Some applications of the Stochastic Maximum Principle

Boualem Djehiche
KTH Royal Institute of Technology
Stockholm

Warwick, December 2, 2011

Outline

I. The pre-commitment case

II. The non-commitment case

Part I is based on

- Andersson, D. and BD (2010): A maximum principle for SDEs of mean-field type. Applied Math. and Optimization.
- Buckdahn, R., BD and Li, J. (2011): A general stochastic maximum principle for SDEs of mean-field type. Applied Math. and Optimization

Part II is based on

- Björk, T., and BD (2011?): Optimal control under non-commitment (in preparation).

  (Competing papers are by Y. Hu, H. Jin and X. Y. Zhou (2011) , and J. Yong (2010,2011))

Part I. The pre-commitment case

The dynamics of the controlled SDE of mean-field type (through the expected value) on $\mathbb{R}$ is

$$\begin{cases} dX(t) = b(t, X(t), \mathbb{E}[X(t)], u(t))dt + \sigma(t, X(t), \mathbb{E}[X(t)], u(t))dB(t). \\ X(0) = x_0, \end{cases} \tag{1}$$

The cost functional is also of mean-field type:

$$J(u) = \mathbb{E}\left[\int_0^T h\left(t, X(t), \mathbb{E}[X(t)], u(t)\right) dt + g\left(X(T), \mathbb{E}[X(T)]\right)\right]. \tag{2}$$

We want to "find" or characterize (through a Maximum Principle)

$$u^* = \arg \min_{u \in \mathcal{U}[0,T]} J(u). \tag{3}$$

For $0 \leq t \leq T$, $\mathcal{U}[t, T]$ is the class "admissible controls": measurable, adapted processes $u : [t, T] \times \Omega \longrightarrow U$ (non-convex in general) satisfying some integrability conditions.

Time inconsistent control problem

The fact that g is **nonlinear in** $\mathbb{E}(X_T)$ makes the problem time inconsistent.

The classical **Bellman optimality principle** based on the law of iterated expectations on J **does not hold**.

A classical example: Mean-variance portfolio selection

The dynamics of the self-financing portfolio is

$$\begin{cases} dX(t) = (\rho X(t) + (\alpha - \rho)\, u(t))\, dt + \sigma u(t) dB(t), \\ X(0) = x_0. \end{cases} \tag{4}$$

All the coefficients are constant.

The control $u(t)$ denotes the amount of money invested in the risky asset at time t.

The cost functional, to be minimized, is given by

$$J(u) = \frac{\gamma}{2} \mathrm{Var}(X(T)) - \mathbb{E}[X(T)]. \tag{5}$$

By rewriting it as

$$J(u) = \mathbb{E}\left(\frac{\gamma}{2} X^2(T) - X(T) \right) - \frac{\gamma}{2} \left(\mathbb{E}[X(T)]\right)^2,$$

it becomes nonlinear in $\mathbb{E}[X(T)]$.

The mean-field SDE is obtained as an L^2-limit of an interacting particle system of the form

$$dx^{i,n}(t) = b^{i,n}(t,\omega,u(t))dt + \sigma^{i,n}(t,\omega,u(t))dB^i_t,$$

when $n \to \infty$, where, the B^i's are independent Brownian motions, and

$$b^{i,n}(t,\omega,u(t)) := b\left(t, x^{i,n}(t), \tfrac{1}{n}\sum_{j=1}^n x^{j,n}(t), u(t)\right)$$
$$\sigma^{i,n}(t,\omega,u(t)) := \sigma\left(t, x^{i,n}(t), \tfrac{1}{n}\sum_{j=1}^n x^{j,n}(t), u(t)\right).$$

The classical example is the McKean-Vlasov model, in which the coefficients are linear in the law of the process. (see e.g. Sznitman (1989) and the references therein).

For the nonlinear case, see Jourdain, Mèlèard and Woyczynski (2008).

Extending the HJB equation to the mean-field case

(1) Ahmed and Ding (2001) express the value function in terms of the Nisio semigroup of operators and derive a (very complicated) HJB equation.

(2) Huang *et al.* (2006) use the Nash Certainty Equivalence Principle to solve an extended HJB equation.

(3) Lasry and Lions (2007) suggest a new class of nonlinear HJB involving the dynamics of the probability laws $(\mu_t)_t$.

(4) Björk and Murgoci (2008), Björk, Murgoci and Zhou(2011) use the notion of Nash equilibrium to transform the time inconsistent control problem into a standard one and derive an "extended" HJB equation.

Assumptions

- The action space U is a subset of $\mathbb{R}$ (not necessarily **convex**!).
- All the involved functions are sufficiently smooth:
  - b, σ, g, h are twice continuously differentiable with respect to (x, y).
  - b, σ, g, h and all their derivatives with respect to (x, y) are continuous in (x, y, v), and bounded.

We let $\hat{u}$ denote an optimal control, and $\hat{x}$ the corresponding state process. Also, denote

$$\delta\varphi(t) = \varphi(t, \hat{x}(t), E[\hat{x}(t)], u(t)) - \varphi(t, \hat{x}(t), E[\hat{x}(t)], \hat{u}(t)),$$

$$\varphi_x(t) = \frac{\partial \varphi}{\partial x}(t, \hat{x}(t), \mathbb{E}[\hat{x}(t)], \hat{u}(t)),$$

and similarly for higher derivatives.

The Stochastic Maximum Principle Approach

Following Peng (1990), we derive the variational inequality (13) below, from the fact that

$$J(u^{\epsilon}(\cdot)) - J(\hat{u}(\cdot)) \geq 0,$$

where, $u^{\epsilon}(\cdot)$ is the so-called spike variation of $\hat{u}(\cdot)$, defined as follows.
For $\epsilon > 0$, pick a subset $E_{\epsilon} \subset [0, T]$ such that $|E_{\epsilon}| = \epsilon$ and consider the control process (spike variation of u)

$$u^{\epsilon}(t) := \begin{cases} u(t), & t \in E_{\epsilon}, \\ \hat{u}(t), & t \in E_{\epsilon}^{c}, \end{cases}$$

where, $u(\cdot) \in \mathcal{U}$ is an arbitrary admissible control.
Denote $x^{\epsilon}(\cdot) := x^{u^{\epsilon}}(\cdot)$ the corresponding state process which satisfies (1).

The key relation between perfromance J and the Hamiltonian H is

$$J(u^{\epsilon}) - J(\hat{u}) = -\mathbb{E}\left[\int_0^T \left(\delta H(t) + \frac{1}{2}P(t)(\delta\sigma(t))^2\right) 1\!\!1_{E_\epsilon}(t)\, dt\right] + R(\epsilon), \quad (6)$$

where,

$$|R(\epsilon)| \le \epsilon\bar{\rho}(\epsilon),$$

for some function $\bar{\rho} : (0,\infty) \to (0,\infty)$ such that $\bar{\rho}(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$.

This is a **finer estimate** than the standard one related to the original Peng's Stochastic Maximum Principle.

Associated Hamiltonian

The Hamiltonian associated with the r.v. X:

$$H(t,X,u,p,q) := b(t,X,\mathbb{E}[X],u)p + \sigma(t,X,\mathbb{E}[X],u)q + h(t,X,\mathbb{E}[X],u); \quad (7)$$

Denote

$$\begin{aligned} \delta H(t) &:= p(t)\delta b(t) + q(t)\delta\sigma(t) + \delta h(t), \\ H_x(t) &= b_x(t)p + \sigma_x(t)q + h_x(t), \\ H_{xx}(t) &= b_{xx}(t)p + \sigma_{xx}(t)q + h_{xx}(t). \end{aligned} \quad (8)$$

Adjoint equations

(a) The first-order adjoint equation is of mean-field type:

$$\begin{cases} dp(t) = & -\left(b_x(t)p(t) + \sigma_x(t)q(t) + h_x(t)\right) dt + q(t)dB(t) \\ & -\left(\mathbb{E}[b_y(t)p(t)] + \mathbb{E}[\sigma_y(t)q(t)] + \mathbb{E}[h_y(t)]\right) dt, \\ p(T) = & g_x(T) + \mathbb{E}[g_y(T)]. \end{cases} \tag{9}$$

Under our assumptions, this is a linear mean-field backward SDE with bounded coefficients. It has a unique adapted solution such that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |p(t)|^2 + \int_0^T |q(t)|^2\, dt\right] < +\infty. \tag{10}$$

(See Buckdahn *et al.* (2009), Theorem 3.1)

(b) The second-order adjoint equation is "standard":

$$\begin{cases} dP(t) = -\left(2b_x(t)P(t) + \sigma_x^2(t)P(t) + 2\sigma_x(t)Q(t) + H_{xx}(t)\right) dt \\ \qquad\qquad +Q(t)\, dB(t), \\ \\ P(T) = g_{xx}(T). \end{cases} \tag{11}$$

This is a standard linear backward SDE, whose unique adapted solution (P, Q) satisfies the following estimate

$$E\left[\sup_{t\in[0,T]} |P(t)|^2 + \int_0^T |Q(t)|^2\, dt\right] < \infty. \tag{12}$$

Necessary Conditions for Optimality

Theorem. Assume the above assumption hold. If $(\hat{X}(\cdot), \hat{u}(\cdot))$ is an optimal solution of the control problem (1)-(3), then there are pairs of $\mathcal{F}$-adapted processes $(p(\cdot), q(\cdot))$ and $(P(\cdot), Q(\cdot))$ that satisfy (9)-(10) and (11)-(12), respectively, such that

$$
\begin{aligned}
&H(t,\hat{X}(t),u,p(t),q(t)) - H(t,\hat{X}(t),\hat{u}(t),p(t),q(t)) \\
&\quad + \tfrac{1}{2}P(t)\left(\sigma(t,\hat{X}(t),\mathbb{E}[\hat{X}(t)],u) - \sigma(t,\hat{X}(t),\mathbb{E}[\hat{X}(t)],\hat{u}(t))\right)^2 \le 0, \\
&\quad \forall u \in U, \;\; a.e.\ t \in [0,T], \;\; \mathbb{P}-a.s.
\end{aligned} \tag{13}
$$

Sufficient Conditions for Optimality

Assuming convexity of the action space U and the coefficients, Condition (13) is also sufficient (without the third term).

In this case Condition (13) is equivalent to

$$\partial_u H(t, \hat{X}(t), \hat{u}(t), p(t), q(t)) = 0, \qquad 0 \leq t \leq T, \quad \mathbb{P}-a.s. \tag{14}$$

A worked out example- Mean-variance portfolio selection

The state process equation is

$$dX(t) = (\rho X(t) + (\alpha - \rho)\, u(t))\, dt + \sigma u(t) dB(t), \quad X(0) = x_0. \tag{15}$$

The cost functional, to be minimized, is given by

$$J(u) = \mathbb{E}\left(\frac{\gamma}{2} X(T)^2 - X(T)\right) - \frac{\gamma}{2}\left(\mathbb{E}[X(T)]\right)^2,$$

The Hamiltonian for this system is

$$H(t, x, u, p, q) = (\rho x + (\alpha - \rho)\, u)\, p + \sigma u q.$$

The adjoint equation becomes

$$\begin{cases} dp(t) & = -\rho p(t)dt + q(t)dB_t, \\ p(T) & = \gamma(X(T) - \mathbb{E}[X(T)]) - 1, \end{cases}$$

Try a solution of the form

$$p_t = A_t(X(t) - \mathbb{E}[X(t)]) - C_t,$$

with A_t, C_t deterministic functions such that

$$A_T = \gamma, \quad C_T = 1.$$

After easy manipulations, together with the first order condition for minimizing the Hamiltonian yielding

$$(\alpha - \rho)\, p(t) + \sigma q(t) = 0, \tag{16}$$

and

$$q(t) = A_t \sigma u(t), \tag{17}$$

we get

$$\begin{cases} (\rho-\alpha)^2 A_t - (2\rho A_t + A_t')\,\sigma^2 = 0, & A_T = \gamma, \\ \rho C_t + C_t' = 0, & C_T = 1. \end{cases}$$

The solutions to these equations are

$$\begin{cases} A_t = \gamma e^{(2\rho-\Lambda)(T-t)}, \\ C_t = e^{\rho(T-t)}, \end{cases} \tag{18}$$

where,

$$\Lambda = \frac{(\rho-\alpha)^2}{\sigma^2}.$$

The optimal control becomes

$$\hat{u}(t, \hat{X}(t)) = \frac{\alpha - \rho}{\sigma^2} \left(x_0 e^{\rho(T-t)} + \frac{1}{\gamma} e^{(\Lambda - \rho)(T-t)} - \hat{X}(t) \right), \qquad (19)$$

which is identical to the optimal control found in Zhou and Li (2000), obtained by embedding the problem into a stochastic LQ problem.

Part II. The non-commitment case

The dynamics of the controlled SDE is

$$\begin{cases} dX^{t,x}(s) = b(t,x,s,\omega)ds + \sigma(t,x,s,\omega)dB(s), \quad s > t, \\ X^{t,x}(t) = x, \end{cases} \tag{20}$$

where,

$$\begin{cases} b(t,s,u) := b(s, X^{t,x}(s), \mathbb{E}[X^{t,x}(s)], u(s)) \\ \\ \sigma(t,s,u) := \sigma(s, X^{t,x}(s), \mathbb{E}[X^{t,x}(s)], u(s)). \end{cases}$$

The cost functional is

$$J(t,x,u) = \mathbb{E}\left[\int_t^T h\left(s, X^{t,x}(s), \mathbb{E}[X^{t,x}(s)], u(s)\right) ds + g\left(X^{t,x}(T), \mathbb{E}[X^{t,x}(T)]\right)\right]. \tag{21}$$

We want to "find" or characterize (through a Maximum Principle) the t-optimal policy

$$u^*(t,x,\cdot) := \text{" arg} \min_{u \in \mathcal{U}[t,T]} \text{"} J(t,x,u). \tag{22}$$

Failure to remain optimal across time!

A key-observation made by Ekeland, Lazrak and Pirvu (2007)-(2008) is that:

Time inconsistent optimal solutions (although they exist mathematically) are irrelevant in practice, since the t-optimal policy may not be optimal after t:

$$u^*(t, x, \cdot) \neq \arg\min_u J(t', x', u)$$

The decision-maker would not implement the t-optimal policy at a later time, if he/she is not force to do so.

Game theoretic approach

Following Ekeland, Lazrak and Pirvu (2007)-(2008), and Björk and Murgoci (2008), we may view the problem as a game and look for a **Nash subgame perfect equilibrium point** $\hat{u}$ in the following sense:

- Assume that all players (selves) s, such that $s > t$, use the control $\hat{u}(s)$.
- Then it is optimal for player (self) t to also use $\hat{u}(t)$.

To characterize the equilibrium point $\hat{u}$, Ekeland *et al.* suggest the following definition that uses a "local" spike variation in a natural way:

- Fix (t, x) and define the control law u^{ϵ} as the "local" spike variation of $\hat{u}$ over the set $E_{t,\epsilon} := [t, t+\epsilon]$, (note that $|E_{t,\epsilon}| = \epsilon$),

$$u^{\epsilon}(s) := \begin{cases} u(s), & s \in E_{t,\epsilon}, \\ \hat{u}(s), & s \in [t, T] \setminus E_{t,\epsilon}, \end{cases}$$

where, $u(\cdot) \in \mathcal{U}$ is an arbitrary admissible control (or simply any real number).

Definition. The control law $\hat{u}$ is an **equilibrium point** if

$$\lim_{\epsilon \downarrow 0} \frac{J(t, x, u^{\epsilon}) - J(t, x, \hat{u})}{\epsilon} \geq 0, \tag{23}$$

for all choices of t, x and u.

A small adaptation of the results in Part I, by keeping track of the dependence on (t, x), yields the key relation between the perfromance J and the Hamiltonian H associated with (20):

$$J(t, x, u^{\epsilon}) - J(t, x, \hat{u}) = -\mathbb{E}\left[\int_t^{t+\epsilon} \delta H(t, s) + \frac{1}{2} P(t, s)(\delta\sigma(s))^2\, ds\right] + R(\epsilon), \tag{24}$$

where,

$$|R(\epsilon)| \leq \epsilon\bar{\rho}(\epsilon),$$

for some function $\bar{\rho} : (0, \infty) \to (0, \infty)$ such that $\bar{\rho}(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$.

$$\delta H(t, s) := H(t, s, \hat{X}^{t,x}(s), u, p, q) - H(t, s, \hat{X}^{t,x}(s), \hat{u}, p, q),$$

where, the Hamiltonian associated with (20) reads

$$H(t,s,X^{t,x}(s),u,p,q) := b(t,s,u)p + \sigma(t,s,u)q + h(t,s,u); \qquad (25)$$

where, for $\varphi = b, \sigma$ or h, we use the notation

$$\varphi(t,s,u) := \varphi(s, X^{t,x}(s), \mathbb{E}[X^{t,x}(s)], u(s)), \quad s \geq t.$$

In particular,

$$\varphi(t,t,u) := \varphi(t,x,x,u(t)),$$

and

$$\bar{H}(t,u,p,q) := H(t,t,x,u,p,q) = b(t,x,x,u)p + \sigma(t,x,x,u)q + h(t,x,x,u); \qquad (26)$$

The first-order adjoint equation is similar to the previous one:

$$\begin{cases} dp(t,s) = & -\left(b_x(t,s)p(t,s) + \sigma_x(t,s)q(t,s) + h_x(t,s)\right) ds + q(t,s)dB(s) \\ & -\left(\mathbb{E}[b_y(t,s)p(t,s)] + \mathbb{E}[\sigma_y(t,s)q(t,s)] + \mathbb{E}[h_y(t,s)]\right) ds, \\ p(t,T) = & g_x(t,T) + \mathbb{E}[g_y(t,T)], \end{cases} \tag{27}$$

which has a unique adapted solution such that

$$\mathbb{E}\left[\sup_{s\in[t,T]} |p(t,s)|^2 + \int_t^T |q(t,s)|^2 \, ds\right] < +\infty, \tag{28}$$

The second-order adjoint equation is "standard":

$$\begin{cases} dP(t,s) = -\left(2b_x(t,s)P(t,s) + \sigma_x^2(t,s)P(t,s) + 2\sigma_x(t,s)Q(t,s) + H_{xx}(t,s)\right) ds \\ \qquad\qquad +Q(t,s)\, dB(s), \\ \\ P(t,T) = g_{xx}(t,T), \end{cases} \tag{29}$$

which is a standard linear backward SDE, whose unique adapted solution (P,Q) satisfies the following estimate

$$E\left[\sup_{s\in[t,T]} |P(t,s)|^2 + \int_t^T |Q(t,s)|^2\, ds\right] < \infty. \tag{30}$$

Characterization of the equilibrium point- Necessary Conditions

Theorem. Assume the above assumption hold. If $(\hat{X}(\cdot), \hat{u}(\cdot))$ is an equilibrium solution of the problem (20)-(22), then there are pairs of $\mathcal{F}$-adapted processes $(p(\cdot), q(\cdot))$ and $(P(\cdot), Q(\cdot))$ that satisfy (27)-(28) and (29)-(30), respectively, such that

$$\bar{H}(t, u, p(t,t), q(t,t)) - \bar{H}(t, \hat{u}(t), p(t,t), q(t,t))$$
$$+\frac{1}{2} P(t,t) \left(\sigma(t,x,x,u) - \sigma(t,x,x,\hat{u}(t))\right)^2 \leq 0, \qquad (31)$$

$$\forall u \in U, \quad \mathbb{P}-a.s.$$

Characterization of the equilibrium point- Sufficient Conditions

Assuming convexity of the action space U and the coefficients, Condition (31) is also sufficient (without the third term).

In this case Condition (31) is equivalent to

$$\partial_u \bar{H}(t, \hat{u}(t), p(t,t), q(t,t)) = 0, \quad \mathbb{P} - a.s. \tag{32}$$

A worked out example- Mean-variance portfolio selection

The state process equation is

$$dX^{t,x}(s) = \left(\rho X^{t,x}(s) + (\alpha - \rho)\, u(s)\right) ds + \sigma u(s) dB(s), \;\; X^{t,x}(t) = x, \quad (33)$$

The cost functional, to be minimized, is given by

$$J(t,x,u) = \mathbb{E}\left(\frac{\gamma}{2}\left(X^{t,x}(T)\right)^2 - X^{t,x}(T)\right) - \frac{\gamma}{2}\left(\mathbb{E}[X^{t,x}(T)]\right)^2,$$

The Hamiltonian for this system is

$$H(t,s,X^{t,x}(s),u,p,q) := \left(\rho X^{t,x}(s) + (\alpha - \rho)\, u\right) p + \sigma u q;$$

Hence,

$$\bar{H}(t,u,p,q) = (\rho x + (\alpha - \rho)\, u)\, p + \sigma u q.$$

The adjoint equation becomes

$$\begin{cases} dp(t,s) &= -\rho p(t,s)ds + q(t,s)dB(s), \\ p(t,T) &= \gamma(\hat{X}^{t,x}(T) - \mathbb{E}[\hat{X}^{t,x}(T)]) - 1, \end{cases} \tag{34}$$

Try a solution of the form

$$p(t,s) = A_s(\hat{X}^{t,x}(s) - \mathbb{E}[\hat{X}^{t,x}(s)]) - C_s, \tag{35}$$

with A_s, C_s deterministic functions such that

$$A_T = \gamma, \quad C_T = 1.$$

Identifying the coefficients in (33) and (34), we get, for $s \geq t$,

$$\begin{array}{l}(2\rho A_s + A_s')(\hat{X}^{t,x}(s) - \mathbb{E}[\hat{X}^{t,x}(s)]) + (\alpha - \rho)(\hat{u}(s) - \mathbb{E}[\hat{u}(s)]) \\ \qquad = C_s' + \rho C_s, \end{array} \tag{36}$$

$$q(t,s) = A_s \sigma \hat{u}(s). \tag{37}$$

Now, the first order condition (32) for minimizing the Hamiltonian yields

$$(\alpha-\rho)p(t,t)+\sigma q(t,t)=0, \tag{38}$$

But, from (35), we have

$$p(t,t)=-C_t, \tag{39}$$

which is deterministic. Therefore, we get

$$q(t,t)=\frac{\rho-\alpha}{\sigma}C_t. \tag{40}$$

In view of 37, the value of the equilibrium point at time t is the **deterministic** function

$$\hat{u}(t)=\frac{\rho-\alpha}{\sigma^2}\frac{C_t}{A_t}. \tag{41}$$

Now, this suggests that at $s > t$, the player s, uses the **deterministic** function

$$\hat{u}(s) = \frac{(\rho - \alpha)}{\sigma^2} \frac{C_s}{A_s}. \tag{42}$$

(**this is the main point of the game theoretic approach!**)

Hence, (36) reduces to

$$(2\rho A_s + A_s')(\hat{X}^{t,x}(s) - \mathbb{E}[\hat{X}^{t,x}(s)]) = C_s' + \rho C_s, \tag{43}$$

suggesting that

$$\begin{cases} (2\rho A_s + A_s') = 0, & A_T = \gamma, \\ \rho C_s + C_s' = 0, & C_T = 1. \end{cases}$$

The solutions to these equations are

$$\begin{cases} A_t = \gamma e^{2\rho(T-t)}, \\ C_t = e^{\rho(T-t)}, \end{cases} \tag{44}$$

The equilibrium point is then

$$\hat{u}(t) = \frac{1}{\gamma} \frac{\rho - \alpha}{\sigma^2} e^{-\rho(T-t)}. \tag{45}$$

which is identical to the equilibrium found in Björk and Murgoci (2008), obtained by solving an extended HJB equation.

Appendix

Variational equations

Let $y^\epsilon(\cdot)$ and $z^\epsilon(\cdot)$ be respectively the solutions of the following SDEs:

$$\begin{cases} dy^\epsilon(t) = \{b_x(t)y^\epsilon(t) + b_y(t)E[y^\epsilon(t)] + \delta b(t)\mathbb{1}_{E_\epsilon}(t)\}\, dt \\ \qquad + \{\sigma_x(t)y^\epsilon(t) + \sigma_y(t)E[y^\epsilon(t)] + \delta\sigma(t)\mathbb{1}_{E_\epsilon}(t)\}\, dB(t), \\ y^\epsilon(0) = 0, \end{cases} \tag{46}$$

$$\begin{cases} dz^\epsilon(t) = \{b_x(t)z^\epsilon(t) + b_y(t)E[z^\epsilon(t)] + \mathcal{L}_t(b, y^\epsilon) + \delta b_x(t)y^\epsilon(t)\mathbb{1}_{E_\epsilon}(t)\}\, dt \\ \qquad + \{\sigma_x(t)z^\epsilon(t) + \sigma_y(t)E[z^\epsilon(t)] + \mathcal{L}_t(b, y^\epsilon) + \delta b_x(t)y^\epsilon(t)\mathbb{1}_{E_\epsilon}(t)\}\, dB(t), \\ z^\epsilon(0) = 0. \end{cases} \tag{47}$$

$$\mathcal{L}_t(\varphi, y) = \frac{1}{2}\varphi_{xx}(t, \hat{x}(t), E[\hat{x}(t)], \hat{u}(t))y^2.$$

Duality

Lemma. We have

$$E\left[p(T)y^{\epsilon}(T)\right] = E\left[\int_0^T y^{\epsilon}(t)\left(h_x(t)+E[h_y(t)]\right)dt\right] + E\left[\int_0^T \left(p(t)\delta b(t)+q(t)\delta\sigma(t)\right)I_{E_\epsilon}(t)dt\right], \tag{48}$$

and

$$E\left[p(T)z^{\epsilon}(T)\right] = E\left[\int_0^T z^{\epsilon}(t)(h_x(t)+E[h_y(t)])dt\right] + E\left[\int_0^T \left(p(t)\delta b_x(t)+q(t)\delta\sigma_x(t)\right)y^{\epsilon}(t)I_{E_\epsilon}(t)dt\right] + E\left[\int_0^T \left(p(t)\mathcal{L}(b,y^{\epsilon}(t))+q(t)\mathcal{L}_t(\sigma,y^{\epsilon})\right)dt\right]. \tag{49}$$

Taylor expansions and estimates

Let

$$dx^{\epsilon}(t) = b(t, x^{\epsilon}(t), E\left[x^{\epsilon}(t)\right], u_t^{\epsilon})dt + \sigma(t, x^{\epsilon}(t), E\left[x^{\epsilon}(t)\right], u_t^{\epsilon})dB(t),$$

$$d\hat{x}(t) = b(t, \hat{x}(t), E\left[\hat{x}(t)\right], \hat{u}_t)dt + \sigma(t, \hat{x}(t), E\left[\hat{x}(t)\right], \hat{u}_t)dB(t),$$

whith $x^{\epsilon}(0) = \hat{x}(0) = x_0$.

Proposition. For any $k \geq 1$,

$$E\left[\sup_{t\in[0,T]} \left|x^{\epsilon}(t) - \hat{x}(t)\right|^{2k}\right] \leq C_k \epsilon^k, \tag{50}$$

$$E\left[\sup_{t\in[0,T]} \left|y^{\epsilon}(t)\right|^{2k}\right] \leq C_k \epsilon^k, \tag{51}$$

$$E\left[\sup_{t\in[0,T]} |z^{\epsilon}(t)|^{2k}\right] \leq C_k \epsilon^{2k}, \tag{52}$$

$$E\left[\sup_{t\in[0,T]} |x^{\epsilon}(t) - (\hat{x}(t) + y^{\epsilon}(t))|^{2k}\right] \leq C_k \epsilon^{2k}, \tag{53}$$

$$E\left[\sup_{t\in[0,T]} |x^{\epsilon}(t) - (\hat{x}(t) + y^{\epsilon}(t) + z^{\epsilon}(t))|^{2k}\right] \leq C_k \epsilon^{2k} \rho_k(\epsilon), \tag{54}$$

where, $\rho_k : (0,\infty) \to (0,\infty)$ is such that $\rho_k(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$.

$$\sup_{t\in[0,T]} |E[y^{\epsilon}(t)]|^2 \leq \epsilon\rho(\epsilon), \;\; \epsilon > 0, \tag{55}$$

for some function $\rho : (0,\infty) \to (0,\infty)$ such that $\rho(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$.

Estimate (55) is derived as a consequence of the following result.

Lemma. For any progressively measurable process $(\Phi(t))_{t\in[0,T]}$ for which, for all $p \geq 1$, there exists a positive constant C_p, such that

$$E[\sup_{t\in[0,T]} |\Phi(t)|^p] \leq C_p, \tag{56}$$

there exists a function $\tilde{\rho} : (0,\infty) \to (0,\infty)$ with $\tilde{\rho}(\epsilon) \to 0$ as $\epsilon \downarrow 0$, such that

$$|E[\Phi(T)y^\epsilon(T)]|^2 + \int_0^T |E[\Phi(s)y^\epsilon(s)]|^2 \, ds \leq \epsilon\tilde{\rho}(\epsilon), \quad \epsilon > 0. \tag{57}$$