Two cases of stochastic maximum principle in the optimal control of SPDEs

Marco Fuhrman

Politecnico di Milano

Ying Hu

Université de Rennes 1

Gianmario Tessitore

Università di Milano-Bicocca

Rennes 24th of May 2013

Structure of the talk

We prove Pontryagin maximum principle (necessary conditions for optimality) for a controlled stochastic PDE in the following situations:

1. Part I: Infinite dimensional (white) noise - Special case (No second variation needed)
   - Stochastic parabolic equations in an interval $[0,1]$
   - driven by space-time white noise (cylindrical Wiener process) (W_t)
   - with convex set of controls (control affects noise)
   - non-linearities are Nemytskii operators $F(x) = f(\cdot, x(\cdot))$, $x \in L^p(\mathcal{O})$

2. Part II: Finite dimensional noise, General case
   - Stochastic parabolic equations in a domain $\mathcal{O} \subset \mathbb{R}^d$
   - driven by a finite dimensional Wiener process $W_t = (\beta_t^1, \ldots, \beta_t^m)$
   - with non convex set of controls (control affects noise)
   - non-linearities are Nemytskii operators:

Very incomplete history of SMP in ∞ dimensions

- *Bensoussan, J. Frank. Inst. (1983)* and *Hu-Peng, Stochastics (1990):* Special case, State ∞-dim., noise has trace class covariance.
- *Peng SICON (1990):* General case, state and noise fin. dim.,
- *Zhou, SICON (1993):* State ∞-dim. noise fin. dim., Linear state equat. and cost.,
- *Tang-Li LNPAM (1994):* General case, state ∞-dim. noise fin. dim., noise can have jumps, second derivatives of the coefficients are Hilbert-Schmidt.
- *Fuhrman-Hu-T. CRAS 2012, AMO 2012 (electronic):* General case, state ∞-dim., noise fin. dim., Specific framework to cover stochastic parabolic PDEs
- *Lu-Zhang, Preprint 2012:* General case, state ∞-dim., noise fin. dim., P_t characterized as "transposition solution" of a BSEE. Nonlinearities regular in functional spaces.
- *Du-Meng, Preprints 2012:* General case, state ∞-dim. noise either fin. dim. or trace class, Leading operator A can depend on t; unbounded linear term affecting noise. Some regularity required for the nonlinearities.
- *Fuhrman-Hu-T. :* Special case, state ∞-dim. noise ∞-dim. and cylindrical.

PART I: INFINITE DIMENSIONAL - (WHITE) NOISE - Special Case

Formulation of the optimal control problem

Let $(\mathcal{W}(t,x))$, $t \geq 0$, $x \in [0,1]$ be a space time white noise
$(\mathcal{F}_t)_{t\geq 0}$ denotes its natural (completed) filtration.

The set of admissible control actions U is a convex subset of $L^\infty([0,1])$.
A control u is a (progressive) process with values in U.

The controlled state equation is the following SPDE: for $t \in [0,T]$, $x \in [0,1]$,

$$\begin{cases} dX_t(x) &= \dfrac{\partial^2}{\partial x^2} X_t(x)\,dt + b(x, X_t(x), u_t(x))\,dt + \sigma(x, X_t(x), u_t(x)) d\mathcal{W}(t,x), \\ X_t(0) &= X_t(1) = 0, \quad t \in [0,T] \\ X_0(x) &= x_0(x), \quad x \in [0,1] \end{cases}$$

where $b(x,r,u), \sigma(x,r,u) : [0,1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are given,
we assume they are C^1 and Lipschitz with respect to r and u;
for fixed r and u we suppose $b(\cdot,r,u) \in L^2([0,1])$, $\sigma(\cdot,r,u) \in L^\infty([0,1])$ bdd.

We also introduce the cost functional:

$$J(u) = \mathbb{E}\int_0^T \int_{\mathcal{O}} l(x, X_t(x), u_t(x))\,dx\,dt + \mathbb{E}\int_{\mathcal{O}} h(x, X_T(x))\,dx$$

where $l(x,r,u) : \mathcal{O} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $h(x,r) : \mathcal{O} \times \mathbb{R} \to \mathbb{R}$ are given bounded functions,
we assume that they are C^1 with bounded derivatives with respect to r and u;

Abstract reformulation

The noise is reformulated as a $L^2([0,1])$ valued cylindrical Wiener process (W_t)

$$\mathbb{E} < W_t, x >_{L^2} < W_s, y >_{L^2} = (t \wedge s) < x, y >_{L^2}, \quad \forall x, y \in H = L^2([0,1])$$

A is the realization of the second derivative operator in H with Dirichlet boundary conditions. So it is an unbounded operator with domain $H^2_0([0,1]) \subset H = L^2([0,1])$.

For all $X, V \in H$, $x \in [0,1]$ the non linearities are defined by

$$F(X,u)(x) = b(x, X(x), u(x)), \qquad [G(X,u)V](x) = \sigma(x, X(x), u(x))V(x),$$

$$L(X,u)(x) = \int_{\mathcal{O}} l(X(x), u(x))dx, \quad \Phi(X)(x) = \int_{\mathcal{O}} h(X(x))dx$$

The state equation written in abstract form becomes

$$d_t X_t = AX_t dt + F(X_s, u_s)ds + G(X_s, u_s)dW_s, \quad X_0 = x_0$$

where $x_0 \in H$ and the solution will evolve in H.

The cost becomes

$$J(x,u) = \mathbb{E} \int_0^T L(X_s, u_s)ds + \mathbb{E}\Phi(X_T)$$

Standing Framework

(i) A is the generator of a C_0 semigroup e^{tA}, $t \geq 0$, in H. Moreover $\forall s > 0$:

$$e^{sA} \in L_2(H) \text{ with } |e^{sA}|_{L_2(H)} \leq L s^{-\gamma}; \text{ for suitable } L > 0,\ \gamma \in [0, 1/2).$$

where $L_2(H)$ is the (Hilbert) space of Hilbert Schmidt operators in H.

(ii) U is a bounded convex subset of a separable Banach space U_0

(iii) $F : H \times U \to H$ is lipschitz in both variables

(iv) $G : H \times U \to L(H)$ verifies for all $s > 0$, $t \in [0, T]$, $X, Y \in H$, $u, v \in U$,

$$\begin{aligned} &|e^{sA} G(t, 0, u)|_{L_2(H)} \leq L\, s^{-\gamma}, \\ &|e^{sA} G(t, X, u) - e^{sA} G(t, Y, v)|_{L_2(H)} \leq L\, s^{-\gamma}(|X - Y| + |u - v|), \end{aligned} \tag{1}$$

for some constants $L > 0$ and $\gamma \in [0, 1/2)$.

(v) $F(\cdot, \cdot)$ is Gateaux differentiable $H \times U \to H$,
for all $s > 0$, $e^{sA} G(\cdot, \cdot)$ is Gateaux differentiable $H \times U \to L_2(H)$.

(vi) $L(\cdot, \cdot)$ and $\Phi(\cdot)$ are bounded lipschitz and differentiable

(vii) For all $\Xi \in H$ the map $u \to G(X, u)\Xi$ is Gateaux differentiable and

$$|\nabla_u G(X, u)\Xi|_{\mathcal{L}(U_0, H)} \leq \text{cost} |\Xi|_H \qquad \text{recall } (U_0 \subset L^\infty)$$

Under the above assumptions the state equation (formulated in mild sense):

$$X_t = e^{tA}x_0 + \int_0^t e^{(t-s)A}F(X_s, u_s)ds + \int_0^t e^{(t-s)A}G(X_s, u_s)dW_s$$

admits a unique solution $X \in L^p_W(\Omega, C([0,T], H))$ see [Da Prato, Zabczyk '92]

Remark: If we perturb the control by spike variation that is we consider solution of

$$X_t^\epsilon = e^{tA}x_0 + \int_0^t e^{(t-s)A}F(X_s, u_s^\epsilon)ds + \int_0^t e^{(t-s)A}G(X_s, u_s^\epsilon)dW_s$$

where $u_s^\epsilon = u_s I_{[t_0,t_0+\epsilon]^c}(s) + v_0 I_{[t_0,t_0+\epsilon]}(s)$ for fixed $t_0 \in [0,T]$, $v_0 \in U$ then

$$|X^\epsilon(t_0+\delta) - X(t_0+\delta)|_{L^2(\Omega,\mathbb{P},H)} \approx \delta^{(1/2-\gamma)}$$

First Variation Equation

Let $(\bar{X}, \bar{u})$ be an optimal pair, fix any other bdd. U-valued progressive control v

Let $u^\epsilon_t = \bar{u}_t + \epsilon(v_t - \bar{u}_t)$ and X^ϵ_t the corr. solution of the state equation.

Finally denote $(\delta u)_t = v_t - \bar{u}_t$

Since we are not considering spike variations things are easy at this level

$$X^\epsilon_t = \bar{X}_t + \epsilon Y_t + o(\epsilon).$$

$$\begin{cases} dY_t &= \left[AY_t + \nabla_X F(\bar{X}_t, \bar{u}_t)Y_t + \nabla_u F(\bar{X}_t, \bar{u}_t)(\delta u)_t\right] dt \\ & \quad \nabla_X G(\bar{X}_t, \bar{u}_t)Y_t\, dW_t + \nabla_u G(\bar{X}_t, \bar{u}_t)(\delta u)_t,\, dW_t \\ Y^\epsilon_0 &= 0 \end{cases}$$

By [Da Prato Zabczyk] the above equation admits an unique mild solution with

$$\mathbb{E}(\sup_{t\in[0,T]} |Y_t|^2) < +\infty,$$

Moreover

$$J(x, u^\epsilon) = J(x, \bar{u}) + \epsilon I(v) + o(\epsilon)$$

with

$$I(v) = \mathbb{E}\int_0^T \left[\langle \nabla_X L(\bar{X}_t, \bar{u}_t), Y_t\rangle + \langle \nabla_u L(\bar{X}_t, \bar{u}_t), (\delta u)_t\rangle\right] dt + \mathbb{E}\langle \nabla_X \Phi(\bar{X}_T), Y_T\rangle$$

We fix a basis $(e_i)_{i\in\mathbb{N}} \in H$ and assume that for all $i \in \mathbb{N}$ the map $X \to G(X,u)e_i$ is Gateaux differentiable $H \to H$.

We notice that in our concrete case for all $V \in H$

$$[\nabla_X(G(X,u)e_i)V](x) = \frac{\partial}{\partial X}\sigma(x, X_t(\xi), u_t(\xi))e_i(\xi)V(\xi)$$

So it is enough to choose $e_i \in L^\infty([0,1])$

We denote $\nabla_X(G(\bar{X}_t, \bar{u}_t)e_i)V = C_i(t)V$.

Recall that gradients ∇_X are with respect to variables $X \in H = L^2([0,1])$

For simplicity we let $F = 0$ from now on:.

The equation for the first variation becomes

$$\begin{cases} dY_t(x) &= AY_t dt + \sum_{i=1}^\infty C_i(t)Y_t\, d\beta_t^i + \nabla_u G(\bar{X}_t, \bar{u}_t)(\delta u)_t\, dW_t \\ Y_0 &= 0 \end{cases}$$

where $\beta_t^i = \langle e_i, W_t\rangle$ and we have:

- $|C_i(t)|_{L(H)} \le c, \quad \mathbb{P}-$a.s. for all $t \in [0,T]$
- $\sum_{i=1}^\infty |e^{tA}C_i(s)v|^2 \le ct^{-2\gamma}|v|_H^2 \quad$ for all $t \ge 0$, $s \ge 0$, $\quad (\gamma < 1/2)$
- $\sum_{i=1}^\infty |e^{tA}e_i|^2 \le ct^{-2\gamma} \quad$ for all $t \ge \quad (\gamma < 1/2)$,

We also take into account that A and C_i are self adjoint (although not essential).

Adjoint equation

The adjoint equation is (at least formally)

$$\begin{cases} -dp_t(x) &= \left[Ap_t + \nabla_X L(\bar{X}_t, \bar{u}_t) + \sum_{i=1}^{\infty} C_i(t) Q_t e_i\right] dt + Q_t dW_t \\ p_T &= \nabla_x \Phi(\bar{X}_T) \end{cases}$$

We expect a solution with $p_t \in H$ and $Q_t \in L_2(H)$ but we notice that the term $\sum_{i=1}^{\infty} C_i(t) Q_t e_i$ does not converge for $Q_t \in L_2(H)$.

We can rewrite the above equation in the mild form

$$\begin{aligned} p_t &= e^{(T-t)A} \nabla_X \Phi(\bar{X}_T) + \int_t^T e^{(s-t)A} \nabla_X L(\bar{X}_s, \bar{u}_s) ds + \\ &+ \int_t^T \sum_{i=1}^{\infty} e^{(s-t)A} C_i(s) Q_s e_i ds + \int_t^T Q_s dW_s \end{aligned}$$

but still $\sum_{i=1}^{\infty} e^{(s-t)A} C_i(s) Q e_i$ doesn't converge if $Q \in L_2(H)$. Indeed if $V \in H$

$$\sum_{i=1}^{\infty} \langle e^{(s-t)A} C_i(s) Q e_i, V \rangle = \sum_{i=1}^{\infty} \langle Q e_i, C_i(s) e^{(s-t)A} V \rangle \leq |Q|_{L_2(H)} \left(\sum_{i=1}^{\infty} |C_i(s) e^{(s-t)A} V|^2 \right)^{1/2} \textbf{?}$$

On he contrary

$$\sum_{i=1}^{\infty} \langle e^{(s-t)A} C_i(s) Q e_i, V \rangle \leq \left(\sum_{i=1}^{\infty} |Q e_i| \right) \sup_{i \in \mathbb{N}} |C_i(s) e^{(s-t)A} V|$$

Easy Facts on Schatten - von Neumann classes

We denote by $L_2(H)$ the Hilbert space of Hilbert Schmidt operators $H \to H$ endowed with the scalar product $\langle L, M\rangle_2 = \sum_{i=1}^{\infty} \langle Le_i, Me_i\rangle_H$

Given $L \in L_2(H)$ there exists a sequence $(a_n^L)_{n\in\mathbb{N}} \in \ell_2$ and a couple of orthonormal bases $(e_n^L)_{n\in\mathbb{N}}$, $(f_n^L)_{n\in\mathbb{N}}$ in H such that

$$L = \sum_{n=1}^{\infty} a_n^L f_n^L \langle e_n^L, \cdot\rangle \text{ and } |L|_2 = \sum_n (a_n^L)^2.$$

If $t \to L_t$ is a L_2 valued process then the above objects can be selected with the same measurability properties as L.

Define $L_1(H) = \{L \in L_2(H) : |L|_1 < \infty\}$ where

$$|L|_1 := \sup\left\{\langle B, L\rangle_2 : B \in L_2(H),\, |B|_{\mathcal{L}(H)} \le 1\right\}$$

- If $B \in \mathcal{L}(H)$ and $L \in L_1(H)$ then LB, BL are in $L_1(H)$ moreover $|LB|_1 \le |L|_1 |B|_{\mathcal{L}(H)}$, $|BL|_1 \le |L|_1 |B|_{\mathcal{L}(H)}$
- If $L \in L_1(H)$ the trace $\mathrm{Tr}(L) := \sum_{i=1}^{\infty} \langle e_i, Le_i\rangle$ converges absolutely and its value is independent on the choice of the basis $(e_i)_{i\in\mathbb{N}}$
- $|L|_1 = \sum_{n=1}^{\infty} |a_n^L|$, $\mathrm{Tr}(L) = \sum_{n=1}^{\infty} a_n^L$ consequently $|\mathrm{Tr}(L)| \le |L|_1$

Coming back to our bad term if $Q \in L_1(H)$ there exist two ONB such that $Q = \sum_j a_j f_j \langle \cdot, g_j \rangle$, and

$$\sum_i \left| e^{(s-t)A} C_i(s) Q e_i \right| = \sum_i \left| \sum_j e^{(s-t)A} C_i(s) a_j f_j \langle e_i, g_j \rangle \right|$$

$$\leq \sum_j |a_j| \sum_i |e^{(s-t)A} C_i(s) f_j| \; |\langle e_i, g_j \rangle| \quad \text{by Cauchy}$$

$$\leq \sum_j |a_j| c (s-t)^{-\gamma} = c |Q|_{L_1} (s-t)^{-\gamma}.$$

Moreover we formally compute $d_t \langle Y_t, p_t \rangle$ we get

$$\mathbb{E} \langle Y_T, \nabla_X \Phi(\bar{X}_T) \rangle + \mathbb{E} \int_0^T \langle Y_s, \nabla_X L(\bar{X}_s, \bar{u}_s) \rangle ds = \mathbb{E} \int_0^T \text{Tr} \left[\left(\nabla_u G(\bar{X}_s, \bar{u}_s)(\delta u)_s \right) Q_s \right] ds$$

- The multiplication operator $\nabla_u G(\bar{X}_s, \bar{u}_s)(\delta u)_s$ is at most bounded in H thus

  $$\text{Tr} \left[\left(\nabla_u G(\bar{X}_s, \bar{u}_s)(\delta u)_s \right) Q_s \right]$$

  is not well defined for $Q \in L_2(H)$ but is well defined for $Q \in L_1(H)$.

- We can not bypass the above term since it will remain in the final formulation of the maximum principle.

Conclusion The non-hilbertian space $L_1(H)$ has something to do here

Existence of a solution by approximations

Let us denote $\eta := \nabla_X \Phi(\bar{X}_T) \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}, H)$ and $f := \nabla_X L(\bar{X}, \bar{u}) \in L^2_W(\Omega \times [0,T], H)$.

Consider the approximating BSDEs

$$dp_t^N = e^{(T-t)A}\eta + \int_t^T e^{(s-t)A} f_s ds + \int_t^T \sum_{i=1}^{N} e^{(s-t)A} C_i(s) Q_s^N e_i ds + \int_t^T e^{(s-t)A} Q_s^N dW_s$$

By the standard theory (see [Hu-Peng 91]) there exists a unique solution with

$$\sup_{t\in[0,T]} E|p_t^N|_H^2 + E\int_0^T |Q_t^N|^2_{L_2(H)} dt \le c\left(\mathbb{E}|\eta|^2 + \mathbb{E}\int_0^T |f_t|^2 dt\right)$$

The idea is to exploit the duality relation with a forward equation in order to obtain good estimates in the L_1 norm.

First we show that the same duality relation easily implies weak convergence of the solutions of the sequence (p^N, Q^N)

The perturbed forward equation

Consider the perturbed (forward) equation

$$\begin{cases} d\widehat{Y}_t^{\Gamma,\xi} &= A\widehat{Y}_t^{\Gamma,\xi}dt + \left[\sum_{i=1}^{\infty} C_i(t)\widehat{Y}_t^{\Gamma,\xi} + \Gamma_t\right] dW_t \\ \widehat{Y}_s^{\Gamma,\xi} &= \xi \in L^2(\Omega, \mathcal{F}_s, \mathbb{P}, H) \end{cases}$$

Proposition 1 *Given* $\Gamma \in L^2_W(\Omega \times [0,T], L_2(H))$ *the above equation admits a unique mild solution that verifies*

$$\mathbb{E}|\widehat{Y}_t^{\Gamma,\xi}|^2 \le c\mathbb{E}\int_s^t |\Gamma_\ell|^2_{L_2(H)} d\ell + \mathbb{E}\xi^2$$

$$\mathbb{E}|\widehat{Y}_t^{\Gamma,\xi}|^2 \le c\mathbb{E}\int_s^t (t-\ell)^{-2\gamma}|\Gamma_\ell|^2_{\mathcal{L}(H)} d\ell + \mathbb{E}\xi^2$$

The same estimates hold (with independent constant) for the solutions $\widehat{Y}^{\Gamma,N}$ of the approximating equations

$$\begin{cases} d\widehat{Y}_t^{\Gamma,\xi,N} &= A\widehat{Y}_t^{\Gamma,\xi,N}dt + \left[\sum_{i=1}^{N} C_i(t)\widehat{Y}_t^{\Gamma,\xi,N} + \Gamma_t\right] dW_t \\ \widehat{Y}_s^{\Gamma,\xi,N} &= \xi \in L^2(\Omega, \mathcal{F}_s, H) \end{cases}$$

Moreover $\mathbb{E}\int_s^T |\widehat{Y}_t^{\Gamma,\xi,N} - \widehat{Y}_t^{\Gamma,\xi}|^2 dt \to 0$ and $\mathbb{E}|\widehat{Y}_t^{\Gamma,\xi,N} - \widehat{Y}_t^{\Gamma,\xi}|^2 \to 0$ for all $t \in [s,T]$

If we compute (by introducing Yosida approximations of A) $d_t\langle \widehat{Y}_t^{\Gamma,N}, p_t^N\rangle$ we obtain

$$\mathbb{E}\int_s^T \langle \Gamma_t, Q_t^N\rangle_2 dt + \mathbb{E}\langle \xi, p_s^N\rangle_2 = \mathbb{E}\int_s^T \langle f_t, \widehat{Y}_t^{\Gamma,\xi,N}\rangle dt + \mathbb{E}\langle \eta, \widehat{Y}_T^{\Gamma,\xi,N}\rangle_H$$

and taking into account the convergence of $\widehat{Y}^{\Gamma,\xi,N}$ towards $\widehat{Y}^{\Gamma,\xi}$

Corollary 2 *There exists a couple of adapted processes $Q \in L^2(\Omega\times[0,T], L_2(H))$, $p \in \mathcal{C}([0,T], L^2(\Omega, H))$, such that*

$$Q^N \rightharpoonup Q \text{ in } L^2(\Omega \times [0,T], L_2(H)) \qquad p_t^N \rightharpoonup p_t \text{ in } L^2(\Omega, H)\ \forall t \in [0,T]$$

Moreover since for all $t \in [0,T]$ the stochastic integral $\int_t^T e^{(s-t)A} Q_s^N dW_s$ converges weakly to $\int_t^T e^{(s-t)A} Q_s dW_s$ we immediately deduce that, by difference, for all $t \in [0,T]$ there exists Ξ_t in $L^2(\Omega, \mathcal{F}_t, H)$ such that

$$\sum_{i=1}^N \int_t^T e^{(s-t)A} C_i(s) Q_s^N e_i ds \rightharpoonup \Xi_t \text{ weakly in } L^2(\Omega, \mathcal{F}_T, H)$$

The mild BSDE for the couple (p, Q) at this point reeds:

$$p_t = \Xi_t + e^{(T-t)A}\eta + \int_t^T e^{(s-t)A} f_s ds + \int_t^T e^{(s-t)A} Q_s dW_s$$

The above is not satisfactory in the sense that we want at least to obtain a mild BSDE. To start with we notice that the convergence of the bad term holds for the limit process Q itself namely

Proposition 3

$$\sum_{i=1}^{N}\int_t^T e^{(s-t)A}C_i(s)Q_s e_i ds \rightharpoonup \Xi_t \qquad \textit{in } L^2(\Omega,\mathcal{F}_T,\mathbb{P},H)$$

Proof: Given $\xi \in L^2(\Omega,\mathcal{F}_T,\mathbb{P},H)$

$$\begin{aligned}\mathbb{E}\langle\sum_{i=1}^{N}\int_t^T e^{(s-t)A}C_i(s)Q_s e_i ds,\xi\rangle &= \sum_{i=1}^{N}\mathbb{E}\int_t^T\langle Q_s e_i, C_i(s)e^{(s-t)A}\mathbb{E}(\xi|\mathcal{F}_s)\rangle ds \\ &= \lim_{M\to\infty}\sum_{i=1}^{N}\mathbb{E}\int_t^T\langle Q_s^M e_i, C_i(s)e^{(s-t)A}\mathbb{E}(\xi|\mathcal{F}_s)\rangle ds\end{aligned}$$

If $\gamma_i(s) = C_i(s)e^{(s-t)A}\mathbb{E}(\xi|\mathcal{F}_s)$ and $\widehat{Y}^{M,N}$ is the solution of the forward mild SDE

$$\widehat{Y}_\zeta^{M,N} = \sum_{i=1}^{M}\int_t^\zeta e^{(s-t)A}C_i(s)\widehat{Y}_s^{M,N}d\beta_s^i + \sum_{i=1}^{N}\int_t^\zeta e^{(s-t)A}\gamma_i(s)d\beta_s^i$$

then

$$\mathbb{E}\langle\sum_{i=1}^{N}\int_t^T e^{(s-t)A}C_i(s)Q_s^M e_i ds,\xi\rangle = \mathbb{E}\langle\eta,\widehat{Y}_\zeta^{M,N}\rangle + \mathbb{E}\int_t^T\mathbb{E}\langle f_\zeta,\widehat{Y}_\zeta^{M,N}\rangle d\zeta$$

Noticing that for all $\rho > t$

$$\sum_{i=1}^{\infty} \mathbb{E} \int_t^{\rho} |e^{(\rho-s)A}\gamma_i(s)ds|^2 = \sum_{i=1}^{\infty} \mathbb{E} \int_t^{\rho} |e^{(\rho-s)A}C_i(s)e^{(s-t)A}\mathbb{E}(\xi|\mathcal{F}_s)ds|^2 \leq c\mathbb{E}|\xi|^2$$

we can show that, for all fixed $t \in [0, T]$, as $N, M \to \infty$

$$\mathbb{E}|\widehat{Y}_\zeta^{M,N} - \widehat{Y}_\zeta^{\infty}|^2 \to 0, \qquad \int_t^T \mathbb{E}|\widehat{Y}_\zeta^{M,N} - \widehat{Y}_\zeta^{\infty}|^2 d\zeta \to 0$$

where $\widehat{Y}^\infty$ is the mild solution of the forward SDE

$$\begin{cases} d\widehat{Y}_t^{\infty} &= A\widehat{Y}_\zeta^{\infty}d\zeta + \sum_{i=1}^{\infty} C_i(\zeta)\widehat{Y}_\zeta^{M,N}d\beta_\zeta^i + \sum_{i=1}^{\infty}\gamma_i(\zeta)d\beta_\zeta^i \\ \widehat{Y}_t^{\infty} &= 0 \end{cases}$$

In conclusion

$$\mathbb{E}\langle\sum_{i=1}^{N}\int_t^T e^{(s-t)A}C_i(s)Q_s e_i ds, \xi\rangle \to \mathbb{E}\langle\eta, \widehat{Y}_t^{\infty}\rangle + \mathbb{E}\int_t^T \mathbb{E}\langle f_\zeta, \widehat{Y}_\zeta^{\infty}\rangle d\zeta$$

In an identical way we can show that

$$\mathbb{E}\langle\sum_{i=1}^{N}\int_t^T e^{(s-t)A}C_i(s)Q_s^N e_i ds, \xi\rangle \to \mathbb{E}\langle\eta, \widehat{Y}_t^{\infty}\rangle + \mathbb{E}\int_t^T \mathbb{E}\langle f_\zeta, \widehat{Y}_\zeta^{\infty}\rangle d\zeta$$

and this concludes the proof $\square$

Estimates of Q in the L_1 norm

Proposition 4 $\mathbb{E}\int_0^T (T-s)^{2\gamma}|Q_s|^2_{L_1(H)}ds \le c\mathbb{E}|\eta|^2 + c\mathbb{E}\int_0^T |f_s|^2 ds$

Remembering the representation $Q_t^N = \sum_{n=1}^\infty a_n(t)f_n(t)\langle e_n(t),\cdot\rangle$ we choose

$$\Gamma_t^N := \alpha(t)\sum_{n=1}^N \text{sgn}(a_n(t))f_n(t)\langle e_n(t),\cdot\rangle \quad \text{with } \alpha : [0,T]\to\mathbb{R}.$$

We notice that $|\Gamma_t^N|_{\mathcal{L}(H)} \le \alpha(t)$ and that $\langle Q_t,\Gamma_t^N\rangle_{L_2(H)} = \alpha(t)\sum_{n=1}^N |a_n^Q(t)|$ so that:

$$\mathbb{E}\int_0^T |Q_t|_{L_1(H)}\alpha(t)dt = \mathbb{E}\int_0^T \sup_N \langle Q_t,\gamma_t^N\rangle_2 dt \le \sup_N \left[\mathbb{E}\langle \eta,\widehat{Y}_T^{\Gamma^N}\rangle_H + \int \langle f_s,\widehat{Y}_s^{\Gamma^N}\rangle_H ds\right]$$

where

$$\begin{cases} d\widehat{Y}_t^\Gamma &= A\widehat{Y}_t^\Gamma dt + \left[\sum_{i=1}^\infty C_i(t)\widehat{Y}_t^\Gamma + \Gamma_t\right]dW_t \\ \widehat{Y}_s^\Gamma &= 0 \end{cases}$$

recalling the estimate of $\widehat{Y}^\Gamma$ with respect to the $\mathcal{L}(H)$ norm of Γ we get.

$$E\int_0^T |Q_t|_1\alpha(t)dt \le c_{\eta,f}\left(\int_0^T (T-s)^{-2\gamma}\alpha^2(s)ds\right)^{1/2}$$

and the proof follows letting $\tilde{\alpha}(s) = (T-s)^{-\gamma}\alpha(s)$ and rewriting the above as

$$E\int_0^T |Q_t|_1(T-t)^\gamma\tilde{\alpha}(t)dt \le c_{\eta,f}|\tilde{\alpha}|_{L^2[0,T]}$$

Corollary 5 *The sequence* $\sum_{i=1}^{\infty}\int_t^T e^{(s-t)A}C_i(s)Q_s e_i ds$ *converges in* $L^1(\Omega,\mathbb{P},H)$ *and the BSDE is satisfied in proper sense that is for all* $t\in[0,T]$ *it holds* $\mathbb{P}$*-a.s.*

$$p_t = e^{(T-t)A}\eta + \int_t^T e^{(s-t)A} f_s ds + \sum_{i=1}^{\infty}\int_t^T e^{(s-t)A}C_i(s)Q_s e_i ds + \int_t^T e^{(s-t)A}Q_s dW_s$$

Proof: *Recall the estimate*

$$\sum_{i=1}^{\infty}\left|e^{(s-t)A}C_i(s)Qe_i\right| \le C|Q|_{L_1}(s-t)^{-\gamma}.$$

Then for all N

$$\mathbb{E}\sum_{i=1}^{N}\left|\int_t^T e^{(s-t)A}C_i(s)Q_s e_i ds\right| \le c\mathbb{E}\int_t^T |Q_s|_{L_1}(s-t)^{-\gamma}ds$$

$$\le \left(\mathbb{E}\int_t^T |Q_s|^2_{L_1}(T-s)^{2\gamma}ds\right)^{1/2}\left(\int_t^T (T-s)^{-2\gamma}(s-t)^{-2\gamma}ds\right)^{1/2}$$

and the claim follows since this last integral converges.

Conclusion

Passing to the limit the duality relation holding for (p^N, Q^N) we get (recalling the expansion of the cost)

$$J(x,u^\epsilon) - J(x,\bar{u}) \;=\; \epsilon\,\mathbb{E}\int_0^T \langle (\delta u)_s, \left[\nabla_u F(\bar{X}_s,\bar{u}_s)\right]^* p_s\rangle ds + \epsilon\mathbb{E}\int_0^T \langle \nabla_u L(\bar{X}_s,\bar{u}_s), (\delta u)_s\rangle ds$$
$$+\epsilon\mathbb{E}\int_0^T \mathrm{Tr}\left[\left(\nabla_u G(\bar{X}_s,\bar{u}_s)(\delta u)_s\right) Q_s\right] ds + o(\epsilon)$$

And we now know that all the terms in the above formula are well defined.

Recall the we are assuming that $|\nabla_u G(\bar{X}_s,\bar{u}_s)v_s|_{\mathcal{L}(H)} \leq$ cost and we have just proved that $Q \in L_1(H)$, $\mathbb{P}\otimes dt$, a.s..

So we con conclude (by the usual localization - Lebesgue differentiation procedure) that $\forall v \in U$ it holds $\mathbb{P}\otimes dt$ a.s.

$$\langle \nabla_u L(\bar{X}_s,\bar{u}_s), v - \bar{u}_s\rangle + \mathrm{Tr}\left[\left(\nabla_u G(\bar{X}_s,\bar{u}_s)v_s\right) Q_s\right] \geq 0$$

Uniqueness of the mild BSDE ?

PART II: FINITE DIMENSIONAL NOISE

Formulation of the optimal control problem

Let $(\beta_t^1, \dots, \beta_t^m)$, $t \geq 0$, be a standard m-dimensional Wiener process. $(\mathcal{F}_t)_{t\geq 0}$ denotes its natural (completed) filtration.

The set of control actions U is a separable metric space not necessarily convex. A control u is a process in U.

$\mathcal{O} \subset \mathbb{R}^n$ is a bounded open set with regular boundary. The controlled state equation is an SPDE of the following semi abstract form: for $t \in [0, T]$, $x \in \mathcal{O}$,

$$\begin{cases} dX_t(x) &= AX_t(x)\, dt + b(x, X_t(x), u_t)\, dt + \sum_{j=1}^{m} \sigma_j(x, X_t(x), u_t)\, d\beta_t^j, \\ X_0(x) &= x_0(x), \end{cases}$$

where

$b(x, r, u), \sigma_j(x, r, u) : \mathcal{O} \times \mathbb{R} \times U \to \mathbb{R}$ are given (all difficulties are already present if b and σ_j are very regular in r and independent on x).

$H = L^2(\mathcal{O})$ is the state space, with usual scalar product $\langle \cdot, \cdot \rangle$.
We assume $x_0 \in H$. The solution X_t, $t \in [0, T]$, will be a process in H.

A is the realization of a partial differential operator, with appropriate boundary conditions.

Standing assumptions

1) *Regular coefficients*

The functions $b(x,r,u), \sigma_j(x,r,u), l(x,r,u), h(x,r)$ are measurable and
a) continuous in u;
b) of class C^2 in $r \in \mathbb{R}$;
c) bounded together with their first and second derivative w.r.t. r,

2) *L^p-boundedness of the semigroup*

A is a generator of a strongly continuous semigroup e^{tA}, $t \geq 0$, in $H = L^2(\mathcal{O})$. Moreover, for every $p \in [2, \infty)$ and $t \in [0, T]$,

$$e^{tA}(L^p(\mathcal{O})) \subset L^p(\mathcal{O}), \qquad \|e^{tA} f\|_{L^p(\mathcal{O})} \leq C_p \|f\|_{L^p(\mathcal{O})}$$

for some constants C_p independent of t and f.

3) *Compactness in L^4 of the semigroup*

the restriction of e^{tA}, $t \geq 0$, to $L^4(\mathcal{O})$ is an analytic semigroup with domain of the infinitesimal generator compactly imbedded in $L^4(\mathcal{O})$.

Statement of the stochastic maximum principle

For $u \in U$ and $X, p, q_1, \ldots, q_m \in H = L^2(\mathcal{O})$ denote

$$\mathcal{H}(u, X, p, q_1, \ldots, q_m) = \int_{\mathcal{O}} \left[l(x, X(x), u) + b(x, X(x), u)p(x) + \sigma_j(x, X(x), u) q_j(x) \right] dx$$

Theorem. *Let $(\bar{X}_t, \bar{u}_t)$ be an optimal pair. Then there are (suitably characterized):*

1) $(m+1)$ $L^2(\mathcal{O})$-valued adapted processes $p_t, q_{1t}, \ldots, q_{mt}$, $t \in [0, T]$,
2) one operator-valued process P_t, $t \in [0, T]$;

for which the following inequality holds $\mathbb{P}$-a.s. for a.e. $t \in [0, T]$:
for every $v \in U$,

$$\mathcal{H}(v, \bar{X}_t, p_t, q_{1t}, \ldots, q_{mt}) - \mathcal{H}(\bar{u}_t, \bar{X}_t, p_t, q_{1t}, \ldots, q_{mt}) \\ + \frac{1}{2} \langle P_t[\sigma_j(\cdot, \bar{X}_t(\cdot), v) - \sigma_j(\cdot, \bar{X}_t(\cdot), \bar{u}_t)], \sigma_j(\cdot, \bar{X}_t(\cdot), v) - \sigma_j(\cdot, \bar{X}_t(\cdot), \bar{u}_t) \rangle \geq 0.$$

The first adjoint processes p_t, q_{jt} are characterized as the unique solutions in H of an appropriate BSPDE and satisfy

$$\sup_{t \in [0,T]} \mathbb{E} \|p_t\|_H^2 + \mathbb{E} \int_0^T \|q_{jt}\|_H^2 \, dt < \infty.$$

The second adjoint process P_t takes values in the space of linear bounded operators $L^4(\mathcal{O}) \to L^4(\mathcal{O})^* = L^{4/3}(\mathcal{O})$ and also admits a suitable unique characterization.

Preliminaries to the proof of the maximum principle

Let $(\bar{X}, \bar{u})$ be an optimal pair. We introduce the spike variation:
we fix an arbitrary interval $[\bar{t}, \bar{t}+\epsilon] \subset (0, T)$ and an arbitrary $v \in U$ and define

$$u_t^\epsilon = \begin{cases} \bar{u}_t & \text{if } t \notin [\bar{t}, \bar{t}+\epsilon], \\ \\ v & \text{if } t \in [\bar{t}, \bar{t}+\epsilon]. \end{cases}$$

Let

$$\begin{array}{rcl} \delta l_t(x) & = & l(x, \bar{X}_t(x), u_t^\epsilon) - l(x, \bar{X}_t(x), \bar{u}_t) \\ \delta b_t(x) & = & b(x, \bar{X}_t(x), u_t^\epsilon) - b(x, \bar{X}_t(x), \bar{u}_t) \\ \delta \sigma_{jt}(x) & = & \sigma_j(x, \bar{X}_t(x), u_t^\epsilon) - \sigma_j(x, \bar{X}_t(x), \bar{u}_t) \\ \delta b_t'(x) & = & b'(x, \bar{X}_t(x), u_t^\epsilon) - b'(x, \bar{X}_t(x), \bar{u}_t) \\ \delta \sigma_{jt}'(x) & = & \sigma_j'(x, \bar{X}_t(x), u_t^\epsilon) - \sigma_j'(x, \bar{X}_t(x), \bar{u}_t) \end{array}$$

Let $(\bar{X}, \bar{u})$ be an optimal pair, u_t^ϵ the spike variation, and X_t^ϵ the corr. solution:

$$\begin{cases} dX_t^\epsilon(x) = AX_t^\epsilon(x)\, dt + b(x, X_t^\epsilon(x), u_t^\epsilon)\, dt + \sigma_j(x, X_t^\epsilon(x), u_t^\epsilon)\, d\beta_t^j, \\ X_0^\epsilon(x) = x_0(x) \end{cases}$$

We wish to represent in the form

$$X_t^\epsilon = \bar{X}_t + Y_t^\epsilon + Z_t^\epsilon + \text{remainder term}$$

where the remainder has to be $o(\epsilon)$.

Equation for Y_t^ϵ (to be understood in a mild sense):

$$\left\{\begin{array}{lcl} dY_t^\epsilon(x) & = & \left[AY_t^\epsilon(x) + b'(x,\bar{X}_t(x),\bar{u}_t)\cdot Y_t^\epsilon(x)\right]dt \\ & & +\sigma_j'(x,\bar{X}_t(x),\bar{u}_t)\cdot Y_t^\epsilon(x)\,d\beta_t^j + \delta b_t(x)\,dt + \delta\sigma_{jt}(x)\,d\beta_t^j \\ Y_0^\epsilon(x) & = & 0 \end{array}\right.$$

Equation for Z_t^ϵ (to be understood in a mild sense):

$$\left\{\begin{array}{lcl} dZ_t^\epsilon(x) & = & \left[AZ_t^\epsilon(x) + b'(x,\bar{X}_t(x),\bar{u}_t)\cdot Z_t^\epsilon(x)\right]dt + \sigma_j'(x,\bar{X}_t(x),\bar{u}_t)\cdot Z_t^\epsilon(x)\,d\beta_t^j \\ & & +\left[\frac{1}{2}b''(x,\bar{X}_t(x),\bar{u}_t)\cdot Y_t^\epsilon(x)^2 + \delta b_t'(x)\cdot Y_t^\epsilon(x)\right]dt \\ & & +\left[\frac{1}{2}\sigma_j''(x,\bar{X}_t(x),\bar{u}_t)\cdot Y_t^\epsilon(x)^2 + \delta\sigma_{jt}'(x)\cdot Y_t^\epsilon(x)\right]d\beta_t^j \\ Z_0^\epsilon(x) & = & 0 \end{array}\right.$$

Proposition. For all $p \geq 2$,

$$\sup_{t\in[0,T]}\left(\mathbb{E}\|Y_t^\epsilon\|^p_{L^p(\mathcal{O})}\right)^{1/p} = \sup_{t\in[0,T]}\left(\mathbb{E}\int_{\mathcal{O}}|Y_t^\epsilon(x)|^p dx\right)^{1/p} \leq C_p\sqrt{\epsilon}.$$

$$\sup_{t\in[0,T]}\left(\mathbb{E}\|Z_t^\epsilon\|^p_{L^p(\mathcal{O})}\right)^{1/p} = \sup_{t\in[0,T]}\left(\mathbb{E}\int_{\mathcal{O}}|Z_t^\epsilon(x)|^p dx\right)^{1/p} \leq C_p\,\epsilon.$$

$$\sup_{t\in[0,T]}\left(\mathbb{E}\|X_t^\epsilon - \bar{X}_t - Y_t^\epsilon - Z_t^\epsilon\|_H^2\right)^{1/2}$$
$$= \sup_{t\in[0,T]}\left(\mathbb{E}\int_{\mathcal{O}}|X_t^\epsilon(x) - \bar{X}_t(x) - Y_t^\epsilon(x) - Z_t^\epsilon(x)|^2 dx\right)^{1/2} = o(\epsilon).$$

Expansion of the cost functional

Let $(\bar{X}, \bar{u})$ be an optimal pair for the cost

$$J(u) = \mathbb{E} \int_0^T \int_{\mathcal{O}} l(x, X_t(x), u_t)\, dx\, dt + \mathbb{E} \int_{\mathcal{O}} h(x, X_T(x))\, dx$$

Let u_t^ϵ be the spike variation, and $J(u^\epsilon)$ the corresponding cost. Then clearly

$$J(u^\epsilon) - J(\bar{u}) \geq 0.$$

Recall

$$\delta l_t(x) = l(x, \bar{X}_t(x), u_t^\epsilon) - l(x, \bar{X}_t(x), \bar{u}_t)$$

Proposition. We have

$$0 \leq J(u^\epsilon) - J(\bar{u}) = \mathbb{E} \int_0^T \int_{\mathcal{O}} \delta l_t(x)\, dx\, dt + \Delta_1^\epsilon + \Delta_2^\epsilon + o(\epsilon),$$

where

$$\begin{aligned} \Delta_1^\epsilon &= \mathbb{E} \int_0^T \int_{\mathcal{O}} l'(x, \bar{X}_t(x), \bar{u}_t)(Y_t^\epsilon(x) + Z_t^\epsilon(x))\, dx\, dt \\ &\quad + \mathbb{E} \int_{\mathcal{O}} h'(x, \bar{X}_T(x))(Y_T^\epsilon(x) + Z_T^\epsilon(x))\, dx, \\ \Delta_2^\epsilon &= \frac{1}{2} \mathbb{E} \int_0^T \int_{\mathcal{O}} l''(x, \bar{X}_t(x), \bar{u}_t)\, Y_t^\epsilon(x)^2\, dx\, dt + \frac{1}{2} \mathbb{E} \int_{\mathcal{O}} h''(x, \bar{X}_T(x))\, Y_T^\epsilon(x)^2\, dx. \end{aligned}$$

The first adjoint processes

Consider the backward SPDE

$$\begin{cases} -dp_t(x) & = & -q_{jt}(x)\,d\beta_t^j + \left[A^* p_t(x) + b'(x, \bar{X}_t(x), \bar{u}_t) \cdot p_t(x) \right. \\ & & \left. +\sigma_j'(x, \bar{X}_t(x), \bar{u}_t) \cdot q_{jt}(x) + l'(x, \bar{X}_t(x), \bar{u}_t)\right] dt \\ p_T(x) & = & h'(x, \bar{X}_T(x)) \end{cases}$$

By Hu-Peng, Stoch Anal Appl ('91) there exists of a unique $(m+1)$-uple of adapted processes $(p, q_1, ..., q_m)$ solving the above in a mild sense and verifying

$$\sup_{t \in [0,T]} \mathbb{E} \int_{\mathcal{O}} |p_t(x)|_H^2 dx + \mathbb{E} \int_0^T \int_{\mathcal{O}} |q_{jt}(x)|_H^2 dx\, dt < \infty$$

Computing $d\int_{\mathcal{O}} Y_t^\epsilon(x) p_t(x)\,dx$, $d\int_{\mathcal{O}} Z_t^\epsilon(x) p_t(x)\,dx$, and joining what one obtains with the expression for Δ_2^ϵ we get

$$0 \le J(u^\epsilon) - J(u) = \mathbb{E} \int_0^T \int_{\mathcal{O}} \left[\delta l_t(x) + \delta b_t(x) p_t(x) + \delta\sigma_{jt}(x) q_{jt}(x)\right] dx\, dt + \frac{1}{2} \Delta_3^\epsilon + o(\epsilon),$$

where $\Delta_3^\epsilon =$

$$\mathbb{E}\int_0^T \int_{\mathcal{O}} \left[l''(x, \bar{X}_t(x), \bar{u}_t) + b''(x, \bar{X}_t(x), \bar{u}_t) p_t(x) + \sigma_j''(x, \bar{X}_t(x), \bar{u}_t) q_{jt}(x)\right] Y_t^\epsilon(x)^2 dx dt$$
$$+\mathbb{E}\int_{\mathcal{O}} h''(x, \bar{X}_T(x)) Y_T^\epsilon(x)^2 dx.$$

The second adjoint processes

Consider again

$$\Delta_3^\epsilon = \mathbb{E}\int_0^T\int_{\mathcal{O}} \bar{H}_t(x)\, Y_t^\epsilon(x)^2 dx dt + \mathbb{E}\int_{\mathcal{O}} \bar{h}(x)\, Y_T^\epsilon(x)^2\, dx = \mathbb{E}\int_0^T \langle \bar{H}_t Y_t^\epsilon, Y_t^\epsilon\rangle\, dt + \mathbb{E}\langle \bar{h} Y_T^\epsilon, Y_T^\epsilon\rangle$$

where

$$\begin{array}{rcl} \bar{H}_t(x) & = & l''(x,\bar{X}_t(x),\bar{u}_t) + b''(x,\bar{X}_t(x),\bar{u}_t)p_t(x) + \sigma_j''(x,\bar{X}_t(x),\bar{u}_t)q_{jt}(x), \\ \bar{h}(x) & = & h''(x,\bar{X}_T(x)). \end{array}$$

Here and below, by $\bar{H}_t$ and $\bar{h}$ we denote multiplication operators by $\bar{H}_t(\cdot)$ and $\bar{h}(\cdot)$, acting on H:

$$\bar{H}_t : f(\cdot) \mapsto \bar{H}_t(\cdot) f(\cdot), \qquad \bar{h} : f(\cdot) \mapsto \bar{h}(\cdot) f(\cdot), \qquad f \in H = L^2(\mathcal{O}).$$

Note that

$$|\bar{h}(x)| \le C := \sup |h''| < \infty, \qquad \mathbb{E}\int_0^T \int_{\mathcal{O}} |\bar{H}_t(x)|^2 dx\, dt < \infty,$$

due to the occurrence of $q_{jt}(x)$, so

$$\bar{h}(\cdot) \in L^\infty(\mathcal{O}) \quad \mathbb{P}-a.s., \qquad \bar{H}_t(\cdot) \in L^2(\mathcal{O}), \quad \mathbb{P}\times dt - a.e.$$

In particular, $\bar{h}$ is bounded but $\bar{H}_t$ is not a (bounded) linear operator on H.

To finish our argument we have to compute $\lim_{\epsilon\to 0} \epsilon^{-1}\Delta_3^\epsilon$

Characterization of P

For fixed $t \in [0,T]$ and $f \in L^4$, we consider the equation

$$\begin{cases} d\mathcal{Y}_s^{t,f}(x) = A\mathcal{Y}_s^{t,f}(x)\,ds + b'(x,\bar{X}_s(x),\bar{u}_s)\mathcal{Y}_s^{t,f}(x)\,ds + \sigma_j'(x,\bar{X}_s(x),\bar{u}_s)\mathcal{Y}_s^{t,f}(x)\,dW_s^j, \\ \mathcal{Y}_t^{t,f}(x) \quad = f(x), \end{cases}$$

We define a progressive process $(P_t)_{t\in[0,T]}$ with values in the space of bounded linear operators $L^4 \to (L^4)^* = L^{4/3}$ setting for $t \in [0,T]$, $f, g \in L^4$

$$\langle P_t f, g\rangle = \mathbb{E}^{\mathcal{F}_t}\int_t^T \langle \bar{H}_s \mathcal{Y}_s^{t,f}, \mathcal{Y}_s^{t,g}\rangle\,ds + \mathbb{E}^{\mathcal{F}_t}\langle \bar{h}\mathcal{Y}_T^{t,f}, \mathcal{Y}_T^{t,g}\rangle, \qquad \mathbb{P}-a.s.$$

The process $(P_t)_{t\in[0,T]}$ enjoys the following properties

Boundedness $\sup_{t\in[0,T]} \mathbb{E}\|P_t\|_{\mathcal{L}}^2 < \infty$,

Continuity $\mathbb{E}|\langle (P_{t+\epsilon} - P_t)f, g\rangle| \to 0, \qquad$ as $\epsilon \to 0, \quad f, g \in L^4(\mathcal{O})$

Regularization For every $\eta \in (0, 1/4)$ there exists a constant C_η such that

$$\mathbb{E}\sup_{f,g} |\langle P_t(-A)^\eta f, (-A)^\eta g\rangle|^2 \le C_\eta (T-t)^{-4\eta}\mathbb{E}\left[\int_0^T \|\bar{H}_s\|_{L^2(\mathcal{O})}^2 ds + \|\bar{h}\|_{L^2(\mathcal{O})}^2\right].$$

where $D(-A)^\eta$ is the domain of the fractional power of A in $L^4(\mathcal{O})$ and the sup is taken over all $f, g \in D(-A)^\eta, \|f\|_{L^4(\mathcal{O})} \le 1, \|g\|_{L^4(\mathcal{O})} \le 1$.

Conclusion of the proof

By the Markov property and suitable estimates
(recalling that, for all $p \geq 1$, $\mathbb{E}\|Y_t^\epsilon\|_{L^p(\mathcal{O})}^{2p} \leq C_p\,\epsilon^p$ for all $t \in [0,T]$.)

$$\begin{aligned}\Delta_3^\epsilon = \mathbb{E}\int_0^T \langle \bar{H}_s Y_s^\epsilon, Y_s^\epsilon\rangle\,ds + \mathbb{E}\langle \bar{h} Y_T^\epsilon, Y_T^\epsilon\rangle &= \mathbb{E}\int_{t_0}^T \langle \bar{H}_s Y_s^\epsilon, Y_s^\epsilon\rangle\,ds + \mathbb{E}\langle \bar{h} Y_T^\epsilon, Y_T^\epsilon\rangle \\ &= o(\epsilon) + \mathbb{E}\int_{t_0+\epsilon}^T \langle \bar{H}_s \mathcal{Y}_s^{t_0+\epsilon, Y^\epsilon_{t_0+\epsilon}}, \mathcal{Y}_s^{t_0+\epsilon, Y^\epsilon_{t_0+\epsilon}}\rangle\,ds + \mathbb{E}\langle \bar{h}\mathcal{Y}_T^{t_0+\epsilon, Y^\epsilon_{t_0+\epsilon}}, \mathcal{Y}_T^{t_0+\epsilon, Y^\epsilon_{t_0+\epsilon}}\rangle \\ &= o(\epsilon) + \mathbb{E}\langle P_{t_0+\epsilon} Y^\epsilon_{t_0+\epsilon}, Y^\epsilon_{t_0+\epsilon}\rangle,\end{aligned}$$

The argument is then concluded by proving the two following two relations:

$$\mathbb{E}\langle (P_{t_0+\epsilon} - P_{t_0})\, Y^\epsilon_{t_0+\epsilon}, Y^\epsilon_{t_0+\epsilon}\rangle = o(\epsilon),$$

$$\mathbb{E}\langle P_{t_0} Y^\epsilon_{t_0+\epsilon}, Y^\epsilon_{t_0+\epsilon}\rangle = \mathbb{E}\int_{t_0}^{t_0+\epsilon} \langle P_s \delta^\epsilon \sigma_j(s,\cdot), \delta^\epsilon \sigma_j(s,\cdot)\rangle\,ds + o(\epsilon)$$

since in that case we obtain

$$\Delta_3^\epsilon = \mathbb{E}\int_{t_0}^{t_0+\epsilon} \langle P_s \delta^\epsilon \sigma_j(s,\cdot), \delta^\epsilon \sigma_j(s,\cdot)\rangle\,ds + o(\epsilon)$$