Arbitrage Consistent Option Surfaces with Bid Ask Bounds
Surface Repair via Convex Optimization and Heston Calibration

Shan Kunzru

Abstract

This project details my implementation of a volatility surface calibrator that takes real option quotes with bid and ask bounds, detects static arbitrage violations, and then repairs/shifts the surface by making the minimum optimal adjustments while adhering to the bid ask interval bounds. After repair, I convert prices into implied volatilities and calibrate the Heston stochastic volatility model both before and after repair. The goal is to implement this pipeline in Python, and detail their mathematical foundations and proofs using the no-arbitrage postulate. I will also explore how we can apply discrete linear constraints on real data, and why a logically consistent surface improves calibration stability. The key mathematical idea is a projection onto an arbitrage consistent poly-tope using convex optimization.

1 Project motivation

Theoretically, under the efficient market hypothesis, we would expected the implied volatility surface to look like a nice smooth object with a clear skew. In reality, market quotes are noisy, especially in illiquid strikes and short dated contracts. If I blindly compute implied vols from bid-ask mid prices, I often get contradictions that violate basic no arbitrage facts. Those contradictions may lead to downstream models fitting to nonsense, thus requiring careful pruning

The pipeline I am building follows a simple philosophy:

1. Treat each option bid-ask quote spread as a price interval bound, not a single number.
2. Detect where the intervals violate the no-arbitrage constraint.
3. Repair by finding adjusted prices that satisfy no-arbitrage while staying within bid ask bounds whenever possible.
4. Use the repaired surface as a clean target for implied volatility extraction and Heston calibration.

Mathematically, the repair step is a convex quadratic program. That is the core of the project.

2 Market data and Preprocessing

2.1 CBOE options data summary format

The data I am using for this project is the NBBO (National Best Bid Offer) data on SPY index options for October 2025. Each file corresponds to one quote date, totalling 31 files. Each row describes the bid and ask data of a single contract given its strike and time-to-expiry, with fields including

- `quote_date`, `expiration`, `strike`, `option_type`
- bid and ask at the 15:45 snapshot: `bid_1545`, `ask_1545`
- underlying bid and ask at the 15:45 snapshot: `underlying_bid_1545`, `underlying_ask_1545`
- end of day bid and ask, volume, open interest

I base the arbitrage checks and repair on the 15:45 snapshot rather than the end of day quotes in accordance to CBOE's own suggestion that they better reflect liquidity. Nonetheless, I still store end of day fields for reporting and sanity checks.

2.2 Data Cleaning

Before any math, I apply practical filters to remove any garbage quotes

- drop quotes with missing bid or ask or with ask less than bid
- drop quotes with non-positive strikes

These are not theoretical steps, but they reduce the number of false alarms before the real arbitrage logic.

2.3 Spot, Calls, Puts, and Parity

I estimate the spot input to Black Scholes and parity checks using the underlying mid at 15:45

$$S_0 \approx \frac{1}{2}\Big(\text{underlying bid}_{1545} + \text{underlying ask}_{1545}\Big).$$

Recall the put-call parity formula:

$$C(K,T) - P(K,T) = S_0 e^{-qT} - K e^{-rT}.$$

For the surface repair, put-call parity allows us to compute the price of a put from the corresponding call price of equal strike and TTE. To prevent redundancy, we will only be repairing the calls, while the puts will be used for consistency checks and forward estimation.

In early experiments I treat rates and dividends as given inputs. If I want to infer a forward from the chain itself, I can rearrange parity in forward form

$$C(K,T) - P(K,T) = D(0,T)\big(F_{0,T} - K\big),$$

where the discount rate $D(0,T) = e^{-rT}$. Using near the money strikes, this provides a forward estimate and a parity consistency check.

3 Pricing framework and assumptions

I work under standard no arbitrage assumptions for European options. I assume the existence of a risk neutral measure under which discounted traded assets are martingales. Rates and dividends may be treated as deterministic inputs. If I take a simplified case with zero rates and zero dividends, the main inequalities become especially clean, and the discrete repair logic remains the same.

4 Static no arbitrage conditions for calls

Let $C(K,T)$ denote the time zero price of a European call with payoff $(S_T - K)^+$. Let $D(0,T) = e^{-rT}$. Under continuous dividend yield q, the forward price is $F_{0,T} = S_0 e^{(r-q)T}$.

4.1 Basic bounds

Proposition 1 (Lower bound)**.** *For any K, T,*

$$C(K,T) \geq \max\big(S_0 e^{-qT} - Ke^{-rT}, 0\big).$$

Proof. For any outcome, $(S_T - K)^+ \geq S_T - K$. Under risk neutral pricing,

$$C(K,T) = D(0,T)\mathbb{E}\big[(S_T - K)^+\big] \geq D(0,T)\big(\mathbb{E}[S_T] - K\big).$$

Under risk neutral dynamics with yield q, $\mathbb{E}[S_T] = F_{0,T} = S_0 e^{(r-q)T}$. Therefore

$$C(K,T) \geq e^{-rT}\big(S_0 e^{(r-q)T} - K\big) = S_0 e^{-qT} - Ke^{-rT}.$$

Also $C(K,T) \geq 0$ since payoffs are nonnegative. Taking the maximum gives the lower bound result. Lower bound still holds for American options as the right of early exercise would only raise the call value □

Proposition 2 (Upper bound)**.** *For any K, T,*

$$C(K,T) \leq S_0 e^{-qT}.$$

Proof. For any outcome, $(S_T - K)^+ \leq S_T$. Then

$$C(K,T) = D(0,T)\mathbb{E}\big[(S_T - K)^+\big] \leq D(0,T)\mathbb{E}[S_T] = e^{-rT} S_0 e^{(r-q)T} = S_0 e^{-qT}.$$

We can illustrate this bound using the no-arbitrage argument. Suppose instead that $C(K,T) > S_0 e^{-qT}$. Then an arbitrageur could:

- Short one call for $C(K,T)$,
- Buy one share of stock for S_0,
- Finance the purchase by borrowing $S_0 e^{-rT}$ and simultaneously lending the present value of dividends $S_0(1 - e^{-qT})$.

The upfront cost of this replicating portfolio is negative since $C(K,T) > S_0 e^{-qT}$while at time T, the replicating portfolio value is at least $(S_T - K)^+ - S_T + S_0 e^{(r-q)T}$, which is nonpositive in all terminal states of S_T.

□

4.2 Monotonicity in strike

Theorem 1 (Monotonicity)**.** *For any two call options in which $K_1 < K_2$, then*

$$C(K_1,T) \geq C(K_2,T).$$

Proof. For all S_T, $(S_T - K_1)^+ \geq (S_T - K_2)^+$. If the inequality in prices were reversed, there exists a static arbitrage opportunity: buy the dominating payoff and sell the dominated payoff for a nonpositive cost and a nonnegative payoff with positive probability of gain. Therefore the price must satisfy the stated inequality. □

4.3 Convexity in strike

Theorem 2 (Convexity)**.** *Fix T. The mapping $K \mapsto C(K,T)$ is convex.*

Proof. Fix strikes $K_1 < K_2 < K_3$. Let $\lambda = \frac{K_3-K_2}{K_3-K_1} \in (0,1)$, so $K_2 = \lambda K_1 + (1-\lambda)K_3$. For any S_T, the function $K \mapsto (S_T - K)^+$ is convex because it is the maximum of two affine functions 0 and $S_T - K$. Therefore

$$(S_T - K_2)^+ \leq \lambda(S_T - K_1)^+ + (1-\lambda)(S_T - K_3)^+.$$

So a portfolio of λ calls at K_1 and $1-\lambda$ calls at K_3 dominates one call at K_2. No arbitrage implies the dominating portfolio costs at least as much as the dominated payoff:

$$C(K_2,T) \leq \lambda C(K_1,T) + (1-\lambda)C(K_3,T).$$

This is convexity. □

Remark 1 (Breeden Litzenberger link)**.** *Under regularity conditions,*

$$\frac{\partial^2 C}{\partial K^2}(K,T) = D(0,T)\, f_{S_T}(K),$$

where f_{S_T} is the risk neutral density of S_T. Convexity is exactly the nonnegativity of this implied density.

4.4 Calendar monotonicity

In the simplified setting with zero carry, call prices are nondecreasing in maturity for fixed K. With nonzero carry, the clean version is easiest to state for forward adjusted prices and under deterministic rates and dividends. In my implementation I treat calendar constraints as configurable. I either

- enforce the simplified calendar monotonicity when the carry effect is negligible over the maturities I include, or
- enforce a forward adjusted version using inputs r and q or a forward estimate from parity.

5 Discrete grid and constraints used in code

For each maturity T_j, I sort the strikes

$$K_{1,j} < K_{2,j} < \cdots < K_{n_j,j}.$$

Let $C_{i,j}$ denote the call price at strike $K_{i,j}$ and maturity T_j. In the repair step, the unknowns are the adjusted prices $\widetilde{C}_{i,j}$.

5.1 Discrete monotonicity

For each maturity j,

$$\widetilde{C}_{i,j} \geq \widetilde{C}_{i+1,j} \quad \text{for } i = 1, \ldots, n_j - 1.$$

5.2 Discrete convexity

For each maturity j and interior index $i = 2, \dots, n_j - 1$,

$$\widetilde{C}_{i-1,j} - 2\widetilde{C}_{i,j} + \widetilde{C}_{i+1,j} \geq 0.$$

5.3 Calendar constraints

When strikes match between maturities T_j and T_{j+1},

$$\widetilde{C}_{i,j+1} \geq \widetilde{C}_{i,j}.$$

I apply this on the intersection of strikes across maturities unless I explicitly add interpolation.

6 Bid ask intervals and definite violation tests

The market gives intervals

$$C(K_{i,j}, T_j) \in \left[C^{\text{bid}}_{i,j}, C^{\text{ask}}_{i,j}\right].$$

A violation is definite if no selection of prices inside these intervals can satisfy the constraint.

6.1 Definite monotonicity violation

The monotonicity requirement is $C_{i,j} \geq C_{i+1,j}$. The best case to satisfy this is to take $C_{i,j}$ as large as possible and $C_{i+1,j}$ as small as possible. Therefore a sufficient definite violation condition is

$$C^{\text{ask}}_{i,j} < C^{\text{bid}}_{i+1,j}.$$

6.2 Definite convexity violation

Convexity requires $C_{i-1,j} - 2C_{i,j} + C_{i+1,j} \geq 0$. The left side is maximized by choosing

$$C_{i-1,j} = C^{\text{ask}}_{i-1,j}, \quad C_{i,j} = C^{\text{bid}}_{i,j}, \quad C_{i+1,j} = C^{\text{ask}}_{i+1,j}.$$

So a sufficient definite violation condition is

$$C^{\text{ask}}_{i-1,j} - 2C^{\text{bid}}_{i,j} + C^{\text{ask}}_{i+1,j} < 0.$$

6.3 Definite calendar violation

For a matched strike across maturities $j < j+1$, the calendar requirement is $C_{i,j+1} \geq C_{i,j}$. The best case to satisfy it is $C_{i,j+1}$ as large as possible and $C_{i,j}$ as small as possible. Therefore a sufficient definite violation condition is

$$C^{\text{ask}}_{i,j} > C^{\text{bid}}_{i,j+1}.$$

7 Surface repair as convex optimization

7.1 Feasible set

The repair step chooses $\widetilde{C}_{i,j}$ such that

1. Bid ask bounds for every quote
$$C_{i,j}^{\text{bid}} \le \widetilde{C}_{i,j} \le C_{i,j}^{\text{ask}}.$$
2. Discrete monotonicity in strike
$$\widetilde{C}_{i,j} \ge \widetilde{C}_{i+1,j}.$$
3. Discrete convexity in strike
$$\widetilde{C}_{i-1,j} - 2\widetilde{C}_{i,j} + \widetilde{C}_{i+1,j} \ge 0.$$
4. Optional calendar constraints on matched strikes
$$\widetilde{C}_{i,j+1} \ge \widetilde{C}_{i,j}.$$

All constraints are linear inequalities, so the feasible region is a convex polytope.

7.2 Objective and weights

I want the smallest possible change from the market, and the natural anchor is the mid price $C_{i,j}^{\text{mid}}$. To find the optimal adjustment, we can minimize the weighted least squares objective function

$$\min_{\widetilde{C}} \sum_{i,j} w_{i,j} \big(\widetilde{C}_{i,j} - C_{i,j}^{\text{mid}}\big)^2.$$

Proof. This weighted least squares objective can be motivated statistically as a maximum likelihood estimator under Gaussian noise assumptions. Assume the observed midprice $C_{i,j}^{\text{mid}}$ is an unbiased estimate of the true arbitrage-free price, perturbed by independent Gaussian noise: $C_{i,j}^{\text{mid}} = \hat{C}_{i,j} + \epsilon_{i,j}$, where $\epsilon_{i,j} \sim \mathcal{N}(0, \sigma_{i,j}^2)$. The variance $\sigma_{i,j}^2$ is assumed inversely proportional to quote tightness, i.e., $\sigma_{i,j}^2 \propto \text{spread}_{i,j} + \epsilon$, reflecting that tighter bid-ask spreads indicate lower uncertainty (more liquidity and reliability).

Under this model, the log-likelihood of the data given the adjusted prices $\hat{\mathbf{C}}$ is

$$\ell(\hat{\mathbf{C}}) = -\frac{1}{2} \sum_{i,j} w_{i,j} (\hat{C}_{i,j} - C_{i,j}^{\text{mid}})^2 + \text{constant},$$

where $w_{i,j} = 1/\sigma_{i,j}^2$. Maximizing the likelihood is equivalent to minimizing the weighted squared residuals, justifying the L2 norm over alternatives (e.g., L1 for Laplace noise). This approach ensures the repair is not only minimal but also statistically principled, prioritizing adjustments to noisier quotes.

For this log-likelihood, a simple and interpretable weight is

$$w_{i,j} = \frac{1}{\text{spread}_{i,j} + \varepsilon}, \quad \text{spread}_{i,j} = C_{i,j}^{\text{ask}} - C_{i,j}^{\text{bid}}.$$

This encodes the belief that tight quotes are more trustworthy as they reflect higher liquidity thus less uncertainty in the price bounds

The resulting solution space is convex surface, so a global optimum exists and is efficiently computable. □

7.3 Infeasibility and soft constraints

Sometimes the bid ask intervals are themselves inconsistent with arbitrage. To avoid a hard failure, I optionally introduce slack variables $s \geq 0$ on selected constraints, for example on convexity

$$\widetilde{C}_{i-1,j} - 2\widetilde{C}_{i,j} + \widetilde{C}_{i+1,j} + s_{i,j} \geq 0, \quad s_{i,j} \geq 0,$$

and penalize slack in the objective

$$\min_{\widetilde{C},s} \sum_{i,j} w_{i,j} \big(\widetilde{C}_{i,j} - C^{\mathrm{mid}}_{i,j}\big)^2 + \Lambda \sum_{i,j} s_{i,j}.$$

This keeps the repair as faithful as possible while still producing a usable surface and a clear report of where the market quotes were fundamentally contradictory.

8 Implied volatility extraction and smile construction

Once I have repaired prices $\widetilde{C}(K,T)$, I invert Black Scholes to obtain implied volatility.

8.1 Black Scholes call price

For spot S_0, strike K, maturity T, rate r, dividend yield q, volatility $\sigma > 0$,

$$C_{\mathrm{BS}} = S_0 e^{-qT} \Phi(d_1) - K e^{-rT} \Phi(d_2),$$

where

$$d_1 = \frac{\ln(S_0/K) + (r - q + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T}.$$

8.2 Implied volatility

For each quote I solve for σ such that

$$C_{\mathrm{BS}}(\sigma) = C_{\mathrm{obs}},$$

where C_{obs} is either C^{mid} or $\widetilde{C}$. This is a one dimensional root finding problem. The existence and uniqueness of implied volatility for reasonable prices follows from the fact that Black Scholes call price is strictly increasing in σ when $T > 0$.

A smile is the curve $K \mapsto \sigma_{\mathrm{imp}}(K, T_j)$ for each maturity T_j. Later in the project I fit a smooth parameterization such as SVI, but only after repair so I am not smoothing contradictions.

9 Risk neutral density check from repaired prices

A nice extra consequence of the convexity constraints is that they give me a stable implied density estimate.

On a discrete grid, the second difference

$$\Delta^2 \widetilde{C}_{i,j} = \widetilde{C}_{i-1,j} - 2\widetilde{C}_{i,j} + \widetilde{C}_{i+1,j}$$

is nonnegative by construction. Up to scaling by strike spacing and discounting, this approximates the risk neutral density at strike $K_{i,j}$. The practical check I use is simply that all discrete densities are nonnegative and that their mass roughly sums to one after accounting for truncation at the wings.

10 Heston model and calibration

10.1 Risk neutral dynamics

Under the risk neutral measure,

$$dS_t = (r - q)S_t\,dt + \sqrt{v_t}S_t\,dW_t^S,$$

$$dv_t = \kappa(\theta - v_t)\,dt + \xi\sqrt{v_t}\,dW_t^v,$$

with correlation $d\langle W^S, W^v\rangle_t = \rho\,dt$. The parameters are

$$\Theta = (v_0, \kappa, \theta, \xi, \rho),$$

with $v_0 > 0$, $\kappa > 0$, $\theta > 0$, $\xi > 0$, and $\rho \in [-1, 1]$.

10.2 Feller condition as a plausibility test

The variance process is a CIR type process. A classical sufficient condition for strict positivity is

$$2\kappa\theta > \xi^2.$$

In calibration, I report whether fitted parameters satisfy this. I do not always hard enforce it, but I treat violations as a warning sign of instability.

10.3 Calibration objective

Let $\sigma_{i,j}^{\mathrm{mkt}}$ denote implied volatility from market mids, and $\sigma_{i,j}^{\mathrm{rep}}$ implied volatility from repaired prices. Let $\sigma_{i,j}^{\mathrm{hes}}(\Theta)$ denote the implied volatility produced by Heston with parameters Θ.

I calibrate by minimizing weighted squared error in implied vol space

$$\min_{\Theta} \sum_{i,j} w_{i,j} \big(\sigma_{i,j}^{\mathrm{hes}}(\Theta) - \sigma_{i,j}^{\mathrm{target}}\big)^2,$$

where σ^{target} is either σ^{mkt} or σ^{rep}.

I run the calibration twice, once on the raw target and once on the repaired target, then compare fit error and parameter stability.

11 Why repair should improve calibration

There is a simple conceptual argument. If the observed surface violates static arbitrage, then it cannot be the surface of any true risk neutral model that prices European options consistently. A parametric model like Heston is being asked to fit data that is not even internally coherent. The optimizer compensates by pushing parameters to extremes, which can look like overfitting but is really a symptom of contradictory targets.

After repair, the target surface is inside the arbitrage consistent region, so it is at least logically compatible with some risk neutral distribution. That usually makes the calibration landscape smoother and produces parameters that are more stable from day to day.

12 Implementation choices and libraries

This section is where I connect the math to the code decisions.

12.1 Data handling

- `pandas` for reading daily files, filtering, grouping by expiration, and sorting strikes
- `numpy` for vectorized computations, finite differences, and building constraint matrices

12.2 Convex repair

I use `cvxpy` to express the quadratic program directly in math form. This is a big decision for me because it reduces the chance of algebra bugs.

- decision variables: a tensor of $\widetilde{C}_{i,j}$ over all strikes and maturities
- constraints: linear inequalities for bounds, monotonicity, convexity, and calendar
- objective: weighted least squares relative to mid prices
- solver: `OSQP` or another QP solver supported by `cvxpy`

I like `cvxpy` here because the code mirrors the mathematical formulation, and it is easier to audit.

12.3 Implied volatility

For implied vol inversion I use `scipy.optimize` root finding, typically bracketing methods because they are robust. The monotonicity of Black Scholes price in σ gives a reliable bracket.

12.4 Heston pricing and calibration

For Heston pricing I use `QuantLib`. The important point for this report is that the calibration objective depends on prices, not on how I compute them. Using a trusted engine lets me focus on the calibration math.

For optimization I again use `scipy.optimize` methods with parameter bounds. I store results for both raw and repaired targets and compare metrics.

13 Evaluation plan and outputs

For each quote date, I output

- counts of definite violations by type and by maturity
- total adjustment magnitude, for example average $|\widetilde{C} - C^{\mathrm{mid}}|$ and max adjustment
- implied vol RMSE for Heston calibration on raw versus repaired targets
- fitted parameters Θ and whether $2\kappa\theta > \xi^2$
- basic stability statistics across days for each parameter

This is the part that turns the project from a math exercise into an empirical story about why data quality matters.

14 References

- J C Hull, Options Futures and Other Derivatives
- J Gatheral, The Volatility Surface
- S L Heston, A Closed Form Solution for Options with Stochastic Volatility, 1993
- Breeden and Litzenberger, Prices of State Contingent Claims Implicit in Option Prices, 1978