# Reconstruction by optimal transport: applications in cosmology and finance

G Loeper

School of Mathematical Sciences
Centre for Quantitative Finance and Investment Strategies
Monash University

Joint work with Ivan Guo and Shiyi Wang

November, 2017

# ODE's, SDE's, PDE's

$X(t,x) \in \mathbb{R}^n, X(0,x) = x, t \in [0,T]$

ODE $\partial_t X(t,x) = v(t, X(t,x))$

SDE $dX(t,x) = \mu dt + \sigma(t, X(t,x)) dW_t$

In both cases, the mass distribution $\rho(t,x)$ is defined by

$$\begin{array}{rcl} X(0,\cdot) & \sim & \rho_0 \\ \forall \varphi, & & \int_{\mathbb{R}^n} \varphi(x) d\rho(t,x) = \mathbb{E}(\varphi(X(t,x))) \end{array}$$

Corresponding pde's (mass conservation)

| | | |
|---|---|---|
| ODE: | $\partial_t \rho + \nabla \cdot (\rho v) = 0$ | continuity equation |
| SDE: | $\partial_t \rho + \nabla \cdot (\rho \mu) - \partial_{ij}(\frac{1}{2}\Sigma_{ij}\rho) = 0$ | forward Kolmogorov equation |

For the stochastic case let $\mu = 0, v = \frac{1}{2}\sigma^2, n = 1$

# General reconstruction problem

Find $v(t,x)$ and $\rho(t,x))$ to minimize

$$\mathcal{A}(\rho, v) = \int_0^T \mathcal{L}(\rho, v) dt$$

under the constraints

$$\rho(t=0) = \rho_0$$
$$\rho(t=T) = \rho_T$$
$$\partial_t \rho + \partial_x(\rho v) = 0 \text{ or } \partial_t \rho - \partial_{xx}(\rho v) = 0$$

Example:

$$\mathcal{L} = \int_{\mathbb{R}^n} \frac{1}{2} \rho F(v) dx + G(\rho)$$

for $F, G$ convex.

For $G = 0$, equivalent to $\mathcal{L} = \mathbb{E} \int_0^T F(v(t))$

# Convexification

The problem can be made convex:

$$\frac{1}{2}\rho|v|^2 = \frac{|J|^2}{2\rho} = \sup_{c+|m|^2/2\leq 0}\{\rho c + J\cdot m\}$$

is convex in $(\rho, J)$

More generally

$$\rho F(v) = \rho F\left(\frac{J}{\rho}\right) = \sup_{c+F^*(m)\leq 0}\{\rho c + J\cdot m\}$$

where $F^*$ is the Legendre transform of $F$

$\mathcal{L}$ and $\mathcal{A}$ become now convex in $(\rho, J)$. The constraints (initial and final density, conservation of mass) define a convex set among all pairs $\rho, J$.

Convex functional under linear constraints allows the use of classical tool of convex analysis.

# The reconstruction problem in cosmology

We know today the density of matter in the Universe (at least partially).

We also know that after the baryons/photons decoupling (just after Big Bang) the density was quasi-uniform

**Question :** From that, can we infer the initial positions of "particles" (galaxies ...) and their initial and current speed? (Peebles, 1989)

the motion is described in co-moving coordinates : deviations from a uniformly expanding motion.

$\rightarrow$ notions of initial position and velocity therefore make sense.

**Answer :** One can answer mathematically to this question. The reconstruction is unique, except in collapsed areas, where particles are indistinguishable.

# The gravitational Euler-Poisson system

$$\partial_t \rho + \nabla \cdot (\rho v) = 0 \quad (1)$$

$$\partial_t (\rho v) + \nabla \cdot (\rho v \otimes v) = -\rho \nabla p \quad (2)$$

$$\Delta p = \rho - 1 \quad (3)$$

**The reconstruction problem consists in finding solutions to this system knowing $\rho(t = 0)$ and $\rho(t = T)$.**

The system (1, 2, 3) is Hamiltonian,

$$\mathcal{H}(\rho, v) = \frac{1}{2} \int_{\mathbb{T}^d} \rho(t, x)|v(t, x)|^2 - |\nabla p(t, x)|^2 dx.$$

The Lagrangien is

$$\mathcal{L}(\rho, v) = \int_{\mathbb{T}^d} \rho(t, x)|v(t, x)|^2 + |\nabla p(t, x)|^2 \, dx.$$

The critical points for the the action of the Lagrangian1

$$\mathcal{A} = \int_0^T \mathcal{L}(\rho(t), v(t))\ dt$$

under constraints of mass conservation, initial and final densities will be solutions of $(EP)$.

The Lagrangian is **convex** by choosing wisely the variables (i.e. $\rho, J$ and not $\rho, v$

Critical points of $\mathcal{A}$ will be minimisers.

# The minimisation problem

Minimise the action

$$\mathcal{A}(\rho, v, p) = \frac{1}{2} \int_0^T \int_{\mathbb{T}^d} \rho(t,x)|v(t,x)|^2 + |\nabla p(t,x)|^2 \, dxdt,$$

## Problem

*Find* $\bar{\rho}, \bar{v}, \bar{p}$ *such that*

$$\mathcal{A}(\bar{\rho}, \bar{v}, \bar{p}) = \inf \mathcal{A}(\rho, v, p)$$

*among all* $\rho, v, p$ *satisfying*

$$\partial_t \rho + \nabla \cdot (\rho v) = 0,$$
$$\Delta p = \rho - 1$$
$$\rho|_{t=0} = \rho_0, \;\; \rho|_{t=T} = \rho_T.$$

# The dual problem (Fenchel-Rockafellar duality theorem)

Assume one can find an admissible triplet $(\rho, J, p)$ *i.e.* such that

$$\partial_t \rho + \nabla \cdot J = 0, \qquad \Delta p = \rho - 1$$
$$\rho|_{t=0} = \rho_0, \ \rho|_{t=T} = \rho_T$$
$$\mathcal{A}(\rho, v, p) < +\infty$$

- One can find such a triplet if $\rho_0, \rho_T \in L^{\frac{2d}{d+2}}$. Otherwise ???

## The dual problem

$$\sup_{\phi,q}\{\int_{\mathbb{T}^d} \rho_T \phi(T) - \rho_0 \phi(0) \ dx + \int_{\mathbb{T}^d \times [0,T]} \rho q + \nabla p \cdot \nabla q - |\nabla q|^2/2 \ dtdx\}$$

among pairs $(\phi, q)$ such that $\partial_t \phi + \frac{|\nabla \phi|^2}{2} + q \leq 0$.

# Properties of the optimizer

The infimum is attained for some pair $\rho, J = \rho v$

$\phi$ and $q$ are the Lagrange multipliers of the constraints of mass conservation and Poisson coupling

If $(\rho, J = \rho v, p)$ is an optimal solution, then for all maximizing sequence $\phi_\varepsilon, q_\varepsilon$

$$\int_{[0,T]\times\mathbb{T}^d} \frac{1}{2}\rho|v - \nabla\phi_\varepsilon|^2 + \frac{1}{2}|\nabla p - \nabla q_\varepsilon|^2 \; dtdx \to 0$$

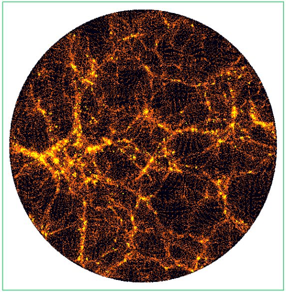
Implies uniqueness

# MAK Reconstruction

Densities are represented by $N$ points of equal masses. numerical solution by assignment algorithm (Bertseakas)

complexity: $O(N^3)$

MAK reconstruction tested on N-body simulation. Yellow points indicate a failed reconstruction. 60 % of points accurately reconstructed.



Alternative approach (Rapetti, Loeper (2005)): direct solution of the Monge-Ampere equation

$\Rightarrow$ optimal cost $O(N \log N)$, $N$ number of grid points

# Financial modelling (j.w. with Ivan Guo and Shiyi Wang, Monash)

- Stock price modelled by random process $S_t(\omega), t \geq 0, \omega \in \Omega$ living on a probability space $\Omega, \mathbb{F}, \mathbb{P}$
- Option: contract that pays $\Phi(S_T)$ at time $T$
- Call: $(S_T - K)^+$, Put $(K - S_T)^+$
- Fundamental theorem of asset pricing: Options can be priced by taking expectations under a risk neutral measure:

$$P(s,t) = \mathbb{E}^{\mathbb{Q}}(\Phi(S_T))$$

  where, under the probability $\mathbb{Q}$, $S_t$ follows

$$\frac{dS_t}{S_t} = \sigma_t dW_t^{\mathbb{Q}}$$

  .

- This work is about calibrating $\sigma$

# Calibration

- Calibration: if $\sigma$ follows a model with parameters, find the parameters from observed market prices
- Observed market prices: Calls and puts at maturities $T_i$ (monthly) and strikes not too far from the money:
- We assume that we know the law of $S_t$ at time $T_i$ under the risk neutral measure (equiv. to assuming that we observe calls and puts of all strikes)
- Choice of the model: local volatility model

$$\sigma_t = \sigma(S_t, t)$$

- Simplest model that allows to match any admissible surface of call and put prices

# Dupire's formula

- Knowing all prices of "vanilla options" (all strikes and maturities) one can find directly (Dupire's formula) the local volatility $\sigma(s,t)$ that would lead to those prices...
- industry practice is to
  - reduce to a finite dimensional description of vanilla prices surfaces
  - use it extrapolate a sparse set of option prices to construct a price *surface*...
  - and then apply this formula...
- not entirely satisfactory as it can lead to unstable local volatilities (either $\sigma = 0$ or $+\infty$)...
- still it is used extensively by option traders.

# An alternative variational approach

Look for $\sigma(s,t)$ that realises

$$\inf \mathbb{E} \int_{T_i}^{T_{i+1}} \mathcal{L}(\sigma(S_t^\sigma, t))dt$$

where the infimum is taken over surfaces $\sigma(s,t)$ such that the solution to

$$\frac{dS_t^\sigma}{S_t^\sigma} = \sigma(S_t^\sigma, t)dW_t$$

with $\mathrm{Law}(S_{T_i}^\sigma) = \mu_i$ satisfies $\mathrm{Law}(S_{T_{i+1}}^\sigma) = \mu_{i+1}$.

# Dual formulation

Let $\mathcal{L}(t, s, \sigma)$ above be equal to $F(t, s, v)$ for $v = \frac{1}{2}\sigma^2$, $F$ convex in $v$

## Dual problem

In this case, on $[0, T]$, the dual problem becomes

$$\sup_{\phi, \partial_t \phi + F^*(\partial_{xx}\phi) \leq 0} \int_{\mathbb{R}} d\rho_T \phi(T) - d\rho_0 \phi(0)$$

The infimum is therefore taken over $\phi$ supersolutions of $\partial_t \phi + F^*(\partial_{xx}\phi) = 0$
Formally on can assume equality

# Numerical solution after Benamou & Brenier

In the quadratic case define the Lagrangian:

$$L(\phi, \rho, m) = \int_D \int_0^1 \frac{|m|^2}{2\rho} + \phi(\partial_t \rho - \partial_{xx} m) \; dt dx,$$

where $m = \rho v$. $\rho_0, \rho_T$ given.

Saddle-point problem:

$$\inf_{\rho, m} \sup_{\phi} L(\phi, \rho, m)$$

If the optimal solution is $(\phi^*, \rho^*, m^*)$, we have:

$$\partial_\phi L|_{(\phi^*, \rho^*, m^*)} = 0 \Leftrightarrow \begin{cases} \partial_t \rho^* - \partial_{xx} m^* = 0 \\ \rho^*(0, \cdot) = \rho_0, \;\; \rho^*(1, \cdot) = \rho_1 \end{cases}$$

$$\partial_\rho L|_{(\phi^*, \rho^*, m^*)} = 0 \Leftrightarrow \partial_t \phi^* + \frac{|m^*|^2}{2{\rho^*}^2} = 0,$$

$$\partial_m L|_{(\phi^*, \rho^*, m^*)} = 0 \Leftrightarrow \frac{m^*}{\rho^*} = \partial_{xx} \phi^*.$$

Stopping Criteria: $\partial_t \phi + \frac{1}{2} |\partial_{xx} \phi|^2 = 0$ .

# Augmented Lagrangian

Use that

$$\frac{|m|^2}{2\rho} = \sup\{a\rho + b\cdot m, a + |b|^2/2 \leq 0\}$$

which makes $L$ convex w,r,t $\rho, m$.

Augmented Lagrangian:

$$L_r(\phi, q, \mu) = F(q) + G(\phi) + < \mu, \nabla_{t,xx}\phi - q >$$
$$+ \frac{r}{2} < \nabla_{t,xx}\phi - q, \nabla_{t,xx}\phi - q >$$

where

- $\mu = \{\rho, m\}$, $q = \{a, b\}$, $r > 0$ is the penalty parameter,
- $F(q) = 0$, if $a + |b|^2/2 \leq 0$, $+\infty$ otherwise,
- $G(\phi) = \int_D \rho_0\phi_0 - \rho_T\phi_T$,
- $\nabla_{t,xx}\phi = \{\partial_t\phi, \partial_{xx}\phi\}$

# Numerical solution

## ADMM Algorithm

ADMM algorithm (alternating direction method of multipliers):

- Given $(\phi^{n-1}, q^{n-1}, \mu^n)$, errTol
- Step A: Find $\phi^n$:

$$L_r(\phi^n, q^{n-1}, \mu^n) \leq L_r(\phi, q^{n-1}, \mu^n), \quad \forall \phi.$$

- Step B: Find $q^n$:

$$L_r(\phi^n, q^n, \mu^n) \leq L_r(\phi^n, q, \mu^n), \quad \forall q.$$

- Step C: Pointwise update $\mu^n$:

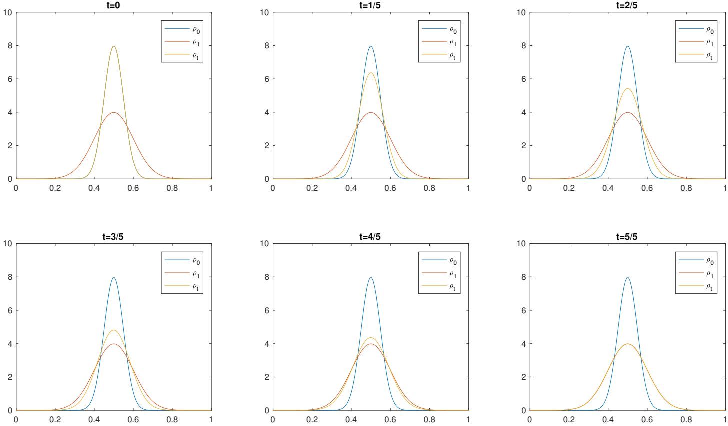
$$\mu^{n+1} = \mu^n + r(\nabla_{t,xx}\phi^n - q^n).$$

- If $res^n > errTol$, go back to step A, where

$$res^n = \max \text{abs}\big\{\partial_t \phi + \frac{1}{2}|\partial_{xx}\phi|^2\big\}$$

# Numerical result

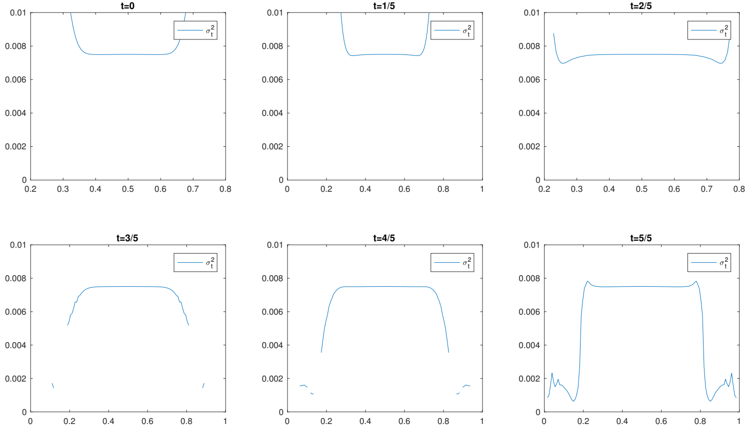
## Example




$$\rho_0 = N(0.5, 0.05),\ \rho_1 = N(0.5, 0.10)$$

# Numerical result

## Example




$$\rho_0 = N(0.5, 0.05),\ \rho_1 = N(0.5, 0.10)$$

# Numerical solution: other approach

## Semimartingale Transport Problem Tan, Touzi (2013)

Go back to dual problem

$$\sup_{\phi, \partial_t \phi + F^*(\partial_{xx}\phi) \le 0} \int_{\mathbb{R}} d\rho_T \phi(T) - d\rho_0 \phi(0)$$

which is solved by a descent algorithm

To compute the functional for a given $\phi(T)$ solve the HJB equation $\partial_t \phi + F^*(\partial_{xx}\phi) = 0$

# “Robust” hedging

Robust Hedging (see Soner, Touzi, Zhang (2011), L. (2017)): The solution to

$$\sup_{\sigma} \mathbb{E}\Big(\Phi(S^{\sigma}) - \int_0^T F(\sigma_t^2)dt\Big).$$

is given by solving $\partial_t \phi + F^*(\partial_{xx}\phi) = 0$, and taking $\sigma^2 = \nabla_{\sigma^2} F^*(\partial_{xx}\phi)$

Example :

$$F = \begin{cases} \frac{1}{\lambda}(a^{1/2} - \hat{\sigma})^2 \text{ if } a \geq 0, \\ +\infty \text{ otherwise.} \end{cases}$$

$$\text{Volmin-Volmax } F = \begin{cases} 0 \text{ if } a \in [\underline{a}, \bar{a}], \\ +\infty \text{ otherwise.} \end{cases}$$

# "Robust" Volatility Calibration

Extend the Benamou Brenier approach when we only know prices of a discrete set of options

Let $\hat{\sigma}$ be a given "target" volatility profile.

$$\inf_{\sigma} \mathbb{E}\left( \int_0^T F_{t,S_t^\sigma,\hat{\sigma}_t^2}(\sigma_t^2)dt \right), \quad \text{subject to:} \quad \Phi_i(S^\sigma) = c_i, \forall i.$$

Saddle Point Problem:

$$\sup_{\sigma} \inf_{\lambda} \mathbb{E}\left( \sum_i \lambda_i(\Phi_i(S^\sigma) - c_i) - \int_0^T F_{t,S_t^\sigma,\hat{\sigma}_t^2}(\sigma_t^2)dt \right),$$

# The algorithm

Saddle Point Problem (cf. Avellaneda et al. (1997)):

$$\sup_{\sigma} \inf_{\lambda} \mathbb{E}\bigg( \sum_i \lambda_i(\Phi_i(S^\sigma) - c_i) - \int_0^T F_{t,S_t^\sigma,\hat{\sigma}_t^2}(\sigma_t^2)dt \bigg),$$

By duality equal to

$$\inf_{\lambda} \sup_{\sigma} \mathbb{E}\bigg( \sum_i \lambda_i(\Phi_i(S^\sigma) - c_i) - \int_0^T F_{t,S_t^\sigma,\hat{\sigma}_t^2}(\sigma_t^2)dt \bigg),$$

Solve the $\sup_\sigma(\ldots)$ part by solving HJB (same problem as in robust hedging)

Gradient descent with respect to $\lambda$

Gradient of $\sup_\sigma(\ldots)$ with respect to $\lambda$ to given by

$$\partial_{\lambda_i} \sup_{\sigma}(\ldots) = \mathbb{E}(\Phi_i(S^\sigma)) - c_i$$

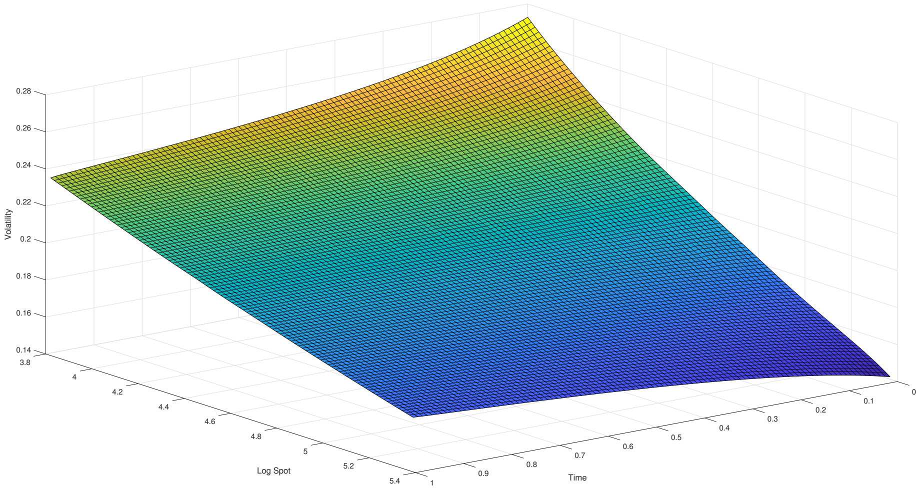
# Simulations

Let $F(t,s,\hat{\sigma},x) = (a(x-\hat{\sigma}(t,s)))^{(1-p)} + (b(x-\hat{\sigma}(t,s)))^{p+1}$

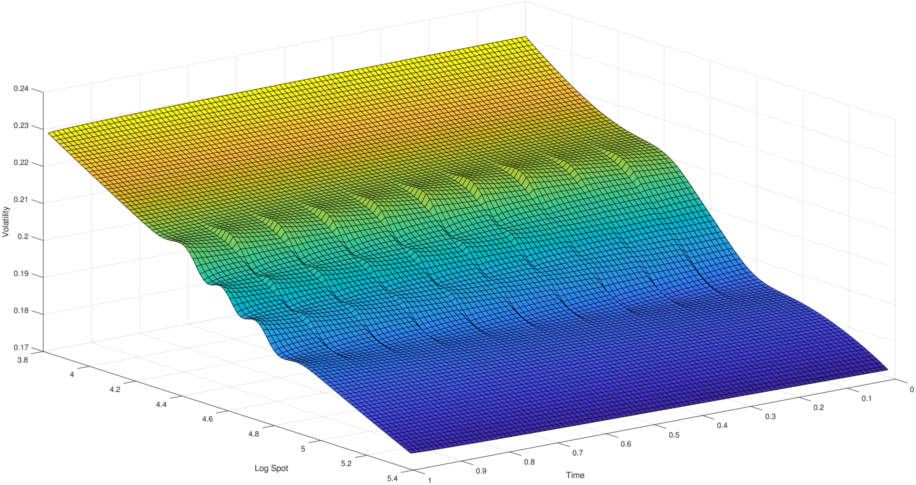
We generate put option prices for 10 maturities $\times$ 11 strikes, then try to calibrate the local volatility with different volatility profiles $\hat{\sigma}$.

Prototype run in Matlab.

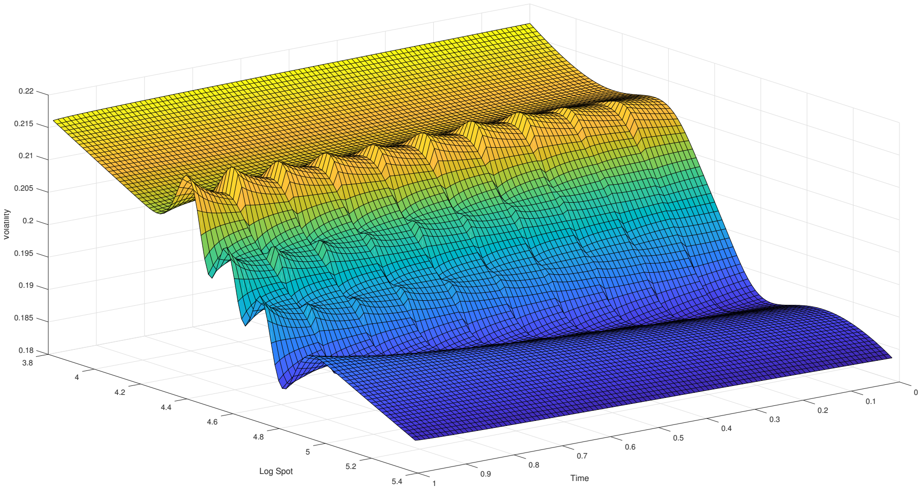
# Using Input Local Vol as Profile



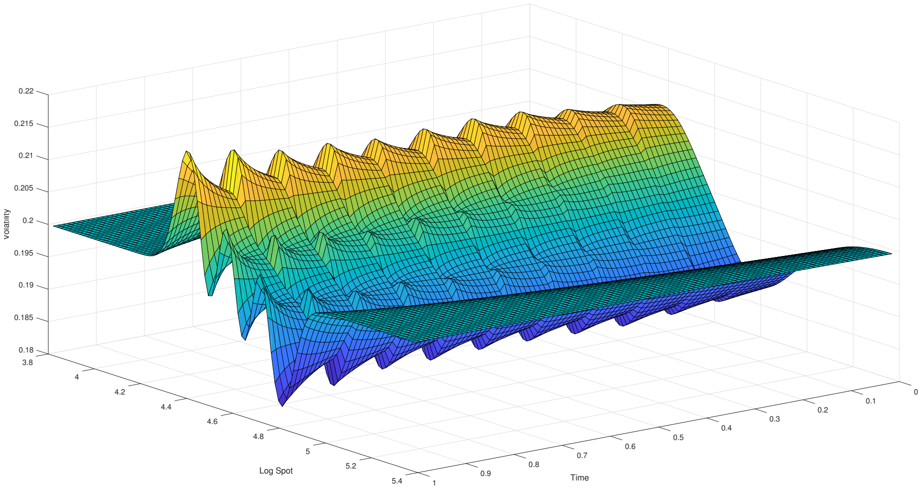
# Constant Skew



# Smaller Constant Skew



# Constant Profile



# Local Stochastic Volatility Calibration

Given

$$dS_t = \sigma_t S_t dW_t, \quad d[W,B]_t = \rho_t dt.$$

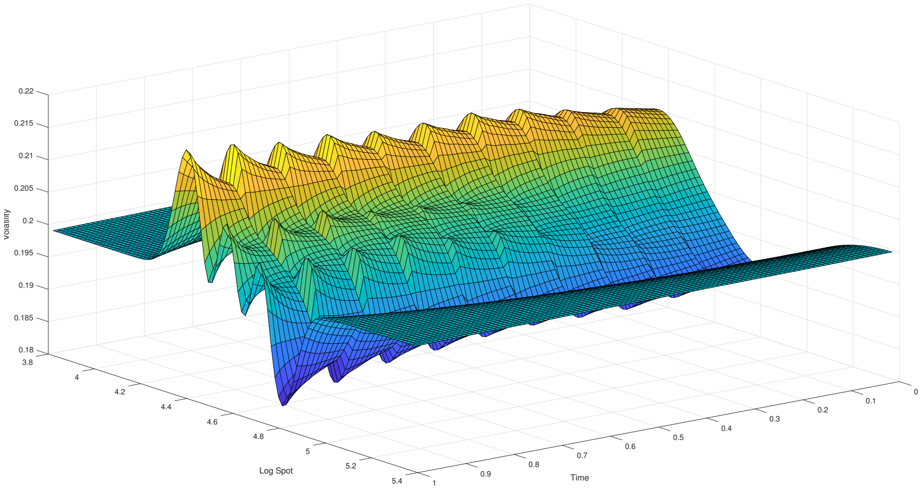
and $V_t$ driven by $B_t$ (ex. Heston). Consider the problem

$$\sup_{\sigma,\rho} \mathbb{E}\left( \Phi(S_T^{\sigma,\rho}) - \int_0^T F_{t,S_t^\sigma,V_t}(\sigma_t^2, \sigma_t\rho_t)dt \right).$$

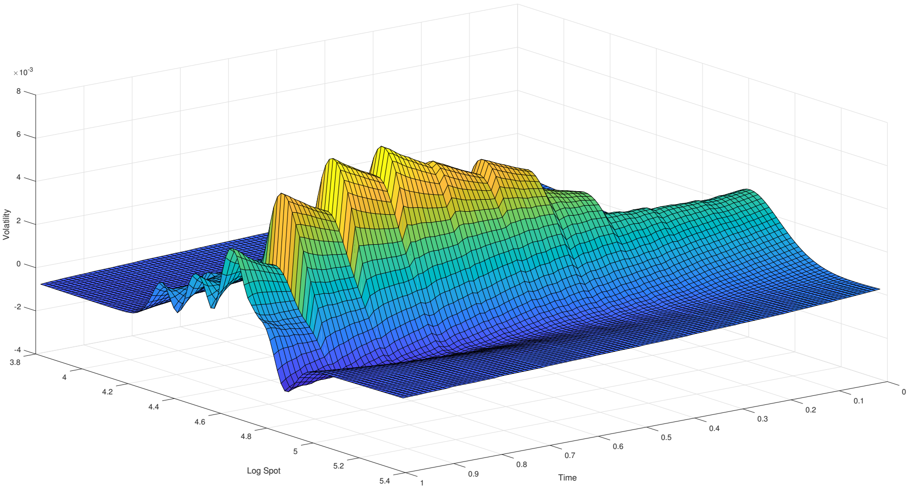
This can be extended to the *stochastic volatility calibration problem*:

$$\sup_{\sigma,\rho} \inf_{\lambda} \mathbb{E}\left( \sum_i \lambda_i(\Phi_i(S^\sigma) - c_i) - \int_0^T F_{t,S_t^\sigma,V_t}(\sigma_t^2, \sigma_t\rho_t)dt \right).$$

# Average Volatility Surface



# Difference with Local Volatility Results



# Some references


- Tan and Touzi (2013):
  *Optimal Transportation under controlled stochastic dynamics*
- Benamou and Brenier (2000):
  *A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem*
- L. (2005):
  *The reconstruction problem for the Euler-Poisson system in cosmology*
- Brenier, Frisch, Hénon, Loeper, Matarrese, Mohayaee, Sobolevski (2003):
  *Reconstruction of the early Universe as a convex optimization problem*
- L. (2017): Option pricing with linear market impact and nonlinear Black Scholes equations
- Guo, Loeper, Wang (2017):
  *Local Volatility Calibration by Optimal Transport*
- Backhoff, Beiglbock, Hesmann, Allbad (2017):
  Martingale Benamou Brenier: A probabilistic perspective
- Huesmann, Trevisan (2017):
  A Benamou Brenier formulation of martingale optimal transport
- Avellaneda et al., (1997):
  Calibration of the volatility surface via minimum relative entropy

# Mixed PDE-MC for Options Pricing

Given a model with stochastic rates and stochastic volatility

$$dS_t = r(t, S_t, V_t)S_t dt + \sigma(t, S_t, V_t)S_t dW_t,$$

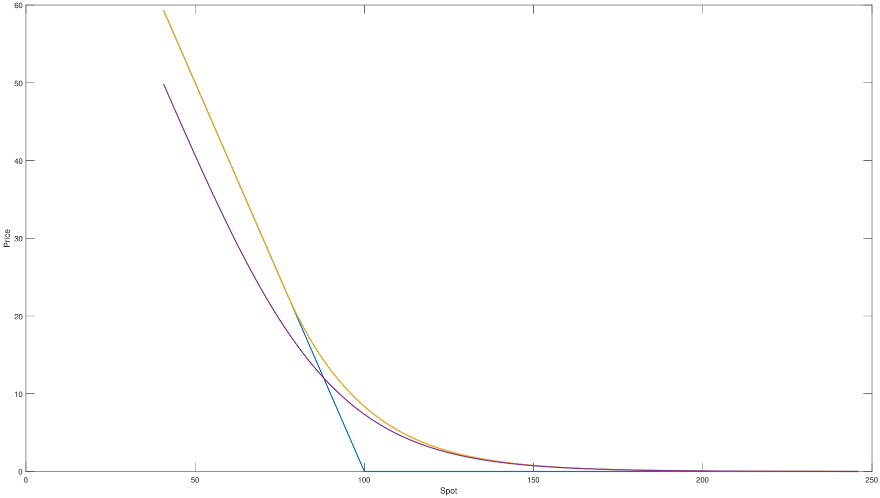
where $V_t$ is driven by $B_t$, we can compute the option prices as follows:

$$\mathbb{E}(\Phi(S_T) \mid S_t, V_t) = \underbrace{\mathbb{E}_t\Bigg( \underbrace{\mathbb{E}\big(\Phi(S_T) \,\big|\, S_t, V_{[t,T]}\big)}_{\text{Finite Difference to solve SPDE}} \Bigg)}_{\text{Monte Carlo simulations}}.$$

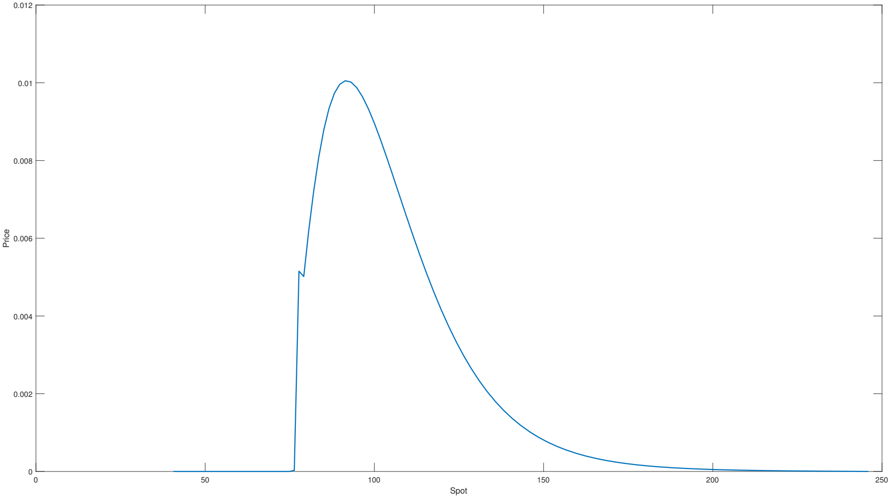
Combined with Least Square Monte Carlo and Stochastic Duality (upper bound) for American options:

$$\min_{\beta,\gamma} \int_x \Bigg(U(t,x,v) - \sum_j \beta_j f_j(x,v) - \sum_j \gamma_j g_j(x,v)' \Delta B_t\Bigg)^2 w(x)dx.$$

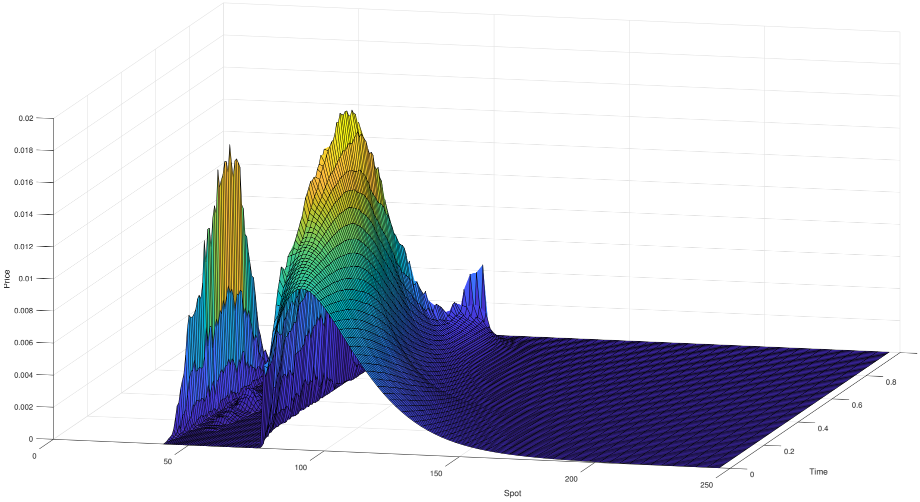
# American Put Price



# Difference Between Upper and Lower Bounds



# Difference Between Upper and Lower Bounds



# Mixed PDE-MC for LSV Adjustment

For the LSV model

$$dS_t = r(t, S_t)S_t dt + \psi(t, S_t)\sqrt{V_t}S_t dW_t,$$

where $V_t$ is driven by $B_t$, let $p_B(t, s)$ be the density of $S_t$ given $B$:

$$\mathbb{E}(\Phi(S_t) \,|\, B) = \int_x \Phi(x)p_B(t, x)dx, \quad \mathbb{E}(p_B(t, x)) = p(t, x).$$

Then derive the conditional forward SPDE:

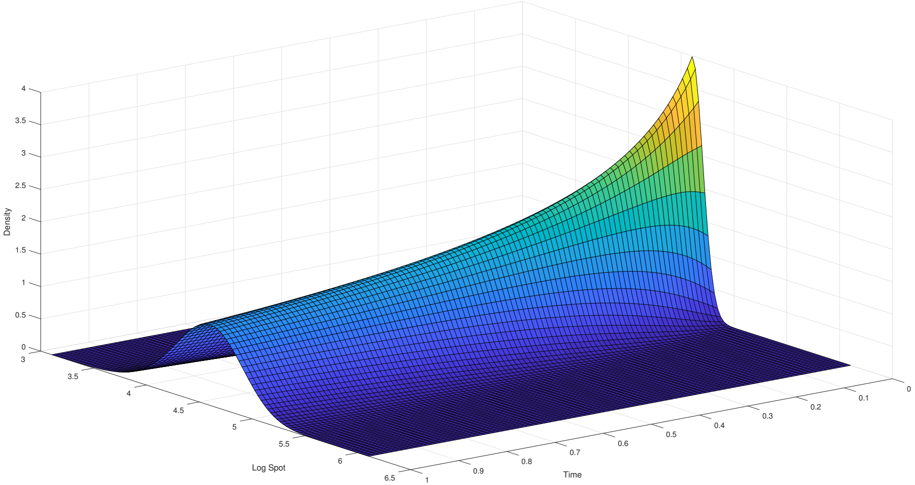
$$dp_B = -\big(\partial_x(rp_B)dt + \partial_x(\psi p_B)\rho\sqrt{V_t}dB_t\big) + \frac{1}{2}\partial_{xx}(\psi^2 p_B)V_t dt,$$

with mimicking formula

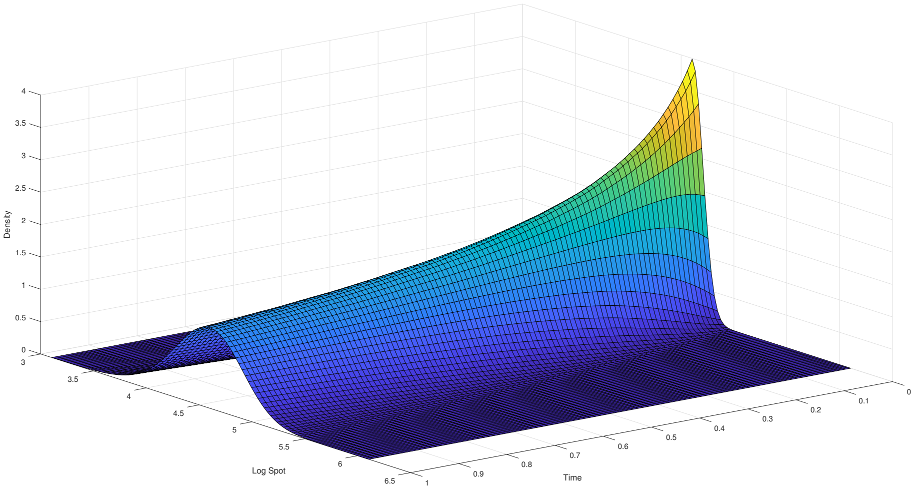
$$\psi^2(t, x)\mathbb{E}(p_B(t, x)V_t) = \sigma_{LV}^2(t, x)\rho(t, x).$$

Given a local volatility function $\sigma_{LV}(t, x)$ or a density function $p(t, x)$, we can recursively find $\psi$.

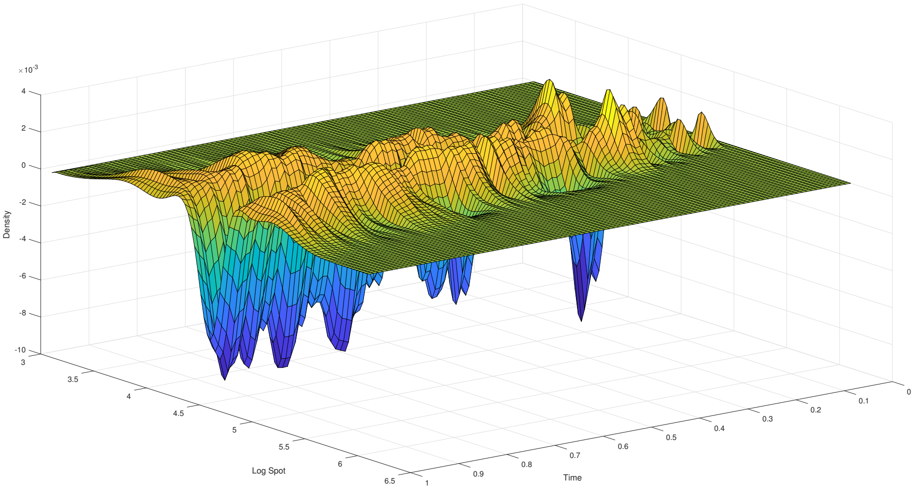
# Target Density



# Fitted Density



# Difference


×10^-3
4
2
0
-2
-4
-6
-8
-10
Density
3
3.5
4
4.5
5
5.5
6
6.5
Log Spot
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Time

# Mixed PDE-MC for Stochastic Control

Given

$$dS_t = \sigma_t S_t dW_t, \quad d[W, B]_t = \rho_t dt.$$

and $V_t$ driven by $B_t$. Consider the problem

$$\sup_{\sigma,\rho} \mathbb{E}\bigg( \Phi(S_T^{\sigma,\rho}) - \int_0^T F_{t, S_t^\sigma, V_t}(\sigma_t^2, \sigma_t \rho_t) dt \bigg).$$

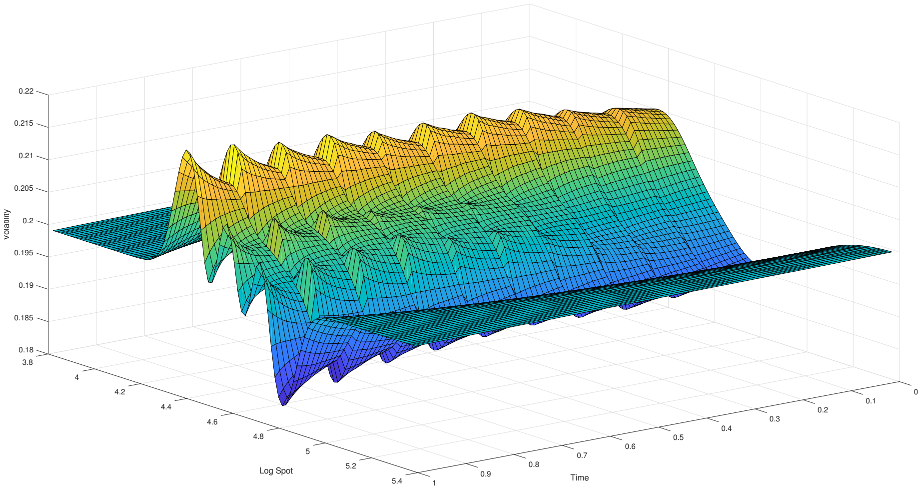
The solution to the HJB equation can be written as

$$du_t + F^* \bigg( \frac{1}{2} x^2 \partial_{xx} \mathbb{E}_t(u_{t+dt}(x, V_{t+dt}))dt, x\partial_x \mathbb{E}_t(u_{t+dt}(x, V_{t+dt}) dB_t) \bigg) = 0,$$

which is solved by the mixed method and least squares Monte Carlo.
This can be extended to the *stochastic volatility calibration problem*:

$$\sup_{\sigma,\rho} \inf_{\lambda} \mathbb{E}\bigg( \sum_i \lambda_i (\Phi_i(S^\sigma) - c_i) - \int_0^T F_{t, S_t^\sigma, V_t}(\sigma_t^2, \sigma_t \rho_t) dt \bigg).$$

# Average Volatility Surface



# Difference with Local Volatility Results