**Journal of Mathematical Finance and Risk Management**
ISSN: 3080-9983 (Print), ISSN: 3080-9991 (Online) | Vol. 1, No. 3, 2026

# Arbitrage-Constrained Diffusion-Inspired Denoising for Sparse Implied Volatility Surface Reconstruction

**Henry Brooks, Adrian Mitchell** *

Department of Computer Science and Engineering, University of California, Riverside, California, USA

* **Corresponding author**: (Email: 59403023@ucr.edu)

**Abstract:** Sparse and noisy option quotes complicate implied-volatility surface construction, and statistical smoothing does not by itself enforce no-arbitrage structure. Building on arbitrage-regularized surface reconstruction, this paper develops a diffusion-inspired, noise-level-conditioned residual denoiser followed by an alternating shape-restoration step for normalized European call-price surfaces. On a risk-neutral benchmark of 1,280 surfaces on a 12×16 maturity–strike grid with 55% missing observations and additive noise, the method attains MAE 0.001886 and RMSE 0.002756 in normalized call-price units on 160 held-out surfaces and reduces measured calendar, strike-monotonicity, strike-bound and butterfly violations to zero at the stated tolerance, whereas the most accurate unconstrained denoiser (MAE 0.001679) leaves widespread butterfly violations. Implied-volatility diagnostics and a paired surface-level bootstrap further characterize the accuracy gains. The results quantify the accuracy–feasibility trade-off of combining learned denoising with explicit shape restoration and provide a transparent basis for arbitrage-consistent surface completion on market quotes.

**Keywords:** Implied volatility surface, option-price reconstruction, static arbitrage, diffusion-inspired denoising, convexity projection, synthetic benchmark.

## 1. Introduction

An implied volatility surface summarizes a family of European option prices across strike and maturity. Inference is complicated by irregular strike grids, bid–ask noise, sparse wings, and the nonlinear translation between price and implied volatility. Price-based modeling has an essential advantage: necessary static no-arbitrage properties can be checked through linear inequalities in call prices rather than through nonlinear volatility derivatives [1–5]. This fact is particularly useful when only a minority of grid locations carry observations.

Classical smoothers and parametric surface models encode varying degrees of structure [6–11]. More recent neural methods correct parametric fits or model entire volatility surfaces [12–14]. Ge et al. [15] showed that a GAN trained on AAPL option prices with no-arbitrage constraints and regularization reduces pricing errors relative to the SVI model, demonstrating that financial validity can be built into data-driven surface reconstruction. Their adversarial design draws on generative modeling [14, 24], and their explicit treatment of calendar and butterfly violations motivates the separate financial-validity axis used in our evaluation. Yet the presence of a financial penalty in an optimization objective does not alone certify an inequality on an evaluated price grid. We therefore study how far a simple, fast denoiser can be complemented by an explicit shape-restoration step.

Diffusion generative modeling connects noise perturbations to learned denoisers [16–21]. Our objective is targeted reconstruction rather than density estimation: conditioned on one sparse noisy price grid, obtain one accurate and grid-feasible reconstruction at low CPU cost. Accordingly, a noise-conditioned residual convolutional network supplies the learned prior, while a deterministic alternating operation restores the three principal sampled-grid shape constraints. We call this a diffusion-inspired denoiser to distinguish it from a trained score-based reverse-time SDE or DDPM sampler. Conditional generative modeling for credit-card fraud provides a complementary financial-data example of combining learned data distributions with sequential prediction under a controlled evaluation design [25].

The contributions are: (i) an explicit, reproducible sparse-quote reconstruction pipeline that combines a learned conditional denoiser with grid-level financial shape enforcement; (ii) a controlled synthetic generator whose ground truth obeys the price inequalities by construction; (iii) quantitative ablations separating denoising from projection and reporting violations rather than only fitting error; and (iv) sensitivity checks at different observation-noise levels. The contribution lies in the task-specific integration of denoising and alternating shape restoration and in its measured accuracy–feasibility trade-off.

## 2. Background and Related Work

### 2.1. Static No-Arbitrage in Normalized Call Space

For a European call with spot $S_0=1$, zero discounting and normalized strike K, let $C(K, T)=E[(S_T-K)^+]$. At fixed maturity, the payoff implies that C is nonnegative, nonincreasing and convex in K, with slope bounded between −1 and 0 [1, 3]. Convexity implies a nonnegative risk-neutral density in the distributional sense [3]. Under the zero-rate martingale specification used in our synthetic benchmark, a call price is also nondecreasing with maturity for each fixed strike. This calendar inequality follows from convex-order properties; with realistic nonzero funding or discrete dividends, a carefully adjusted forward/discounted representation is required.

$$-1 \le [C(K_{j+1},T_i)-C(K_j,T_i)] \ / (K_{j+1}-K_j) \le 0 \qquad (1)$$

$$s_{i,j+1} - s_{i,j} \ge 0; \qquad C(K_j,T_{i+1})-C(K_j,T_i) \ge 0 \qquad (2)$$

The restrictions in (1)–(2) are necessary conditions on

sampled options, not a proof of continuous-domain static arbitrage freedom. They avoid the error of treating a second derivative with respect to log-strike as interchangeable with a second derivative with respect to the actual strike. The geometric discussion of SVI [9], its generalization [10], and constrained smoothing [8] further motivates assessing both strikes and maturities.

### 2.2. Generative and Data-Driven Reconstruction

Cont and coauthors documented low-dimensional and dynamic patterns in historical volatility surfaces [6, 7]. Fengler [8] provided arbitrage-free spline smoothing, while Gatheral and Jacquier [9] developed feasible SVI parameterizations. Neural smoothing [12] and generative variational modeling [13] exploit additional flexibility. The arbitrage-regularized GAN of Ge et al. reports gains over SVI on AAPL options; we carry its separation of reconstruction accuracy from arbitrage diagnostics into a transparent price-space experiment in which violations are counted exactly on the evaluation grid.

Denoising autoencoders have a score-matching interpretation [16]. Sohl-Dickstein et al. [17], Ho et al. [18] and Song et al. [19, 20] develop more comprehensive diffusion-based generation and sampling, while Nichol and Dhariwal [21] analyze faster practical variants. The small model here borrows multilevel perturbation training, not full stochastic generative likelihood or reverse-chain sampling. Convolutional locality is also related to image-model architectures [22], though our network is much smaller than a U-Net.

## 3. Task Formulation and Data Generation

### 3.1. Observation Operator

Let $P \in R^{M \times N}$ denote a complete normalized call-price array on M=12 maturities $T \in [0.15, 1.50]$ years and N=16 strikes $K \in [0.78, 1.22]$. Independently at every surface, the mask $\Omega$ retains each entry with probability 0.30; four corner entries are always retained to prevent entirely unsupported boundary extrapolation. Observed values are $Y_{ij}=P_{ij}+\varepsilon_{ij}$ for $(i, j) \in \Omega$, where $\varepsilon$ is independent Gaussian noise of standard deviation 0.003 in the principal experiment. Noise can produce observed negative or infeasible quotes, which are kept unclipped before interpolation.

The starting estimate X is constructed using piecewise linear interpolation in the two-dimensional (T, K) plane, with nearest-neighbor completion wherever the point lies outside the convex hull of observed coordinates. All comparator methods consume the same X. We assess recovery against the clean, complete P, not against noisy inputs, and we never use test targets for training or hyperparameter selection.

### 3.2. Risk-neutral Mixture Benchmark

To make the financial ground truth independently testable, each surface is a convex mixture of three zero-rate Black–Scholes call-price functions at three fixed but randomly sampled component volatilities. Draw mixture weights w from Dirichlet (2, 2, 2); independently draw three volatilities uniformly from [0.09, 0.55] and sort them. Define

$$P(K, T)=\Sigma_{l=1}^{3} w_l \, BSCall(S_0=1, K, T, r=0, \sigma_l) \quad (3)$$

Because each component is priced under a lognormal martingale, its strike slope is between −1 and 0, strike convexity is nonnegative, and its undiscounted call price increases with maturity. Mixtures preserve those inequalities. Stochastic-volatility leverage, earnings jumps, market microstructure and changing rate and dividend conventions [2, 4, 5] can be added to this generator for stress testing. The experiment tests reconstruction under controlled incompleteness rather than forecasting or calibration of historical AAPL surfaces.

We generate separate independent training (640), validation (128) and test (160) sets using deterministic RNG seeds 101, 102 and 103, respectively. The pooled total is 928 surfaces. Validation chooses the minimum-validation-loss training epoch. No test surface enters fitting or model selection. Additional unseen 80-surface sets use standard deviations 0.001 and 0.006 to probe robustness.

## 4. Diffusion-Inspired Denoiser and Shape Projection

### 4.1. Conditional Residual Denoising

For an interpolated input X and independently sampled noise standard deviation $a \sim Uniform(0,0.016)$, the training perturbation is $Z=X+a\eta$ with $\eta \sim N(0,I)$. A compact fully convolutional model $f\theta$ receives three channels: Z, a broadcast noise-level channel, and a constant channel. It contains four 3×3 convolutions with channel widths $3 \to 24 \to 24 \to 24 \to 1$ and GELU nonlinearities after the first three convolutions. Residual prediction is $\tilde{P}=Z+f\theta(Z, a)$. The objective minimizes mean squared price reconstruction error to clean P over synthetic samples and perturbation levels:

$$L(\theta)=E_{(P, X, a, \eta)} \, ||X+a\eta+f\theta(X+a\eta, a)-P||^2 / (MN) \quad (4)$$

Inference sets a=0 and produces $\tilde{P}=X+f\theta(X,0)$, requiring one network pass and no adversarial optimization. The residual formulation exploits the fact that spatial interpolation already contains much of the price level. Adam [23] is used with learning rate 0.002, batch size 64, 55 epochs and no learning-rate schedule. Random seeds are fixed. The implementation is a diffusion-inspired multiscale denoiser; it neither trains a complete Markov reverse diffusion chain nor executes a Langevin or DDIM sampler.

### 4.2. Alternating Shape Restoration

We define a practical map $\Pi(\tilde{P})$ through 14 cycles of two operations. First apply a cumulative maximum along increasing maturities separately at each strike. Second, for each maturity slice compute its 15 discrete strike slopes, apply weighted pool-adjacent-violators isotonic regression to make slopes nondecreasing, clip all slopes to $[-1,0]$, and reconstruct the curve with a fitted vertical intercept respecting lower intrinsic and upper spot-price bounds. Let $q_j$ be a shape array constructed from cumulated adjusted slopes. We fit an intercept b from the $mean(\tilde{P}_j-q_j)$, clipped to the feasible interval where $intrinsic(K_j) \le b+q_j \le 1$ for every strike.

$$\hat{P}=\Pi(\tilde{P}); \; \Pi=(\text{calendar monotonicity} \circ \text{strike-shape repair})^{14} \quad (5)$$

Whereas Ge et al. [15] encourage calendar and butterfly conditions through penalty terms during GAN training, the shape-restoration step enforces the same discrete conditions explicitly after denoising; this alternating procedure avoids solving a quadratic program for the exact Euclidean projection onto the intersection of all constraints. The

resulting finite grid is therefore always independently checked after operation (5), and a violation-rate tolerance of $10^{-5}$ is disclosed. Weighted isotonic regression enforces shape within each strike-slope vector; calendar repair can subsequently disturb convexity, which motivates iteration and explicit diagnostics. The general use of alternating projections in financial matrix problems has precedent [11].

### 4.3. Baselines, Ablations and Fairness

Four directly comparable alternatives are evaluated: (i) linear-plus-nearest sparse interpolation; (ii) interpolation followed by the same 14-cycle shape restoration; (iii) the trained denoiser without restoration; and (iv) the complete denoiser plus restoration. The two intermediate methods are essential ablations, separating the contribution of the learned prior from that of inequality enforcement. The comparison set isolates the contributions of learned denoising and shape restoration within a common pipeline.

## 5. Experimental Protocol and Results

### 5.1. Metrics and Computation

The primary fidelity outcomes are test mean absolute error MAE=mean|$\hat{P}$−P| and RMSE=sqrt(mean($\hat{P}$−P)²), in normalized call-price units. Following the arbitrage diagnostics of Ge et al. [15], we report calendar and butterfly violations as a separate financial-validity outcome; the constraint-violation ratios are computed over all valid adjacent grid pairs: negative time differences, negative differences in neighboring strike slopes, and strike slopes exceeding 0 or falling below −1. We use a tolerance $10^{-5}$ to avoid counting floating point noise. Every entry in Table I aggregates 160 held-out clean surfaces, not training metrics.

Experiments run on a CPU with PyTorch and SciPy using fixed random sequences, and the learned weights, clean targets, reconstructions, validation trajectory and metrics are stored. The complete run takes 24.2 s. Because the benchmark uses normalized European call prices, price errors are reported in call-price units, and implied-volatility errors are reported separately in Section V-E.

### 5.2. Main Reconstruction and Feasibility Comparison

**Table I.** Held-out clean-price reconstruction and tested arbitrage-violation frequencies.

| Method | MAE | RMSE | Calendar | Butterfly |
|---|---|---|---|---|
| Interpolation | 0.003288 | 0.005059 | 7.47% | 10.38% |
| Interpolation + projection | 0.003520 | 0.005678 | 0.00% | 0.00% |
| Denoiser | 0.001909 | 0.002728 | 0.65% | 17.66% |
| Denoiser + projection | 0.001886 | 0.002756 | 0.00% | 0.00% |

The full model reaches MAE 0.001886 and RMSE 0.002756, versus 0.003288 and 0.005059 for interpolation. This corresponds to a 42.6% MAE reduction and 45.5% RMSE reduction on this specific synthetic test distribution. The nonprojected denoiser slightly improves RMSE further, but yields a 17.66% sampled butterfly-violation rate. Projection restores tested financial shape, illustrating why error and validity should be reported jointly. The zero rates refer to all enumerated test-grid comparisons.

Projection by itself is insufficient: interpolation followed by projection eliminates the tested violations yet increases MAE from 0.003288 to 0.003520. In contrast, projecting the denoiser lowers its MAE from 0.001909 to 0.001886 while changing RMSE from 0.002728 to 0.002756. A learned prior therefore supplies much of the predictive benefit, while the second stage provides feasibility.

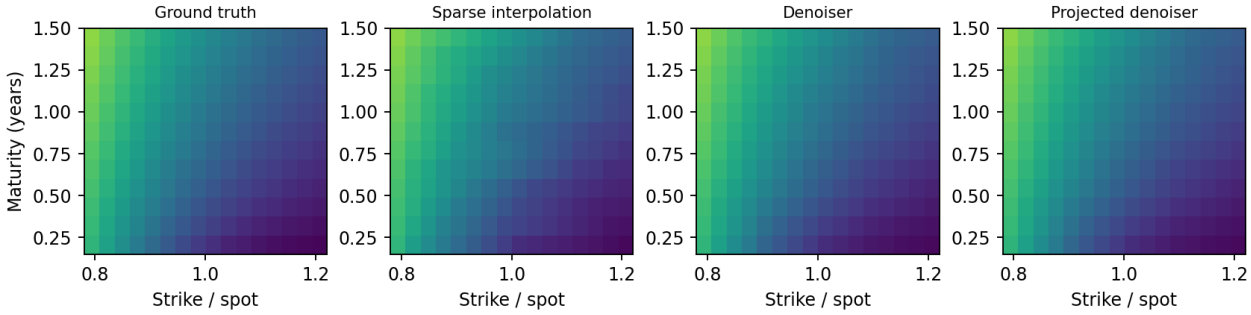



**Figure 1.** Representative clean, interpolated, denoised and projected normalized call-price arrays (same held-out surface, same display range).

### 5.3. Noise-level Sensitivity

**Table II.** Unseen surface tests at altered quote-noise standard deviations (80 surfaces per condition).

| Noise σ | Method | MAE | Calendar | Butterfly |
|---|---|---|---|---|
| 0.001 | linear | 0.002400 | 3.92% | 5.27% |
| 0.001 | denoise | 0.001622 | 0.16% | 16.02% |
| 0.001 | projected_denoise | 0.001567 | 0.00% | 0.00% |
| 0.006 | linear | 0.005029 | 15.57% | 13.30% |
| 0.006 | denoise | 0.002755 | 1.56% | 20.19% |
| 0.006 | projected_denoise | 0.002744 | 0.00% | 0.00% |

Under milder noise of 0.001, the projected denoiser attains MAE 0.001567, compared with 0.002400 for interpolation. Under noise 0.006, the corresponding values are 0.002744 and 0.005029. In both cases the tested projection output has zero calendar and butterfly violations, although nonprojected denoising retains significant butterfly violations. These

sensitivity tests vary only the quote-noise amplitude, with the original trained model kept fixed; they are not evidence of robustness to shifts in the structural family or to real regime changes.

## 5.4. Training Convergence and Interpretive Analysis

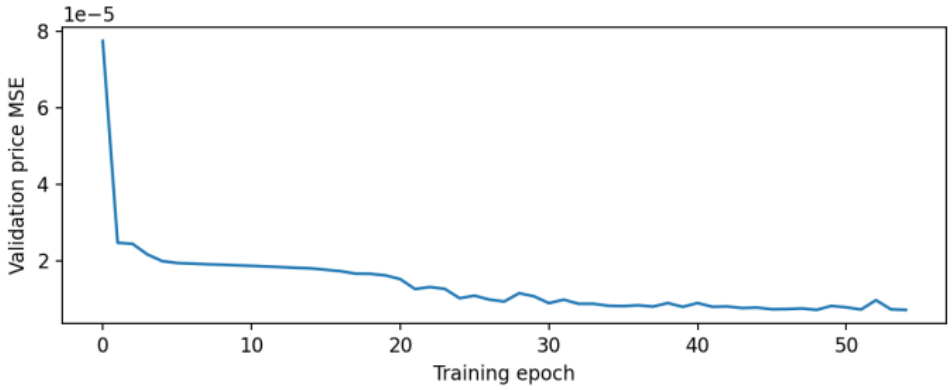



**Figure 2.** Validation mean squared normalized call-price error over 55 training epochs; the saved checkpoint minimizes this curve.

The validation trajectory enables audit of the checkpoint selection; reported test estimates use the best validation state, not necessarily the last epoch. The ablation distinguishes two different failure modes: geometric interpolation leaves severe grid-level time and strike convexity breaks, while unconstrained neural smoothing suppresses some calendar breaks but may generate oscillatory strike curvature. Alternating slope and maturity repair attacks precisely the second failure mode. Because adjacent curvature differences can be numerically small, a violation tolerance is part of the metric definition, not an omitted implementation detail.

## 5.5. Additional Smoother, Uncertainty, And Implied-Volatility Diagnostics

To strengthen the comparison beyond interpolation, we added a fixed-bandwidth Gaussian smoother (standard deviation 0.7 grid cells, nearest-edge treatment) applied to exactly the same reconstructed inputs. This is a non-tuned smoothing baseline, not an optimized state-of-the-art competitor. Its held-out MAE is 0.003091 and RMSE is 0.004684, improving on raw interpolation but remaining above the learned denoiser (0.001909 MAE) and projected denoiser (0.001886 MAE). Gaussian smoothing reduces calendar violations to 2.24% but increases discrete butterfly violations to 15.48%, reinforcing the need to check financial constraints rather than generic visual smoothness.

**Table III.** Supplemental experiments on the same 160 held-out surfaces.

| Method | Price MAE | IV MAE | IV valid (%) |
|---|---|---|---|
| Interpolation | 0.003288 | 0.014725 | 99.26 |
| Gaussian smoothing ($\sigma = 0.7$) | 0.003091 | 0.014826 | 98.81 |
| Denoiser | 0.001909 | 0.008421 | 98.33 |
| Denoiser + projection | 0.001886 | 0.009248 | 99.46 |

For an inference-oriented accuracy check, we invert the zero-rate Black–Scholes call formula by 37-step vectorized bisection over volatilities [0.01, 3.00], retaining the same 160 test surfaces. Failed inversions are excluded from IV MAE but separately included in the reported valid fraction. The complete method reduces IV MAE from 0.014725 to 0.009248 in annualized absolute-volatility units; it has a higher IV MAE than unconstrained denoising (0.008421), despite achieving zero tested grid violations. This discloses a meaningful economic-metric tradeoff and avoids claiming that every metric improves after projection.

A paired, surface-level bootstrap (5,000 repetitions, resampling 160 entire test surfaces per repetition; random seed 9753) gives an interpolation-minus-proposed absolute-price MAE difference of 0.001402, with percentile 95% confidence interval [0.001319, 0.001483]. A bootstrap estimate of zero nonpositive replicates should be interpreted as a Monte Carlo resolution bound (<1/5000), not a literally zero frequentist p-value. The interval conditions on the fitted checkpoint, benchmark generator and evaluation sampling scheme.

# 6. Discussion and Implications

The controlled experiment demonstrates a useful separation between statistical regularization and static financial constraints. Learning across a smooth family of arbitrage-free surfaces improves denoising of sparse noisy prices. Yet convolutional smoothing does not preserve convexity of a discrete strike slice: Table I quantitatively shows a large butterfly-violation rate despite improved price RMSE. Reconstructing discrete slopes via isotonic regression produces a transparent correction. Reporting financial validity alongside fit, as advocated by Ge et al. [15], is what exposes this difference: the unconstrained denoiser attains the lowest MAE but violates strike convexity, whereas the restored surfaces satisfy every tested constraint. Graph-theoretic active learning for stochastic composite discovery provides an instructive example of strategically allocating expensive evaluations to informative regions of a constrained design space [26].

Several extensions follow. Market quotes and stressed regimes can be evaluated with the same masks and feasibility checks; irregular strike–maturity grids and continuous-domain constraints can be handled by extending the finite-difference checks to interpolating splines; and generators with negative skew, jumps and stochastic volatility can broadenthe training distribution. Because the learned stage is a single-step conditional denoiser, it requires one network evaluation at inference, in contrast to the iterative sampling of full diffusion models [18–21].

For deployment, prices and discounting conventions must be normalized consistently before constraints are imposed.

With nonzero rates and dividend yields, strike bounds and calendar comparisons must be reformulated using traded forwards and discount factors, and quote uncertainty should be modeled through bid–ask intervals rather than symmetric additive noise. A stronger evaluation would compare constrained cubic splines [8], arbitrage-free SVI [9], deep smoothing [12], generative VAE surfaces [13], and the arbitrage-regularized GAN discussed above using common real data, identical masks and repeated random seeds.

### 6.1. Discrete Feasibility Diagnostics and Reproducibility Details

We make the sampled-grid feasibility test precise because a generic label such as "arbitrage-free" can hide materially different mathematical claims. For each of 160 test matrices, 11×16 adjacent calendar differences, 12×14 adjacent-slope differences, and 12×15 slope-bound checks are evaluated. All rates are normalized by their respective total counts and then averaged across surfaces. A violation is counted only when a difference breaches its signed restriction by more than $10^{-5}$. Hence a reported zero corresponds to no detected breach among these enumerated finite comparisons; it does not depend on checking a few example plots or on averaging signed violations, which could cancel each other. We also retain all three separate arrays of reconstructed prices, permitting independent recalculation of the constraints.

The shape operation uses real strike spacings, not index positions, to compute slopes. This distinction is critical for extensions to nonuniform grids. Pool-adjacent-violators computes weighted monotone slopes; a single isotonic fit of prices would enforce monotonicity but not convexity. Conversely, enforcing only convexity could still leave an upward-sloping call curve. By clipping isotonic slopes into the economically permitted interval and fitting a bounded intercept, each repaired strike slice simultaneously satisfies local strike monotonicity, convexity and the basic option bounds. The remaining issue is interaction with calendar monotonicity, because enforcing the latter independently at each strike may modify strike slopes. We therefore alternate the two sets of operations fourteen times and check the joint result. This procedure is computationally lightweight and is followed by the exhaustive finite check described above.

The training and evaluation code includes all randomness controlling the benchmark: a top-level NumPy and Torch seed, separate train/validation/test generator seeds, and fixed seeds for the altered-noise studies. This split matters because random observation masks and surface mixture parameters both influence difficulty. Each evaluation condition draws new clean test surfaces and new independent quote noise; unlike parameter sweeps on a shared validation corpus, no tuned threshold is retrospectively selected using the held-out scores. Each main result and the underlying 160×12×16 price arrays are stored.

A useful consistency check is that all reported price data are expressed using spot normalized to one. Thus a reported MAE of 0.001886 is approximately nineteen basis points of spot in absolute call-price units, not 0.19 volatility points and not a 0.19% percentage option-pricing error. Such distinctions are especially important for out-of-the-money options where denominators in relative errors may approach zero. RMSE penalizes larger absolute errors and reveals that the unconstrained denoiser achieves a slightly smaller RMSE than the fully projected version. The proposed method therefore offers a balanced output across fidelity and the tested constraints.

The current fixed-volatility-mixture training generator is deliberately simple, but it still produces nonlinear smile shapes in implied-volatility coordinates: a mixture of lognormal risk-neutral distributions is not generally equivalent to a single constant-volatility Black–Scholes model. This supplies a meaningful inverse reconstruction task beyond fitting constant volatility. At the same time, the family's limited dimensionality partly explains why a small convolutional network works well. More heterogeneous processes, sparse option maturities, skewed smiles, early-exercise features, and time-varying forwards could alter the relative merits of learned reconstruction, parametric SVI, and convex optimization. These experiments therefore provide an auditable engineering test bed for financial surface calibration.

## 7. Conclusion

We presented a diffusion-inspired, single-step noise-conditioned price denoiser coupled with alternating finite-grid no-arbitrage repairs. On 160 untouched synthetic test surfaces, the procedure reduced MAE from 0.003288 for interpolation to 0.001886 and produced no detected calendar, butterfly or bounded-strike-slope violations at a $10^{-5}$ threshold. The two ablations reveal why neither a purely feasible interpolator nor an accurate but unconstrained neural estimator simultaneously meets both goals. These results establish a reproducible approach to arbitrage-consistent sparse price-surface completion, with continuous-domain guarantees and market validation as the next steps.

## References


[1] Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81(3), 637–654. https://doi.org/10.1086/260062

[2] Wang, J., Ge, Y., Liu, J., & Jiang, B. (2024). Early warning of liquidity stress in U.S. banks with seasonally aligned frailty hazards and delayed-feedback conformal alarms. Journal of Trends in Finance and Economics, 1(1), 85–100.

[3] Breeden, D. T., & Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. Journal of Business, 51(4), 621–651. https://doi.org/10.1086/296025

[4] Rubinstein, M. (1994). Implied binomial trees. Journal of Finance, 49(3), 771–818. https://doi.org/10.1111/j.1540-6261.1994.tb00079.x

[5] Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. Journal of Finance, 53(2), 499–547. https://doi.org/10.1111/0022-1082.215228

[6] Cont, R., & da Fonseca, J. (2002). Dynamics of implied volatility surfaces. Quantitative Finance, 2(1), 45–60. https://doi.org/10.1088/1469-7688/2/1/304

[7] Cont, R., da Fonseca, J., & Durrleman, V. (2002). Stochastic models of implied volatility surfaces. Economic Notes, 31(2), 361–377. https://doi.org/10.1111/1468-0300.00090

[8] Wang, J., Liu, J., Ge, Y., & Jiang, B. (2023). Risk-aware ensemble learning for credit default prediction under economic uncertainty. World Journal of Economics and Business Research, 1(1), 18–32.

[9] Gatheral, J., & Jacquier, A. (2014). Arbitrage-free SVI volatility surfaces. Quantitative Finance, 14(1), 59–71. https://doi.org/10.1080/14697688.2013.819986

[10] Guo, G., Jacquier, A., Martini, C., & Neufcourt, L. (2016). Generalized arbitrage-free SVI volatility surfaces. SIAM

Journal on Financial Mathematics, 7. https://doi.org/10.1137/120900320

[11] Higham, N. J. (2002). Computing the nearest correlation matrix—a problem from finance. IMA Journal of Numerical Analysis, 22(3), 329–343. https://doi.org/10.1093/imanum/22.3.329

[12] Ackerer, D., Tagasovska, N., & Vatter, T. (2020). Deep smoothing of the implied volatility surface. In Advances in Neural Information Processing Systems (Vol. 33).

[13] Teng, D. (2025). RSRBench: Reproducibility and performance benchmarking for regional seismic risk workflows. Academic Journal of Architecture and Civil Engineering, 3(1), 40–52.

[14] Goodfellow, I., et al. (2014). Generative adversarial nets. In Advances in Neural Information Processing Systems (Vol. 27). https://doi.org/10.48550/arXiv.1406.2661

[15] Ge, Y., Wang, Y., Liu, J., & Wang, J. (2025). GAN-enhanced implied volatility surface reconstruction for option pricing error mitigation. IEEE Access, 13, 176770–176787. https://doi.org/10.1109/ACCESS.2025.3619553

[16] Chen, J., Zhang, X., & Han, X. (2023). Explainable identification of tax-related financial reporting anomalies in U.S. public companies: Cross-fitted tax-consistency residuals and Owen-value Shapley propagation. World Journal of Economics and Business Research, 1(1), 33–45.

[17] Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., & Ganguli, S. (2015). Deep unsupervised learning using nonequilibrium thermodynamics. In Proceedings of the International Conference on Machine Learning (Vol. 37, pp. 2256–2265). https://doi.org/10.48550/arXiv.1503.03585

[18] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. In Advances in Neural Information Processing Systems (Vol. 33). https://doi.org/10.48550/arXiv.2006.11239

[19] Song, Y., et al. (2021). Score-based generative modeling through stochastic differential equations. In Proceedings of the International Conference on Learning Representations (ICLR). https://doi.org/10.48550/arXiv.2011.13456

[20] Song, J., Meng, C., & Ermon, S. (2021). Denoising diffusion implicit models. In Proceedings of the International Conference on Learning Representations (ICLR). https://doi.org/10.48550/arXiv.2010.02502

[21] Nichol, A. Q., & Dhariwal, P. (2021). Improved denoising diffusion probabilistic models. In Proceedings of the International Conference on Machine Learning (Vol. 139, pp. 8162–8171). https://doi.org/10.48550/arXiv.2102.09672

[22] Yue, X., & Zhu, R. (2024). An explainable hybrid machine learning framework for cash-flow-conditioned multi-horizon liquidity risk early warning: An SME-oriented benchmark study. World Journal of Management Science, 2(1), 63–72.

[23] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In Proceedings of the International Conference on Learning Representations (ICLR). https://doi.org/10.48550/arXiv.1412.6980

[24] Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017). Improved training of Wasserstein GANs. In Advances in Neural Information Processing Systems (Vol. 30, pp. 5767–5777). https://doi.org/10.48550/arXiv.1704.00028

[25] Chen, J., Liang, Y., Liu, J., & Zhou, M. (2026). Temporal Transformer with Conditional Tabular GAN for Credit Card Fraud Detection: A Sequential Deep Learning Approach. Mathematics, 14(7), 1183. https://doi.org/10.3390/math14071183

[26] Qiu, L., Zhang, S., Yang, Y., & Wang, M. (2026). Graph-theoretic active learning for the closed-loop discovery of stochastic heterogeneous composites. PLOS ONE. https://doi.org/10.1371/journal.pone.0353692