Learning To Trade

Visiting Prof. Dr. Hans Buehler

University of Oxford, Institute for Mathematics
hans.buehler@maths.ox.ac.uk

May 21st, 2026

Agenda

Quant Finance 2.0

- Instead of focusing on solvable low-parametric “interpolation“ models solve the hedging problem in real markets.
- As we do not usually have sufficient market data, we build market simulators which capture broad features of the market
- Then we learn to trade and risk manage in the simulated environments with reinforcement learning techniques.
- We present the broad outline of the approach, applications, and challenges in implementation … and active areas of research.

Deep Bellman Hedging

Framework

Environment

- We operate in discrete time $t \in \mathrm{N}$ under the statistical measure P.
- We observe the market **state** s_t: time, prices, news, historical trading events etc and assume it generates our filtration.
- Any publicly observable market quantity X_t can be written for some measurable function X as $X_t = X(s_t)$.
- We also assume here that our trading causes no impact → current MSc project on Deep Hedging with Impact.

Framework

UNIVERSITY OF OXFORD

Trading Instruments

- We assume that we are trading in a market with liquid **tradable instruments**.
  - These include primary assets such as spot, FX as well as derivatives such as options on equity, indices etc. The one-day deterministic **discount factor** is $\beta_t \in (0, \beta^*]$ with $\beta^* < 1$.

- At each time step we can trade a different set of n_t instruments $H^{(t)}$.
  - We denote by $H_t^{(t)} \in R^{n_t}$ their **book value** at time t, and by $\gamma_t \in R^{n_t}_{\geq 0}$ their bid/ask spreads. The book value is the accounting value[1] which reflects the value of future cashflows.
  - We assume that we always have a book value. In particular the value of $H^{(t)}$ tomorrow is denoted by $H_{t+1}^{(t)}$.

$$d\widetilde{H}_t^{(t)} := \beta_t\, H_{t+1}^{(t)} - H_t^{(t)}.$$

  - We also denote by $h_t^{(t)}$ the **cash flows** arising from holding $H_t^{(t)}$ over $(t, t+1]$ and define the self-financing

$$dH_t^{(t)} := d\widetilde{H}_t^{(t)} + h_t^{(t)}$$

  - We do not require that an instrument that was tradable in some t is also tradable at a later time.

(1) For liquid asset that can be a mid-price; for OTC derivatives this would be the classic derivatives price.

Framework

Trading Instruments

- Our **trading cost** to trade $a \in R^{n_t}$ units of $H^{(t)}$ in excess of their book values is given by a non-negative and convex function with $c_t(0) = 0$.[1]
- That implies c_t which is increasing in all directions, i.e. $\partial_\epsilon c_t(\epsilon a) \geq 0$.
- Recall that $c_t(a) \equiv c(a; s_t)$.

(1) Convexity excludes fixed fee cost which are, in fact, common.

Framework

Risk Limits and other Trading Restrictions

- Convex transaction cost allow defining convex limits to trading capacity by setting $c_t(\neg A) = \infty$ outside a convex set A.

- That means we can use trading cost to impose a wide variety of convex **trading restrictions** of the following type:
  - Maximum liquidity: $A = \{a: \text{askcapacity}^i \leq a^i \leq \text{bidcapacity}^i\}$
  - Total Vega Traded: $A = \{a: |\sum_i a_i \text{Vega}_i| \leq \text{Limit}\}$

- Trading restrictions which refer to the current portfolio are not always be feasible, as the available capacity in the market might not be sufficient to hedge all our risk:
  - Total Vega Held: $A = \{a: |\text{PortfolioVega} + \sum_i a_i \text{Vega}_i| \leq \text{Limit}\}$

Statistical Hedging

- Assume that we are given a portfolio $Z^{(t)}$ with book-value $Z_t^{(t)}$ today. Here book value excludes past cash flows.
  - Tomorrow's (unknown) book value of today's portfolio is $Z_{t+1}^{(t)}$.
  - Let

$$d\tilde{Z}_t^{(t)} := \beta_t Z_{t+1}^{(t)} - Z_t^{(t)}$$

  - We denote by $z_t^{(t)}$ cash flows from our portfolio.
  - As before

$$dZ_t^{(t)} := d\tilde{Z}_t^{(t)} + z_t^{(t)}$$

Statistical Hedging

- If we trade $a \in R^{n_t}$ units of $H^{(t)}$ then tomorrow's portfolio is formally defined on the space of book-price processes as:

$$Z^{(t+1)} := Z^{(t)} \oplus a' H^{(t)}$$

- Our **rewards** are the mark-to-book values

$$dZ_t^{(t)} + a' dH_t^{(t)} - c_t(a)$$

- We measure success [1] with **Monetary Utilities**, essentially negative Convex Risk Measures given in terms of a utility function as:

$$U(X) := \sup_y E[u(X+y) - y]$$

  - For $u(x) := ((1 - \exp(-\lambda x))/\lambda$ we obtain the entropy (approximately "mean-variance").
  - For $u(x) := (1+\lambda)\min\{x, 0\}$ we obtain CVaR.

[1] Statistical Hedging, Buehler 2017 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2913250

Statistical Hedging

Statistical Hedging

- Assume we have a time series of $dZ_t^{(t)}$ and $dH_t^{(t)}$:
- The statistical hedging problem is given as [1,2]:

$$W^*\left(Z^{(t)}, s_t\right) = \sup_{\text{a}} U_t\left[dZ_t^{(t)} + a' dH_t^{(t)} - c_t(a)\right]$$

solves for the locally optimal risk-adjusted hedge.

- Basically Markoviz optimization for non-linear assets.
- Very intuitive: find statistically best hedge for your portfolio using regression.
- Key point: we are using derivative returns … not just spot/FX as usual.

[1] Statistical Hedging, Buehler 2017 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2913250
[2] Rama's Dynamic Hedging of Portfolio Credit Derivatives SIAM J. FINANCIAL MATH. 2011 Vol. 2, pp. 112–140

Statistical Hedging

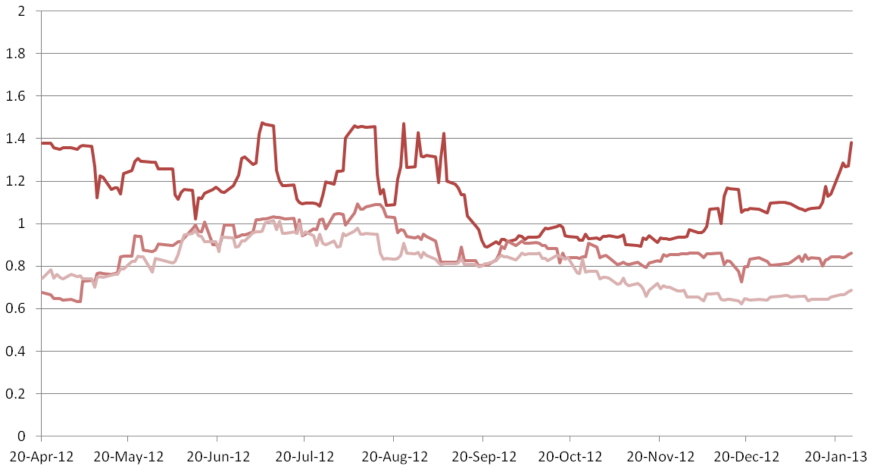
Statistical Hedging

- "Assume we have a time series of $dZ_t^{(t)}$ and $dH_t^{(t)}$"…. is a bit more difficult than it sounds:
- That means time series of the *relative same* instrument:
  - For options, historic returns of options with the same time-to-expiry and moneyness. Even for options this requires interpolation on the surface.
  - For exotics, same relative situation w.r.t. to barriers and other product parameters.
  - Particularly tricky for variance-related products.

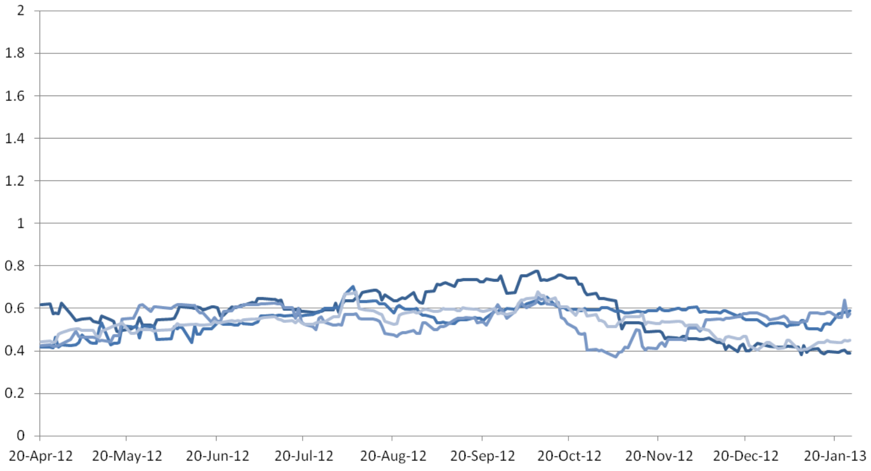
Statistical Hedging

Less vol is better

STOXX50E DBLBR_PUT_100_DKI_80D_UKO_120D 2Y_1
Greeks Hedging: Robust Vol

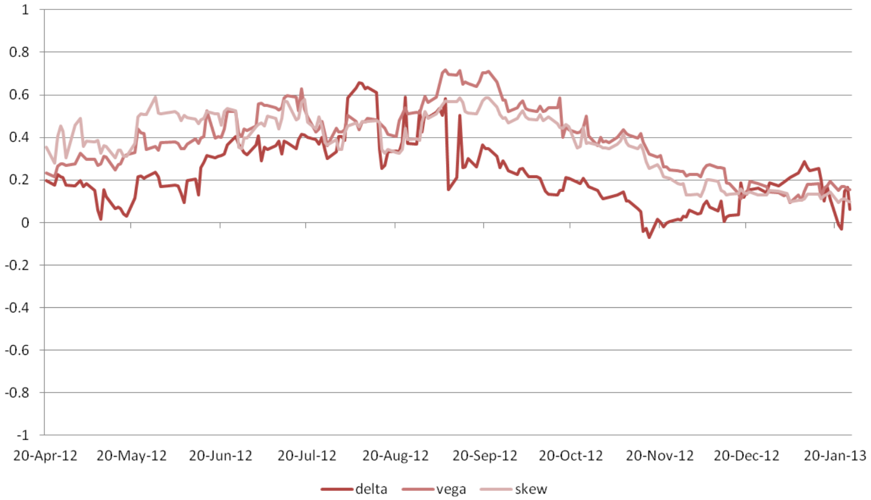


STOXX50E DBLBR_PUT_100_DKI_80D_UKO_120D 2Y_1
Statistical Hedging: Robust Vol

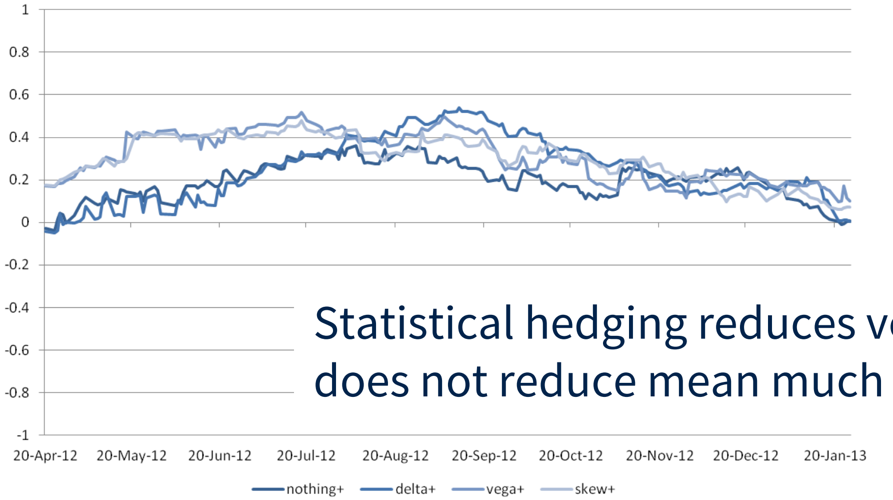


Mean should be zero

STOXX50E DBLBR_PUT_100_DKI_80D_UKO_120D 2Y_1
Greeks Hedging: Robust Mean



STOXX50E DBLBR_PUT_100_DKI_80D_UKO_120D 2Y_1
Statistical Hedging: Robust Mean



Statistical hedging reduces volatility but does not reduce mean much (CVaR 90%)

Statistical Hedging

UNIVERSITY OF OXFORD

Parameter Hedging

- Simplistic alternative which does not require time series of prices, just market data $x_t \subset s_t$ (spots, implied vols, rates etc).
  - Expand $\mathrm{X} = Z^{(t)}, H^{(t)}$ with Taylor in all relevant market data $x_t \subset s_t$:

$$dX_t \approx (\partial_x X) dx_t + \frac{1}{2}(\partial_{xx}^2 X)\, dx_t^2 + \partial_t X dt$$

  - Estimate normal parameters $dx_t = \mu dt + \sigma dW_t$ and write

$$dX_t \approx \left(\partial_t X + (\partial_x X)\mu + \frac{1}{2}(\partial_{xx}^2 X)\sigma^2\right) dt + (\partial_x X)\sigma dW_t$$

  - Solve mean-variance optimal portfolio $dZ_t^{(t)} + a' dH_t^{(t)}$ with transaction cost.
  - Without transaction this yields the *regression hedge*

$$a = \left[\left(\partial_x H_t^{(t)}\right)\sigma^2\left(\partial_x H_t^{(t)}\right)\right]^{-1}\left[\left(\partial_x Z_t^{(t)}\right)\sigma^2\left(\partial_x H_t^{(t)}\right) + \frac{1}{\lambda}\mathrm{drift}\ H\right]$$

[1] Statistical Hedging, Buehler 2017 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2913250
[2] Rama's Dynamic Hedging of Portfolio Credit Derivatives SIAM J. FINANCIAL MATH. 2011 Vol. 2, pp. 112–140

Statistical Hedging

Statistical Hedging

- Very intuitive – baseline for “RoboTrader”.
- However, can only locally account for pricing errors.
- Manages risk, does not price.

- Fun fact: if we were under the risk-neutral measure then statistical hedging without cost being essentially local regression would yield the risk-neutral hedge.

Deep Bellman Hedging

- Natural extension: Bellman approach [1]:

$$V^*\left(Z_t^{(t)}, s_t\right) \coloneqq \sup_{a} U_t\left[\beta_t V^*\left(Z^{(t)} \oplus a' H^{(t)}, s_{t+1}\right) + dZ_t^{(t)} + a' dH_t^{(t)} - c_t(a)\right]$$

  - V^* is the value function of the residual model error.
  - If we are already under a risk-neutral pricing measure then absent cost $V^* \equiv 0$, and we are back to "regression".

—

[1] Deep Bellman Hedging https://arxiv.org/abs/2207.00932

Deep Bellman Hedging

- Theorem [1]: define the Bellman Operator T for candidate value functions V such that

$$TV\left(Z^{(t)}, s_t\right) := \sup_{a} U_t \left[\beta_t V\left(Z^{(t)} \oplus a' H^{(t)}, s_{t+1}\right) + dZ_t^{(t)} + a' dH_t^{(t)} - c_t(a)\right]$$

Assume that there is **no statistical arbitrage** such that $T0 < \infty$.

- For any monetary utility U the sequence $V^{n+1} := TV^n$ starting in $V^0 := 0$ converges to a finite optimal solution V^*.
- The solution V^* then satisfies the Belman equation.
  - Scetch of proof: T is a contraction operator because of monotonicity and cash-invariance of U:
$$(Tg)(s) \leq T(f + |f-g|) = T(f)(s) + \beta^*|f-g|$$
$$(Tf)(s) \leq T(h + |f-g|) = T(g)(s) + \beta^*|f-g|$$
hence $|T(g) - T(f)| \leq \beta^*|f-g|$.

[1] Deep Bellman Hedging, Buehler et al, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4151026

Deep Bellman Hedging

- Numerical solution via Actor-Critic:

$$TV\left(Z_t^{(t)}, s_t\right) := \sup_{a,y} E_t\left[u\left(y + \beta_t V\left(Z^{(t)} \oplus a' H^{(t)}, s_{t+1}\right) + dZ_t^{(t)} + a' dH_t^{(t)} - c_t(a)\right) - y\right]$$

- **Actor:** Start in $V^0 := 0$; given V^{n-1} maximize a^n and y^n which also yields samples of TV^n.

- **Critic:** given samples of TV^{n-1} solve for a neural network V^n to satisfy

$$V^n(w, s_t) \equiv TV^{n-1}(w, s_t)$$

- This interpolation program may also be solved by simpler, classic methods such as kernel interpolators.

Deep Bellman Hedging

- Implementation with fixed T: worked [1]

- Actual actor/critic: pretty unstable with results so far for only simple cases (such as portfolios of vanillas)

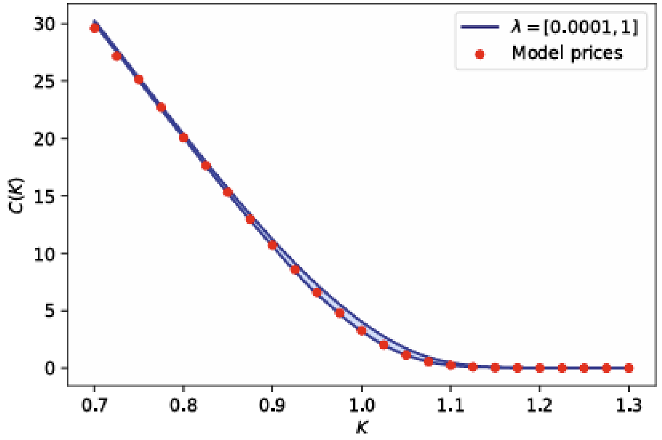


Figure 1: Call prices obtained by the value function for maturity $T = 30$ compared with risk-neutral model prices. Low λ produces prices very close to the model price, and higher λ results in additional risk premium added.

---
[1] Deep Hedging: Continuous Reinforcement Learning for Hedging of General Portfolios across Multiple Risk Aversions, Murray et al, https://arxiv.org/abs/2207.07467

Deep Bellman Hedging

Deep Bellman Hedging

$$V^*\left(Z_t^{(t)}, s_t\right) := \sup_{\mathrm{a}} U_t\left[\beta_t V^*\left(Z^{(t)} \oplus a' H^{(t)}, s_{t+1}\right) + dZ_t^{(t)} + a' dH_t^{(t)} - c_t(a)\right]$$

- How do we represent our portfolio → use what trading has been using for years: the greeks and risk scenarios available to them as well as any cashflows.
  - Typically, we have these values historically at individual hedging and OTC instrument level.
  - We can therefore sample over historic periods and solve for V^*.

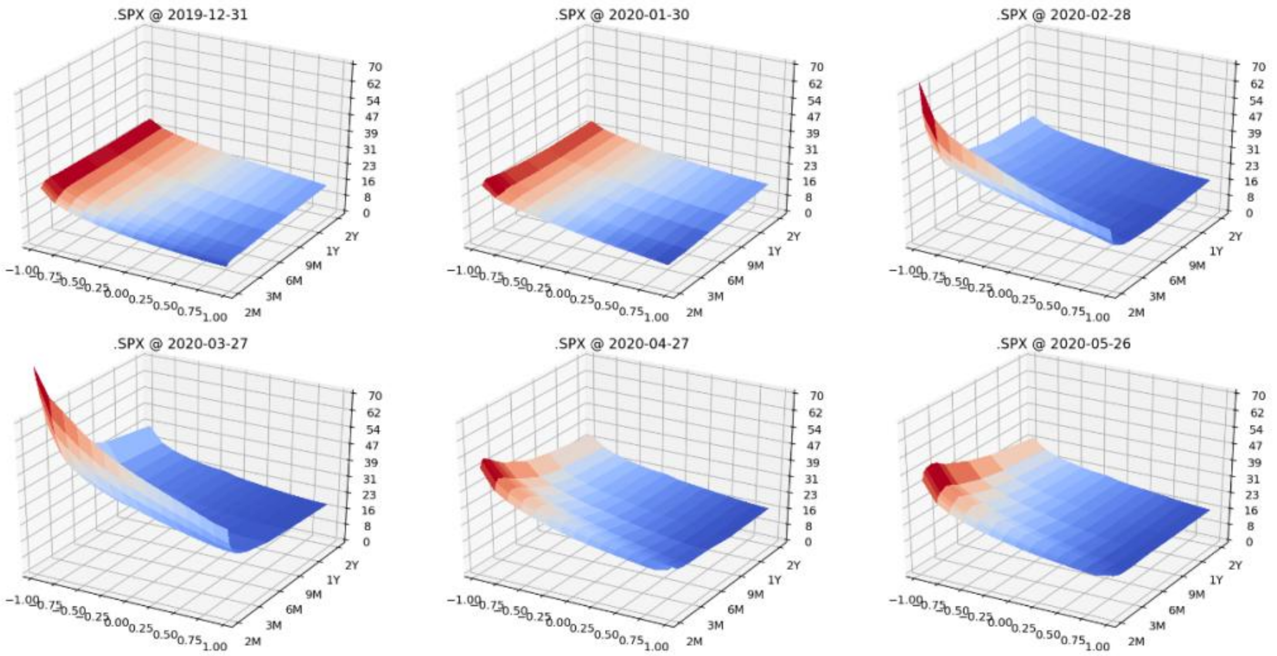
---
[1] Deep Bellman Hedging https://arxiv.org/abs/2207.00932

Market Simulation

Market Simulation

- Ideally simulate against purely historic data.
- Lack of data: for example, S&P 500 index spot and option prices.
  - 10Y of data ~ 2500 data points
  - Typically index option surface has 1000’s of options

→ Market Simulator



Market Simulation

- There are several machine learning methods, all discussed in various papers
  - PCA [1]
  - Autoencoders [2], Variational Autoencoders [2], [3]
  - Multi-asset versions [4]
  - SDE methods [5]
  - Neural SDEs [6]

__

[1] Deep hedging: learning to remove the drift, Buehler et al 2022 https://www.risk.net/cutting-edge/banking/7932226/deep-hedging-learning-to-remove-the-drift and https://arxiv.org/abs/2111.07844
[2] Deep Hedging: Learning to Simulate Equity Option Markets, Wiese et al 2020 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3470756
[3] VolGAN: a generative model for arbitrage-free implied volatility surfaces 2023 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4617536
[4] Multi-Asset Spot and Option Market Simulation, Wiese et al 2021 https://arxiv.org/abs/2112.06823
[5] Arbitrage-free neural-SDE market models, Cohen et al, 2021, https://arxiv.org/abs/2105.11053
[6] Neural SDEs as Infinite-Dimensional GANs, Kidger et al, 2021, https://arxiv.org/abs/2102.03657

Market Simulation

Wish List

- Absence of Static Arbitrage ← SANOS
  - Trader should not fall for false arbitrage opportunities.
  - Can arise from real static arbitrage and also from sampling (e.g. S&P never fell by 25% … can I therefore just sell lots of short term gap puts?)

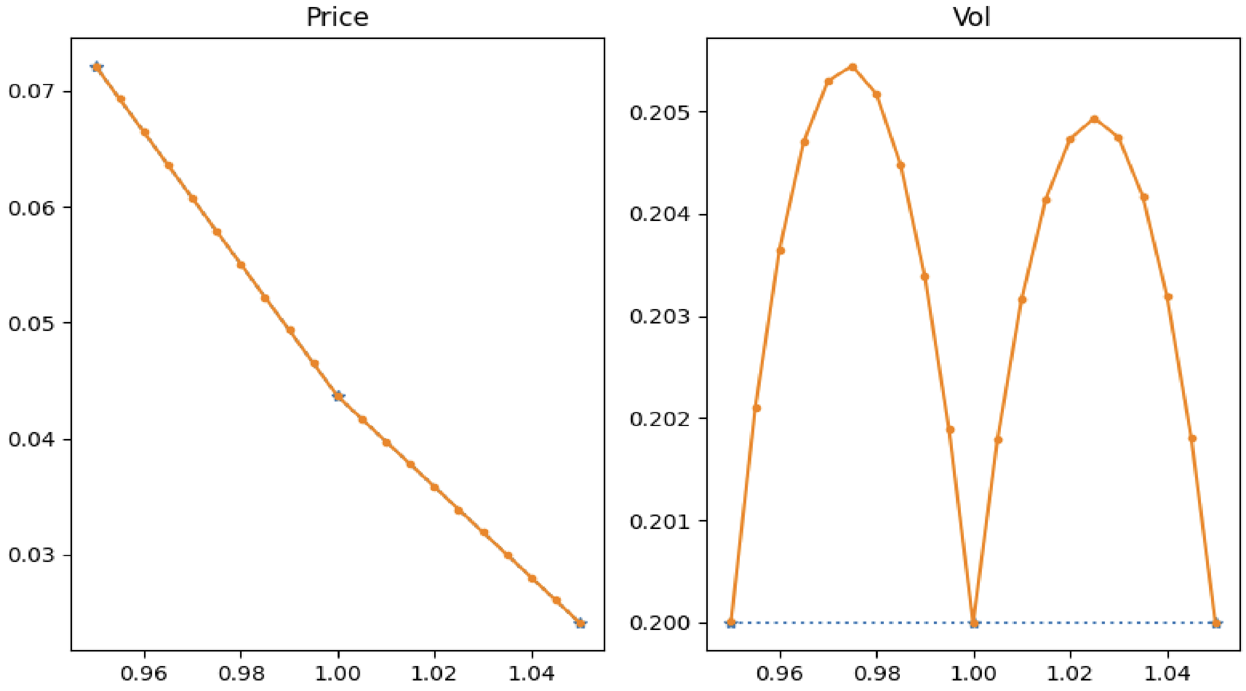
- Does it matter
  - Should only matter if bigger than bid/ask and/or drowned in sampling noise.
  - Practical idea in [1] was to simply remove paths which have static arbitrage (as it does not happen often).

[1] VolGAN: a generative model for arbitrage-free implied volatility surfaces 2023 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4617536

Market Simulation

Wish List

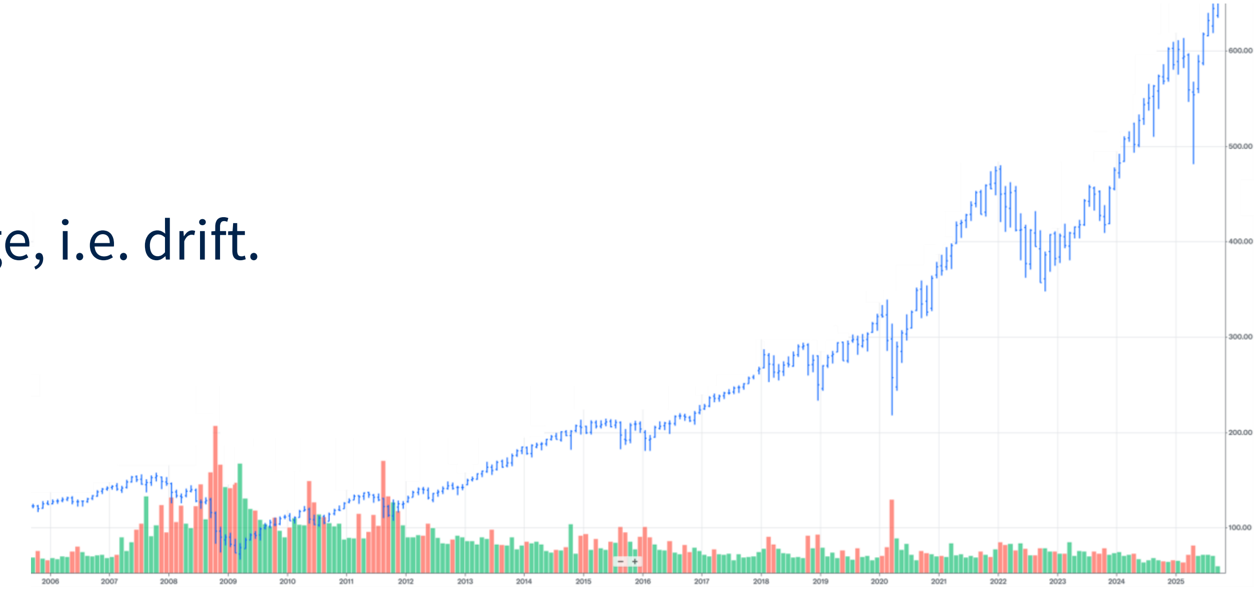
- Price all Options ← SANOS
  - Able to price every option (continuous surface) … realistically.
  - Linear interpolation is "too expensive": example 20% BS prices interpolated linearly.
  - Basically it selling interpolated options is near-arbitrage.



Market Simulation

Wish List

- Manage Statistical Arbitrage
  - Absence/Control of statistical arbitrage, i.e. drift.
  - Drift is a fact in real life
  - High model uncertainty for the drift.



- Approaches
  - Practical: minimize risk, not risk-adjusted return. Drops monotonicity, though.
  - Change to a risk-neutral measure [1, 2] ← have some slides if interested.
  - Optimize under model uncertainty [3].

[1] Arbitrage-free neural-SDE market models, Cohen et al, 2021, https://arxiv.org/abs/2105.11053
[2] Deep hedging: learning to remove the drift, Buehler et al 2022 https://www.risk.net/cutting-edge/banking/7932226/deep-hedging-learning-to-remove-the-drift and https://arxiv.org/abs/2111.07844
[3] Uncertainty-Aware Strategies: A Model-Agnostic Framework for Robust Financial Optimization through Subsampling, Buehler et al, 2025, https://arxiv.org/abs/2506.07299

Market Simulation

Ideal setup:

- Model something with trivial no-arb constraints, e.g. local volatility
- Good idea: **Discrete Local Volatility** [1] defined over model expiries $0 < \tau_1 < \cdots < \tau_m$ and non-homogeneous coarse strikes $0 < k_j^1 < \cdots 1 \cdots < k_j^{n_j}$ as:

$$\Sigma_j^i := \sqrt{\frac{\left(k_j^i\right)^2 (\tau_j - \tau_{j-1})}{k_j^{i+1} - k_j^{i-1}} \frac{dC_j^i}{q_j^i}}$$

- 1:1 map between arbitrage-free call prices and DLVs.
- Can be simulated, too: [2, 3, 4]
- However, how do we interpolate between strikes? Linear → "too expensive".

—

[1] Hans Buehler and Evgeny Ryskin. Discrete local volatility for large time steps(extended version). https://papers.ssrn.com/abstract=2642630 2015.
[2] Deep Hedging: Learning to Simulate Equity Option Markets, Wiese et al 2019 https://arxiv.org/abs/1911.01700
[2] Multi-Asset Spot and Option Market Simulation, Wiese et al 2021 https://arxiv.org/abs/2112.06823
[3] Arbitrage-free neural-SDE market models, Cohen et al, 2021, https://arxiv.org/abs/2105.11053

SANOS

Strictly Arbitrage-free Non-parametric Option Surfaces

SANOS

UNIVERSITY OF OXFORD

- SANOS [1] provides a smooth, strictly arbitrage-free option surface interpolation.
- Basic definition:
  - Let $0 < \tau_1 < \cdots < \tau_M$ be model expiries with strikes $0 < k_j^1 < \cdots 1 \cdots < k_j^{n_j}$.
  - Let $V(t)$ be a variance curve (e.g. ATM variance) and $\eta \geq 0$ a smoothing parameter.
  - Assume that $q_1, \ldots, q_M$ is a sequence of martingale densities over the model strikes.
  - Finally assume $\alpha_j(T) \coloneqq (T - \tau_{j-1})/(\tau_j - \tau_{j-1})$ or a smoother version thereof.
  - Define then

$$C(T,K) \coloneqq \begin{array}{l} \alpha_j(T) \displaystyle\sum_{i=1}^{N_j} q_j^i \, BSCall\left(k_j^i, K, \eta V(T)\right) + \\ (1 - \alpha_j(T)) \displaystyle\sum_{i=1}^{N_{j-1}} q_{j-1}^i \, BSCall\left(k_{j-1}^i, K, \eta V(T)\right) \end{array}$$

  - Smooth version of "Discrete Local Vol" [1] which is given by $\eta = 0$.

__
[1] Hans Buehler and Evgeny Ryskin. Discrete local volatility for large time steps(extended version). https://papers.ssrn.com/abstract=2642630 2015.

SANOS

Absence of Arbitrage in Continuous Space and Time

- Theorem [1]: Let $C(T,K)$ be a pure call price function for $T \geq 0$ and $K \geq 0$. Then there exist martingale X under some measure P such that

$$C(T,K) = \mathrm{E}[(X_T - K)^+]$$

if and only if(*)

1. $C(T,0) = 1, \lim_{K\uparrow\infty} C(T,K) = 0,$
2. $C(T,\cdot)$ is non-increasing with $\partial_K C(T,0) \in [-1,0),$
3. $C(T,\cdot)$ is convex, and
4. $C(\cdot,K)$ is non-decreasing.
5. Moreover, $X > 0$ if and only if $\partial_K C(T,0) = -1.$

- These are minimal; none of these conditions can be dropped [1].

6. In particular, $\partial_K C(T,K) \leq 0$ and $\partial_{KK} C(T,K) \geq 0$ alone are not sufficient.
7. On the flip side, $\partial_K C(T,K) \in [-1,0]$ and $C \geq \max\{1-K,0\}$ follow from above.

—

[1] Hans Buehler. Expensive martingales. Quantitative Finance, 6(3):207–218, 2006.

(*) All derivatives are right hand derivatives, e.g. $\partial_K C(T,K) := \lim_{\epsilon\downarrow 0} \frac{1}{\epsilon}(C(T,K+\epsilon) - C(T,K))$

SANOS

- SANOS [1] provides a smooth, strictly arbitrage-free option surface interpolation.

$$C(T,K) := \begin{matrix} \alpha_j(T) & \displaystyle\sum_{i=1}^{N_j} q_j^i \, BSCall\left(k_j^i, K, \eta V(T)\right) + \\ (1-\alpha_j(T)) & \displaystyle\sum_{i=1}^{N_{j-1}} q_{j-1}^i \, BSCall\left(k_{j-1}^i, K, \eta V(T)\right) \end{matrix}$$

- Direct proof:
  - $C(T,\cdot) \in C^\infty$ for all $\eta > 0$
  - $C(\cdot,K)$ is as smooth as α and V are. Using monotone splines makes it C^1.
  - Being a convex sum of arbitrage-free operators means C is arbitrage-free; main step is to prove that $C(\tau_j, K) \geq C(\tau_{j-1}, K)$.

SANOS

- Let Y be the log-normal process with variance ηV. Then

$$C(\tau_j, K) := \sum_{i=1}^{n_j} \mathrm{E}\left[\left(Y_j\, k_j^i - K\right)^+\right] q_j^i \geq \sum_{i=1}^{k_j} \mathrm{E}\left[\left(Y_{j-1}\, k_j^i - K\right)^+\right] q_j^i = (*)$$

- Now we use the martingale property of q which means for any convex function f

$$\sum_{i=1}^{n_j} f(k_j^i) q_j^i \geq \sum_{i=1}^{n_{j-1}} f(k_{j-1}^i) q_{j-1}^i$$

indeed:

$$(*) = E\left[Y_{j-1} \sum_{i=1}^{n_j} \left(k_j^i - \frac{K}{Y_j}\right)^+ q_j^i\right] \geq E\left[Y_{j-1} \sum_{i=1}^{n_{j-1}} \left(k_{j-1}^i - \frac{K}{Y_{j-1}}\right)^+ q_{j-1}^i\right] = C(\tau_{j-1}, K)$$

SANOS

- In fact, the last step is a consequence of the observation that a time-discrete process which is consistent with C is

$$Z_j := Y_j X_j$$

  where X is the time-and-space discrete process implied by q, independent of Y.

- Other processes such as Heston for Y are possible, but for our experiments so far log-normal seems to perform surprisingly well.

Fitting to the Market

- Assume now we are given **market expiries** $0 < T_1 < \cdots < T_M$ with **market strikes** $0 = K_\ell^0 < K_\ell^1 < \cdots 1 \cdots < K_\ell^{N_\ell}$ with observed bids $B_\ell^i > 0$ and asks $A_\ell^i > B_\ell^i$.
- We note that given $v_j \coloneqq V(T_j)$ the model prices $C_\ell^i = C(T_\ell, K_\ell^i)$ are linear functions of the martingale density q:

$$C(T,K) \coloneqq \alpha_j(T) \sum_{i=1}^{N_j} q_j^i \, \mathrm{BS}Call\left(k_j^i, K, v_T\right) + (1 - \alpha_j(T)) \sum_{i=1}^{N_{j-1}} q_{j-1}^i \, \mathrm{BS}Call\left(k_{j-1}^i, K, v_T\right)$$

Fitting to the Market

UNIVERSITY OF OXFORD

- Then [1] we may find a closest arbitrage-free linear interpolation of discrete call prices C_j^i jointly across all strikes and expiries by solving the **linear** problem

$$\min_q : \sum_{\ell=1}^{M} \sum_{i=1}^{N_j} \frac{1}{\gamma_\ell^i} \left\{ \left(C_\ell^i - A_\ell^i \right)^+ + \left(B_\ell^i - C_\ell^i \right)^+ + \epsilon \left| C_\ell^i - M_\ell^i \right| \right\}$$

for mid prices $M_j^i := (A_j^i + B_j^i)/2$ and bid/ask spreads $\gamma_j^i := A_j^i - B_j^i > 0$ with linear conditions of the form:

- $C := \Xi\, q$
- $q_j \geq 0, 1' q_j = 1, k_j' q_j = 1$
- $q_j \left(k_j - k_j^\ell \right)^+ \geq q_{j-1} \left(k_{j-1} - k_j^\ell \right)^+$

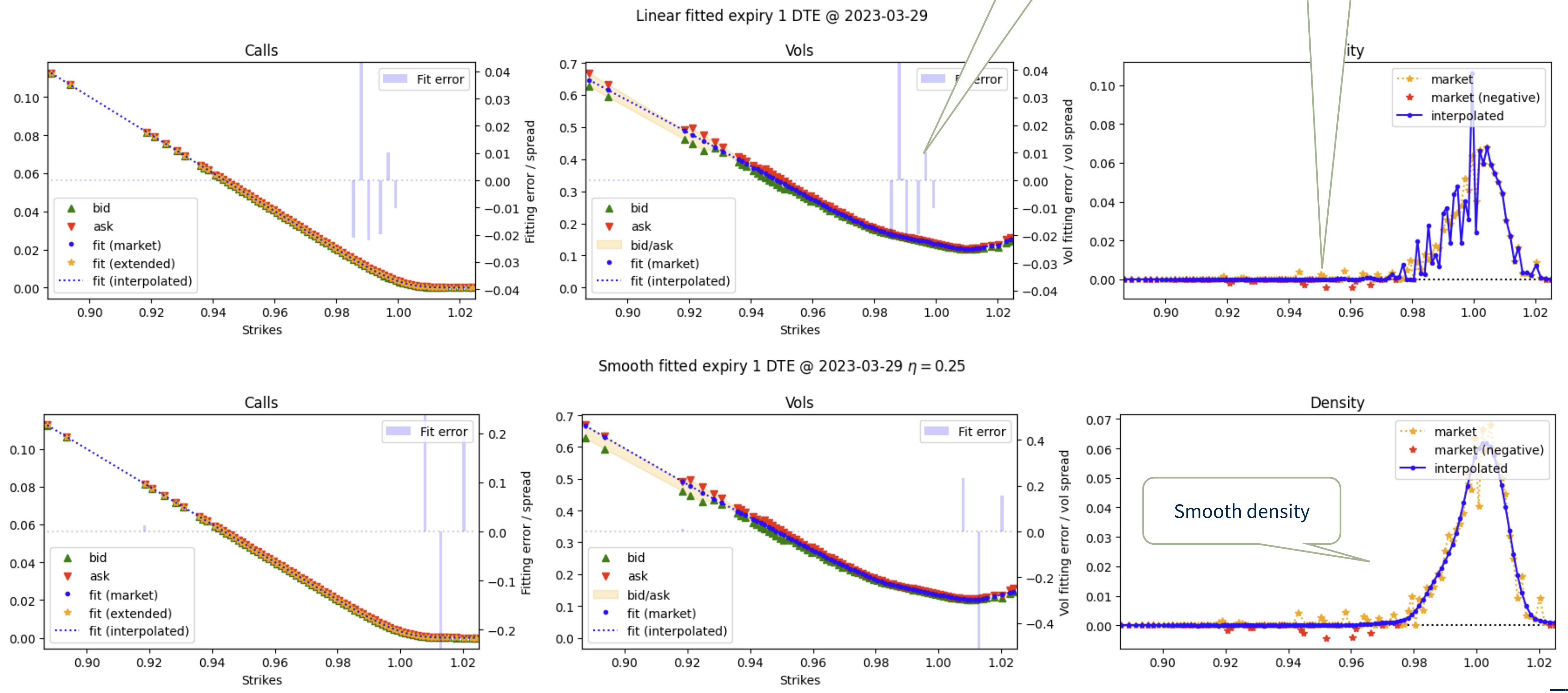
Straight forward linear programming for the global fit.

[1] Buehler et al, 2026 SANOS https://arxiv.org/pdf/2601.11209

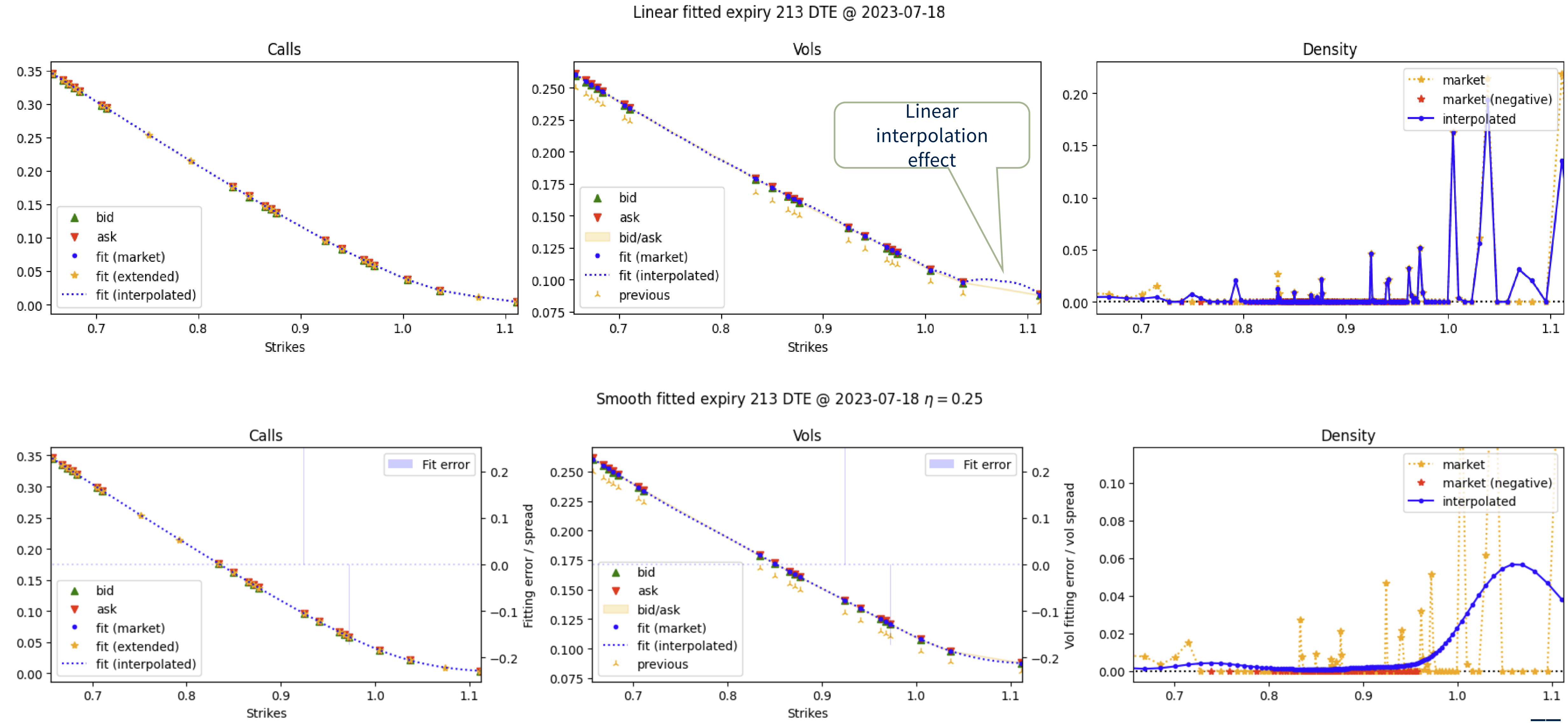
Fitting to the Market

Market data has arbitrage

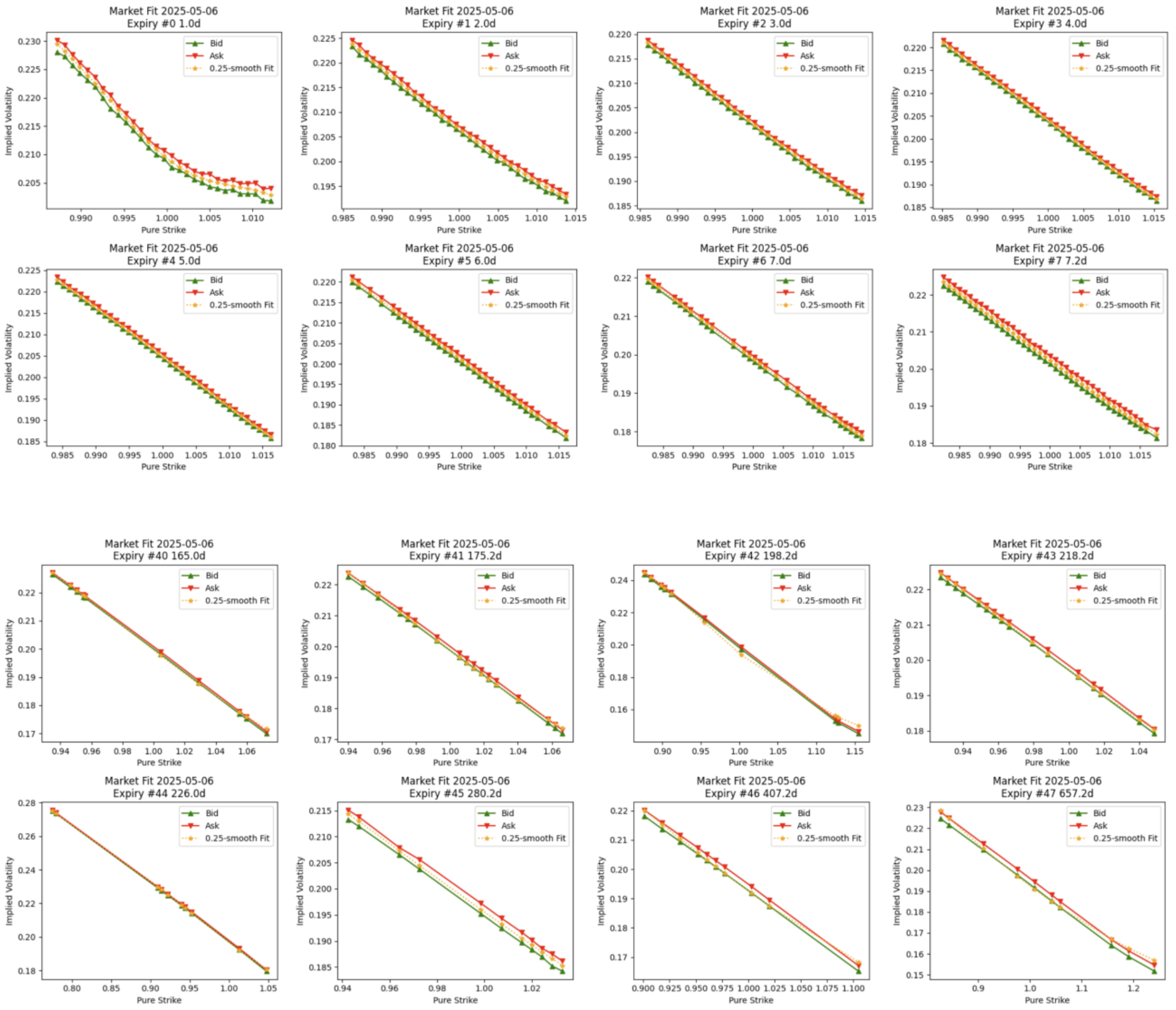
Market mid has arbitrage



Fitting to the Market



Fitting to the Market

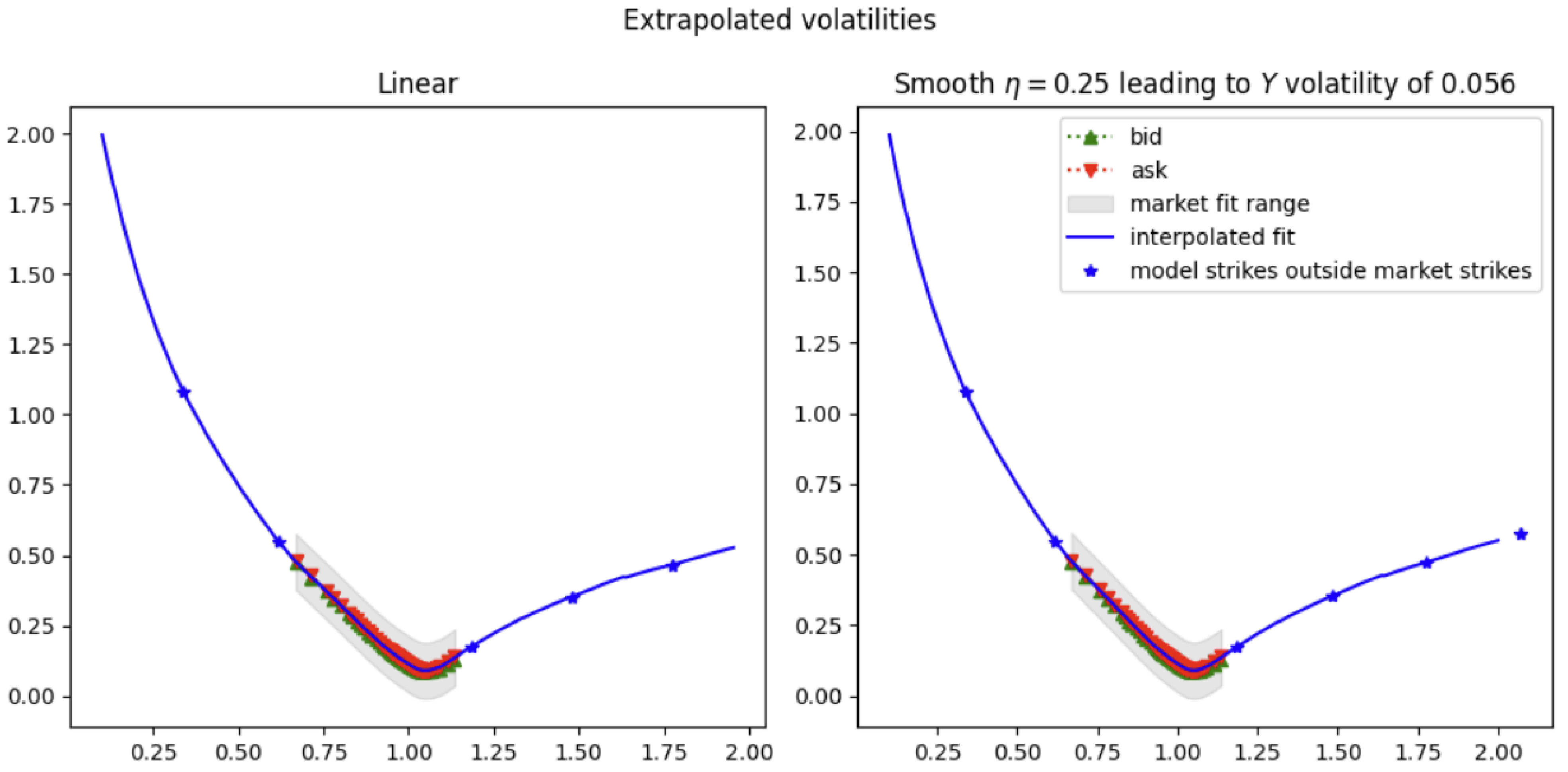


Fitted 1000 options within 2 ATM implied volatility standard deviations which had a Vega/sqrtT of at least 0.1% on 2025-05-06 across all 48 expiries from 1D to 657 business days.

Options were chosen by closeness to ATM. The model fitted 91.4% of all options within bid/ask. Of those options not fitted the median error is just 21% of half spread.

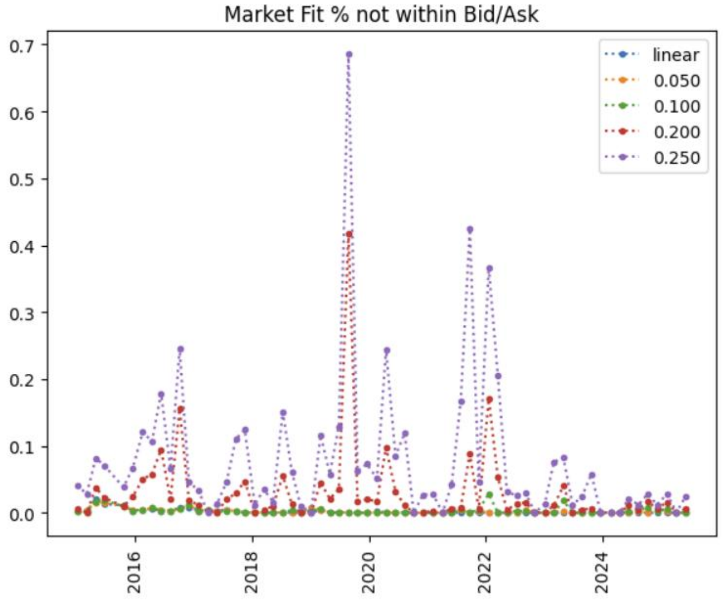
The fit took sub-seconds on a desktop PC.

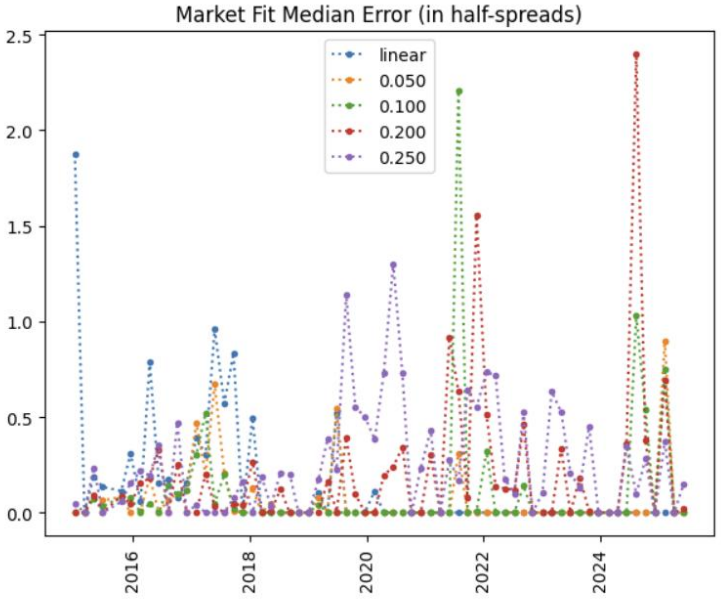
Fitting to the Market

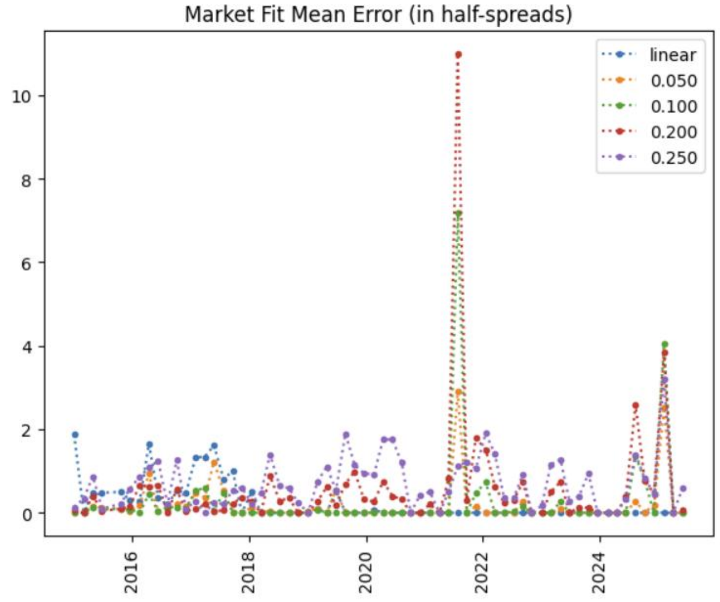


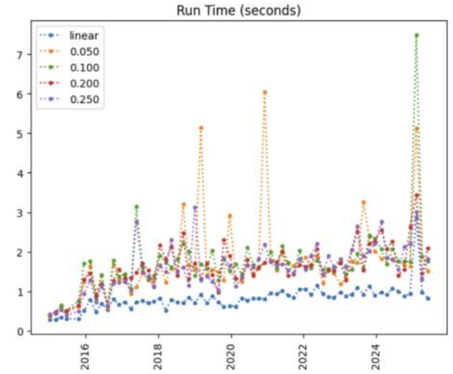
Fitting to the Market

- Fitting 1000 options within 2 ATM implied volatility standard deviations which had a Vega/sqrtT of at least 0.1% across expiries









Fitting to the Market

UNIVERSITY OF OXFORD

- Similar to [1] we may also attempt to "hard fit" our model within bid/ask:

$$\min_q : \sum_{\ell=1}^{M} \sum_{i=1}^{N_j} \frac{1}{\gamma_\ell^i} \left| C_\ell^i - M_\ell^i \right|$$

  with linear conditions of the form:
  - $C := \Xi\, q$
  - $q_j \geq 0, 1' q_j = 1, k_j' q_j = 1$
  - $q_j \left(k_j - k_j^\ell\right)^+ \geq q_{j-1}\left(k_{j-1} - k_j^\ell\right)^+$
  - $C_\ell^i \leq A_\ell^i, B_\ell^i \leq C_\ell^i$
- Assuming the strike grid satisfies $k_j \subset k_{j+1}$ then existence of q for the linear case $\eta = 0$ is equivalent to absence of actionable arbitrage [2].
- Hard fit across the surface often fails. Useful as per-expiry pre-processor.

[1] Cohen et al, 2020 Detecting and Repairing Arbitrage in Traded Option Prices https://arxiv.org/abs/2008.09454

Fitting to the Market

- Similar to [1] we may also attempt to "hard fit" our model within bid/ask:

$$\min_q : \sum_{\ell=1}^{M} \sum_{i=1}^{N_j} \frac{1}{\gamma_\ell^i} \left| C_\ell^i - M_\ell^i \right|$$

  with linear conditions of the form:
  - $C \coloneqq \Xi\, q, c_j := U_j q_j, \bar{c}_{j|j-1} = R_j q_{j-1}$
  - $q_j \geq 0, 1' q_j = 1, k_j' q_j = 1$
  - $q_j \left(k_j - k_j^\ell\right)^+ \geq q_{j-1} \left(k_{j-1} - k_j^\ell\right)^+$
  - $C_\ell^i \leq A_\ell^i, B_\ell^i \leq C_\ell^i$
- Assuming the strike grid satisfies $k_j \subset k_{j+1}$ then existence of q for the linear case $\eta = 0$ is equivalent to absence of actionable arbitrage [2].
- Hard fit across the surface often fails. Useful as per-expiry pre-processor.

[1] Cohen et al, 2020 Detecting and Repairing Arbitrage in Traded Option Prices https://arxiv.org/abs/2008.09454

Discrete Local Volatility

Discrete Local Volatility

- SANOS is a smooth call price function whose model parameters are the martingale densities q and variances V.
- The former is not a nicely constrained parameter space for machine learning.
- *Next step:* reparametrize SANOS in easily modelled parameters.
- Assume for simplicity here strikes are the same across expires (we cover the general case in [1])

Discrete Local Volatility

- Define

$$w_j^{i+} := \frac{\left(k_j^i\right)^2}{\left(k_j^{i+1} - k_j^i\right)\left(k_j^{i+1} - k_j^i\right)} \qquad w_j^{i-} := \frac{\left(k_j^i\right)^2}{\left(k_j^{i+1} - k_j^i\right)\left(k_j^i - k_j^{i-1}\right)}$$

and then

$$\Omega_j := \begin{pmatrix} 0 & -w_j^{2-} & & & \\ & w_j^{2+} + w_j^{2-} & & & \\ & -w_j^{2+} & \ddots & -w_j^{(N_j-1)-} & \\ & & & w_j^{(N_j-1)+} + w_j^{(N_j-1)-} & \\ & & & -w_j^{(N_j-1)+} & 0 \end{pmatrix} \in R^{N_j \times N_j}$$

Discrete Local Volatility

UNIVERSITY OF OXFORD

Discrete Local Volatility

- This yields the tri-band matrix

$$Q_j^{-1} = E_j + \Omega_j \cdot \hat{\Sigma}_j^2 (\tau_j - \tau_{j-1}) \in R^{n_j \times n_j}$$

  where "·" denotes row-wise multiplication and where E_j is the $R^{n_j \times n_j}$ unit matrix.

- **Theorem [1], see also [2]:** the inverse of Q_j is a martingale transition matrix for any discrete local volatility $\Sigma_j \geq 0$ for states k_j, i.e.
  - $Q_j \geq 0$
  - $1' Q_j = 1'$
  - $k_j' \, Q_j = k_j'$

---

[1] Hans Buehler and Evgeny Ryskin. Discrete local volatility for large time steps(extended version). https://papers.ssrn.com/abstract=2642630 2015.
[2] Buehler et al: SANOS Smooth strictly Arbitrage-free Non-parametric Option Surfaces, https://arxiv.org/abs/2601.11209, 2026

Discrete Local Volatility

- It remains to construct a matrix $L_j \in R^{n_j \times n_{j-1}}$ which transitions from k_{j-1} to k_j.
- Define:

$$L_j^{\ell,i} := \frac{\left(k_{j-1}^i - k_j^{\ell+1}\right)^+ - \left(k_{j-1}^i - k_j^{\ell}\right)^+}{k_j^{\ell+1} - k_j^{\ell}} - \frac{\left(k_{j-1}^i - k_j^{\ell}\right)^+ - \left(k_{j-1}^i - k_j^{\ell-1}\right)^+}{k_j^{\ell} - k_j^{\ell-1}}$$

Discrete Local Volatility

- The matrix

$$M_j := \left(E_j + \Omega_j \cdot \Sigma_j^2 (T_j - T_{j-1})\right)^{-1} L_j \in R^{N_j \times N_{j-1}}$$

is a martingale transition matrix for *any* set of discrete local volatilities $\Sigma_j \geq 0$, i.e. when starting in $q_1^0 = (1)$ we obtain a sequence of marginal densities $q_j := M_j q_{j-1}$ defined over k_j which are the marginal densities of a martingale.

- Importantly:
  - Ω_j and L_j can be computed separately from Σ_j ahead of training.
  - The inverse in above is a tri-band matrix which can be computed efficiently on a GPU using Thomas' method.
  - It is also trivial to compute DLVs from a fitted density q, c.f. [1].
  - It is straight forward to compute DLVs from a given set of densities q, c.f. [1].

[1] Buehler et al: SANOS Smooth strictly Arbitrage-free Non-parametric Option Surfaces, https://arxiv.org/abs/2601.11209, 2026

Learning To Trade ?

Learning To Trade

UNIVERSITY OF OXFORD

- Theory in place
- Market Simulator in progress
- Next step is *benchmarking*.

- Several alternative methods:
  - Classic model-based greek hedging (with models calibrated daily to simulated data) and modern "martingale transport" [1]
  - Regression/Parameter/Statistical Hedging
  - Deep Hedging, Deep Bellman Hedging, Deeper Hedging [2]
  - Hedging under model uncertainty
  - …...

[1] Robust pricing and hedging of options on multiple assets and its numerics, Eckstein et al 2019, https://arxiv.org/abs/1909.03870
[2] Deeper Hedging: A New Agent-based Model for Effective Deep Hedging , Gao et al, 2023, https://dl.acm.org/doi/fullHtml/10.1145/3604237.3626913

Learning To Trade

Learning from the AI community

- Access to data and standardized, clean, normalized data sets.
  Oxford just go access to massive.com; pipeline work in progress; please reach out if you want to help.

- Compute (GPUs)
  UK Isembard cluster; Nvidia startup program; …

- Benchmark problems for such data sets with clearly defined success metrics
  First trial is realized volatility and volume forecasting.

- Live competitions, leader boards etc.
  Nothing yet…

Learning To Trade

UNIVERSITY OF OXFORD

Data

- Access to data
- Standardized processing
- Shared pipelines

Compute

- Access to modern GPUs
- Ability to use them well

Benchmarks

- Well defined problems with data and success metrics
- Leaderboards
- Competitions

Models for Finance

- For non-stationary time series
- For sparse, wide, noisy data

Please ask questions

hans.buehler@maths.ox.ac.uk

UNIVERSITY OF
OXFORD

Market Simulation

UNIVERSITY OF OXFORD

Risk-Neutral Measures within Bid/Ask

- Changing to risk-neutral measure [1] using the minimal entropy martingale measure MEMM:
    1. Let $m_t^{\pm i} \geq 0$ be the marginal cost of buying (+) or selling (-) the *i*th instrument.
    2. Find for each step the optimal prop trading strategy

$$a_t^0 \equiv a^0(s_t) \coloneqq \sup_{y \in R, a \in R^{n_t}} : E\left[u\left(a' dH_t^{(t)} - a^+ m_t^+ - a^- m_t^- + y^0\right) - y^0\right]$$

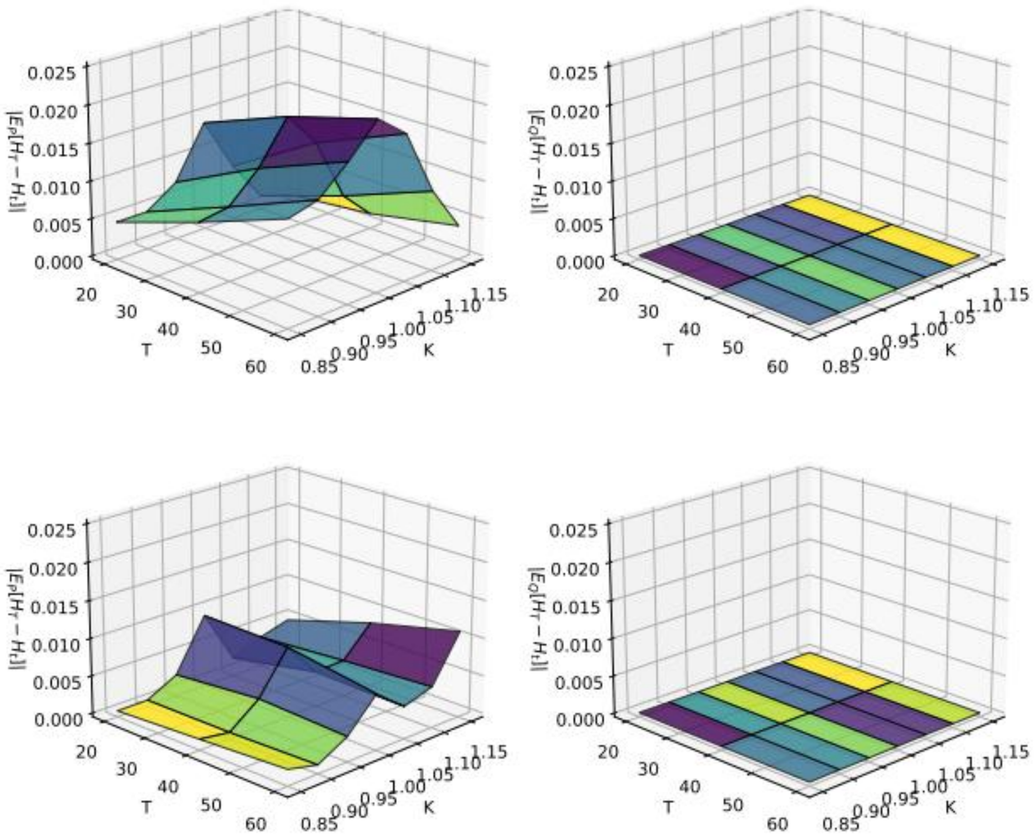
    3. Define the near-martingale measure Q as

$$\frac{dQ}{dP} = u'(a^0 dH_t^{(t)} + y^0)$$

    4. Then $m_t^{-i} \leq E_t^Q\left[dH_t^{(t)}\right] \leq m_t^{+i}$.

[1] Deep hedging: learning to remove the drift, Buehler et al 2022 https://www.risk.net/cutting-edge/banking/7932226/deep-hedging-learning-to-remove-the-drift and https://arxiv.org/abs/2111.07844

Market Simulation



Full market simulation results [1]: left are expected returns under P, right under Q under transaction cost

---
[1] Deep hedging: learning to remove the drift, Buehler et al 2022 https://www.risk.net/cutting-edge/banking/7932226/deep-hedging-learning-to-remove-the-drift and https://arxiv.org/abs/2111.07844