Strong Noise Limits / Filtering

Clément Pellegrini

Institut de Mathématiques de Toulouse,
Université Paul Sabatier

Angers / June 2022

Dream Team

- **Work in collaboration, in progress, to be continued**
  - Cedric Bernardin (Nice)
  - Reda Chhaibi (Toulouse)
  - Raphael Chetrite (Nice)
  - Joseph Najnudel (Bristol)
- **Spiking and collapsing in large noise limits of SDEs**
- **To spike or not to spike: the whims of the Wonham filter in the strong noise regime**

Dream Team

- **Work in collaboration, in progress, to be continued**
  - Cedric Bernardin (Nice)
  - Reda Chhaibi (Toulouse)
  - Raphael Chetrite (Nice)
  - Joseph Najnudel (Bristol)
- **Spiking and collapsing in large noise limits of SDEs**
- **To spike or not to spike: the whims of the Wonham filter in the strong noise regime**

Plan

- **I) The model**
- **II) The result**

- I) Wonham Shiryaev Filtering Model

General setting

- Consider a *Hidden Markov process* $\mathbf{x} = (\mathbf{x}_t \; ; \; t \geq 0)$. Typically a continuous time Markov chain with finite state space.
- At hand we have an *Observation Process* $\mathbf{y}$ correlated to $\mathbf{x}$: $\mathbf{y} = (\mathbf{y}_t \; ; \; t \geq 0)$.
- Most simple setup: "signal plus noise"

$$d\mathbf{y}_t^\gamma = \mathbf{x}_t dt + \frac{1}{\sqrt{\gamma}} dB_t,$$

  with (B_t) is a standard Brownian motion.
- The parameter γ is going to ∞.
- In the whole talk, we concentrate on a Hidden Markov process which is a continuous time Markov chain valued in two points $\{0, 1\}$.

General setting

- Consider a *Hidden Markov process* $\mathbf{x} = (\mathbf{x}_t \; ; \; t \geq 0)$. Typically a continuous time Markov chain with finite state space.
- At hand we have an *Observation Process* $\mathbf{y}$ correlated to $\mathbf{x}$: $\mathbf{y} = (\mathbf{y}_t \; ; \; t \geq 0)$.
- Most simple setup: "signal plus noise"

$$d\mathbf{y}_t^\gamma = \mathbf{x}_t dt + \frac{1}{\sqrt{\gamma}} dB_t,$$

  with (B_t) is a standard Brownian motion.
- The parameter γ is going to ∞.
- In the whole talk, we concentrate on a Hidden Markov process which is a continuous time Markov chain valued in two points $\{0, 1\}$.

Simulation

- A simulation for a Markov chain with 2 points.

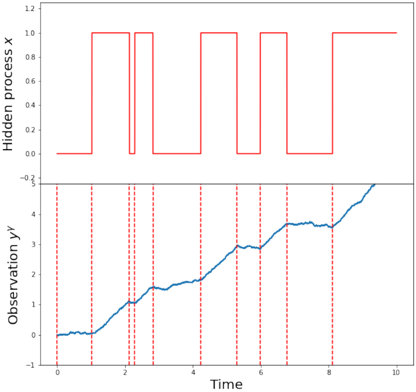


Figure: The process in red is the Hidden process **x**. The process in blue is the observation process **y**

Estimation

- Framework: Observe **y** and take a decision.
- Imagine a classical bit 0 or 1 subject to modification either by computation instructions or errors and follows the Markov chain **x**.
- Then you have access to an electric current described by **y**.
- Taking into account the electric current you have to take decision: for example correct an error.
- Then you have to design an estimator of **x** knowing **y**.

Estimation

- Framework: Observe $\mathbf{y}$ and take a decision.
- Imagine a classical bit 0 or 1 subject to modification either by computation instructions or errors and follows the Markov chain $\mathbf{x}$.
- Then you have access to an electric current described by $\mathbf{y}$.
- Taking into account the electric current you have to take decision: for example correct an error.
- Then you have to design an estimator of $\mathbf{x}$ knowing $\mathbf{y}$.

Estimation

- Framework: Observe **y** and take a decision.
- Imagine a classical bit 0 or 1 subject to modification either by computation instructions or errors and follows the Markov chain **x**.
- Then you have access to an electric current described by **y**.
- Taking into account the electric current you have to take decision: for example correct an error.
- Then you have to design an estimator of **x** knowing **y**.

Estimation

- In mean square sense the best estimator of $\mathbf{x}$ valued in $\{0,1\}$ and $\mathcal{F}_t^{\mathbf{y}} = \sigma\{\mathbf{y}_s, s \leqslant t\}$ mesurable is

$$\hat{\mathbf{x}}_t = \mathbf{1}_{\pi_t^\gamma > \frac{1}{2}}$$

where

$$\pi_t^\gamma = \mathbb{P}[x_t = 1 | \mathcal{F}_t^{\mathbf{y}}]$$

- Note that the optimal filter π_t^γ is defined by

$$\pi_t^\gamma = \underset{(c_t,)\mathcal{F}_t^{\mathbf{y}}\, mesurable}{argmax} \mathbb{E}[(c_t - x_t)^2]$$

and remark that $x_t = \mathbf{1}_{x_t=1}$

- Main concern: what is the behaviour when γ goes to infinity.

Estimation

- In mean square sense the best estimator of $\mathbf{x}$ valued in $\{0,1\}$ and $\mathcal{F}_t^{\mathbf{y}} = \sigma\{\mathbf{y}_s, s \leqslant t\}$ mesurable is

$$\hat{\mathbf{x}}_t = \mathbf{1}_{\pi_t^\gamma > \frac{1}{2}}$$

  where

$$\pi_t^\gamma = \mathbb{P}[x_t = 1 | \mathcal{F}_t^{\mathbf{y}}]$$

- Note that the optimal filter π_t^γ is defined by

$$\pi_t^\gamma = \underset{(c_t,)\mathcal{F}_t^{\mathbf{y}} \, mesurable}{argmax} \mathbb{E}[(c_t - x_t)^2]$$

  and remark that $x_t = \mathbf{1}_{x_t=1}$
- Main concern: what is the behaviour when γ goes to infinity.

Estimation

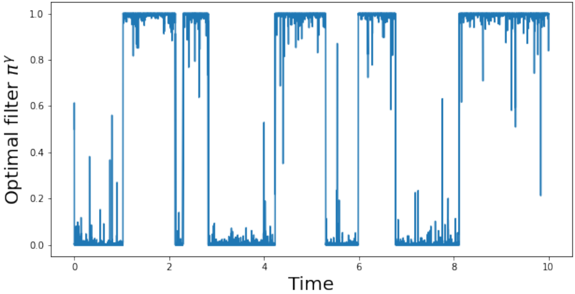


Figure: "The whims of the Wonham filter"

- Denote $(\mathbb{X}_t)$ the process describing the above figure which is that he limit of π_t^γ for γ large.
- If you take decision with $\mathbf{1}_{\mathbb{X}_t > \frac{1}{2}}$, you will make errors.

Estimation

- Solution: do not make instantaneous decision (a general philosophy : −))
- Take or not take a decision: that is the question ?
- Consider

$$\pi_t^{\delta,\gamma} := \mathbb{P}\left(\mathbf{x}_{t-\delta} = 1 \mid (\mathbf{y}_s^{\gamma})_{s\leq t}\right) \ , \tag{1}$$

and

$$\hat{\mathbf{x}}_t^{\delta,\gamma} = \mathbf{1}_{\left\{\pi_t^{\delta,\gamma} > \frac{1}{2}\right\}} \ .$$

- Of course if you take $\delta > 0$ fixed you will take good decision
- Sometimes you have to take quick decision. What happens if δ depends on γ.

Estimation

- Solution: do not make instantaneous decision (a general philosophy : $-$))
- Take or not take a decision: that is the question ?
- Consider

$$\pi_t^{\delta,\gamma} := \mathbb{P}\left(\mathbf{x}_{t-\delta} = 1 \mid (\mathbf{y}_s^{\gamma})_{s\leq t}\right) \ , \tag{1}$$

and

$$\hat{\mathbf{x}}_t^{\delta,\gamma} = \mathbf{1}_{\left\{\pi_t^{\delta,\gamma} > \frac{1}{2}\right\}} \ .$$

- Of course if you take $\delta > 0$ fixed you will take good decision
- Sometimes you have to take quick decision. What happens if δ depends on γ.

- II) Limit result

Estimation

- First let us describe the process $(\mathbb{X}_t)$ limit of π_t^γ.
- Derive the equation for π_t^γ

$$d\pi_t^\gamma = -\lambda\,(\pi_t^\gamma - p)\,dt + \sqrt{\gamma}\pi_t^\gamma\,(1 - \pi_t^\gamma)\,dW_t\ ,$$

  where $\lambda > 0$ and $p \in (0,1)$ are parameters that are linked to the unobserved Markov chain $\mathbf{x}$.
- (W_t) is a Brownian motion.
- What can be the limit of such process when γ goes to infinity?
- Which topology ? Usual ? Exotic ?

Estimation

- First let us describe the process $(\mathbb{X}_t)$ limit of π_t^γ.
- Derive the equation for π_t^γ

$$d\pi_t^\gamma = -\lambda \left(\pi_t^\gamma - p\right) dt + \sqrt{\gamma}\pi_t^\gamma \left(1 - \pi_t^\gamma\right) dW_t \ ,$$

  where $\lambda > 0$ and $p \in (0,1)$ are parameters that are linked to the unobserved Markov chain $\mathbf{x}$.
- (W_t) is a Brownian motion.
- What can be the limit of such process when γ goes to infinity?
- Which topology ? Usual ? Exotic ?

Candidate for the limit ?

- $(\mathbf{x}_t\ ;\ t \geq 0)$ which is a pure jump Markov process on $\{0,1\}$ with càdlàg trajectories. More precisely λp (resp. $\lambda(1-p)$) are the jump rate between 0 and 1 (resp. between 1 and 0), with $p \in (0,1)$ and $\lambda > 0$. $(\mathbf{x}_t\ ;\ t \geqslant 0)$ The initial position is sampled according to

$$\mathbb{P}\left(\mathbf{x}_0 = 1\right) = 1 - \mathbb{P}\left(\mathbf{x}_0 = 0\right) = x_0\ .$$

- Sample a random initial segment $\mathbb{X}_0$ as

$$\mathbb{X}_0 = \begin{cases} [Y,1] \text{ when } \mathbf{x}_0 = 1, & \mathbb{P}\left(Y \in dy \mid \mathbf{x}_0 = 1\right) = \frac{1-x_0}{x_0}\mathbf{1}_{\{0<y<x_0\}}\frac{dy}{(1-y)^2}\ , \\ [0,Y] \text{ when } \mathbf{x}_0 = 0, & \mathbb{P}\left(Y \in dy \mid \mathbf{x}_0 = 0\right) = \frac{x_0}{1-x_0}\mathbf{1}_{\{x_0<y<1\}}\frac{dy}{y^2}\ . \end{cases}$$

Candidate for the limit ?

- Sample $\left(t, \widetilde{M}_t\right)$ following a Poisson point process on $\mathbf{R}_+ \times [0,1]$ with intensity

$$\left(dt \otimes \frac{dm}{m^2} \mathbf{1}_{\{0 \leqslant m < 1\}}\right) .$$

Then, by progressively rescaling time for $\left(t, \widetilde{M}_t\right)$ by

$$\begin{cases} \frac{1}{\lambda p} & \text{when} \quad \mathbf{x}_t = 0 \ , \\ \frac{1}{\lambda(1-p)} & \text{when} \quad \mathbf{x}_t = 1 \ , \end{cases}$$

we obtain a Poisson point process with random intensity which we denote by (t, M_t).

- Finally

$$\mathbb{X}_t = \begin{cases} [0, M_t] & \text{if} \quad \mathbf{x}_t = \mathbf{x}_{t^-} = 0 \ , \\ [1 - M_t, 1] & \text{if} \quad \mathbf{x}_t = \mathbf{x}_{t^-} = 1 \ , \\ [0, 1] & \text{if} \quad \mathbf{x}_t \neq \mathbf{x}_{t^-} \ . \end{cases}$$

Sketch of the proof

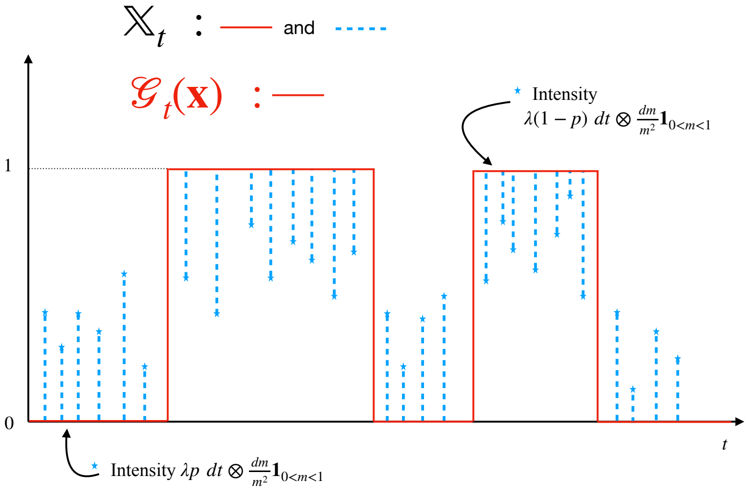


Figure:

Sketch of the proof

- Any one dimensional diffusion is a Brownian motion up to change in time and change in space
- Harmonic function h_γ such that $h_\gamma(\pi_t^\gamma)$ is a martingale

$$h_\gamma : -\lambda(x-p)h_\gamma'(x) + \frac{\gamma}{2}x^2(1-x)^2 h_\gamma''(x) = 0$$

- One can find an explicit solution.
- Dambis Dubin Schwarz

$$h_\gamma(\pi_t^\gamma) = \beta_{T_t}$$

where (β_t) is a Brownian motion and

$$T_t^\gamma = \gamma \int_0^t (\pi_s^\gamma)^2 (1-\pi_s^\gamma)^2 ds$$

Sketch of the proof

- Any one dimensional diffusion is a Brownian motion up to change in time and change in space
- Harmonic function h_γ such that $h_\gamma(\pi_t^\gamma)$ is a martingale

$$h_\gamma : -\lambda(x-p)h_\gamma'(x) + \frac{\gamma}{2}x^2(1-x)^2 h_\gamma''(x) = 0$$

- One can find an explicit solution.
- Dambis Dubin Schwarz

$$h_\gamma(\pi_t^\gamma) = \beta_{T_t}$$

where (β_t) is a Brownian motion and

$$T_t^\gamma = \gamma \int_0^t (\pi_s^\gamma)^2(1-\pi_s^\gamma)^2 ds$$

Sketch of the proof

- Any one dimensional diffusion is a Brownian motion up to change in time and change in space
- Harmonic function h_γ such that $h_\gamma(\pi_t^\gamma)$ is a martingale

$$h_\gamma : -\lambda(x-p)h'_\gamma(x) + \frac{\gamma}{2}x^2(1-x)^2 h''_\gamma(x) = 0$$

- One can find an explicit solution.
- Dambis Dubin Schwarz

$$h_\gamma(\pi_t^\gamma) = \beta_{T_t}$$

where (β_t) is a Brownian motion and

$$T_t^\gamma = \gamma \int_0^t (\pi_s^\gamma)^2(1-\pi_s^\gamma)^2 ds$$

Sketch of the proof

- Now we have

$$\pi_t^\gamma = h_\gamma^{(-1)} h_\gamma(\pi_t^\gamma) = h_\gamma^{(-1)}(\beta_{T_t})$$

- First convergence: almost surely, and uniformly on all compact of the form $[0, L]$

$$T_t^\gamma \xrightarrow[\gamma]{} \sigma_t, \tag{2}$$

where

$$\sigma_t = \inf\left\{\ell \,\middle|\, \frac{L_0^\ell(\beta)}{\lambda p} + \frac{L_0^\ell(\beta)}{\lambda(1-p)} > t\right\}$$

where $L_a^\ell(\beta)$ is the local time of the brownian motion at time ℓ in a.

- To prove this convergence, it is easier to look at the inverse $T_t^{(-1)}$ and use the occupation formula

$$T_\ell^{\langle -1\rangle} = \int_0^\ell \varphi_\gamma(\beta_u)\, du = \int_{\mathbb{R}} \varphi_\gamma(a) L_\ell^a(\beta)\, da\ .$$

The Theorem

Theoreme

It is possible to couple the processes $(\mathbf{x}, \mathbb{X})$ *and* π^γ *for all values of* $\gamma > 0$ *on the same probability space, so that the following limits hold almost surely.*

- *Upon smoothing via a continuous function with compact support* $f : \mathbf{R}_+ \times \mathbf{R} \to \mathbf{R}$, *we have the almost sure convergence:*

$$\lim_{\gamma \to \infty} \int_0^\infty f(t, \pi_t^\gamma)\, dt = \int_0^\infty f(t, \mathbf{x}_t)\, dt \ . \tag{3}$$

- *In the sense of Hausdorff convergence of closed sets, for all* $H > 0$, *we have the almost sure convergence of graphs:*

$$\lim_{\gamma \to \infty} (\pi_t^\gamma;\ 0 \leqslant t \leqslant H) = (\mathbb{X}_t;\ 0 \leqslant t \leqslant H) \ . \tag{4}$$

The Theorem

Theoreme

As long as $\delta_\gamma \to 0$, we have the convergence in probability, as in the first item of Theorem **??***:*

$$\left(\pi_t^{\delta_\gamma,\gamma}\ ; 0 \leqslant t \leqslant H\right) \overset{\gamma\to\infty}{\longrightarrow} (\mathbf{x}_t\ ; 0 \leqslant t \leqslant H)\ . \tag{5}$$

The Theorem

Theoreme

However, in the stronger topologies, there exists a sharp transition when writing:

$$\delta_\gamma = C\frac{\log \gamma}{\gamma} \ .$$

The following convergences hold in the Hausdorff topology.

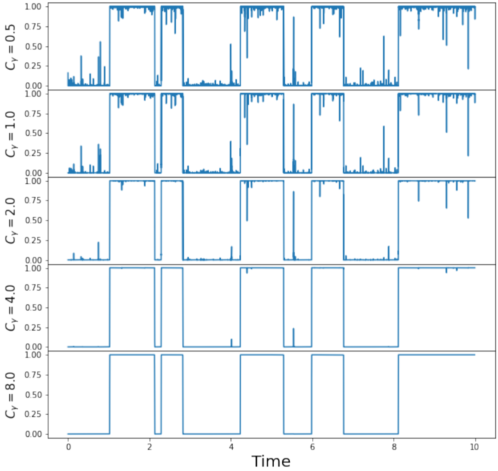
- *(Fast feedback regime) If $C < 2$, smoothing does not occur and we have convergence in law to the spike process:*

$$\lim_{\gamma \to \infty} \pi^{\delta_\gamma, \gamma} = \mathbb{X} \ .$$

- *(Slow feedback regime) If $C > 2$, smoothing occurs and we have convergence:*

$$\lim_{\gamma \to \infty} \pi^{\delta_\gamma, \gamma} = \mathbf{x} \ .$$

The Theorem



$C_\gamma = 0.5$
$C_\gamma = 1.0$
$C_\gamma = 2.0$
$C_\gamma = 4.0$
$C_\gamma = 8.0$
1.00
0.75
0.50
0.25
0.00
0
2
4
6
8
10
Time

Thank you