Optimal Control of Stochastic Reaction-Diffusion Equations

Lukas Wessels

joint work with Wilhelm Stannat

Technische Universität Berlin

www.lukaswessels.org

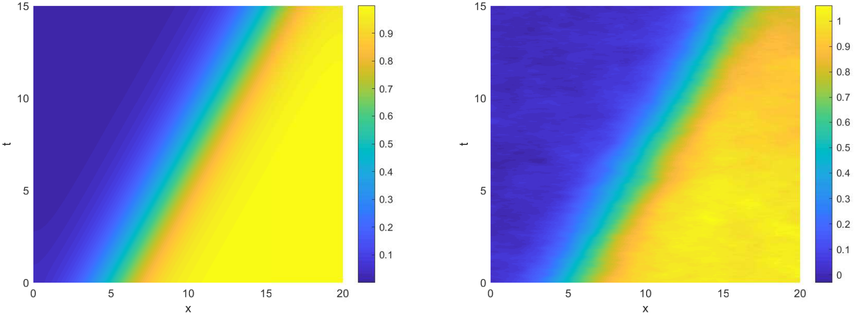
Oberseminar Dynamics, TU München, July 4, 2022

Motivation

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 $a \in (0, 1)$.



(a) Deterministic Case, $\sigma = 0$

(b) Stochastic Case, $\sigma = 0.5$

Figure: Uncontrolled Nagumo Equation

Motivation

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

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

and cost functional

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

where $c_{\overline{\Lambda}}, \nu, c_T \geq 0$.
Goal:

$$\boxed{\text{Minimize} \quad J \quad \text{subject to (1)}}$$

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.

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

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

Taylor expansions: For terminal costs we have

$$\mathbb{E}\left[\int_\Lambda h(x_T^\varepsilon) - h(\bar{x}_T)\mathrm{d}\lambda\right] \approx \mathbb{E}\left[\int_\Lambda h_x(\bar{x}_T)y_T^\varepsilon + \frac{1}{2}h_{xx}(\bar{x}_T)y_T^\varepsilon y_T^\varepsilon \mathrm{d}\lambda\right],$$

where y^ε satisfies linearized state equation

$$\begin{cases} \mathrm{d}y_t^\varepsilon = [\Delta y_t^\varepsilon + b_x(\bar{x}_t, \bar{u}_t)y_t^\varepsilon + b(\bar{x}_t, u_t^\varepsilon) - b(\bar{x}_t, \bar{u}_t)]\,\mathrm{d}t \\ \qquad\qquad + [\sigma_x(\bar{x}_t, \bar{u}_t)y_t^\varepsilon + \sigma(\bar{x}_t, u_t^\varepsilon) - \sigma(\bar{x}_t, \bar{u}_t)]\,\mathrm{d}W_t \\ y_0^\varepsilon = 0. \end{cases}$$

First Order Adjoint State

Consider SPDE

$$\begin{cases} \mathrm{d}y_t^\varepsilon = [\Delta y_t^\varepsilon + b_x(\bar{x}_t, \bar{u}_t) y_t^\varepsilon + \varphi_t]\,\mathrm{d}t + [\sigma_x(\bar{x}_t, \bar{u}_t) y_t^\varepsilon + \psi_t]\,\mathrm{d}W_t \\ y_0^\varepsilon = 0, \end{cases}$$

where $(\varphi, \psi) \in L^2([0,T] \times \Omega; L^2(\Lambda)) \times L^2([0,T] \times \Omega; L_2(\Xi, L^2(\Lambda)))$.
Construct linear functional

$$\mathcal{T}(\varphi, \psi) := \mathbb{E}\left[\int_0^T \int_\Lambda l_x(\bar{x}_t(\lambda), \bar{u}_t) y_t^\varepsilon(\lambda) \mathrm{d}\lambda \mathrm{d}t + \int_\Lambda h_x(\bar{x}_T(\lambda)) y_T^\varepsilon(\lambda) \mathrm{d}\lambda\right].$$

By Riesz's representation theorem, there is a unique pair

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

such that

$$\mathcal{T}(\varphi, \psi) = \mathbb{E}\left[\int_0^T \langle \varphi_t, p_t \rangle_{L^2(\Lambda)} + \langle \psi_t, q_t \rangle_{L_2(\Xi, L^2(\Lambda))} \mathrm{d}t\right]$$

for all $(\varphi, \psi) \in L^2([0,T] \times \Omega; L^2(\Lambda)) \times L^2([0,T] \times \Omega; L_2(\Xi, L^2(\Lambda)))$.

First Order Adjoint Equation

Adjoint state property:

$$\mathbb{E}\left[\int_0^T \langle y_t, l_x(\bar{x}_t, \bar{u}_t)\rangle \mathrm{d}t + \langle h_x(\bar{x}_T), y_T\rangle\right] = \mathbb{E}\left[\int_0^T \langle p_t, \varphi_t\rangle + \langle q_t, \psi_t\rangle \mathrm{d}t\right].$$

Applying Itô's product rule yields

$$\begin{aligned}\mathrm{d}\langle p_t, y_t\rangle_{L^2(\Lambda)} =& \langle p_t, \mathrm{d}y_t\rangle_{L^2(\Lambda)} + \langle y_t, \mathrm{d}p_t\rangle_{L^2(\Lambda)} + \mathrm{d}\langle p, y\rangle_t \\ \overset{!}{=}& \left[\langle p_t, \varphi_t\rangle_{L^2(\Lambda)} + \langle q_t, \psi_t\rangle_{L_2(\Xi, L^2(\Lambda))} - \langle y_t, l_x(\bar{x}_t, \bar{u}_t)\rangle_{L^2(\Lambda)}\right] \mathrm{d}t + dM_t.\end{aligned}$$

for some martingale $(M_t)_{t\geq 0}$. Thus,

$$\begin{cases}\mathrm{d}p_t = -\left[\Delta p_t + b_x(\bar{x}_t, \bar{u}_t)p_t + l_x(\bar{x}_t, \bar{u}_t) + \langle \sigma_x(\bar{x}_t, \bar{u}_t), q_t\rangle_{L_2(\Xi, \mathbb{R})}\right] \mathrm{d}t + q_t \mathrm{d}W_t \\ p_T = h_x(\bar{x}_T).\end{cases}$$

Unique variational solution (p, q), where

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

and

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

Backward Stochastic Differential Equations

Let $(W_t)_{t\geq 0}$ be a Brownian motion and $(\mathcal{F}_t)_{t\geq 0}$ its natural filtration. Consider the BSDE

$$\begin{cases} \mathrm{d}p_t = 0 \\ p_T = \xi, \end{cases}$$

where ξ is $\mathcal{F}_T$-measurable. ⚠ Adapted solution doesn't exist!
Natural candidate: $p_t = \mathbb{E}[\xi|\mathcal{F}_t]$. However, does not solve BSDE. Martingale representation theorem:

$$\mathbb{E}[\xi|\mathcal{F}_t] = \mathbb{E}[\xi] + \int_0^t q_s \mathrm{d}W_s.$$

Restate original problem:

$$\begin{cases} \mathrm{d}p_t = q_t \mathrm{d}W_t \\ p_T = \xi. \end{cases}$$

For $\xi = f(W_T)$:

$$\mathbb{E}\left[\xi|\mathcal{F}_t\right] = \mathbb{E}\left[f(W_T)|\mathcal{F}_t\right] = S_{T-t}f(W_t) = S_T f(W_0) + \int_0^t \nabla S_{T-s} f(W_s)\mathrm{d}W_s.$$

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

$$\begin{aligned} L^2(\Lambda) \otimes L^2(\Lambda) &\cong L^2(\Lambda^2) \\ y \otimes z &\leftrightsquigarrow ((\lambda, \mu) \mapsto y(\lambda) z(\mu)) \,. \end{aligned}$$

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

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

Gradient of the Cost Functional

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;
- Nonlinearity not Lipschitz.

Gradient of the Cost Functional

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

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

Gradient Descent Algorithm

Fix initial control $u_0 \in L^6([0, T] \times \Lambda)$, step size $s_0 > 0$ and stopping criterion $\eta > 0$.

1. Solve state equation.
2. Solve adjoint equation.
3. Repeat Step 1 and Step 2 to approximate

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

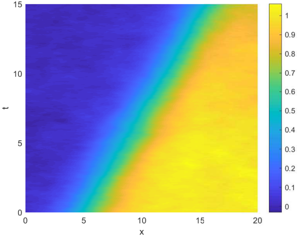
   via a Monte Carlo method.
4. Direction of descent: $D_n = -\nabla J(u_n) + \beta_n D_{n-1}$, where $\beta_n = \frac{\|\nabla J(u_n)\|}{\|\nabla J(u_{n-1})\|}$.
5. Compute the new control via $u_{n+1} = u_n + s_n D_n$.
6. Stop if $\|\nabla J(u_n)\| < \eta$, otherwise reset the step size $s_n = s_0$ and go to step 1.

Simulations

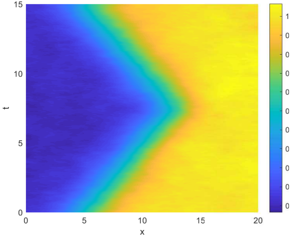
Consider

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

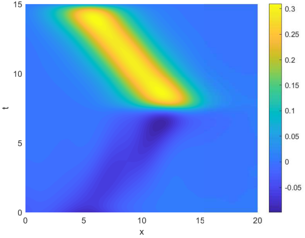
where $a = 39/40$, $\sigma = 1/2$, $T = 15$, $\Lambda = [0, 20]$.



(a) Uncontrolled Solution



(b) Controlled Solution



(c) Optimal Control

Outlook

Consider feedback control $u_t = \mathfrak{u}(t, x_t^u)$. Leads to

$$\mathrm{d}x_t^{\mathfrak{u}} = [\Delta x_t^{\mathfrak{u}} + x_t^{\mathfrak{u}}(x_t^{\mathfrak{u}} - 1)(a - x_t^{\mathfrak{u}}) + \mathfrak{u}(t, x_t^{\mathfrak{u}})]\,\mathrm{d}t + \sigma \mathrm{d}W_t, \quad t \in [0, T].$$

Questions:

- Existence of solution?
- How to approximate $\mathfrak{u}$?

Restrict class of controls to

$$\Phi(t, x, u) = \sum_{i=1}^{N} u^i(t, \cdot) \exp\left(-\|x_t - v^i\|_{L^2(\Lambda)}^2\right).$$

Optimize over *deterministic* u^i, v^i using our algorithm.

References

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

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

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.