Controlled SPDEs: Peng's Maximum Principle and Numerical Methods

Lukas Wessels

joint work with Wilhelm Stannat and Alexander Vogler

www.lukaswessels.org

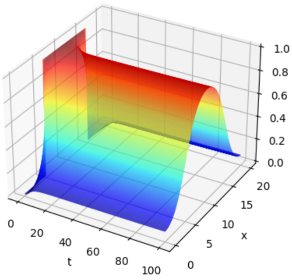
Mathematical Finance Seminar, Columbia University, April 6, 2023

Example

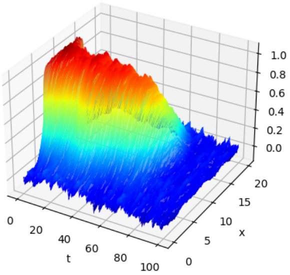
Consider the Nagumo equation

$$\begin{cases} \mathrm{d}u_t = \left[\Delta u_t + u_t(u_t - 1)\left(\frac{1}{2} - u_t\right)\right]\mathrm{d}t + \sigma \mathrm{d}W_t, & t \in [0,100] \\ u_0 = u \in L^2(0,20) \end{cases}$$

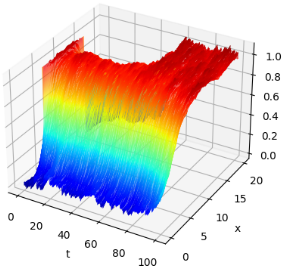
with Neumann boundary conditions and $u = 1_{[5,15]}$.



$\sigma = 0$



$\sigma = 0.05$



$\sigma = 0.05$

Control of the Stochastic Nagumo Equation

Introduce control $g : [0, 100] \times [0, 20] \times \Omega \to \mathbb{R}$

$$\begin{cases} \mathrm{d}u_t^g = \left[\Delta u_t^g + u_t^g(u_t^g - 1)\left(\frac{1}{2} - u_t^g\right) + g_t\right]\mathrm{d}t + \sigma \mathrm{d}W_t, \quad t \in [0, 100] \\ u_0^g = u \in L^2(0, 20) \end{cases} \qquad (\star)$$

and cost functional

$$J(g) = \mathbb{E}\left[\int_0^{100}\int_\Lambda \left(u_t^g(x) - u^{\mathrm{ref}}(t, x)\right)^2 + g^2(t, x)\mathrm{d}x\mathrm{d}t + \int_\Lambda \left(u_{100}^g(x) - u^T(x)\right)^2 \mathrm{d}x\right]$$

where u^{ref}, u^T are desired reference profiles.

Goal:

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

Outline

Setting

Minimize

$$J(g) := \mathbb{E}\left[\int_0^T \int_\Lambda l(u_t^g(x), g_t)\mathrm{d}x\mathrm{d}t + \int_\Lambda m(u_T^g(x))\mathrm{d}x\right]$$

subject to

$$\begin{cases} \mathrm{d}u_t^g = [\Delta u_t^g + b(u_t^g, g_t)]\mathrm{d}t + \sigma(u_t^g, g_t)\mathrm{d}W_t, & t \in [0, T] \\ u_0^g = u_0 \in L^2(\Lambda), \end{cases}$$

where

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

Pontryagin 1956: Controlled ODEs

Minimize

$$J(g) := \int_0^T l(u_t^g, g_t)\mathrm{d}t + m(u_T^g)$$

subject to

$$\begin{cases} \partial_t u_t^g = b(u_t^g, g_t), & t \in [0, T] \\ u_0^g = u_0 \in \mathbb{R}^n. \end{cases}$$

Let $(\bar{u}, \bar{g})$ be optimal. Then

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

for almost every $t \in [0, T]$, where Hamiltonian $\mathcal{H} : [0, T] \times \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}$

$$\mathcal{H}(t, u, g) := l(u, g) + \langle p_t, b(u, g) \rangle$$

where p_t is the adjoint state.

Peng 1990: Controlled SDEs

Minimize

$$J(g) := \mathbb{E}\left[\int_0^T l(u_t^g, g_t)\mathrm{d}t + m(u_T^g)\right]$$

subject to

$$\begin{cases} \mathrm{d}u_t^g = b(u_t^g, g_t)\mathrm{d}t + \sigma(u_t^g, g_t)\mathrm{d}W_t, \quad t \in [0, T] \\ u_0^g = u_0 \in \mathbb{R}^n. \end{cases}$$

Let $(\bar{u}, \bar{g})$ be optimal. Then

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

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

$$\mathcal{G}(t, u, g) := \mathcal{H}(t, u, g) + \frac{1}{2}\mathrm{tr}(\sigma(u, g)^* P_t \sigma(u, g)) + \mathrm{tr}\left(\sigma(u, g)^* \left[q_t - P_t \sigma(\bar{u}_t, \bar{g}_t)\right]\right)$$

where (p_t, q_t) first order adjoint state and $P_t \in \mathbb{R}^{n \times n}$ second order adjoint state.

Now: Controlled SPDEs

Minimize

$$J(g) := \mathbb{E}\left[\int_0^T \int_\Lambda l(u_t^g(x), g_t)\mathrm{d}x\mathrm{d}t + \int_\Lambda m(u_T^g(x))\mathrm{d}x\right]$$

subject to

$$\begin{cases} \mathrm{d}u_t^g = [\Delta u_t^g + b(u_t^g, g_t)]\mathrm{d}t + \sigma(u_t^g, g_t)\mathrm{d}W_t, & t \in [0, T] \\ u_0^g = u_0 \in L^2(\Lambda). \end{cases}$$

- Since Peng's result: Many generalizations to infinite-dimensional stochastic case by Du, Fuhrman, Frankowska, Guatteri, Hu, Li, Lü, Meng, Tang, Tessitore, Zhang, ...
- Major drawback in previous works: Restrictive assumptions on coefficients l and m, in particular excluding quadratic costs.
- Main obstacle: Characterization of second order adjoinst state $P_t \in L(H)$.

Spike Variation

Let $(\bar{u}, \bar{g})$ be optimal. Fix $\tau \in [0, T]$, $\varepsilon > 0$, $h \in \mathcal{U}$, and set

$$g_t^\varepsilon := \begin{cases} h, & \tau \le t \le \tau + \varepsilon, \\ \bar{g}_t & \text{otherwise.} \end{cases}$$

Then

$$0 \le J(g^\varepsilon) - J(\bar{g}) = \mathbb{E}\left[\int_0^T \int_\Lambda l(u_t^\varepsilon, g_t^\varepsilon) - l(\bar{u}_t, \bar{g}_t) \mathrm{d}x \mathrm{d}t + \int_\Lambda m(u_T^\varepsilon) - m(\bar{u}_T) \mathrm{d}x \right].$$

Taylor expansions: For terminal costs we have

$$\mathbb{E}\left[\int_\Lambda m(u_T^\varepsilon) - m(\bar{u}_T) \mathrm{d}x \right] \approx \mathbb{E}\left[\int_\Lambda m_u(\bar{u}_T) v_T^\varepsilon \mathrm{d}x \right],$$

where v^ε satisfies linearized state equation

$$\begin{cases} \mathrm{d}v_t^\varepsilon = \left[\Delta v_t^\varepsilon + b_u(\bar{u}_t, \bar{g}_t) v_t^\varepsilon + b(\bar{u}_t, g_t^\varepsilon) - b(\bar{u}_t, \bar{g}_t) \right] \mathrm{d}t \\ \qquad\qquad + \left[\sigma_u(\bar{u}_t, \bar{g}_t) v_t^\varepsilon + \sigma(\bar{u}_t, g_t^\varepsilon) - \sigma(\bar{u}_t, \bar{g}_t) \right] \mathrm{d}W_t \\ v_0^\varepsilon = 0. \end{cases}$$

Asymptotic of First Order Taylor Expansions

Applying Itô's formula to $\|v_t^\varepsilon\|^2_{L^2(\Lambda)}$ yields

$$\begin{aligned}\|v_T^\varepsilon\|^2_{L^2(\Lambda)} =& 2\int_0^T \langle \Delta v_s^\varepsilon + b_u(\bar{u}_s,\bar{g}_s)v_s^\varepsilon + b(\bar{u}_s,g_s^\varepsilon) - b(\bar{u}_s,\bar{g}_s), v_s^\varepsilon\rangle \mathrm{d}s\\ &+ 2\int_0^T \langle v_s^\varepsilon, (\sigma_u(\bar{u}_s,\bar{g}_s)v_s^\varepsilon + \sigma(\bar{u}_s,g_s^\varepsilon) - \sigma(\bar{u}_s,\bar{g}_s))\mathrm{d}W_s\rangle\\ &+ \int_0^T \|\sigma_u(\bar{u}_s,\bar{g}_s)v_s^\varepsilon + \sigma(\bar{u}_s,g_s^\varepsilon) - \sigma(\bar{u}_s,\bar{g}_s)\|^2_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}s.\end{aligned}$$

Itô correction term

$$\begin{aligned}&\int_0^T \|\sigma_u(\bar{u}_s,\bar{g}_s)v_s^\varepsilon + \sigma(\bar{u}_s,g_s^\varepsilon) - \sigma(\bar{u}_s,\bar{g}_s)\|^2_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}s\\ &\leq 2\int_0^T \|\sigma_u(\bar{u}_s,\bar{g}_s)v_s^\varepsilon\|^2_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}s + 2\int_\tau^{\tau+\varepsilon} \|\sigma(\bar{u}_s,h) - \sigma(\bar{u}_s,\bar{g}_s)\|^2_{L_2(\Xi,L^2(\Lambda))}\mathrm{d}s.\end{aligned}$$

The last term is of order $\mathcal{O}(\varepsilon)$; we need $o(\varepsilon)$. $\rightsquigarrow$ need second order Taylor expansions!

Variational Equations

First order:

$$\begin{cases} \mathrm{d}v_t^\varepsilon = [\Delta v_t^\varepsilon + b_u(\bar{u}_t, \bar{g}_t)v_t^\varepsilon + b(\bar{u}_t, g_t^\varepsilon) - b(\bar{u}_t, \bar{g}_t)]\,\mathrm{d}t \\ \qquad\qquad + [\sigma_u(\bar{u}_t, \bar{g}_t)v_t^\varepsilon + \sigma(\bar{u}_t, g_t^\varepsilon) - \sigma(\bar{u}_t, \bar{g}_t)]\,\mathrm{d}W_t \\ v_0^\varepsilon = 0. \end{cases}$$

Second order:

$$\begin{cases} \mathrm{d}w_t^\varepsilon = \left[\Delta w_t^\varepsilon + b_u(\bar{u}_t, \bar{g}_t)w_t^\varepsilon + \frac{1}{2}b_{uu}(\bar{u}_t, \bar{g}_t)v_t^\varepsilon v_t^\varepsilon + (b_u(\bar{u}_t, g_t^\varepsilon) - b_u(\bar{u}_t, \bar{g}_t))v_t^\varepsilon\right]\mathrm{d}t \\ \qquad\qquad + \left[\sigma_u(\bar{u}_t, \bar{g}_t)w_t^\varepsilon + \frac{1}{2}\sigma_{uu}(\bar{u}_t, \bar{g}_t)v_t^\varepsilon v_t^\varepsilon + (\sigma_u(\bar{u}_t, g_t^\varepsilon) - \sigma_u(\bar{u}_t, \bar{g}_t))v_t^\varepsilon\right]\mathrm{d}W_t \\ w_0^\varepsilon = 0. \end{cases}$$

Lemma

It holds that

$$\sup_{t\in[0,T]} \mathbb{E}\left[\|v_t^\varepsilon\|_{L^2(\Lambda)}^2\right] \le C\varepsilon \qquad \textit{and} \qquad \sup_{t\in[0,T]} \mathbb{E}\left[\|w_t^\varepsilon\|_{L^2(\Lambda)}^2\right] \le C\varepsilon^2.$$

Second Order Expansion of Cost Functional

From

$$0 \leq J(g^{\varepsilon}) - J(\bar{g}) = \mathbb{E}\left[\int_0^T \int_\Lambda l(u_t^{\varepsilon}, g_t^{\varepsilon}) - l(\bar{u}_t, \bar{g}_t) \mathrm{d}x\mathrm{d}t + \int_\Lambda m(u_T^{\varepsilon}) - m(\bar{u}_T)\mathrm{d}x\right]$$

we derive:

Lemma

It holds

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \int_\Lambda l_u(\bar{u}_t(x), \bar{g}_t)(v_t^{\varepsilon}(x) + w_t^{\varepsilon}(x)) + \frac{1}{2} l_{uu}(\bar{u}_t(x), \bar{g}_t) v_t^{\varepsilon}(x) v_t^{\varepsilon}(x)\mathrm{d}x\mathrm{d}t\right] \\
+&\mathbb{E}\left[\int_\Lambda m_u(\bar{u}_T(x))(v_T^{\varepsilon}(x) + w_T^{\varepsilon}(x)) + \frac{1}{2} m_{uu}(\bar{u}_T(x)) v_T^{\varepsilon}(x) v_T^{\varepsilon}(x)\mathrm{d}x\right] \\
+&\mathbb{E}\left[\int_0^T \int_\Lambda l(\bar{u}_t(x), g_t^{\varepsilon}) - l(\bar{u}_t(x), \bar{g}_t)\mathrm{d}x\mathrm{d}t\right] \geq o(\varepsilon).
\end{aligned}$$

Next step: Separate dependence on ε.

First Order Adjoint State

Consider SPDE

$$\begin{cases} \mathrm{d}v_t = [\Delta v_t + b_u(\bar{u}_t, \bar{g}_t)v_t + \varphi_t]\,\mathrm{d}t + [\sigma_u(\bar{u}_t, \bar{g}_t)v_t + \psi_t]\,\mathrm{d}W_t \\ v_0 = 0, \end{cases}$$

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

$$\mathcal{T}(\varphi, \psi) := \mathbb{E}\left[\int_0^T \int_\Lambda l_u(\bar{u}_t(x), \bar{g}_t)v_t(x)\mathrm{d}x\mathrm{d}t + \int_\Lambda m_u(\bar{u}_T(x))v_T(x)\mathrm{d}x\right].$$

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:

$$\mathbb{E}\left[\int_0^T \langle v_t, l_u(\bar{u}_t, \bar{g}_t)\rangle \mathrm{d}t + \langle m_u(\bar{u}_T), v_T\rangle\right] = \mathbb{E}\left[\int_0^T \langle p_t, \varphi_t\rangle + \langle q_t, \psi_t\rangle \mathrm{d}t\right].$$

Applying Itô's product rule yields

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

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

$$\begin{cases} \mathrm{d}p_t = -\left[\Delta p_t + b_u(\bar{u}_t, \bar{g}_t)p_t + l_u(\bar{u}_t, \bar{g}_t) + \langle \sigma_u(\bar{u}_t, \bar{g}_t), q_t\rangle_{L_2(\Xi, \mathbb{R})}\right] \mathrm{d}t + q_t \mathrm{d}W_t \\ p_T = m_u(\bar{u}_T). \end{cases}$$

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

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

Quadratic Terms

Follow same route for quadratic terms

$$\mathbb{E}\left[\int_0^T \int_\Lambda l_{uu}(\bar{u}_t(x),\bar{g}_t)v_t^\varepsilon(x)v_t^\varepsilon(x)\mathrm{d}x\mathrm{d}t + \int_\Lambda m_{uu}(\bar{u}_T(x))v_T^\varepsilon(x)v_T^\varepsilon(x)\mathrm{d}x\right].$$

Peng's idea in finite dimensions: Linearize using tensor product $v_t^\varepsilon \otimes v_t^\varepsilon$ and derive equation on

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

Infinite dimensional analogue:

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

Problem: Quadratic costs require duality analysis between $L_1(H)$ and $L(H)$.

Explicit Tensor Product

Instead, we use explicit representation

$$\begin{aligned} L^2(\Lambda) \otimes L^2(\Lambda) &\cong L^2(\Lambda^2) \\ v \otimes w &\leftrightsquigarrow ((x,y) \mapsto v(x)w(y)) \,. \end{aligned}$$

Rewrite quadratic terms as

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

where

$$\begin{aligned} \delta : H_0^1(\Lambda^2) &\to L^2(\Lambda) \\ V &\mapsto (x \mapsto V(x,x)). \end{aligned}$$

Second Order Adjoint Equation

Theorem (Stannat, W., SICON 2021)

The equation

$$\begin{cases} \mathrm{d}P_t(x,y) = -[\Delta P_t(x,y) + (b_u(\bar{u}_t(x),\bar{g}_t) + b_u(\bar{u}_t(y),\bar{g}_t))P_t(x,y) \\ \qquad\qquad +\langle \sigma_u(\bar{u}_t(x),\bar{g}_t), \sigma_u(\bar{u}_t(y),\bar{g}_t)\rangle_{L_2(\Xi,\mathbb{R})} P_t(x,y) \\ \qquad\qquad +\langle \sigma_u(\bar{u}_t(x),\bar{g}_t) + \sigma_u(\bar{u}_t(y),\bar{g}_t), Q_t(x,y)\rangle_{L_2(\Xi,\mathbb{R})} \\ \qquad\qquad +\delta^*(l_{uu}(\bar{u}_t(x),\bar{g}_t)) + \delta^*(b_{uu}(\bar{u}_t(x),\bar{g}_t)p_t(x)) \\ \qquad\qquad +\delta^*(\langle \sigma_{uu}(\bar{u}_t(x),\bar{g}_t), q_t\rangle_{L_2(\Xi,\mathbb{R})})]\mathrm{d}t + Q_t(x,y)\mathrm{d}W_t \\ P_T(x,y) = \delta^*(m_{uu}(\bar{u}_T(x))) \end{cases}$$

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 $(\bar{u}, \bar{g})$ *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

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

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

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

Feedback Controls

Consider feedback control

$$G : [0, T] \times L^2(\Lambda) \to L^2(\Lambda)$$

and controlled SPDE

$$\begin{cases} du_t^G = [\Delta u_t^G + b(u_t^G) + G(t, u_t^G)]dt + \sigma dW_t, \ t \in [0, T] \\ u_0^G = u \in L^2(\Lambda). \end{cases}$$

Approximate by artificial neural network

$$G(t, u) \approx \Phi(t, u, \alpha) = \sum_{i=1}^{n} \left(B\theta \left(A \begin{pmatrix} t \\ \pi_n u \end{pmatrix} + a \right) \right)_i e_i$$

where

- (e_i) is an ONB of $L^2(\Lambda)$, $\pi_n u = (\langle u, e_1 \rangle, \dots, \langle u, e_n \rangle)^\top$
- $\alpha = (A, a, B) \in \mathbb{R}^{k \times (n+1)} \times \mathbb{R}^k \times \mathbb{R}^{n \times k}$
- θ activator function.

Reduced Control Problem

Leads to controlled SPDE

$$\begin{cases} \mathrm{d}u_t^\alpha = [\Delta u_t^\alpha + b(u_t^\alpha) + \Phi(t, u_t^\alpha, \alpha)]\mathrm{d}t + \sigma \mathrm{d}W_t \\ u_0^\alpha = u \in L^2(\Lambda), \end{cases}$$

where $\alpha \in \mathbb{R}^d$, and cost functional $J : \mathbb{R}^d \to \mathbb{R}$,

$$J(\alpha) = \mathbb{E}\left[\int_0^T \int_\Lambda l(t, x, u_t^\alpha(x))\mathrm{d}x + \frac{\nu}{2}\|\Phi(t, u_t^\alpha, \alpha)\|_{L^2(\Lambda)}^2 \mathrm{d}t + \int_\Lambda m(x, u_T^\alpha(x))\mathrm{d}x\right].$$

Find gradient:

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left(J(\alpha + \varepsilon\beta) - J(\alpha)\right) = \langle \nabla J(\alpha), \beta \rangle.$$

Adjoint Equation

Linearized state equation

$$\begin{cases} \mathrm{d}v_t^\beta = [\Delta v_t^\beta + b_u(u_t^\alpha)v_t^\beta + \Phi_u(t, u^\alpha, \alpha)v_t^\beta + \Phi_\alpha(t, u_t^\alpha, \alpha)\beta]\mathrm{d}t \\ v_0^\beta = 0. \end{cases}$$

Adjoint state equation:

$$\begin{cases} \mathrm{d}p_t = -[(\Delta + b_u(u_t^\alpha) + \Phi_u^*(t, u_t^\alpha, \alpha))p_t + l_u(t, u_t^\alpha) + \nu\Phi_u^*(t, u_t^\alpha, \alpha)\Phi(t, u_t^\alpha, \alpha)]\mathrm{d}t \\ p_T = m_u(u_T^\alpha). \end{cases}$$

Adjoint state property:

$$\begin{aligned} &\mathbb{E}\left[\int_0^T \langle \Phi_\alpha(t, u_t^\alpha, \alpha)\beta, p_t\rangle \mathrm{d}t\right] \\ &= \mathbb{E}\left[\int_0^T \langle l_u(t, u_t^\alpha), v_t^\beta\rangle + \nu\langle \Phi(t, u_t^\alpha, \alpha), \Phi_u(t, u_t^\alpha, \alpha)v_t^\beta\rangle \mathrm{d}t + \langle m_u(u_T^\alpha), v_T^\beta\rangle\right] \end{aligned}$$

Gradient of Cost Functional

Directional derivative of cost functional:

$$
\begin{aligned}
\frac{\partial J(\alpha)}{\partial \beta} =& \mathbb{E}\Big[\int_0^T \langle l_u(t, u_t^\alpha), v_t^\beta \rangle + \nu \langle \Phi(t, u_t^\alpha, \alpha), \Phi_\alpha(t, u_t^\alpha, \alpha)\beta \rangle \\
& \qquad + \nu \langle \Phi(t, u_t^\alpha, \alpha), \Phi_u(t, u_t^\alpha, \alpha) v_t^\beta \rangle \mathrm{d}t + \langle m_u(u_T^\alpha), v_T^\beta \rangle \Big] \\
=& \langle \nabla J(\alpha), \beta \rangle,
\end{aligned}
$$

where

$$
\nabla J(\alpha) = \mathbb{E}\left[\int_0^T \nu \Phi_\alpha^*(t, u_t^\alpha, \alpha) \Phi(t, u_t^\alpha, \alpha) + \Phi_\alpha^*(t, u_t^\alpha, \alpha) p_t \mathrm{d}t \right].
$$

Gradient Descent Algorithm

1) For control α_n, solve state equation

$$\begin{cases} \mathrm{d}u_t^n = [\Delta u_t^n + b(u_t^n) + \Phi(t, u_t^n, \alpha_n)]\mathrm{d}t + \sigma \mathrm{d}W_t \\ u_0^n = u \in L^2(\Lambda) \end{cases}$$

2) Solve adjoint equation

$$\begin{cases} \mathrm{d}p_t^n = -[(\Delta + b_u(u_t^n) + \Phi_u^*(t, u_t^n, \alpha_n))p_t^n + l_u(t, u_t^n) + \nu \Phi_u^*(t, u_t^n, \alpha_n)\Phi(t, u_t^n, \alpha_n)]\mathrm{d}t \\ p_T = m_u(u_T^n) \end{cases}$$

3) Approximate gradient

$$\nabla J(\alpha_n) = \mathbb{E}\left[\int_0^T \nu \Phi_\alpha^*(t, u_t^n, \alpha_n)\Phi(t, u_t^n, \alpha_n) + \Phi_\alpha^*(t, u_t^n, \alpha_n)p_t^n \mathrm{d}t\right]$$

4) For step size s, compute new control

$$\alpha_{n+1} = \alpha_n - s\nabla J(\alpha_n)$$

Simulations

Consider

$$\begin{cases} \mathrm{d}u_t^\alpha = \left[\Delta u_t^\alpha + u_t^\alpha(u_t^\alpha - 1)\left(\frac{1}{2} - u_t^\alpha\right) + \Phi(t, u_t^\alpha, \alpha)\right]\mathrm{d}t + \sigma \mathrm{d}W_t, \quad t \in [0,100] \\ u_0^\alpha = u \in L^2(0,20) \end{cases}$$

with Neumann boundary conditions and $u = 1_{[5,15]}$.

$$J(\alpha)$$
$$= \mathbb{E}\left[\int_0^{100} \|u_t^\alpha - u_t^0\|_{L^2(0,20)}^2 + \frac{1}{2}\|\Phi(t, u_t^\alpha, \alpha)\|_{L^2(0,20)}^2 \mathrm{d}t + \|u_{100}^\alpha - u_{100}^0\|_{L^2(0,20)}^2\right],$$

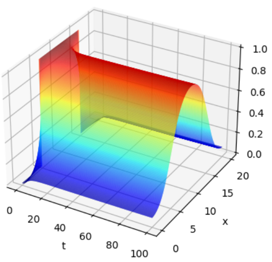
where u^0 is the solution without control and noise.

Simulations

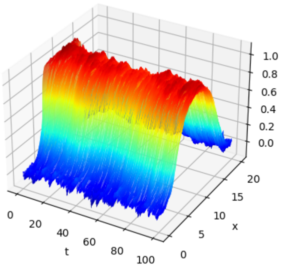
Consider

$$\begin{cases} \mathrm{d}u_t^\alpha = \left[\Delta u_t^\alpha + u_t^\alpha(u_t^\alpha - 1)\left(\frac{1}{2} - u_t^\alpha\right) + \Phi(t, u_t^\alpha, \alpha)\right]\mathrm{d}t + \sigma \mathrm{d}W_t, \quad t \in [0, 100] \\ u_0^\alpha = u \in L^2(0, 20) \end{cases}$$

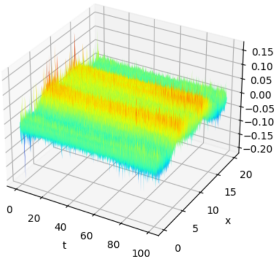
with Neumann boundary conditions and $u = 1_{[5,15]}$.



reference profile
u^0 $(\sigma = 0)$



controlled solution
$\sigma = 0.05$

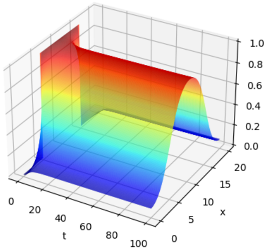


approximated control
$\sigma = 0.05$

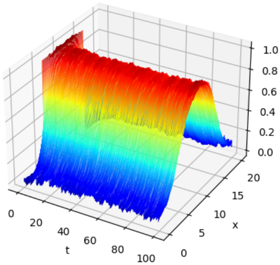
Radial Basis Function Network

Consider feedback control of the type

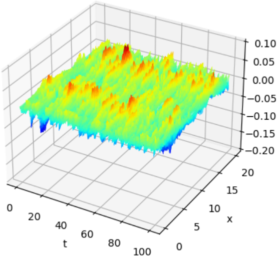
$$G(t,u)(x) \approx \Phi(t,u,\alpha,\tilde{u})(x) = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{r} \alpha_{ijk} 1_{[t_{k-1},t_k]}(t) e^{-\kappa|u(x)-\tilde{u}_j|^2} e_i(x).$$



reference profile
u^0 $(\sigma = 0)$



controlled solution
$\sigma = 0.05$



approximated control
$\sigma = 0.05$

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, A. Vogler, L. Wessels
Neural network approximation of optimal controls for stochastic reaction-diffusion equations
Submitted, arXiv:2301.11926.