Necessary and Sufficient Conditions for Optimal Control of Semilinear SPDEs

Lukas Wessels

joint work with Wilhelm Stannat

Technische Universität Berlin

www.lukaswessels.org

BSDE Annecy, July 1, 2022

SFB 910

Berlin Mathematical School BMS

berlin

Outline

Setting

Minimize

$$J(u) := \mathbb{E}\left[\int_0^T \int_\Lambda l(x_t^u(\lambda), u_t)\mathrm{d}\lambda\mathrm{d}t + \int_\Lambda h(x_T^u(\lambda))\mathrm{d}\lambda\right]$$

subject to

$$\begin{cases} \mathrm{d}x_t^u = [\Delta x_t^u + b(x_t^u, u_t)]\mathrm{d}t + \sigma(x_t^u, u_t)\mathrm{d}W_t, \quad t \in [0, T] \\ x_0^u = x \in L^2(\Lambda), \end{cases}$$

where

- b, σ Nemytskii operators, Lipschitz;
- $(W_t)_{t\geq 0}$ cylindrical Wiener process;
- $\Lambda \subset \mathbb{R}$ bounded interval;
- l, h Nemytskii operators of (at most) quadratic growth;
- control domain U non-convex.

Goal: Derive necessary conditions for optimality.

Setting

Minimize

$$J(u) := \mathbb{E}\left[\int_0^T \int_\Lambda l(x_t^u(\lambda), u_t)\mathrm{d}\lambda\mathrm{d}t + \int_\Lambda h(x_T^u(\lambda))\mathrm{d}\lambda\right]$$

subject to

$$\begin{cases} \mathrm{d}x_t^u = [\Delta x_t^u + b(x_t^u, u_t)]\mathrm{d}t + \sigma(x_t^u, u_t)\mathrm{d}W_t, \quad t \in [0, T] \\ x_0^u = x \in L^2(\Lambda). \end{cases}$$

- Pontryagin's maximum principle (1956): Necessary optimality condition in finite-dimensional, deterministic case ($\sigma \equiv 0$).
- Peng's maximum principle (1990): Generalization to finite-dimensional stochastic case.
- Since then: Many generalizations to infinite-dimensional, stochastic case by Du, Fuhrman, Frankowska, Guatteri, Hu, Li, Lü, Meng, Tang, Tessitore, Zhang, ...
  Major drawback in previous works: Strong assumptions on coefficients l and h, in particular excluding quadratic costs.

Spike Variation

Let $\bar{u}$ be optimal. Fix $\tau \in [0, T)$, $\varepsilon > 0$, $v \in U$, and set

$$u_t^\varepsilon := \begin{cases} v, & \tau \leq t \leq \tau + \varepsilon, \\ \bar{u}_t & \text{otherwise.} \end{cases}$$

Then

$$0 \leq J(u^\varepsilon) - J(\bar{u}) = \mathbb{E}\left[\int_0^T \int_\Lambda l(x_t^\varepsilon, u_t^\varepsilon) - l(\bar{x}_t, \bar{u}_t)\mathrm{d}\lambda \mathrm{d}t + \int_\Lambda h(x_T^\varepsilon) - h(\bar{x}_T)\mathrm{d}\lambda\right].$$

Roadmap: 1) Taylor expand integrands;
2) Divide by ε, send $\varepsilon \to 0$;
3) Identify remaining terms.

Because of stochastic calculus, we have to Taylor expand up to second order.

Quadratic Terms

Taylor expanding the cost functional to second order leads to the quadratic terms

$$\mathbb{E}\left[\int_0^T \int_\Lambda l_{xx}(\bar{x}_t(\lambda), \bar{u}_t) y_t^\varepsilon(\lambda) y_t^\varepsilon(\lambda) \mathrm{d}\lambda \mathrm{d}t + \int_\Lambda h_{xx}(\bar{x}_T(\lambda)) y_T^\varepsilon(\lambda) y_T^\varepsilon(\lambda) \mathrm{d}\lambda\right].$$

Idea: Linearize using tensor product. In finite dimensions, Peng derived equation for $y_t^\varepsilon \otimes y_t^\varepsilon$ on

$$\mathbb{R}^n \otimes \mathbb{R}^n \cong \mathbb{R}^{n\times n}.$$

In existing literature, this is generalized in infinite dimensions to

$$H \otimes H \cong L_2(H).$$

Problem: In order to perform duality analysis, we need to solve equation in $L_1(H)$.

Explicit Tensor Product

Instead, we use explicit representation

$$L^2(\Lambda) \otimes L^2(\Lambda) \cong L^2(\Lambda^2),$$

where a simple tensor $y \otimes z$ is identified with the function $(\lambda, \mu) \mapsto y(\lambda)z(\mu)$. Thus, we can rewrite quadratic terms as

$$\begin{aligned} &\mathbb{E}\left[\int_0^T \int_\Lambda l_{xx}(\bar{x}_t(\lambda), \bar{u}_t) y_t^\varepsilon(\lambda) y_t^\varepsilon(\lambda) \mathrm{d}\lambda \mathrm{d}t\right] \\ =&\mathbb{E}\left[\int_0^T \int_\Lambda l_{xx}(\bar{x}_t(\lambda), \bar{u}_t) \delta(y_t^\varepsilon \otimes y_t^\varepsilon)(\lambda) \mathrm{d}\lambda \mathrm{d}t\right], \end{aligned}$$

where $\delta : H_0^1(\Lambda^2) \to L^2(\Lambda)$ is defined by $\delta(w)(\lambda) := w(\lambda, \lambda)$. This expression is linear in $y_t^\varepsilon \otimes y_t^\varepsilon$.

Second Order Adjoint Equation

Theorem (Stannat, W., SICON 2021)

The equation

$$\begin{cases} \mathrm{d}P_t(\lambda,\mu) = -[\Delta P_t(\lambda,\mu) + (b_x(\bar{x}_t(\lambda),\bar{u}_t) + b_x(\bar{x}_t(\mu),\bar{u}_t))P_t(\lambda,\mu) \\ \qquad +\langle \sigma_x(\bar{x}_t(\lambda),\bar{u}_t), \sigma_x(\bar{x}_t(\mu),\bar{u}_t)\rangle_{L_2(\Xi,\mathbb{R})} P_t(\lambda,\mu) \\ \qquad +\langle \sigma_x(\bar{x}_t(\lambda),\bar{u}_t) + \sigma_x(\bar{x}_t(\mu),\bar{u}_t), Q_t(\lambda,\mu)\rangle_{L_2(\Xi,\mathbb{R})} \\ \qquad +\delta^*(l_{xx}(\bar{x}_t(\lambda),\bar{u}_t)) + \delta^*(b_{xx}(\bar{x}_t(\lambda),\bar{u}_t)p_t(\lambda)) \\ \qquad +\delta^*(\langle \sigma_{xx}(\bar{x}_t(\lambda),\bar{u}_t), q_t\rangle_{L_2(\Xi,\mathbb{R})})]\mathrm{d}t + Q_t(\lambda,\mu)\mathrm{d}W_t \\ P_T(\lambda,\mu) = \delta^*(h_{xx}(\bar{x}_T(\lambda))) \end{cases}$$

has a unique adapted solution (P, Q)*, where*

$$P \in L^2([0,T]\times\Omega; L^2(\Lambda^2)) \cap L^2(\Omega; C([0,T]; H^{-1}(\Lambda^2))),$$

and

$$Q \in L^2([0,T]\times\Omega; L_2(\Xi; H^{-1}(\Lambda^2))).$$

Peng's Maximum Principle for SPDEs

Theorem (Stannat, W., SICON 2021)

Let $(\bar{x}, \bar{u})$ be an optimal pair. Then there exist adapted processes

$$(p, q) \in L^2([0,T] \times \Omega; H_0^1(\Lambda)) \times L^2([0,T] \times \Omega; L_2(\Xi, L^2(\Lambda)))$$

satisfying the first order adjoint equation and adapted processes

$$(P, Q) \in L^2([0,T] \times \Omega; L^2(\Lambda^2)) \times L^2([0,T] \times \Omega; L_2(\Xi, H^{-1}(\Lambda^2)))$$

satisfying the second order adjoint equation such that

$$\inf_{u \in U} \mathcal{G}(t, \bar{x}_t, u) = \mathcal{G}(t, \bar{x}_t, \bar{u}_t),$$

for almost all $(t, \omega) \in [0,T] \times \Omega$, where $\mathcal{G} : [0,T] \times L^2(\Lambda) \times U \to \mathbb{R}$

$$\mathcal{G}(t, x, u) := \int_\Lambda l(x(\lambda), u) \mathrm{d}\lambda + \langle p_t, b(x,u) \rangle_{L^2(\Lambda)} + \frac{1}{2} tr(\sigma(x,u)^* P_t \sigma(x,u)) + tr\left(\sigma(x,u)^* \left[q_t - P_t \sigma(\bar{x}_t, \bar{u}_t)\right]\right).$$

Dynamic Programming Approach

Minimize

$$J(s,x;u) := \mathbb{E}\left[\int_s^T \int_\Lambda l(x_t^u(\lambda), u_t)\mathrm{d}\lambda\mathrm{d}t + \int_\Lambda h(x_T^u(\lambda))\mathrm{d}\lambda\right]$$

over $u \in \mathcal{U}_s$ subject to

$$\begin{cases}\mathrm{d}x_t^u = [\Delta x_t^u + b(x_t^u, u_t)]\mathrm{d}t + \sigma(x_t^u, u_t)\mathrm{d}W_t, & t \in [s, T] \\ x_s^u = x \in L^2(\Lambda). \end{cases}$$

Introduce value function $V : [0,T] \times L^2(\Lambda) \to \mathbb{R}$,

$$V(s,x) := \inf_{u \in \mathcal{U}_s} J(s,x;u).$$

Satisfies dynamic programming principle

$$V(s,x) = \inf_{u \in \mathcal{U}_s} \mathbb{E}\left[\int_s^\tau l(x_t^u, u_t)\mathrm{d}t + V(\tau, x_\tau^u)\right], \qquad \forall \tau \in [s,T].$$

Can be used to derive optimality conditions.

Relationship under Smoothness Assumptions

Theorem (Bismut (1978), Bensoussan (1982))

If $V \in C^{1,3}([0,T] \times \mathbb{R}^d)$ *and* V_{sx} *is continuous, then for almost all* $t \in [0,T]$

1. $\begin{cases} V_x(t, \bar{x}_t) = p_t \\ V_{xx}(t, \bar{x}_t)\sigma(\bar{x}_t, \bar{u}_t) = q_t \end{cases}$
2. $-V_s(t, \bar{x}_t) = \inf_{u \in U} \mathcal{H}(\bar{x}_t, u, V_x(t, \bar{x}_t), V_{xx}(t, \bar{x}_t))$
3. $\inf_{u \in U} \mathcal{H}(\bar{x}_t, u, V_x(t, \bar{x}_t), V_{xx}(t, \bar{x}_t)) = \mathcal{H}(\bar{x}_t, \bar{u}_t, V_x(t, \bar{x}_t), V_{xx}(t, \bar{x}_t))$

where

$$\mathcal{H}(x, u, p, P) := \int_\Lambda l(x(\lambda), u)\mathrm{d}\lambda + \langle p, b(x, u)\rangle_{L^2(\Lambda)} + \frac{1}{2} tr(\sigma(x, u)^* P \sigma(x, u)).$$

Problem: In general, V is not differentiable!

Parabolic Viscosity Superdifferential

If $V \in C^{1,2}([0,T] \times L^2(\Lambda))$, it holds

$$\lim_{\tau \downarrow t, z \to x} \frac{1}{|\tau - t| + \|z - x\|^2} \Big[V(\tau, z) - V(t,x) - \partial_t V(t,x)(\tau - t) - \langle DV(t,x), z - x \rangle_{L^2(\Lambda)} - \frac{1}{2} \langle z - x, D^2 V(t,x)(z - x) \rangle_{L^2(\Lambda)} \Big] = 0.$$

Weaker notion of differentiability: We say $(G, p, P) \in D^{1,2,+}_{t+,x} V(t,x)$ if

$$\limsup_{\tau \downarrow t, z \to x} \frac{1}{|\tau - t| + \|z - x\|^2} \Big[V(\tau, z) - V(t,x) - G(\tau - t) - \langle p, z - x \rangle_{L^2(\Lambda)} - \frac{1}{2} \langle z - x, P(z - x) \rangle_{L^2(\Lambda)} \Big] \leq 0.$$

Theorem (Stannat, W. (2022+))

For almost every $t \in [0,T]$, it holds that

$$[-\langle \Delta \bar{x}_t, p_t \rangle_{H^{-1}(\Lambda) \times H^1_0(\Lambda)} - \mathcal{G}(t, \bar{x}_t, \bar{u}_t), \infty) \times \{p_t\} \times [P_t, \infty) \subset D^{1,2,+}_{t+,x} V(t, \bar{x}_t)$$

$\mathbb{P}$-almost surely.

Viscosity Solutions

Value function formally satisfies HJB equation

$$\begin{cases} V_s + \langle \Delta x, DV \rangle_{L^2(\Lambda)} + \inf_{u \in U} \mathcal{H}(x, u, DV, D^2 V) = 0, & (s,x) \in [0,T] \times L^2(\Lambda) \\ V(T,x) = \int_\Lambda h(x(\lambda)) \mathrm{d}\lambda, & x \in L^2(\Lambda) \end{cases}$$

where

$$\mathcal{H}(x,u,p,P) := \int_\Lambda l(x(\lambda),u) \mathrm{d}\lambda + \langle p, b(x,u) \rangle_{L^2(\Lambda)} + \frac{1}{2} \mathrm{tr} \left(\sigma(x,u)^* P \sigma(x,u) \right).$$

Definition (Viscosity Solution, Bounded Case)

V is viscosity subsolution, if

- $V(T,x) \leq \int_\Lambda h(x(\lambda)) \mathrm{d}\lambda, \quad x \in L^2(\Lambda)$;
- *for every* $(G,p,P) \in D^{1,2,+}_{t,x} V(t,x)$

$$G + \langle Ax, p \rangle + \inf_{u \in U} \mathcal{H}(x,u,p,P) \leq 0.$$

Viscosity Solutions II

It holds:

$$\boxed{(G, p, P) \in D^{1,2,+}_{t,x} v(t,x)}$$

$$\Longleftrightarrow$$

$\exists \phi \in C^{1,2}((s,T) \times L^2(\Lambda))$ such that:

1. $v - \phi$ attains maximum at (t,x),
2. $(\phi(t,x), \partial_t \phi(t,x), D\phi(t,x), D^2\phi(t,x)) = (v(t,x), G, p, P)$.

Equivalent definition of viscosity solution in the bounded case (!):

Definition (Viscosity Solution, Bounded Case)

V is a viscosity subsolution, if

- $V(T,x) \leq \int_\Lambda h(x(\lambda))\mathrm{d}\lambda, \quad x \in L^2(\Lambda)$;
- $\forall \phi \in C^{1,2}((s,T) \times L^2(\Lambda))$ *such that* $V - \phi$ *attains maximum at* (t,x), *it holds*

$$\phi_s(t,x) + \langle Ax, D\phi(t,x)\rangle + \inf_{u \in U} \mathcal{H}(x, u, D\phi(t,x), D^2\phi(t,x)) \leq 0.$$

Generalized Hamiltonian vs. Hamiltonian

Corollary (Stannat, W. (2022+))

It holds for almost all $t \in [s, T]$

$$\mathcal{G}(t, \bar{x}_t, \bar{u}_t) \leq \mathcal{H}(t, \bar{x}_t, \bar{u}_t, p_t, P_t),$$

$\mathbb{P}$*-almost surely, i.e.,*

$$tr(\sigma(\bar{x}_t, \bar{u}_t)(q_t - P_t\sigma(\bar{x}_t, \bar{u}_t))) \leq 0.$$

In the unbounded case, we need to make sense of

$$\langle \Delta x, D\phi(t, x)\rangle_{L^2(\Lambda)}, \qquad x \in L^2(\Lambda).$$

$\rightsquigarrow$ Need to restrict class of test functions.

To circumvent this issue, we use higher regularity of $\bar{x}_t \in H_0^1(\Lambda)$, $\mathrm{d}t \otimes \mathbb{P}$-a.s.

Verification Theorem

Minimize

$$J(s,x;u) := \mathbb{E}\left[\int_s^T L(x_t^u, u_t)\mathrm{d}t + H(x_T^u)\right]$$

over $u \in \mathcal{U}_s$ subject to

$$\begin{cases} \mathrm{d}x_t^u = [\Delta x_t^u + B(x_t^u, u_t)]\mathrm{d}t + \Sigma(x_t^u, u_t)\mathrm{d}W_t, & t \in [s,T] \\ x_s^u = x \in L^2(\Lambda), \end{cases}$$

where

- B, Σ are Lipschitz, linear growth;
- $(W_t)_{t\geq 0}$ cylindrical Wiener process;
- $\Lambda \subset \mathbb{R}$ bounded interval;
- L, H of (at most) quadratic growth.

Verification Theorem

Theorem (Stannat, W. (2022+))

Assume

- $\|\Sigma(x,u)\|_{L_2(\Xi, H_0^1(\Lambda))} \leq C(1 + \|x\|_{H_0^1(\Lambda)})$
- $V(t+\tau, x) - V(t,x) \leq C(1 + \|x\|^2_{H_0^1(\Lambda)})\tau$
- $V(t,\cdot) - C\|\cdot\|^2_{L^2(\Lambda)}$ *is concave.*

Let (x^*, u^*) *be an admissible pair. Suppose there are adapted processes* (G, p, P) *taking values in* $\mathbb{R}$, $H_0^1(\Lambda)$ *and* $L_2(L^2(\Lambda))$, *such that for almost all* $t \in [s, T]$*:*

$$(G_t, p_t, P_t) \in D^{1,2,+}_{t+,x} V(t, x_t^*)$$

$\mathbb{P}$*-almost surely, and*

$$\mathbb{E}\left[\int_s^T G_t + \langle \Delta x_t^*, p_t\rangle_{H^{-1}(\Lambda)\times H_0^1(\Lambda)} + \mathcal{H}(x_t^*, u_t^*, p_t, P_t)\mathrm{d}t\right] \geq 0.$$

Then (x^*, u^*) *is an optimal pair.*

References I

W. Stannat, L. Wessels
Peng's maximum principle for stochastic partial differential equations
SIAM J. Control Optim. 59 (2021), pp. 3552–3573.

W. Stannat, L. Wessels
Necessary and sufficient conditions for optimal control of semilinear SPDEs
Submitted, arXiv:2112.09639.

W. Stannat, L. Wessels
Deterministic control of stochastic reaction-diffusion equations
Evol. Equ. Control Theory 10 (2021), pp. 701–722.

References II

A. Bensoussan
Lectures on stochastic control
in Nonlinear filtering and stochastic control, Springer, Berlin-New York, 1982, pp. 1–62.

J. M. Bismut
An introductory approach to duality in optimal stochastic control
SIAM Rev. 20 (1978), pp. 62–78.

S. Peng
A general stochastic maximum principle for optimal control problems
SIAM J. Control Optim. 28 (1990), pp. 966–979.

L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko
The mathematical theory of optimal processes
Publishers John Wiley & Sons, Inc., New York, 1962.