# Filtering in finance: theory and numerics

Thorsten Schmidt

Technische Universität Chemnitz

www.tu-chemnitz.de/mathematik/fima
thorsten.schmidt@mathematik.tu-chemnitz.de

Vienna, April 2013

This is based on joint work with Rüdiger Frey and Ling Xu.

## 1. Introduction

A credit risky position is characterised by

- A stopping time $\tau$ (default time)
- The loss given default $\delta$

- Simplest model: $\tau$ is deterministic.
- $\tau$ is exponential $(\lambda)$
$$\mathbb{P}(\tau > t) = \exp(-\lambda t).$$
- $\tau$ is the first jumping time of a *renewal process*
$$\mathbb{P}(\tau > t) = G(t).$$

We start with a given filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ (asset information).

## First passage time models (threshold models)

Consider a non-negative $\mathbb{F}$-progressive process $(\lambda_t)_{t\geq 0}$ and an independent exponential (1)- random variable $\Theta$. Set

$$\tau := \inf\{t \geq 0 : \int_0^t \lambda_s ds \geq \Theta\}.$$

Then,

$$\begin{aligned}\mathbb{P}(\tau > t) &= \mathbb{P}\Big( \int_0^t \lambda_s ds < \Theta \Big) \\ &= \mathbb{E}\Big( \exp(-\int_0^t \lambda_s ds) \Big).\end{aligned}$$

Moreover, if $(r_t)_{t\geq 0}$ is independent of $\Theta$, then

$$\begin{aligned}P(0,T) &= \mathbb{E}^Q\Big( \exp(-\int_0^T r_s ds) \mathbb{1}_{\{\tau > T\}} \Big) \\ &= \mathbb{E}^Q\Big( \exp(-\int_0^T (r_s + \lambda_s) ds) \Big).\end{aligned}$$

## The problem

- In general the driving process $\lambda$ is *not* observable.
- On the market, information is available through traded instruments which give noisy information about $\lambda$.
- Defaults are observable which give jump-information on the problem.

We formulate this as as filtering problem.

## Filtering

We consider an unobserved *state process* $X$ on $\mathbb{R}^d$ which is the solution of the SDE

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dV_s, \quad 0 \le t \le T, \tag{1}$$

for a $m$-dimensional $\mathbb{F}$-Brownian motion $V$. We denote the generator of $X$ by

$$\mathscr{L} = \sum_{i=1}^d b_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}.$$

Then for $f \in C_b^2(\mathbb{R}^d)$, the process $f(X_t) - f(X_0) - \int_0^t \mathscr{L} f(X_s)ds$ is a martingale.

## Observation

The observation is given by

- *continuous observations* given by the continuous process

$$Z_t = \int_0^t h(X_s)ds + W_t,$$

  where $W$ is independent of $X$, and

- *jump information* given by a counting process $N$ with intensity $(\lambda(X_t))_{t\geq 0}$.

More formally, Let $\mathcal{F}_t^{Z,N} := \sigma(Z_u, N_u : 0 \leq u \leq t)$ denote the observation filtration. We are interested in the conditional distribution of $X_t$ given the observation which is determined by

$$\pi_t(f) := \mathbb{E}\big(f(X_t)|\mathcal{F}_t^{Z,N}\big), \quad f \in L^\infty(\mathbb{R}^d).$$

## Example

- Consider a portfolio credit risk model driven by factor process $X$.
- Large homogeneous portfolio is given by a doubly stochastic Poisson process $N$ with intensity $\lambda(X_t)$.
- $X$ is not observable, but market instrumtens give noisy observations of $X$, modelled by $Z$.
- Of course $N$ is observable.

We obtain the following connection to nonlinear filtering:

- credit derivative: some $\mathcal{F}_T^N$-measurable payoff $H$ with price

$$H_t = \mathbb{E}^{\mathbb{Q}}(H \mid \mathcal{F}_t^{Z,N})$$

(under full information / no statistical insecurity). We obtain

$$H_t = \mathbb{E}^{\mathbb{Q}}\Big(\mathbb{E}^{\mathbb{Q}}(H \mid \mathcal{F}_t) \mid \mathcal{F}_t^{Z,N}\Big).$$

- As $(X, N)$ is $\mathbb{F}$-Markov, $\mathbb{E}^{\mathbb{Q}}(H|\mathcal{F}_t) = h(t, X_t, N_t)$ such that

$$H_t = \mathbb{E}^{\mathbb{Q}}(h(t, X_t, N_t)|\mathcal{F}_t^{Z,N}).$$

Our aim is to derive the Zakai equation by a change of measure argument. By the Girsanov theorem we can find an equivalent measure $\mathbb{P}_0$ such that $d\mathbb{P} = \Lambda_T d\mathbb{P}_0$.

The Kallianpur-Striebel formula relates the conditional distribution $\pi$ to the unnormalized density $\rho_t(f) := \mathbb{E}^0[f(X_t)|\mathcal{F}_t^{Z,N}]$:

$$\pi_t(f) = \frac{\mathbb{E}^0(f(X_t)\Lambda_t|\mathcal{F}_t^{Z,N})}{\mathbb{E}^0(\Lambda_t|\mathcal{F}_t^{Z,N})} = \frac{\rho_t(f)}{\rho_t(1)}. \tag{2}$$

We explicitely construct a useful change of measure as follows: let

$$\Lambda_t := \prod_{\tau_n \leq t} \lambda(X_{\tau_n -}) \cdot \exp\left( \int_0^t h(X_s)^\top dW_s + \frac{1}{2} \int_0^t \|h(X_s)\|^2 ds - \int_0^t (\lambda(X_s) - 1) ds \right)$$

and define $d\mathbb{P}_0 := \Lambda_T^{-1} d\mathbb{P}$. Then under $\mathbb{P}_0$

- $Z$ is a standard Brownian motion
- $N$ is a standard Poisson process (intensity $= 1$)
- $Z, N, X$ are independent.

Define $Y_t := N_t - t$ such that $Y$ is a $\mathbb{P}_0$-martingale. Then $\rho_t$ satisfies the Zakai equation: for any $f \in C_b^2(\mathbb{R}^d)$, $t \in [0, T]$,

$$\rho_t(f) = \rho_0(f) + \int_0^t \rho_s(\mathscr{L} f)ds + \int_0^t \rho_s(f\, h^\top)dZ_s + \int_0^t \rho_{s-}\big(f(\lambda - 1)\big)dY_s,$$

$\mathbb{P}^0 - a.s.$

To derive the Zakai equation we need the following assumptions:

(**A1**) Assume that the following three conditions hold:

1. $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d \times m}$, and $h : \mathbb{R}^d \to \mathbb{R}^l$ are bounded on $\mathbb{R}^d$. Moreover, $b$ is $C^1$ with bounded derivatives and $\sigma$ is $C^2$ with bounded first and second order derivatives.
2. There exists $\alpha > 0$, such that $z^\top a(x) z \geq \alpha z^\top z$, $\forall x, z \in \mathbb{R}^d$.
3. $\lambda : \mathbb{R}^d \to [\varpi_1, \varpi_2]$ is a continuous function for constants $0 < \varpi_1 < \varpi_2$.

## An SPDE for the conditional density

- Consider the separable Hilbert space $H = L^2(\mathbb{R}^d)$ with scalar product $(\cdot, \cdot)$.
- We are interested in a Lebesgue density, i.e. we look for some $H$-valued process $q$ such that

$$\rho_t(f) = (q_t, f)$$

  for sufficiently many $f$.
- From the Zakai equation,

$$(q_t, f) = (q_0, f) + \int_0^t (\mathscr{L}^* q_s, f) ds + \int_0^t (h^\top q_s, f) dZ_s + \int_0^t ((\lambda - 1) q_{s-}, f) dY_s.$$

## An SPDE for the conditional density

- Consider the separable Hilbert space $H = L^2(\mathbb{R}^d)$ with scalar product $(\cdot,\cdot)$.
- We are interested in a Lebesgue density, i.e. we look for some $H$-valued process $q$ such that
$$\rho_t(f) = (q_t, f)$$
for sufficiently many $f$.
- From the Zakai equation,

$$(q_t, f) = (q_0, f) + \int_0^t (\mathscr{L}^* q_s, f)ds + \int_0^t (h^\top q_s, f)dZ_s + \int_0^t ((\lambda - 1)q_{s-}, f)dY_s.$$

$$\rho_t(f) = \rho_0(f) + \int_0^t \rho_s(\mathscr{L} f)ds + \int_0^t \rho_s(f\, h^\top)dZ_s + \int_0^t \rho_{s-}\Big(f(\lambda - 1)\Big)dY_s,$$

## An SPDE for the conditional density

- Consider the separable Hilbert space $H = L^2(\mathbb{R}^d)$ with scalar product $(\cdot, \cdot)$.
- We are interested in a Lebesgue density, i.e. we look for some $H$-valued process $q$ such that
$$\rho_t(f) = (q_t, f)$$
for sufficiently many $f$.
- From the Zakai equation,

$$(q_t, f) = (q_0, f) + \int_0^t (\mathscr{L}^* q_s, f)ds + \int_0^t (h^\top q_s, f)dZ_s + \int_0^t ((\lambda - 1)q_{s-}, f)dY_s.$$

We extend the generator $\mathscr{L}^*$ and denote the extenden operator by $\mathcal{A}^*$. Define the multiplication-operators $\mathcal{B}\colon H \to H^l$, $\mathcal{B}f := fh^\top$ and $\mathcal{C}\colon H \to H$, $\mathcal{C}f := (\lambda - 1)f$. We look for *mild solution* of the SPDE

$$dq_t = \mathcal{A}^* q_t dt + \mathcal{B}q_t dZ_t + \mathcal{C}q_{t-}dY_t. \tag{3}$$

**Theorem**

*Assume that* **(A1)** *holds. Then for all $q_0 \in V$ there is a unique mild solution $q$ of the SPDE* (3)*. Moreover, $q_t \in H^1(\mathbb{R}^d)$ and for all $f \in L^2(\mathbb{R}^d)$ we have that $\rho_t(f) = (q_t, f)$.*

## The Galerkin approximation

The Galerkin approximation for a (stochastic) PDE essentially projects the equation to a finite-dimensional subspace. In the case of the Zakai equation the projected equation can be characterized in terms of a finite-dimensional system of ordinary stochastic differential equations.

- Let $\{e_1, e_2, \ldots\} \subset D(\mathcal{A}^*) \cap D(\mathcal{A})$ be a basis of the Hilbert-space $H$.
- $H_n$ is the linear subspace spanned by $\{e_1, \ldots, e_n\}$ and by $P_n$ we denote the projection to $H_n$.
- We define the projection of the operator $\mathcal{A}^*$ by

$$(\mathcal{A}^*)^{(n)} := P_n \mathcal{A}^* P_n \, ;$$

and analogously $\mathcal{B}^{(n)}$ and $\mathcal{C}^{(n)}$.

## Definition

The $n$-dimensional *Galerkin approximation* of (3) is the solution of

$$\begin{aligned} dq_t^{(n)} &= (\mathcal{A}^*)^{(n)} q_t^{(n)} dt + \mathcal{B}^{(n)} q_t^{(n)} dZ_t + \mathcal{C}^{(n)} q_{t-}^{(n)} dY_t, \\ q_0^{(n)} &= P_n q_0. \end{aligned} \tag{4}$$

We obtain that $q_t^{(n)}$ can be written as

$$q_t^{(n)}(x) = \sum_{i=1}^{n} \psi_i^{(n)}(t) e_i(x), \qquad t \in [0, T], \tag{5}$$

where $\psi_i^{(n)}$, $1 \leq i \leq n$ are called *Fourier coefficients*.
Define the $n \times n$ matrices $A, C, D$ and $B^\ell, \ell = 1, \dots, l$ by their components:

$$a_{ji} := (e_i, \mathcal{A} e_j),\ b_{ji}^\ell := (e_i, h^\ell e_j),\ c_{ji} := (e_i, (\lambda - 1)e_j), d_{ji} := (e_i, e_j)\,. \tag{6}$$

We obtain the following SDE system for the vector-valued process $\Upsilon^{(n)} := (\psi_1^{(n)}, \dots \psi_n^{(n)})^\top$,

$$\begin{aligned} d\Upsilon_t^{(n)} &= D^{-1}\Big(A\Upsilon_t^{(n)} dt + \sum_{\ell=1}^{l} B^\ell \Upsilon_t^{(n)} dZ_t^\ell + C\Upsilon_{t-}^{(n)} dY_t\Big), \\ \Upsilon_0^{(n)} &= D^{-1} q_0^{(n)}\,. \end{aligned} \tag{7}$$

This SDE system will be the starting point for our numerical analysis.

## Convergence results

- $\mathcal{A}^*$ generates an analytic $C_0$-semigroup $G^*$.
- Then $G^*_t x$ is the solution of the Kolmogorov forward PDE with initial condition $x$.

We have the following convergence result of the Galerkin approximation $q^{(n)}$ to the solution of the Zakai equation $q$.

### Theorem

*Assume that* **(A1)** *holds. Let $q$ be the solution of the Zakai equation in* (3) *and $q^{(n)}$ be the corresponding Galerkin approximation. Then, for any $q_0 \in V$,*

$$\sup_{t\in[0,T]} \mathbb{E}^0(\|q_t^{(n)} - q_t\|_H^2) \to 0, \quad \textit{as} \quad n \to \infty,$$

*if and only if, for any $x \in H$,*

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \left\| \left( \exp(P_n \mathcal{A}^* P_n t) - G_t^* \right) x \right\|_H = 0. \tag{8}$$

# Numerical solution of the Zakai equation

In order to solve the SDE system numerically, we discretize in time.

- We consider the Euler-Maruyama (EM) and the Splitting-Up (SU) methods.
- While the EM method is easier to implement it can be quite unstable if the time step is large. This can be overcome by the SU method.

Rewriting Equation (7) leads to

$$d\Upsilon_t^{(n)} = D^{-1}\Big((A-C)\Upsilon_t^{(n)}dt + \sum_{\ell=1}^{l} B^{\ell}\Upsilon_t^{(n)}dZ_t + C\Upsilon_{t-}^{(n)}dN_t\Big), \quad \Upsilon_0^{(n)} = q_0^{(n)}. \tag{9}$$

## Algorithm (EM method)

*For $k = 1, \dots, K$, compute $\Upsilon_k$ from $\Upsilon_{k-1}$ by*

$$\Upsilon_k = \Upsilon_{k-1} + D^{-1}\Big((A - C)\Upsilon_{k-1}\Delta + \sum_{\ell=1}^{l} B^\ell \Upsilon_{k-1}(Z^\ell_{t_k} - Z^\ell_{t_{k-1}}) + C\Upsilon_{k-1}(N_{t_k} - N_{t_{k-1}})\Big).$$

The splitting-up method (SU method) is a numerical method based on semigroup theory.

## Algorithm (SU method)

*For $k = 1, \dots, K$, compute $\Upsilon_k$ from $\Upsilon_{k-1}$ by*

(1) *Compute* $\Upsilon^1_k := \exp\big((A - C)\Delta\big)\Upsilon_{k-1}$.

(2) *Compute* $\Upsilon^2_k := \exp\big(\sum_{\ell=1}^{l}(B^\ell(Z_{t_k} - Z_{t_{k-1}}) - \frac{1}{2}(B^\ell)^2\Delta)\big)\Upsilon^1_k$.

(3) *Return* $\Upsilon_k := (I_n + C)^{(N_{t_k} - N_{t_{k-1}})}\Upsilon^2_k$

## Numerical results

The numerical experiments consider the following case.

### Example (Kalman filter with point process observations)

Consider a one-dimensional example where

$$X_t = X_0 + \int_0^t bX_s ds + \sigma V_t.$$

Furthermore,

$$Z_t = \int_0^t hX_s ds + W_t$$

and $N$ is a doubly stochastic Poisson process $N$ with intensity

$$(\lambda X_t^2)_{t \geq 0}.$$

The coefficients of the Galerkin approximations can be obtained by direct computation (see paper).

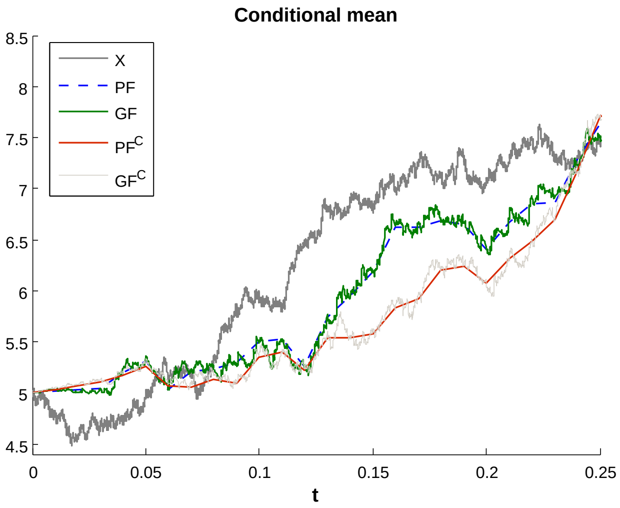
## The adaptive Galerkin approximation

- During the filtering process the conditional distribution $\pi_t(dx)$ typically changes location and scale which can create problems for the Galerkin approximation with a fixed basis.
- We propose an adaptive scheme, called adaptive Galerkin approximation (AGA), which improves the numerical performance of the Galerkin approach significantly.
- Basically, if projected conditional mean or standard deviation deviate too much over time, we adjust the basis to the current mean and standard deviation.

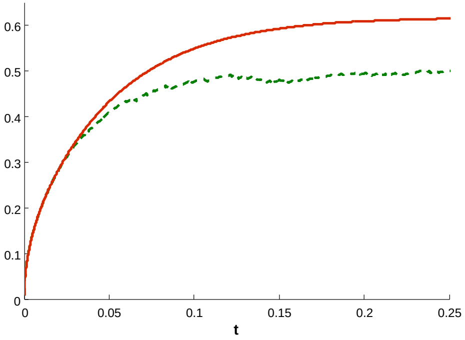
| $N_G/N_P$ | 5/20 | 10/50 | 15/100 |
|---|---|---|---|
| AGAH(EM) | 0.63 (0.1s) | 0.42 (0.1s) | 0.42 (0.1s) |
| AGAH(SU) | 0.65 (2.4s) | 0.43 (3.1s) | 0.43 (3.9s) |
| PF | 0.46 (9s) | 0.46 (22s) | 0.42 (46s) |

### Performance comparison for different filter algorithms

- We plot the RMSE and the computation time for two Galerkin filters and a particle filter (PF).
- The parameter values $h = \lambda = 0.1$ used in the experiment correspond to a relatively uninformative observation filtration.

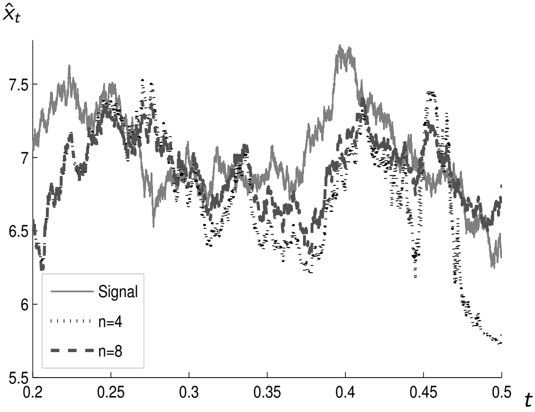



Comparison of Galerkin and particle filtering methods.

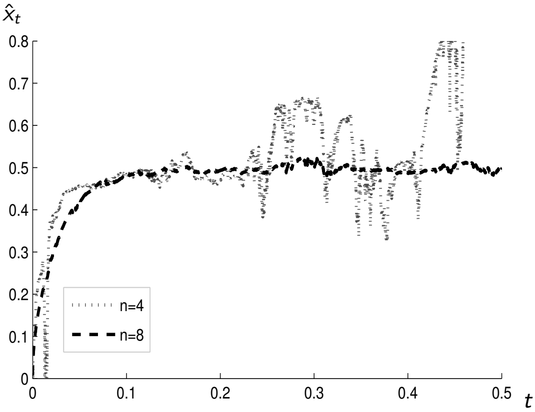



## Value of point process information

- We plot the conditional standard deviation $\hat{\sigma}_t$ for the case with only continuous observation $\lambda = 0$ and with continuous and point process observations ($\lambda = 10$, lower trajectory).
- Clearly, including point process information reduces the conditional standard deviation significantly.




The adaptive Galerkin approximation for different numbers of basis functions.




The corresponding conditional variance can be a useful tool of determing the appropriate number of basis functions.

Many thanks for your attention!

Frey, R., T. Schmidt & L. Xu (2012): On Galerkin Approximations for the Zakai Equation with Diffusive and Point Process Observations. Forthcoming *SIAM Num. Anal.*

Frey, R. and T. Schmidt (2011). Filtering and incomplete information in credit risk. In T. Bielecki, D. Brigo, and F. Patras (Eds.), *Recent Advancements in the Theory and Practice of Credit Derivatives*, pp. 185–218. Bloomberg Press.