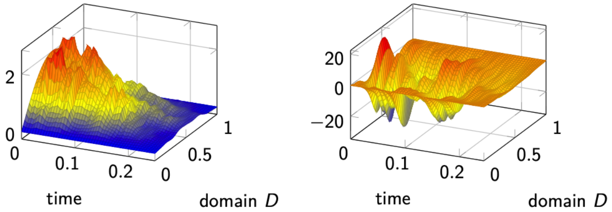
Numerics for SLQ problems with SPDE's

A. Prohl (U Tübingen, Germany), Y. Wang (Southwest U, China)

June 28, 2025



0. Motivation — Numerics for **Backward** SDE's

Forward-Backward SDE Problem:

Let $T > 0$. Find $\mathbb{R}^m \times \mathbb{R}^{m\times d}$-valued (Y, Z) of the **backward** SDE

$$-\mathrm{d}Y_t = f\big(t, X_t, Y_t, Z_t\big)\mathrm{d}t - Z_t \mathrm{d}W_t \qquad t \in [0, T]$$
$$Y_T = \Phi(X_T)\,.$$

where $\mathbb{R}^q$-valued X solves an SDE with Lipshitz functions (b, σ).

Stochastic Analysis:

- **Bismut '73**: 'stoch. version of Pontryagin's max-principle' (**stochastic control**)
- $\exists\,!$ solution tuple: Y (adapted) and Z (progr. meas. — 'steers the system')
- $m = 1$: $v(t,\mathbf{x}) = Y_t^{t,\mathbf{x}} = \mathbb{E}\Big[\Phi(X_T) + \int_t^T f\big(s, X_s, Y_s, Z_s\big)\mathrm{d}s\Big]$ solves **semil. PDE**

$$\big(\partial_t + L_t\big)v(t,\mathbf{x}) + f\Big(t, \mathbf{x}, v(t,\mathbf{x}), \langle \nabla v(t,\mathbf{x})\sigma(t,\mathbf{x})\rangle\Big) = 0$$
$$v(T,\mathbf{x}) = \Phi(X_T)\,,$$

on $[t, T] \times \mathbb{R}^q$, with $\mathbf{X}_t^{t,\mathbf{x}} = \mathbf{x}$ (**'nonlinear Feynman-Kac'**).

Problem: Discretization of BSDE & Guaranteed Convergence!

1. **Convergence** with **Rates** for a **Discretization** of **B**-SDE.
2. **Implementation**: ... role of **Conditional Expectations** for **dim's** q, d **large**!

0. Motivation — Numerics for Backward SDE's [Survey '23: Chessari et al.]

A **'Backward Methods':** solve **backwards** in time & **Conditional Expect's** appear!

Explicit Euler [Chevance '97, Zhang '04, Bouchard & Touzi '04]**:**

Let $\{t_j\}_j \subset [0,]$ be of size τ. Compute iterates $\big(Y^j, Z^j\big)$ via:

$$Z^j = \frac{1}{\tau}\mathbb{E}\big[Y^{j+1}\Delta_{j+1}W\big|\underline{\mathscr{F}_{t_j}}\big]$$
$$Y^j = \mathbb{E}\big[Y^{j+1} + \tau f\big(t_j, X^j, Y^j, Z^j\big)\big|\underline{\mathscr{F}_{t_j}}\big].$$

- **Analysis** by **Zhang '04**, **Bouchard & Touzi '04**: **Convergence with Rates**
- **Implementation**: Simulate **Cond'l Expect's** by **Statist. Learning Meth's**!
  - LS Regression by **Gobet et al. ['05, '06,...,'16]**: get **Estimators**!
  - **accurate computation**: large $\mathscr{M}$-samples of $\{X^{j,m}\}_m$: $\mathscr{M} \approx \frac{1}{\tau^{d+3}}$
  - **Reliable Simulations**: up to $d \leq 10$.

Part A: Statistical methods to Simulate Conditional Expectation's

1. **Statist. Learning: Clever** methods **needed** to simulate in **higher** dimensions!
2. **COD — also for Related Meth's:** Quantiz'n, tree based or cubature meth.'s,...

B 'forward methods': avoid simulation of **conditional** expect's:

- solve PDE above at **fixed** $(t, \mathbf{x})$ to approximate solutions (Y, Z).

C 'deep learning based methods': to allow **high-dimensional** state spaces $\mathbb{R}^q$:

Conclusions drawn for Stochastic Control ...but NOW with SPDE's!

1. **Analysis**: An ∞-dimensional **SDE**: $q = \infty$, and $d \gg 1$
2. **Numerical Analysis**: **Rates** of Convergence for a **Discretization**
3. **Simulation**: How to simulate **Conditional Exp's** — due to **COD**?

Subject of my Talk: Stochastic Linear-Quadratic Problem (SLQ)

1. **Problem**: involves linear heat eqn. **SPDE** with **linear noise** term
2. **NA** with **Rates** — driven by **Efficient Implementability**:
   a) **Gradient descent** algorithm based on **Pontryagin Max. Principle**
   b) **Direct approach** based on **Riccati eqn** — avoiding **Cond'l Exp's**!
3. **Cond'l Exp's** in a) — a new **Recursive Formula** that **avoids SL**! **(A. Chaudhary)**

I. Aim — Numerics for Optimal Control Problem with linear SPDE

SLQ Problem:

Let $T > 0$, and $\alpha, \beta \geq 0$. Find a minimizer $(X^\star, U^\star)$ of

$$J(X,U) = \frac{1}{2}\mathbb{E}\Big[\int_0^T \|X_s - \widetilde{X}_s\|^2_{\mathbb{L}^2} + \|U_s\|^2_{\mathbb{L}^2}\,\mathrm{d}s + \alpha\|X_T - \widetilde{X}_T\|^2_{\mathbb{L}^2}\Big]$$

$$\text{s.t.} \qquad \mathrm{d}X_t = \Big[\Delta X_t + U_t\Big]\mathrm{d}t + \big[\sigma_t + \beta X_t\big]\mathrm{d}W_t, \qquad X_0 = x_0.$$

Data and Objects:

- $\widetilde{X} : \Omega \times D_T \to \mathbb{R}$ 'desired profile'. $\sigma \equiv \{\sigma_t\}_t : \Omega \times D_T \to \mathbb{R}$ given
- $W \equiv \{W_t\}_t$ Wiener process on probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$.
- **Solution:** $\exists\,!\,(X^*, U^*)$ on it: *in particular*, adapted to $\{\mathscr{F}_t\}_{t\geq 0}$!

Question: How to approximate it **numerically**: Role of **complexity** of algorithms!

- **Algo 1** - via **PM**: **a)** Space-time discretization of **optimality conditions**
  b) Gradient descent method gives sequence of controls
  - ☹ involves BSPDE — **(LS-)estimator** for **conditional expectations**
    - **data-dependent** regression **s estimator** for **high-dim'l** state space!
  - ☺ Strong rates for **space-time** discretization error: $\sim O(\sqrt{\tau} + h)$

II. Algo 1 — Numerics based on PM (Complexity & Rates of Convergence)

- **Optimality Conditions** via **Pontryagin Maximum principle**:

$$
\begin{aligned}
\mathrm{d}X_t^* &= \Big[\Delta X_t^* + U_t^*\Big]\mathrm{d}t + \Big[\sigma_t + \beta X_t^*\Big]\mathrm{d}W_t \qquad \forall\, t \in (0,T) \\
\mathrm{d}Y_t &= \Big[-\Delta Y_t - \beta Z_t + X_t^*\Big]\mathrm{d}t + Z_t \mathrm{d}W_t \qquad \forall\, t \in (0,T)\,, \\
U_t^* &= Y_t \qquad \forall\, t \in (0,T) \\
X_0^* &= X_0\,, \qquad Y_T = -\alpha X_T^*\,,
\end{aligned}
$$

with solution $\big(X^*, Y, Z, U^*\big)$ **[Bensoussan, '83]**.

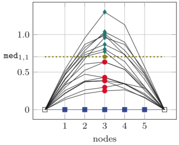
- **Problems**:
  - a **coupled FB-SPDE**: **de**coupling via Gradient descent method
  - 'Solve **B-SPDE** part' is what is **numerically challenging**, since it requires to compute **conditional expectations**!
  - **Motivation** that algorithm is **highly complex**:
    - **Space**: FEM gives FE-space $\mathbb{V}_h \subset \mathbb{H}_0^1$ — of **high** dimension $L \gg 1$.
    - **Time**: Implicit Euler **(Bouchard, Touzi, '04** — where **BSDE**s are simulated)

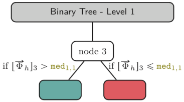
$$
\big(Z_h^j, \phi_h\big) = \frac{1}{\tau}\mathbb{E}\big[\Delta_j W\big(Y_h^{j+1}, \phi_h\big)\big|F_{t_j}\big]
$$

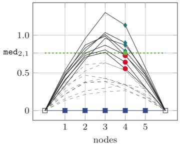
$$
\big(Y_h^j, \phi_h\big) + \tau\big(\nabla Y_h^j, \nabla\phi_h\big) = \underbrace{\mathbb{E}\big[\big(Y_h^{j+1}, \phi_h\big)\big|F_j\big]}_{\text{via LS-regression estimator}} + \beta\tau\big(Z_h^j, \phi_h\big)
$$

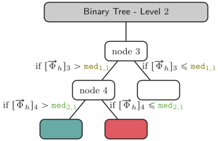
    - **Partitioning est**: **Gobet & al., '05, '14** — **practicable** for $q \leq 3$

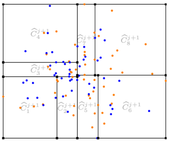
- **Complexity**: $\mathscr{M}$-sample $\big\{\{\mathbf{X}^j_{h,m}\}_{j=0}^J;\ 1 \leq m \leq \mathscr{M}\big\}$!
- **Q**: How **estimate** $\mathbb{E}\big[(\mathbf{Y}^{j+1}_h, \phi_h)\big|\ \underbrace{\mathbf{X}^j}_{\ldots\dim \mathbb{V}_h = q}\ \big]$ via **partitioning** esti. — for q **large**?
- **A**: **data-dependent** partitioning esti. — **'adaptive mesh'** instead of **uniform**!
  - **Dunst, P., '17**: this **SL-Algo** for controlled **SPDE**'s **applicable**, but **Costly** — see *e.g.* page 1 ($d = 10$).
  - Proof of strong consistency **(still) open**...











- **NA of BSPDE — Part 1**: FEM — discretization in space...

$$\begin{aligned} \mathrm{d}Y_h(t) &= \big(-\Delta_h Y_h(t) - \beta Z_h\big)\mathrm{d}t + Z_h \mathrm{d}W(t) \qquad \forall\, t \in [0,T]\,, \\ Y_h(T) &= Y_{T,h}\,. \end{aligned}$$

A) **strong stability** inherited by limiting BSPDE
B) **error estimate**:

$$\sup_{t\in[0,T]} \mathbb{E}\big[\|Y(t)-Y_h(t)\|^2_{\mathbb{L}^2}\big] + \mathbb{E}\Big[\int_0^T \|\nabla[Y(t)-Y_h(t)]\|^2_{\mathbb{L}^2} + \|Z(t)-Z_h(t)\|^2_{\mathbb{L}^2}\,\mathrm{d}t\Big] \le Ch^2\,.$$

- **NA of FBSPDE — Part 2**: FEM —the coupled problem

A) 'control-to-state' map S_h from SPDE: $\mathbf{X}_h^* = S_h(\mathbf{U}_h^*)$.
B) The reduced functional $\widehat{J}(U_h) = J\big(S_h(U_h), U_h\big)$

$$\sup_{t\in[0,T]} \mathbb{E}\big[\|X^*(t) - X_h^*(t)\|^2_{\mathbb{L}^2} + \|Y(t) - Y_h(t)\|^2_{\mathbb{L}^2}\big]$$
$$+\mathbb{E}\Big[\int_0^T \|U^*(t) - U_h^*(t)\|^2_{\mathbb{L}^2} + \|\nabla[Y(t) - Y_h(t)]\|^2_{\mathbb{L}^2} + \|Z(t) - Z_h(t)\|^2_{\mathbb{L}^2}\,\mathrm{d}t\Big] \le Ch^2\,.$$

- **NA of FBSPDE — Part 3**: time discretization for FEM-discretization above
  - **A)** requires **bound**: $\mathbb{E}\big[\|Z_h(t) - Z_h(s)\|^2_{\mathbb{L}^2}\big] \leq C|t-s|$ (**uniform** in h) obtained via Malliavin calculus and Riccati equation

$$P' + \Delta_h P + \beta^2 P + \mathrm{Id} - P^2 = 0 \qquad \forall\, t \in (0,T)\,, \qquad P(T) = \alpha \mathrm{Id}\,.$$

Then: $X_h^* = P_h U_h^* - \phi$, and $Z_h = D_\bullet Y_h$.

  - **B)** modified implicit Euler: based on discretization of Problem
    - **B$_1$)** First write down $(\mathbf{SLQ})_{h,\tau}$ — set $\widetilde{X} = 0$:

$$J_\tau(X_{h,\tau}, U_{h,\tau}) = \frac{1}{2}\tau \sum_{n=1}^{N} \mathbb{E}\big[\|X_{h,\tau}\|^2_{\mathbb{L}^2} + \|U_{h,\tau}\|^2_{\mathbb{L}^2}\big] + \frac{\alpha}{2}\mathbb{E}\big[\|X_{h,\tau}(T)\|^2_{\mathbb{L}^2}\big] \longrightarrow \text{min!}$$

s.t. $$X_{h,\tau}(t_{n+1}) - X_{h,\tau}(t_n) = \tau\Big(\Delta_h X_{h,\tau}(t_{n+1}) + U_{h,\tau}(t_n)\Big) + \Big(X_{h,\tau}(t_n) + \sigma_{t_n}\Big)\Delta_{n+1}W$$

    - **B$_2$)** derive 'discrete PM' — modified implicit Euler

Result for PM-based Space-Time -Discretization of SLQ: [P., Wang, '21]

$$\max_{0\leq n\leq N} \mathbb{E}\Big[\|X^*(t_n) - X^*_{h,\tau}(t_n)\|^2_{\mathbb{L}^2} + \sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \|U^*(t) - U^*_{h,\tau}(t_k)\|^2_{\mathbb{L}^2}\,\mathrm{d}t\|\Big] \leq C(\tau + h^2)$$

III. Algo 2 — Numerics via Riccati eqn (Complexity & Rates: [Wang, P.,'23])

- **Algo 1**: • **Statistical tools** for **high dimensions** to compute cond.'l expect's
  • gradient descent method
- **Tools for Theory**:
  - **Malliavin Calculus** & **Riccati equation** to get **Rates**:

$$P' + \Delta P + P\Delta + \beta^2 P + \mathrm{Id} - P^2 = 0 \quad \forall\, t \in (0,T)\,, \qquad P(T) = \alpha \mathrm{Id}\,.$$

- **Idea for Algo 2**: make Riccati equation relevant part of another Algorithm!

(a) **PDE**$_1$: solve the Riccati equation above to get $\{P(t)\}_{t\geq 0}$. Then
(b) **PDE**$_2$: get $\{\eta(t)\}_{t\geq 0}$ via

$$\eta' = -\Delta\eta + P\,\eta - \beta P\sigma \quad \forall\, t \in (0,T)\,, \qquad \eta(T) = 0\,.$$

(c) Insert in SPDE the **Feedback law** for minimizer U^*:

$$U^* = -P X^* - \eta \quad \forall\, t \in (0,T)\,.$$

☺ **no BSPDE**, **no** minimization! To solve SPDE comparably easy!
☺ Discretization of 2 PDE's and 1 SPDE: we expect order $O(\sqrt{\tau})$
☹ **restricted** now to **SLQ**. **Riccati** has operator-valued solution!

IV. Deterministic LQ: A review of Riccati-based approach

NA of **LQ** — the **deterministic** counterpart of **SLQ**:

- Key is **NA** of **Riccati-equation**: its analysis *e.g.* in **[Lasiecka & Triggiani, '00]**

1. **FEM-Semi**-Discretization of **Riccati equation** to solve **LQ** **[Kroller & Kunisch, '91]**
   - Optimal Rates
   - **Tool: Role of** P in **LQ**: '... **not just PDE sol**: **minimum of** J **representable with** help of P...'
2. **IE**-Semi-Discretization of **Riccati equation** **[Hansen & Stillfjord, '14]**
   - Sub-Optimal Rate $\frac{1}{2}$
   - **Tool:** Use **IE** for **PDE only**: '...use **monotonicity** properties & **[Rulla, '96]...**'
3. **BDF**-based time discretizations: '...heuristic evidence in works by [**Benner et al.**] ...'

Question/Motivation:

- **Construct optimally** convergent space-time Discretization — even for **LQ**!
- **Idea** for construction: properly address in **NA** the role of P in **(S)LQ**!
- **Standpoint**: **Results/NA tools** below for **SLQ** also address **LQ**!

V. ...now SLQ: Constr. via Difference Riccati-eqn [P., Wang, '23]

A) Concepts to get an **optimally convergent** **Difference Riccati equation**:

- **Starting point**: (slightly mod.) **Riccati-equation** also for **SLQ**!
- **Difference** Riccati eqn: **Discretization** for P **matters!** **[Ait Rami, Chen, Zhou, '02]**
  - **Not** just discretize **Riccati equation** by **IE**, but address **SLQ**!
  - **Scheme**: Denote $A_0 = \big(\mathrm{Id} - \tau\Delta_h\big)^{-1}$. Then get $\{P_\ell\}_{\ell=0}^N$ via

$$P_\ell = \Big(1 + \frac{\beta^2\tau}{2}\Big)^2\Big(A_0 P_{\ell+1} A_0 - \tau A_0 P_{\ell+1} A_0 A_0 P_{\ell+1} A_0\Big) + \tau \mathrm{Id} \qquad P_N = \alpha \mathrm{Id}\,. \quad (1)$$

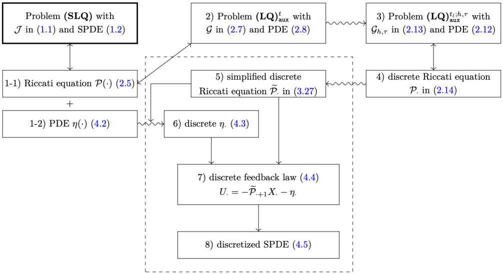
Result 1 for FEM-version of Riccati (1): **[P., Wang, '23]**

Let $\tau \le \tau_0(\alpha, \beta)$. Then

$$\|P(t_\ell) - P_\ell\|_{L(\mathbb{L}^2)} \le C(h^2 + \tau)\Big(\frac{\alpha}{t_N - t_\ell} + \ln\frac{1}{h}\Big) \quad (0 \le \ell \le N-1)\,.$$

  - **Tools**: **[Ait Rami, Chen, Zhou, '02]**, (discrete) semigroup methods, (discrete) stability, induction arguments

B): **Construction** of an **optimally conv.** Discr. for **SLQ**



Result 2 for Space-Time -Discret. of SLQ: **[P., Wang, '23]**

Let $\tau \leq \tau_0(\alpha, \beta)$. Then

$$\max_{0 \leq n \leq N} \mathbb{E}\Big[\|X^*(t_n) - X_n^*\|^2_{\mathbb{L}^2} + \|U^*(t_n) - U_n^*\|^2_{\mathbb{L}^2}\|\Big] \leq C\Big(\Big|(h^2 + \tau)\ln\frac{1}{h}\Big|^2 + \beta^2\tau\Big).$$

- **Tools**: Result 1, stability, tools for SPDE-conv. analysis (no Malliavin calc.!)

Summary of Talk

Numerical Analysis of **SLQ**:

1) Algo 1 — an algorithm that uses **PM**-principle:

- : an iterative method that uses gradient descent and **BSPDE**s
- ☺ : construction: suitable for general minimization methods.
- ☺ : space-time discretization: optimal rate of convergence
- ☹ : Discretization of **B**SPDE costly!
  - ⇒ **A. Chaudhary**: A **Recursive Formula** to get $\{Y^i\}_i$ **Avoiding SL**!

2) Algo 2 — an algorithm that uses **Riccati equation** for **(S)LQ**

- construction: **Key** for optimal scheme: Riccati serves **minimization**!
- ☺ **Simulation**: problem on page 1 — at a fraction of time !
- ☺ : space-time discretization: optimal rate of convergence

3) Whatever way we choose for (**SLQ**): take **optim. viewpoint** for discretization!

- **only** discretization of BSPDE or Riccati might be **sub**-optimal!