Calibration Using Consistent Bayesian Estimators

29th June 2009, University of Calgary

Alok Gupta
D.Phil Mathematical Finance
`alok.gupta@maths.ox.ac.uk`

MCFG, Mathematical Institute, Oxford University
Nomura Bank
EPSRC

Supervised By Christoph Reisinger

Contents

- Introduction
  - Motivation
  - Calibration Problem
  - Bayesian Estimators
- Consistency of Bayian Estimators
  - Single Price Observation with Gaussian Noise
  - Multiple Price Observations with Gaussian Noise
  - Observation(s) with Non-Gaussian Noise
  - Non-Scalar Parameter
- Bayesian Financial Modelling
  - Prior Distribution (Regularisation)
  - Likelihood Function (Calibration Error)
  - Posterior Distribution
- Numerical Examples
- Conclusion

Introduction

Motivation

- Since Black Scholes model proposed in 1973, huge growth in variety of financial models to capture behaviour of different markets e.g. stochastic interest rate models, credit models, etc.
- Agent will typically want to use model to price or hedge an instrument but before she can do this she must *calibrate* model to observable prices to avoid introducing arbitrage.
- Calibration not straight forward: instead of Black Scholes single parameter, now calibrate vectors and functions e.g. Levy density, local volatility.
- Not even clear if perfect calibration is possible. In reality can usually at best only reproduce observable prices to within their *bid-ask* spreads — introducing problem of uniqueness.
- Wealth of literature on local volatility calibration e.g. Lagnado & Osher (1997), Jackson, Suli & Howison (1999) find *best-fit* surfaces.

Motivation

- However, these papers do not attack the problem of uniqueness or robustness. Cont's papers (2005), (2006) do make a start.
- The calibration problem should be viewed as an *inverse problem*: given a model we can explicitly and directly find the prices of instruments, but given the prices we cannot directly recover the model.
- The calibration problem is often *ill-posed* in the sense that it fails ii) and iii) of Hadamard's criteria for well posedness:
  - i) For all admissible data, a solution exists.
  - ii) For all admissible data, the solution is unique.
  - iii) The solution depends continuously on the data.
- For this reason this research uses a Bayesian approach to focus on finding a *distribution* of solutions rather than a best-fit one.
- This allows better decision-making and a fuller understanding of the uncertainty of the calibration procedure.

Calibration Problem

Suppose we observe a price process $S = (S_t)_{t \geq 0}$ and model it as a function of time t, some stochastic process(es) $X = (X_t)_{t \geq 0}$, and finite dimensional parameter $\theta \in \Theta$, i.e.

$$S_t = S(t, X_t, \theta) \tag{1}$$

by abuse of notation of S. Let $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ be the filtration generated by X so S is an $\mathcal{F}$-adapted process.

Now consider an option over a finite time horizon $[0, T]$ written on S and with payoff function h. Let the time t value of this option be written as $f_t(\theta)$, where we include the argument θ to emphasise the dependence of this price on the model parameters. Explicitly,

$$f_t(\theta) = \mathbb{E}[B(t, T)h(S)|\mathcal{F}_t]$$

with respect to some measure $\mathbb{P}$ and where $B(t, T)$ is the discount factor, possibly stochastic.

Calibration Problem

Suppose at time $t \in [0, T]$ we observe a set of such option prices

$$\{f_t^{(i)}(\theta) : i \in I_t\}$$

possibly with noise $\{e_t^{(i)} : i \in I_t\}$. In other words, we observe

$$Y_t^{(i)} = f_t^{(i)}(\theta) + e_t^{(i)} \tag{2}$$

for $i \in I_t$.

Then the calibration problem is to find the value of θ that *best* reproduces the observed prices $\{Y_t^{(i)} : i \in I_t,\, t \in \Upsilon_n([0, T])\}$, for some measurement of *best*. Here

$$\Upsilon_n([0, T]) = \{t_1, \ldots, t_n : 0 = t_1 < t_2 < \ldots < t_n \leq T\}$$

is a partition of the interval $[0, T]$ into n parts.

Bayesian Estimators

Suppose we wish to estimate the value of some parameter θ. Assume we have some prior information for θ (for example that it belongs to a particular space, or is positive, or represents a smooth function), summarised by a *prior* density $p(\theta)$ for θ. And suppose we observe some noisy data $Y = \{Y_t : t \in \Upsilon_n\}$ related to θ by

$$Y_t = f_t(\theta) + e_t$$

for all $t \in \Upsilon_n$ where e_t is some random noise and Υ_n is an index set of size n. Then $p(Y|\theta)$ is the probability of observing the data Y given θ and is called the *likelihood* function.

Application of Bayes rule gives that the *posterior* density of θ is given by

$$p(\theta \mid Y) \propto p(Y|\theta)\, p(\theta).$$

Bayesian Estimators

The *loss function* $L(\theta, \theta')$ gives the deficit incurred by taking θ' as the estimator for θ. It must satisfy

$$\begin{cases} L(\theta, \theta') = 0 & \text{if } \theta' = \theta \\ L(\theta, \theta') > 0 & \text{if } \theta' \neq \theta \end{cases}$$

Given data Y, the corresponding *Bayes estimator* $\theta_L(Y)$ is the value of θ which minimises the expected loss with respect to the posterior i.e.

$$\theta_L(Y) = \arg\min_{\theta'} \left\{ \int L(\theta, \theta')\, p(\theta|Y)\, \mathrm{d}\theta \right\}.$$

Assumption. Since the loss function should penalise estimators which are further from the true value more than those which are closer, it is natural to assume that the larger the error, the greater the loss. So we assume L is a (not necessarily strictly) increasing function of $|\theta - \theta'|$.

Bayesian Estimators

Example. Different choices of loss function give well known Bayesian estimators:

1. $L_1(\theta, \theta') = |\theta - \theta'|^2$ gives Bayes estimator

$$\theta_{L_1}(Y) = \mathbb{E}[\theta|Y]$$

   which is the *mean* value of θ with respect to the Bayesian posterior density $p(\theta|Y)$.

2. $L_2(\theta, \theta') = 1 - 1_{\theta=\theta'}$, known as the uniform loss distribution, gives the Bayes estimator

$$\theta_{L_2}(Y) = \max\{p(\theta|Y)\}$$

   which is the *maximum a posteriori* (MAP) estimator; it is the value which maximises the posterior density.

Remark. The minimiser $\theta_L(Y)$ is not necessarily unique.

Consistency of Bayesian Estimators

Consistency

In order to judge the validy of an estimator, we would like it to satisfy certain properties. For example, we might expect the estimator θ_L to be unbiased. Or instead we might hope that, as we observe more data Y_t, that our estimator converges in some sense to the true value.

Definition. A sequence of estimators $\hat{\theta}_n = \hat{\theta}(\{Y_t : t \in \Upsilon_n, |\Upsilon_n| = n\})$ of unknown θ is said to be *consistent* if $\hat{\theta}_n$ tends in probability to θ (written $\hat{\theta}_n \xrightarrow{P} \theta$) for all $\theta \in \Omega$, i.e.

$$\forall\, \theta \in \Omega \quad \forall\, \delta > 0 \quad \mathbb{P}_\theta[|\hat{\theta}_n - \theta| \geq \delta] \to 0 \quad \text{as} \quad n \to \infty.$$

For what follows we work in the risk-neutral measure $\mathbb{Q}$. The unknown parameter σ is assumed to be constant in time (and scalar for the first 3 sections).

Single Observation

Let $I_t = I = \{1\}$ so we consider observing only one option price each time t. Denote by σ^* the true value of the parameter. Assume that the market noises are given by

$$e_t^{(1)} \sim N(0, \varepsilon_t^2) \quad \varepsilon_t \in [c, C] \subseteq \Re^+ \setminus \{0\}$$

and are independent of each other. Drop the superscripts. Take a nested sequences of partitions $\Upsilon_n \supset \Upsilon_{n-1}$ and define

$$\mathcal{G}_{t_n} = \sigma\left(\{e_s : s \in \Upsilon_n\}\right),$$

i.e. the sigma-field generated by all the noise random variables upto time t_n. So in particular $\mathcal{G}_{t_n} \supset \mathcal{G}_{t_m}$ for all $n \geq m$.

Assumption. $\mathcal{F}_{t_n} \perp \mathcal{G}_{t_m}$ for all (n, m), i.e. the driving process of the underlying is independent of the market noise. This is a reasonable assumption to make since observational errors are unlikely to be correlated to fundamental economic phenomena.

Single Observation

Let $Y = \{Y_t : t \in \Upsilon_\infty\}$ be the (noisy) observations and write $\sigma_n(Y) := \sigma|\mathcal{F}_{t_n} \vee \mathcal{G}_{t_n}$ for shorthand. Define the sequence of Bayes estimators $\hat{\sigma}$ by,

$$\begin{aligned} g_{\sigma'}(\sigma_n(Y)) &= \mathbb{E}[L(\sigma_n(Y), \sigma')] \\ &= \int_\Sigma L(\sigma, \sigma')\, p(\sigma_n(Y))\, \mathrm{d}\sigma \\ \hat{\sigma}_n(Y) &= \arg \min_{\sigma' \in \Sigma} \{g_{\sigma'}(\sigma_n(Y))\} \end{aligned}$$

where $L(\sigma, \sigma')$ is the loss function, and Σ is the support for σ and

$$\begin{aligned} p(\sigma_n(Y)) = p_n(\sigma|Y) &= \frac{p_n(Y|\sigma)\, p(\sigma)}{p_n(Y)} \\ &= \prod_{t \in \Upsilon_n} \tfrac{1}{\sqrt{2\pi}\varepsilon_t} \exp\left\{-\tfrac{1}{2\varepsilon_t^2}(Y_t - f_t(\sigma))^2\right\} \tfrac{p(\sigma)}{p_n(Y)} \end{aligned}$$

is the posterior density after n observations and $p_n(Y)$ is constant wrt σ.

Single Observation

To prove consistency of $\hat{\sigma}_n(Y)$ we make the following assumptions.

Assumption. The prior $p(\sigma)$ with support Σ satisfies:

i) Σ is compact

ii) $p(\sigma)$ is bounded

iii) $\sigma^* \in \Sigma$ i.e. $p(\sigma^*) > 0$

Assumption. For each t, conditional on $\mathcal{F}_t$ the function $f_t(\sigma)$ satisfies:

i) f_t is differentiable everwhere in Σ

ii) for all $\sigma \in \Sigma$, $0 < k \leq f_t'(\sigma) \leq K < \infty$

Single Observation

Lemma. *For all* Y $\sigma_n(Y) \xrightarrow{P} \sigma^*$.

Proof. (Outline)

- Write $p_n(\sigma|Y) = q_n(Y)p(\sigma)e^{-\frac{1}{2}\phi_n(\sigma,Y)}$ then can show

$$\left|\frac{\phi_n(\sigma,Y) - 2u(\sigma-\sigma^*)}{\phi_n(\sigma,Y)} - 1\right| \leq \frac{1}{\alpha_n} \to 0 \qquad n \to \infty$$

- Define the moment generating function

$$\varphi_n(u) = \mathbb{E}[e^{u(\sigma_n - \sigma^*)}]$$

  then it follows that $\varphi_n(u) \to 1$ as $n \to \infty$ i.e. Dirac density $\delta(\sigma - \sigma*)$.

- By Levy's Continuity Theorem this implies that $\sigma_n(Y) \xrightarrow{D} \sigma^*$ where σ^* is a constant almost surely.

- Hence $\sigma_n(Y) \xrightarrow{P} \sigma^*$.

Single Observation

Theorem. *(Main Result) For all L bounded and continuous on Σ the Bayes estimator $\hat{\sigma}_n(Y)$ is consistent.*

Proof. (Outline)

- First observe that we can write $L(\sigma, \sigma') = l(\sigma - \sigma')$ for some function l.
- $\mathbb{P}_{\sigma^*}[|\hat{\sigma}_n(Y) - \sigma^*| \geq \delta]$

$$\leq \mathbb{P}_{\sigma^*}[|\hat{\sigma}_n(Y) - \sigma_n(Y)| \geq \tfrac{1}{2}\delta] + \mathbb{P}_{\sigma^*}[|\sigma_n(Y) - \sigma^*| \geq \tfrac{1}{2}\delta]$$

- But $\mathbb{P}_{\sigma^*}[|\sigma_n(Y) - \sigma^*| \geq \frac{1}{2}\delta] \to 0$ as $n \to \infty$ by above lemma.
- And can show $\mathbb{P}_{\sigma^*}[|\hat{\sigma}_n(Y) - \sigma_n(Y)| \geq \frac{1}{2}\delta] \to 0$ as $n \to \infty$ for L bounded and continuous.
- Hence, for all $\delta > 0$, $\mathbb{P}_{\sigma^*}[|\hat{\sigma}_n(Y) - \sigma^*| \geq \delta] \to 0$ as $n \to \infty$.

$\square$

Single Observation

Example. Working in the Black Scholes model, if we let

$$f_t(\sigma) = e^{-r(T-t)}\mathbb{E}[(S_T(\sigma) - K)^+|\mathcal{F}_t]$$

then we know that f_t is differentiable everywhere on $[\sigma_{\min}, \sigma_{\max}]$ $(0 < \sigma_{\min}$ and $\sigma_{\max} < \infty)$ and the derivative (*vega*) is given by

$$f_t'(\sigma) = S_t\sqrt{T-t}N'(\tfrac{\log(S_t/K)+(r+\sigma^2/2)(T-t)}{\sigma\sqrt{T-t}})$$

for non-dividend paying stock S where $N'(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$.

Note that for all combinations of S_t and K the vega is positive and bounded on the interval $[\sigma_{\min}, \sigma_{\max}]$ so long as $\sigma_{\min}$ is sufficiently large and $\sigma_{\max}$ sufficiently small.

Hence, Black-Scholes European call prices satisfy the assumptions (for suitably restrictive Σ) and can be used as the observation option prices to get a consistent estimator for the Black-Scholes volatility.

Multiple Observations

Suppose $|I_t| > 1$ so we observe several prices $f_t^{(i)}(\sigma)$ for $i \in I_t$ at each time t. Again assume market noises are independent and Gaussian.

Corollary. *For multiple price observation sets I_t and all bounded and continuous loss functions L the Bayes estimator $\hat{\sigma}_n(Y)$ is consistent.*

Proof. Identical to the one for the previous theorem except that the sigma-field generated by all the noise random variables upto time t_n is now

$$\mathcal{G}_{t_n} = \sigma\left(\{e_s^{(i)} : s \in \Upsilon_n, i \in I_t\}\right),$$

and values Y_t, f_t, ε_t get superscripts to become $Y_t^{(i)}, f_t^{(i)}, \varepsilon_t^{(i)}$ respectively. □

Remark. The only real difference in the convergence of the estimator is that for multiple observations the convergence is faster since more information is added at each timestep to update the posterior $p_n(\sigma|Y)$.

Non-Gaussian Noise

Recall that

$$Y_t = f_t(\sigma^*) + e_t$$

where e_t is the additive market noise. Now, suppose that the e_t is given by the density $\xi_t(x)$. Then we can write the conditional density of $Y_t|\sigma$ as

$$p(Y_t|\sigma) = \xi_t(Y_t - f_t(\sigma)).$$

Assumption. For all observation times t, the density functions ξ_t satisfy:

i) ξ_t is bounded and unimodal with mode at 0

ii) $\xi_t(Y_t - f_t(\sigma^*)) > 0$

iii) $\lim_{n\to\infty} \frac{1}{n} \sum_{t\in\Upsilon_n} e_t = 0$

Theorem. *For all L bounded and continuous $\hat{\sigma}_n(Y)$ is consistent.*

Proof. Similar to first proof: we show the mgf is Dirac in the limit and then use Levy Continuity Theorem. □

Non-Scalar Parameter

We generalise to the case when σ is non-scalar but a vector of parameters.

Assumption. For each t, conditional on $\mathcal{F}_t$ the function $f_t(\sigma)$ (which is a mapping $f_t : \Sigma \to \Re$ where $\Sigma \subseteq \Re^m$) satisfies:

i) Σ is convex

ii) $\nabla f_t(\sigma) = \left(\frac{\partial f_t}{\partial \sigma_1}(\sigma), \ldots, \frac{\partial f_t}{\partial \sigma_m}(\sigma) \right)^T$ exists everwhere in Σ

iii) Define $\Upsilon_n(\sigma, \sigma'; k) = \{t \in \Upsilon_n : |f_t(\sigma) - f_t(\sigma')| > k\|\sigma - \sigma'\|\}$. Then for any pair (σ, σ') in $\Sigma \times \Sigma$, there exists a $k > 0$ such that $|\Upsilon_n(\sigma, \sigma'; k)| \to \infty$ as $n \to \infty$.

Theorem. *For all L bounded and continuous on Σ the non-scalar Bayes estimator $\hat{\sigma}_n(Y)$ is consistent.*

Again the above result is easily generalised to the case of multiple observations and non-Gaussian noise.

Bayesian Financial Modelling

Local Volatility Model

Corresponding to the model originally proposed by Black & Scholes, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, (Z_t)_{0 \leq t \leq T})$ be the standard Wiener space i.e. Z_t is Brownian motion, $\mathcal{F}_t$ is the natural filtration of Z_t over Ω and $\mathcal{F} = \mathcal{F}_T$. Then the underlying asset price S is given by

$$dS_t = \mu S_t dt + \sigma S_t dZ_t$$

where μ is the drift and σ the volatility. In the Local Volatility model we choose σ to be a function of both the asset price and the time:

$$\sigma = \sigma(S, t).$$

Although Dupire found an explicit formula to calculate this function using the implied volatility surface, the resulting local volatility surface is unstable and spikey. Furthermore, the formula depends on knowledge of the prices of options for all strikes and maturities, which is usually not available in practice.

Local Volatility Model

Instead, we identify key characteristics expected of the local volatility surface that can be recast into a Bayesian prior. There are three properties we would expect of $\sigma(S,t)$:

Positivity: $\sigma(S,t) > 0$ for all values of S and t; since the price variation squared $\sigma^2 > 0$ we adopt the convention $\sigma > 0$.

Smoothness: there should be no sharp spikes or troughs in the surface; this ensures pricing and hedging is stable.

Consistency: for small values of t especially, σ should be close to today's at-the-money (ATM) volatility σ_{atm}.

The Prior (Regularisation)

For the purposes of introducing the theory we consider the simplest density - the Gaussian density. It is also the second order approximation to any density. In light of the assumptions presented earlier we take for our prior

$$p_{lv}(\sigma) \propto \exp\left\{-\frac{1}{2}\lambda_p \|\log(\sigma) - \log(\sigma_{atm})\|_\kappa^2\right\}$$

where $\|\cdot\|_\kappa$ is a Sobolev norm given by

$$\|u(x,y)\|_\kappa^2 = (1-\kappa)\|u(x,y)\|_2^2 + \kappa\||\nabla u(x,y)|\|_2^2.$$

Working in the logarithmic space guarantees σ is positive and the norm ensures greater prior density is attached to σ that are both smoother and closer to ATM volatility.

λ_p quantifies how strong our prior assumptions are: a higher value of λ_p indicating greater confidence in our assumptions.

Clearly, those θ which better satisfy prior beliefs have greater prior density.

The Likelihood (Calibration)

Let $V_t^{(i)*}$ be the market observed price at time t of a European call and $V_t^{(i)}(\theta)$ the corresponding theoretical price. Then define the basis point square-error functional as

$$G_t(\theta) = \frac{10^8}{S_t^2} \sum_{i \in I} w_i \left| V_t^{(i)}(\theta) - V_t^{(i)*} \right|^2$$

where the w_i are weights summing to one. But only attach positive Bayesian posterior density if paramater reproduces prices to within their bid ask spreads i.e.

$$G(\theta) \leq \delta^2$$

where $\delta^2 = \sum_{i \in I} w_i \delta_i^2$ is the pre-specified tolerance. Hence, for the Bayesian likelihood for non-parametric models we will take

$$p(V^*|\theta) = 1_{G(\theta) \leq \delta^2} \exp\left\{-\tfrac{1}{2} G(\theta)\right\}.$$

So those surfaces σ which reproduce prices closest to the market observed prices V^* have the greatest likelihood values.

The Posterior

Combining the prior and likelihood functions we get the explicit form for the posterior function $p(\theta|V^*)$ as

$$p(\theta|V^*) \propto 1_{G(\theta)\leq\delta^2} \exp\left\{-\frac{1}{2}\left[\lambda_p\|\theta\|^2 + G(\theta)\right]\right\}.$$

Remark. Observe that maximising the posterior is equivalent to minimising the expression

$$\lambda_p\|\theta\|^2 + G(\theta)$$

which is exactly the form of functional authors such as Lagnado & Osher (1997) and Jackson, Suli & Howison (1999) seek to minimise to find their optimal calibration parameter. This is not a coincidence but an insight into how the Bayesian approach reformats traditional Tikhonov and regularisation methods into a unified and rigorous framework.

Numerical Examples

Discretisation

There is no exact method for calibrating a surface or function since they are infinite dimensional. Instead we discretise our functions.

We represent the local volatility surface $\sigma(S,t)$ by a grid of nodes whose positions are given by: $S_{min} = s_1 < \ldots < s_j < \ldots < s_J = S_{max}$ in the spatial direction and $0 = t_1 < \ldots < t_l < \ldots < t_L = t_{max}$ in the temporal direction.

For each time t_l we construct the unique $(J-1)$-dimensional natural cubic spline through the nodes $(s_1, t_l), \ldots, (s_J, t_l)$ to give all values $\sigma(S, t_l)$. Then for $(S,t) \in [s_j, s_{j+1}] \times [t_l, t_{l+1}]$ the value of $\sigma(S,t)$ is found by linear interpolation of the two values $\sigma(S, t_l)$ and $\sigma(S, t_{l+1})$. This is the same approach taken in Jackson, Suli & Howison (1999).

Proposition. *For the local volatility discretisation scheme described above and sufficiently diverse European call option pricing functions $f_t(\sigma)$, the Bayes estimator $\hat{\sigma}_n$ is consistent.*

Metropolis Sampling

We generate samples from $p(\theta|V^*)$ using the Markov Chain Monte-Carlo (MCMC) Metropolis algorithm:

1. Select a starting point θ_0 for which $p(\theta_0|V^*) > 0$.
2. For $t = 1, \ldots, n$ sample a proposal $\theta^\#$ from a symmetric jumping distribution $J(\theta^\#|\theta_{t-1})$ and set

$$\theta_t = \begin{cases} \theta^\# & \text{with probability } \min\{\frac{p(\theta^\#|V^*)}{p(\theta_{t-1}|V^*)}, 1\} \\ \theta_{t-1} & \text{otherwise} \end{cases}$$

Then the iterations $\theta_1, \ldots, \theta_n$ converge to the target distribution $p(\theta|V^*)$.

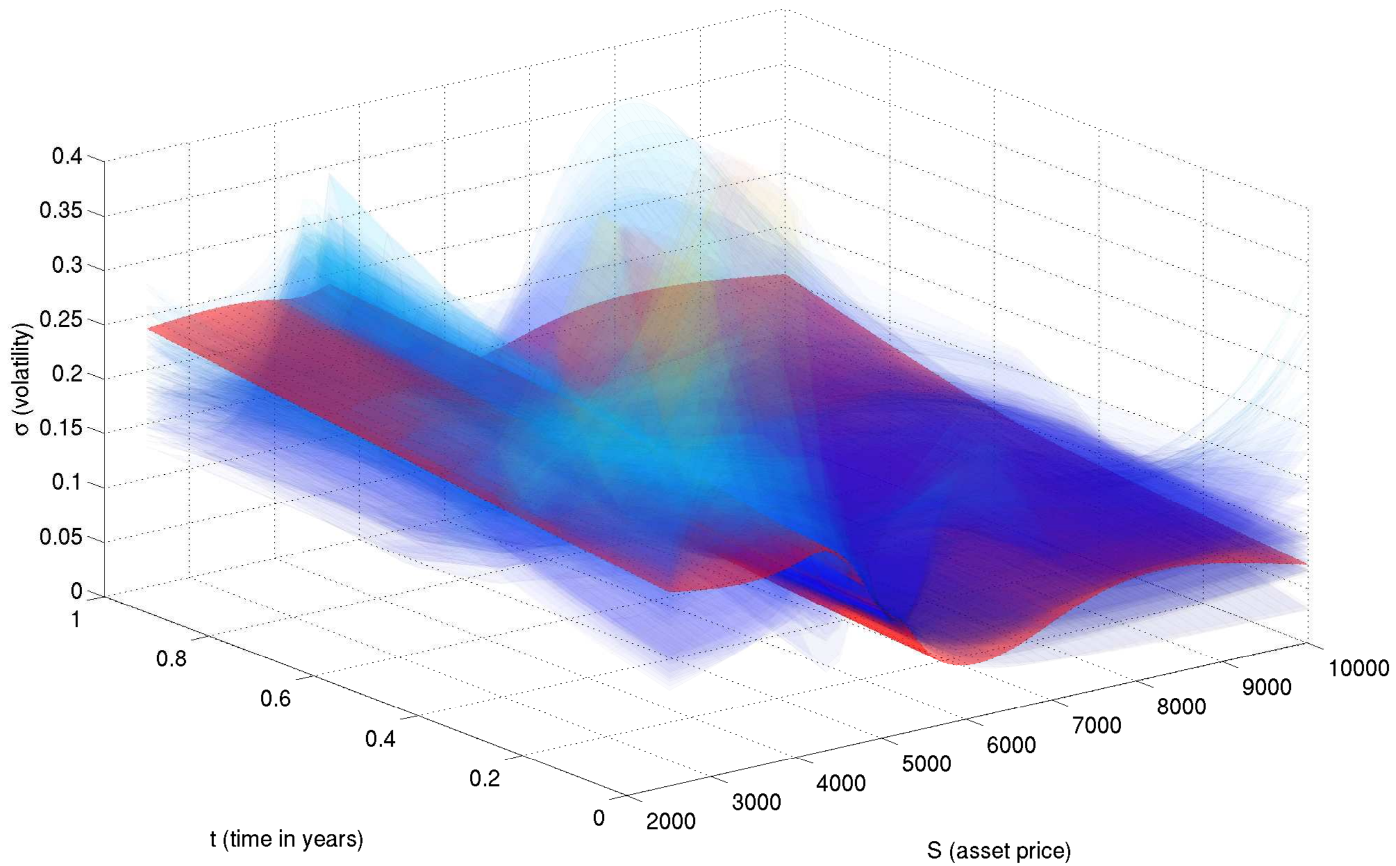
In our case the jump function $J(\theta'|\theta)$ is given by

$$\theta' = \theta + \sqrt{2du}B\xi$$

where B is a matrix corresponding to the norm function $\|\cdot\|_\kappa$, $\xi \sim N(0, I_M)$ and du is the *step size* of the random walk.

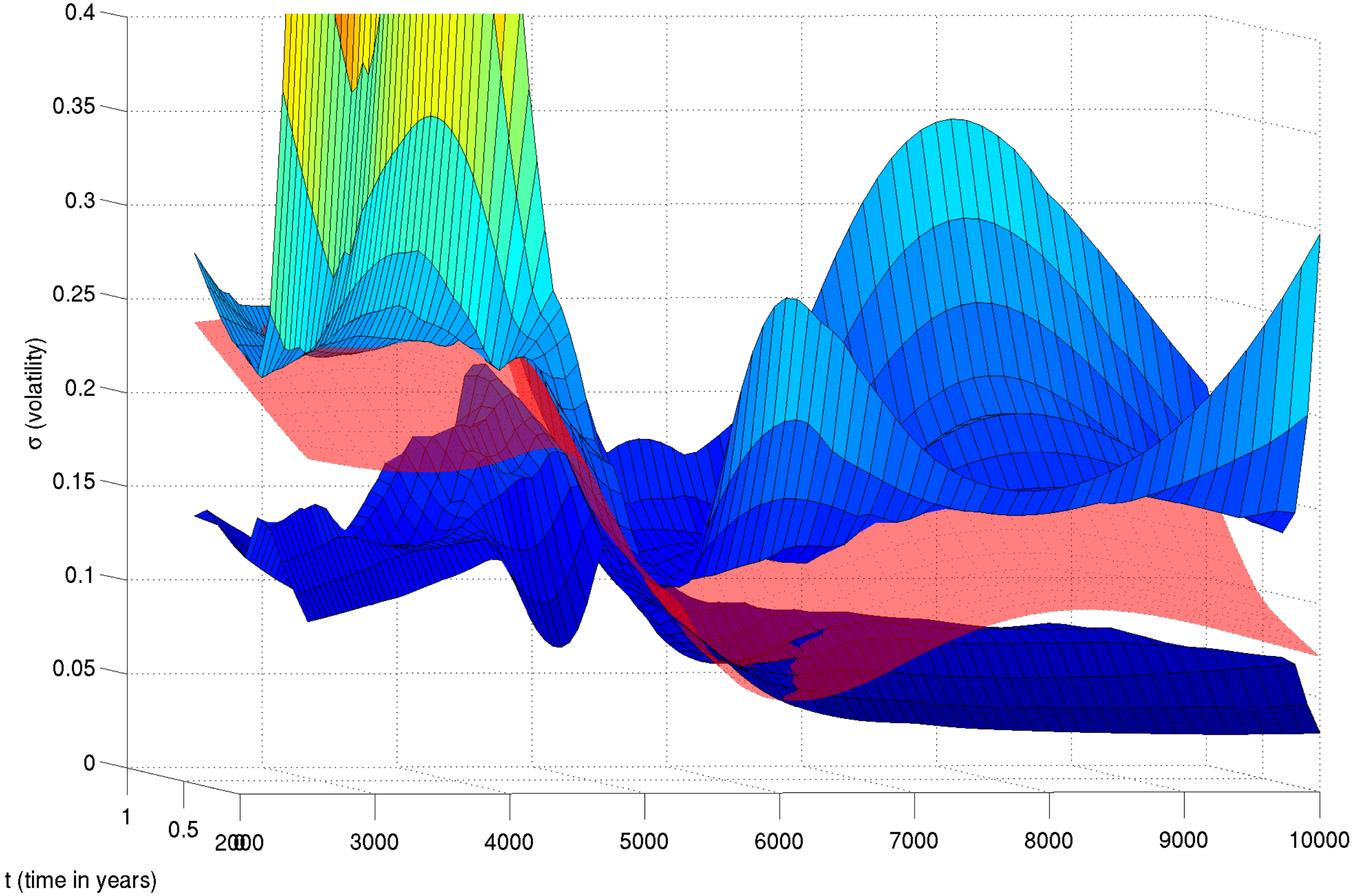
Simulated Dataset

Priced 66 European call options (on a known surface) with 11 strikes and 6 maturities and added Gaussian noise. We take the calibration error tolerance to be $\delta = 3$ basis points and try to calibrate a 27-node surface.



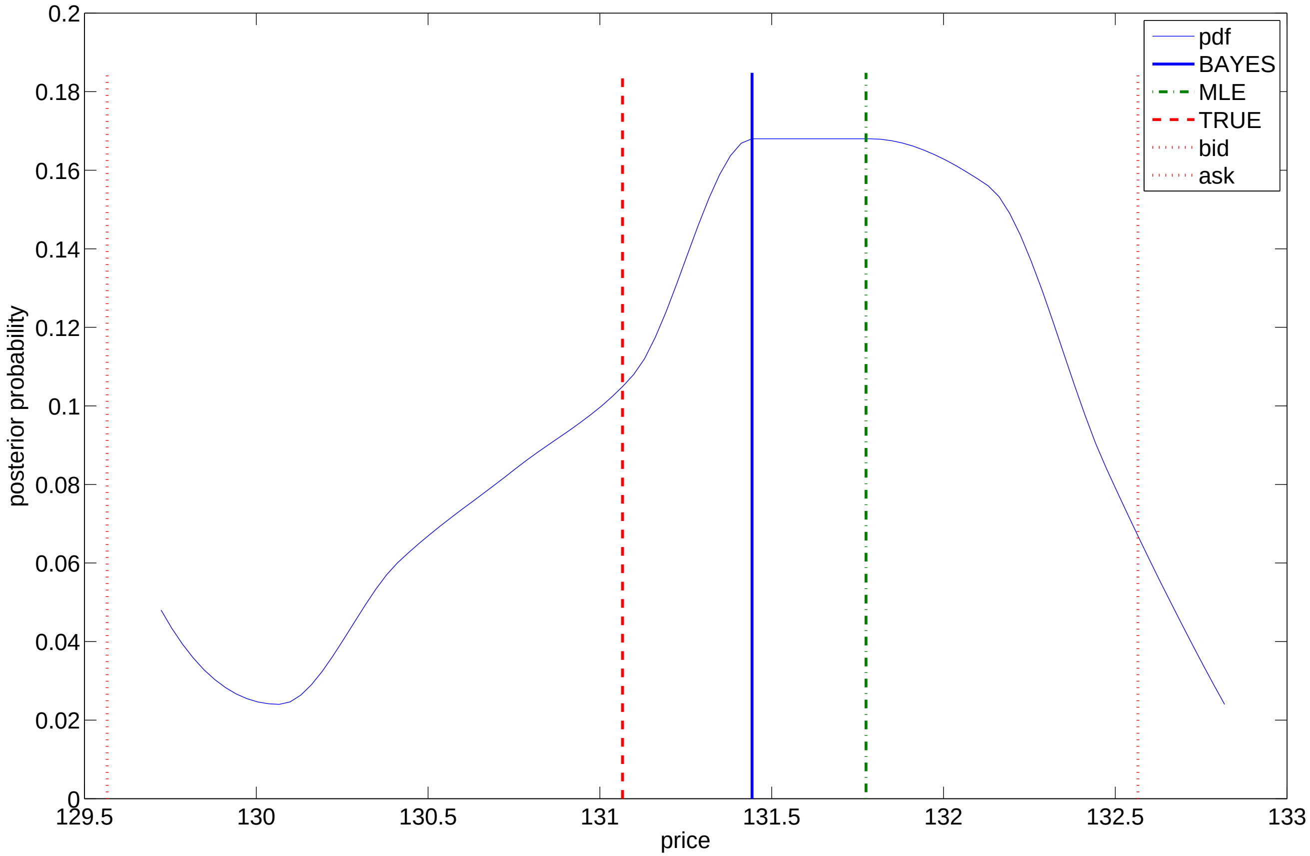
Simulated Dataset

Using this distribution of surfaces we can construct a confidence interval of the value of the local volatility surface $\sigma(S,t)$ at any point (S,t). The figure shows the 95% pointwise confidence interval.



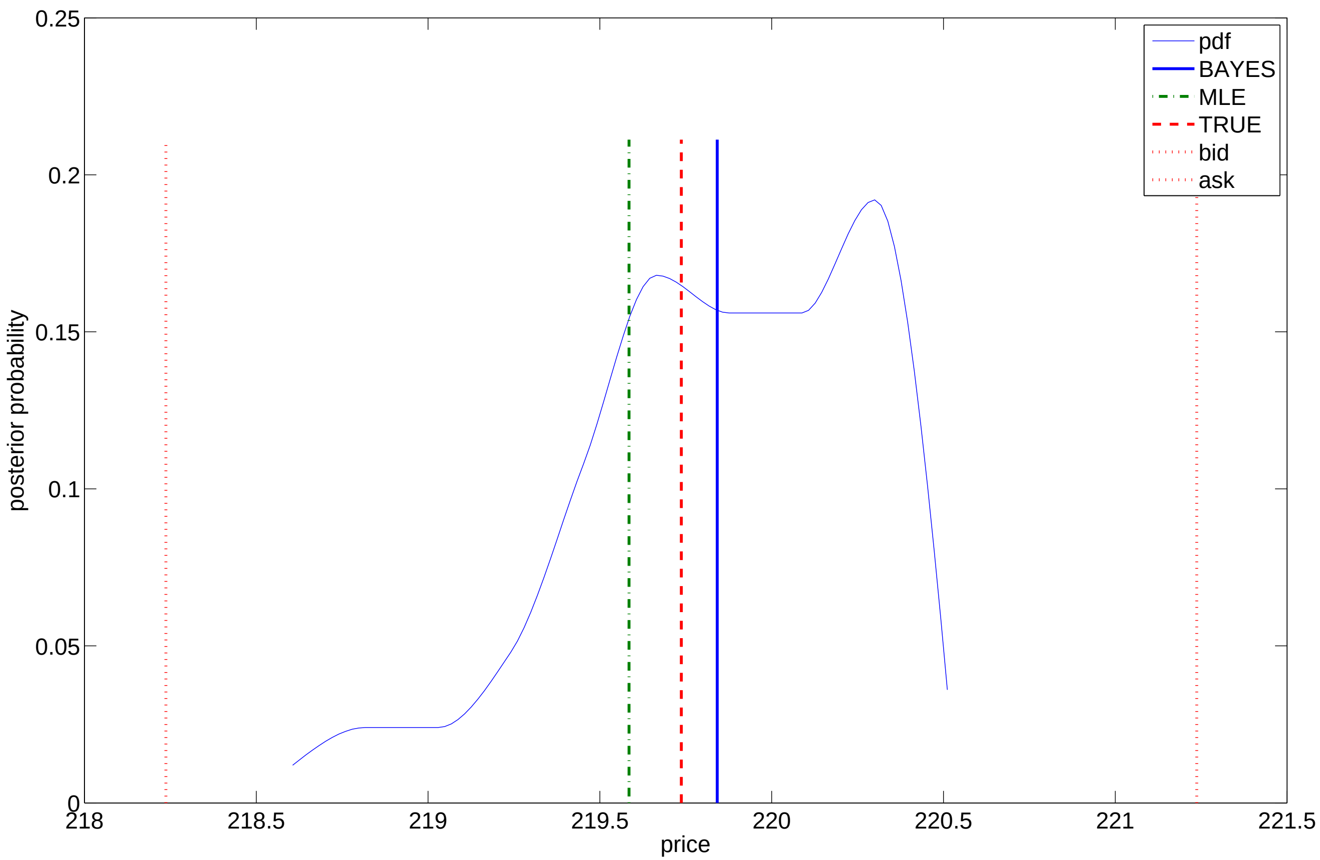
Simulated Dataset

Prices for American put option with strike 5100 ($S_0 = 5000$) and maturity 6 months. Included is the true value (TRUE), maximum likelihood value (MLE), mean of the Bayesian posterior values (BAYES).



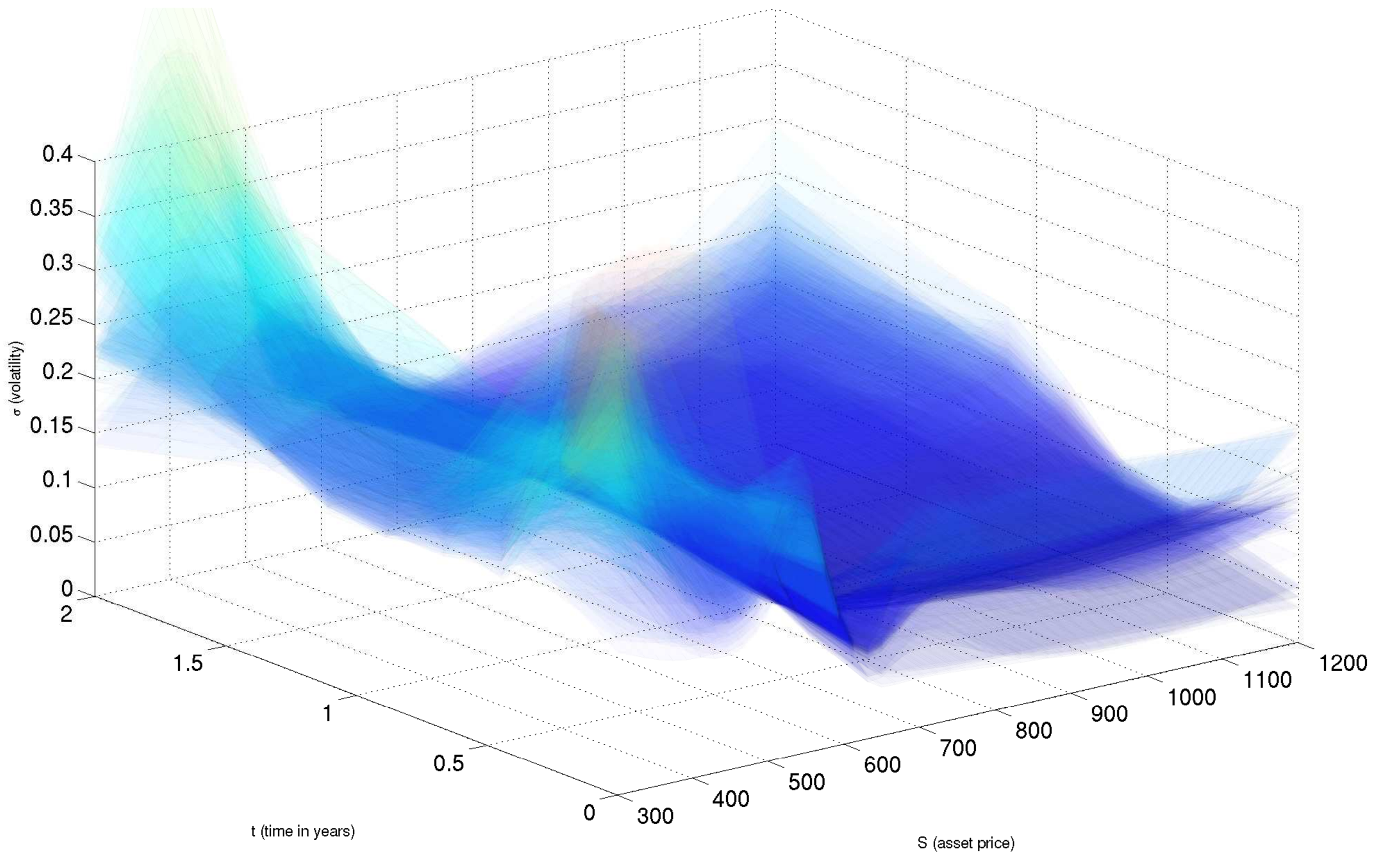
Simulated Dataset

Prices for up-and-out barrier put option: barrier 5300, strike 5200 ($S_0 = 5000$), maturity 1 year. Included is the true value (TRUE), maximum likelihood value (MLE), mean of the Bayesian posterior values (BAYES).



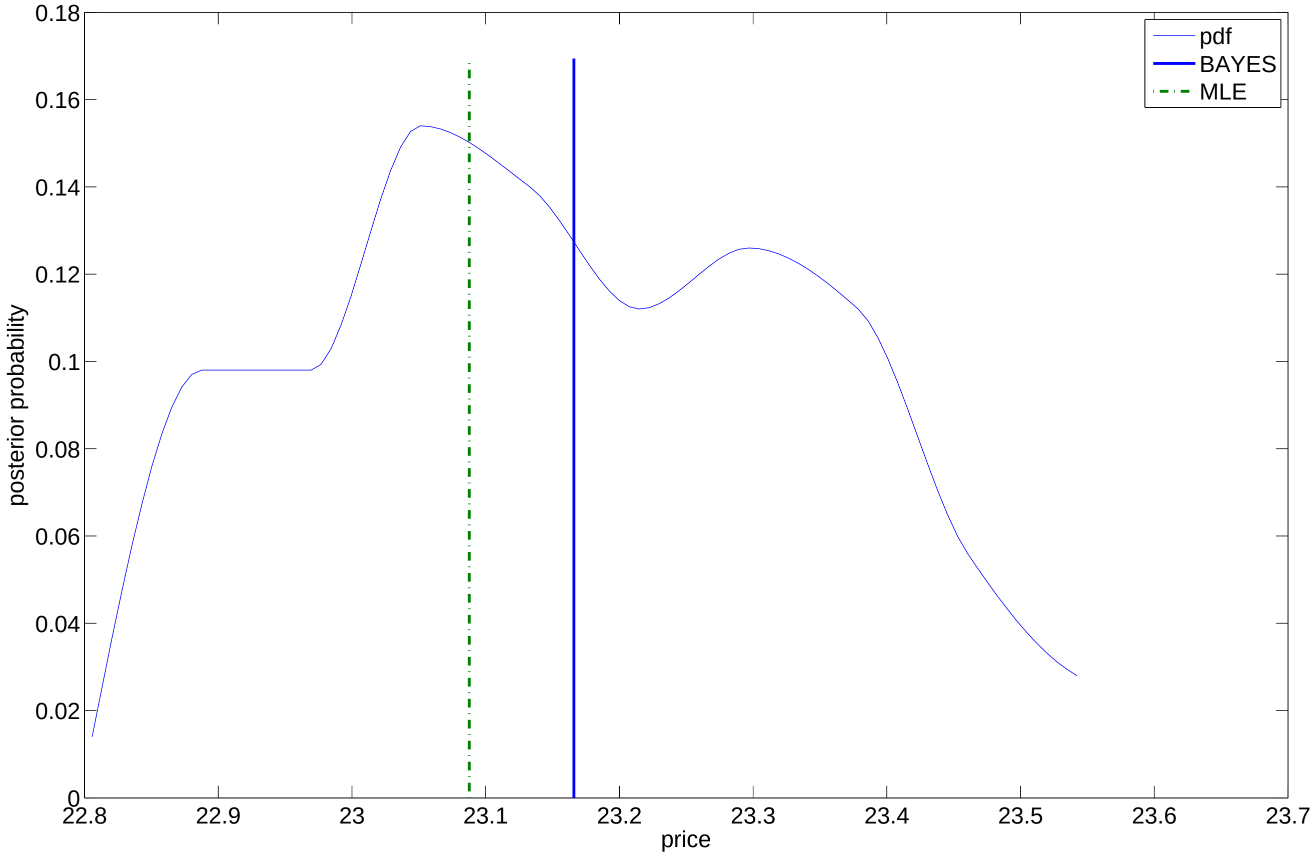
S&P 500

70 S&P 500 European call prices are used (10 strikes and 7 maturities) where S_0 =\$590. We take the calibration error tolerance to be $\delta = 5$ basis points and try to calibrate a 32-node surface.



S&P 500

Prices for American put option with strike 610 ($S_0 = \$590$) and maturity 1 year. Included is the maximum likelihood value (MLE) and mean of the Bayesian posterior values (BAYES).



Conclusion

Summary

- Introduced the Bayesian framework for calibrating the parameters of financial models to market prices.
- Analytic results have been found that prove the Bayesian estimate for the true model paramaters is consistent over time when an underlying constant parameter exists. The results were verified for non-scalar model parameters also.
- Demonstrated a practical method for formulating the prior and likelihood functions necessary for the Bayes procedure.
- Used Local Volatility model as a case study and tested both simulated and real data.
- Saw improvements in pricing when using the Bayesian procedure instead of typical maximum likelihood methods.

Extensions

- The methodology is very general and can be applied to any parametric or non-parametric model. Currently working on applications to non-parametric jump diffusion models.
- Can use Bayesian posterior density $p(\theta|V^*)$ to derive a measure for the model uncertainty of any contingent claim. Such measures would be important for a risk manager or agent trying to decide between different products. (Paper to be submitted).
- Can use Bayesian posterior to develop better hedging strategies. This is more fundamental than pricing as typically a trader will be more interested in the hedging strategy. To this end, the Bayesian loss functions could be designed to correspond to hedging losses so that the Bayes estimator is that parameter θ which minimises the expected hedging loss. (Paper to be submitted).

Thank you for your attention

Questions?