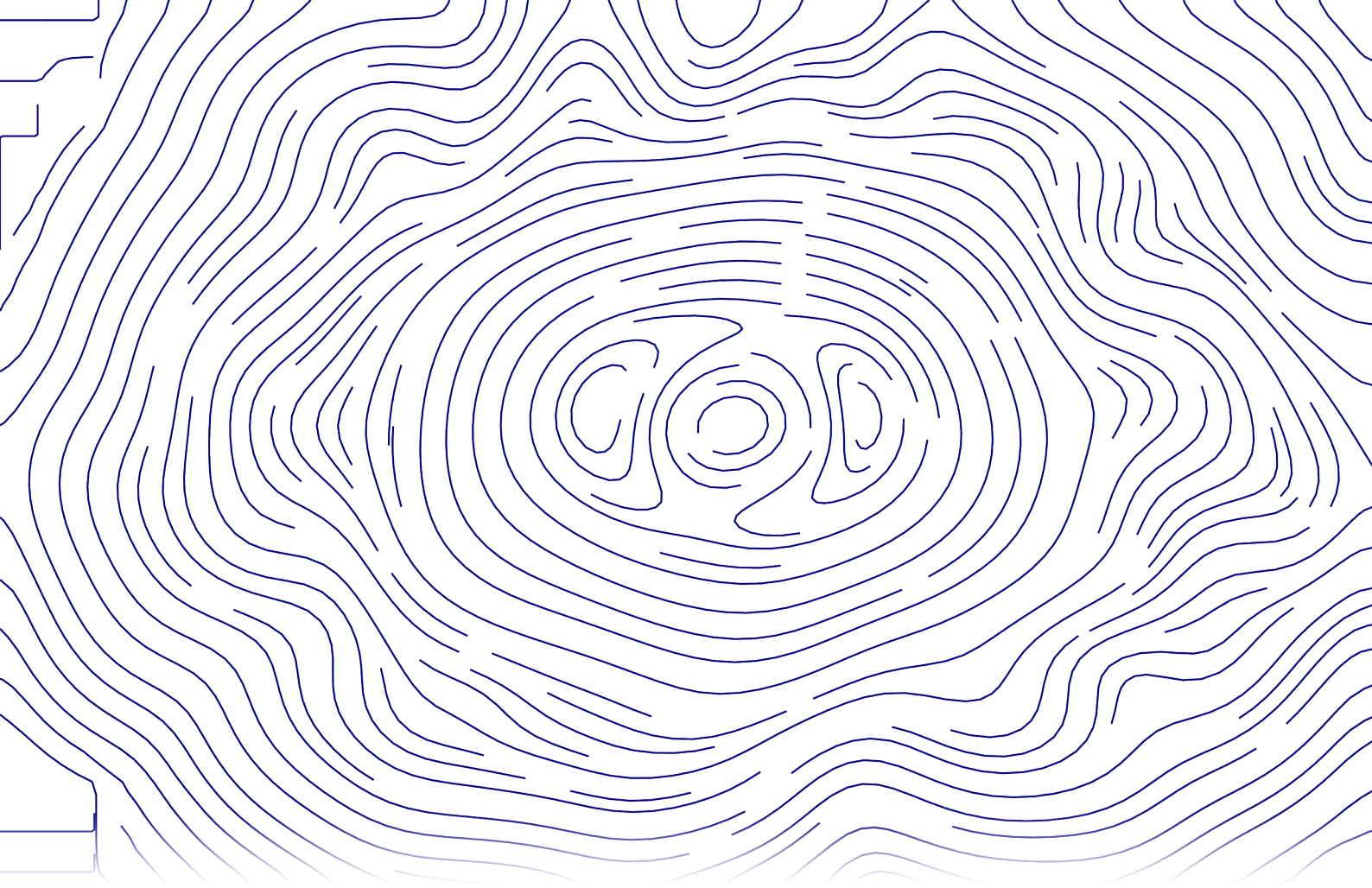
Stochastic analysis is the grammar of every other desk: a price, an order queue, a population or a value function is a functional of a random path, and what can be said about it depends on what can be proved about the path. The Stochastic Analysis desk works on those proofs and on the exact formulas they produce.

Its core questions are the classical ones, asked with care: when a diffusion first reaches a boundary, fixed or moving; how long it spends at each level before it does; and what happens at the boundary itself — reflection, killing, a sticky or skewed point, a reset to the start. The answers come through Itô calculus and martingale methods, local time and Tanaka’s formula, Ray–Knight theorems, time changes, and the Feynman–Kac and Fokker–Planck equations that turn questions about paths into equations for expectations and densities. Wherever the problem allows, the desk derives the law in closed form and then checks it against an independently built simulation.

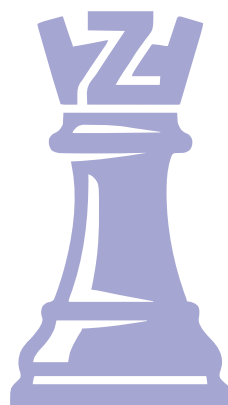
The results travel. The same calculus corrects a barrier that is checked once a day, explains why most stocks lose to the average stock, separates real correlation from the noise of random matrices, sets the acceptance rate of a Metropolis sampler, and models a neuron’s membrane and a day’s rainfall. The desk also studies the numerics that decide whether a simulated path tells the truth: the strong and weak convergence of Euler–Maruyama and Milstein schemes, Brownian-bridge corrections for barriers monitored in discrete time, and the error left when a continuous process is sampled on a grid.

The desk is supervised by **Dr. T. Zamrik**, Head of Zamrik Quant Research, PhD, Imperial College London (2011).





Hamiltonian flow of the Lévy-stable density: $\dot{\mathbf{x}} = J \nabla f(\mathbf{x})$, $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $\varphi(\xi) = \exp[i\mu\xi - |\sigma\xi|^\alpha(1 - i\beta \operatorname{sgn}(\xi)\Phi)]$



ZAMRIK QUANT RESEARCH

The research arm of Zamrik Quant Lab



WP-2026-16039117
zamrik.com