Optimal Control of Stochastic Reaction-Diffusion Equations

Lukas Wessels

Technische Universität Berlin

www.lukaswessels.org

PhD Defense, July 19, 2022

SFB 910

Berlin Mathematical School BMS

Stochastic Reaction-Diffusion Equations

Consider

$$\partial_t x_t(\lambda) = \Delta x_t(\lambda) + b(x_t(\lambda)) + \xi_t(\lambda), \quad \lambda \in \Lambda \subset \mathbb{R}^n,$$

where

- Δ – Laplace operator
- $b : \mathbb{R} \to \mathbb{R}$ local reaction term
- ξ – random fluctuations.

Formulate as stochastic integral equation

$$\mathrm{d}x_t = [\Delta x_t + b(x_t)]\,\mathrm{d}t + \sigma \mathrm{d}W_t \quad \text{in } L^2(\Lambda),$$

where

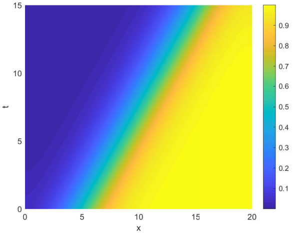
- σ – noise intensity
- $(W_t)_t$ cylindrical Wiener process.

Example

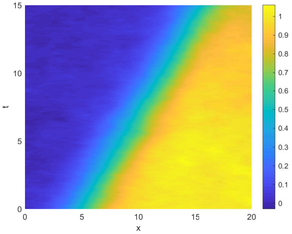
Consider the Nagumo equation

$$\mathrm{d}x_t = [\Delta x_t + x_t(x_t - 1)(a - x_t)]\,\mathrm{d}t + \sigma \mathrm{d}W_t, \quad x_0 = x \in L^2(\Lambda),$$

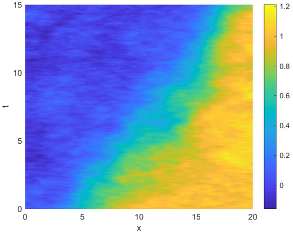
with Neumann boundary conditions, where $\Lambda \subset \mathbb{R}$ is bounded and $a \in (0, 1)$.



$\sigma = 0$



$\sigma = \frac{1}{2}$



$\sigma = 2$

Control of the Stochastic Nagumo Equation

Fix $T > 0$. Introduce control $u : [0, T] \times \Lambda \times \Omega \to \mathbb{R}$

$$\begin{cases} \mathrm{d}x_t^u = [\Delta x_t^u + x_t^u(x_t^u - 1)(a - x_t^u) + u_t]\,\mathrm{d}t + \sigma \mathrm{d}W_t \\ x_0^u = x \in L^2(\Lambda), \end{cases} \qquad (\star)$$

and cost functional

$$J(u) = \mathbb{E}\left[\int_0^T \int_\Lambda \frac{c_{\overline{\Lambda}}}{2} \left(x_t^u(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right)^2 + \frac{\nu}{2} u^2(t,\lambda)\mathrm{d}\lambda \mathrm{d}t + \int_\Lambda \frac{c_T}{2} \left(x_T^u(\lambda) - x^T(\lambda)\right)^2 \mathrm{d}\lambda\right]$$

where

- $x_{\overline{\Lambda}}$, x^T – desired reference profiles
- $c_{\overline{\Lambda}}, \nu, c_T \geq 0$.

Goal:

$$\boxed{\text{Minimize} \quad J \quad \text{subject to } (\star)}$$

Outline

1. Necessary Optimality Conditions
2. Sufficient Optimality Conditions
3. Applications

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

- l, h Nemytskii operators of (at most) quadratic growth
- b, σ Nemytskii operators, Lipschitz
- $\Lambda \subset \mathbb{R}$ bounded interval (also extensions to higher dimensions)
- control domain U non-convex.

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 for controlled ODEs.
- Peng's maximum principle (1990): Generalization to controlled SDEs.
- Since then: Generalizations to controlled SPDEs by Du, Fuhrman, Frankowska, Guatteri, Hu, Li, Lü, Meng, Tang, Tessitore, Zhang, ...
  Previous results require strong assumptions on coefficients l and h 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.

Use representation

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

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* $\mathcal{G} : [0, T] \times L^2(\Lambda) \times U \to \mathbb{R}$

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

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

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

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

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

Value function formally satisfies Hamilton-Jacobi-Bellman equation

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

where

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

Can be used to derive optimality conditions.

Link Between Value Function and Adjoint States

Under smoothness assumptions, it holds

$$\begin{cases} V_s(t,\bar{x}_t) = -\langle \Delta \bar{x}_t, DV(t,\bar{x}_t)\rangle_{L^2(\Lambda)} - \mathcal{H}(\bar{x}_t, \bar{u}_t, DV(t,\bar{x}_t), D^2V(t,\bar{x}_t)) \\ DV(t,\bar{x}_t) = p_t \\ D^2V(t,\bar{x}_t)\sigma(\bar{x}_t,\bar{u}_t) = q_t. \end{cases}$$

Generalizations dropping smoothness assumptions and using viscosity differentials up to first order in infinite dimensions by Cannarsa, Frankowska, Zhou, ...

Theorem (Stannat, W. (2022+))

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

$$[-\langle \Delta \bar{x}_t, p_t\rangle_{H^{-1}(\Lambda)\times H_0^1(\Lambda)} - \mathcal{G}(t,\bar{x}_t,\bar{u}_t), \infty) \times \{p_t\} \times [P_t, \infty) \subset D^{1,2,+}_{t+,x} V(t,\bar{x}_t)$$

$\mathbb{P}$-almost surely.

Link Between Value Function and Adjoint States

Under smoothness assumptions, it holds

$$\begin{cases} V_s(t,\bar{x}_t) = -\langle \Delta \bar{x}_t, DV(t,\bar{x}_t)\rangle_{L^2(\Lambda)} - \mathcal{H}(\bar{x}_t, \bar{u}_t, DV(t,\bar{x}_t), D^2V(t,\bar{x}_t)) \\ DV(t,\bar{x}_t) = p_t \\ D^2V(t,\bar{x}_t)\sigma(\bar{x}_t,\bar{u}_t) = q_t. \end{cases}$$

Generalizations dropping smoothness assumptions and using viscosity differentials up to first order in infinite dimensions by Cannarsa, Frankowska, Zhou, ...

Corollary (Stannat, W. (2022+))

It holds for almost all $t \in [0, T]$

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

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

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

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

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

Smooth verification theorem: Let (x^*, u^*) be admissible such that

$$u_t^* \in \arg\min_{u \in U} \mathcal{H}(x_t^*, u, DV(t, x_t^*), D^2V(t, x_t^*)),$$

for almost every (t, ω). Then (x^*, u^*) is optimal.

Verification theorems in the stochastic case in the framework of viscosity solutions by Fabbri, Gozzi, Li, Święch, Yong, Zhou, ...

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

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

Smooth verification theorem: Let (x^*, u^*) be admissible such that

$$V_s(t, x_t^*) + \langle \Delta x_t^*, DV(t, x_t^*)\rangle_{L^2(\Lambda)} + \mathcal{H}(x^*, u_t^*, DV(t, x_t^*), D^2V(t, x_t^*)) = 0$$

for almost every (t, ω). Then (x^*, u^*) is optimal.

Verification theorems in the stochastic case in the framework of viscosity solutions by Fabbri, Gozzi, Li, Święch, Yong, Zhou, ...

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*

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

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

Approximation of Optimal Controls

Numerical approximations by Dunst, Majee, Prohl, Vallet, ...

Minimize

$$J(u)=\mathbb{E}\left[\int_0^T\int_\Lambda \frac{c_{\overline{\Lambda}}}{2}\left(x_t^u(\lambda)-x_{\overline{\Lambda}}(t,\lambda)\right)^2+\frac{\nu}{2}u^2(t,\lambda)\mathrm{d}\lambda\mathrm{d}t+\int_\Lambda \frac{c_T}{2}\left(x_T^u(\lambda)-x^T(\lambda)\right)^2\mathrm{d}\lambda\right]$$

over $u \in L^2([0,T]\times\Lambda)$ subject to

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

where $b(x) = x(x-1)(a-x)$.

Major changes:

- additive noise
- deterministic controls
- convex control domain
- nonlinearity not Lipschitz.

Approximation of Optimal Controls

Numerical approximations by Dunst, Majee, Prohl, Vallet, ...

Minimize

$$J(u) = \mathbb{E}\left[\int_0^T\int_\Lambda \frac{c_{\overline{\Lambda}}}{2}\left(x_t^u(\lambda) - x_{\overline{\Lambda}}(t,\lambda)\right)^2 + \frac{\nu}{2}u^2(t,\lambda)\mathrm{d}\lambda\mathrm{d}t + \int_\Lambda \frac{c_T}{2}\left(x_T^u(\lambda) - x^T(\lambda)\right)^2\mathrm{d}\lambda\right]$$

over $u \in L^2([0,T]\times\Lambda)$ subject to

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

where $b(x) = x(x-1)(a-x)$.

Theorem (Stannat, W., EECT 2021)

The gradient of the cost functional is given by

$$\nabla J(u)(t,\lambda) = \mathbb{E}\left[p_t(\lambda) + \nu u(t,\lambda)\right],$$

where p is the solution of the adjoint equation

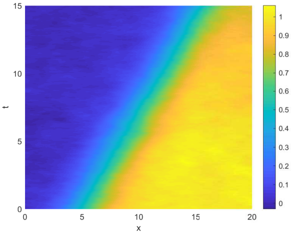
$$\begin{cases} -\partial_t p_t = \Delta p_t + b'(x_t^u)p_t + c_{\overline{\Lambda}}\left(x_t^u - x_{\overline{\Lambda}}(t,\cdot)\right), & t \in (0,T) \\ p_T = c_T\left(x_T^u - x^T\right) \in L^2(\Lambda). \end{cases}$$

Simulations

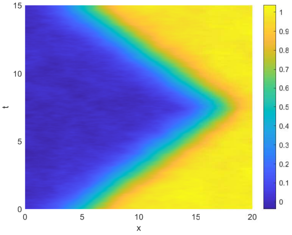
Consider

$$\begin{cases} \mathrm{d}x_t^u = [\Delta x_t^u + x_t^u(x_t^u - 1)(a - x_t^u) + u_t]\,\mathrm{d}t + \sigma \mathrm{d}W_t, & t \in [0, T] \\ x_0^u = x \in L^2(\Lambda), \end{cases}$$

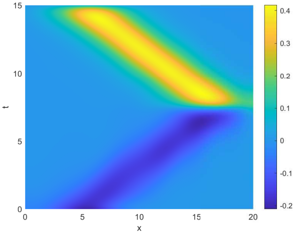
where $a = 39/40$, $\sigma = 1/2$, $T = 15$, $\Lambda = [0, 20]$.



Uncontrolled Solution



Controlled Solution



Optimal Control

Bibliography

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.