Time-Reversal of Stochastic Maximum Principle

Presented at the
IEEE Conference on Decision and Control, Milan, Italy, Dec. 2024

Amirhossein Taghvaei

University of Washington, Seattle

Dec 17, 2024

Presentation overview

Stochastic optimal control (SOC):

$$\min_{U_t \in \mathcal{F}_t} J(U)$$

- $U := \{U_t; 0 \leq t \leq T\}$ is the control input
- $U_t \in \mathcal{F}_t$ means that control input is non-anticipative

Objective: Application of optimization algorithms to solve the SOC, e.g.

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Challenge: Computing the gradient involves numerical solution of the forward-backward stochastic differential equation (FBSDE) in the stochastic maximum principle

This paper: A time-reversal approach to solve the FBSDE

Presentation overview

Stochastic optimal control (SOC):

$$\min_{U_t \in \mathcal{F}_t} J(U)$$

- $U := \{U_t; 0 \leq t \leq T\}$ is the control input
- $U_t \in \mathcal{F}_t$ means that control input is non-anticipative

Objective: Application of optimization algorithms to solve the SOC, e.g.

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Challenge: Computing the gradient involves numerical solution of the forward-backward stochastic differential equation (FBSDE) in the stochastic maximum principle

This paper: A time-reversal approach to solve the FBSDE

Presentation overview

Stochastic optimal control (SOC):

$$\min_{U_t \in \mathcal{F}_t} J(U)$$

- $U := \{U_t; 0 \le t \le T\}$ is the control input
- $U_t \in \mathcal{F}_t$ means that control input is non-anticipative

Objective: Application of optimization algorithms to solve the SOC, e.g.

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Challenge: Computing the gradient involves numerical solution of the forward-backward stochastic differential equation (FBSDE) in the stochastic maximum principle

This paper: A time-reversal approach to solve the FBSDE

Presentation overview

Stochastic optimal control (SOC):

$$\min_{U_t \in \mathcal{F}_t} J(U)$$

- $U := \{U_t; 0 \leq t \leq T\}$ is the control input
- $U_t \in \mathcal{F}_t$ means that control input is non-anticipative

Objective: Application of optimization algorithms to solve the SOC, e.g.

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Challenge: Computing the gradient involves numerical solution of the forward-backward stochastic differential equation (FBSDE) in the stochastic maximum principle

This paper: A time-reversal approach to solve the FBSDE

- Maximum principle in the deterministic setting
- Maximum principle in the stochastic setting
- Numerical solution of FBSDE $\rightarrow$ time-reversal formulation

- Maximum principle in the deterministic setting
- Maximum principle in the stochastic setting
- Numerical solution of FBSDE $\rightarrow$ time-reversal formulation

Objective function: $$\min_{U_t} J(U) = \int_0^T \ell(X_t, U_t)\mathrm{d}t + \ell_f(X_T)$$

Dynamic constraint: $$\frac{\mathrm{d}X_t}{\mathrm{d}t} = a(X_t, U_t), \quad X_0 = x_0$$

- $X_t \in \mathbb{R}^n$ is the state
- $U_t \in \mathbb{R}^m$ is the control input

Solution methodologies:

1. Dynamical programming $\longrightarrow$ HJB eq.
2. First-order optimality condition $\longrightarrow$ Maximum principle

D. Liberzon, Calculus of variations and optimal control theory: a concise introduction. Princeton university press; 2011

Objective function: $$\min_{U_t} J(U) = \int_0^T \ell(X_t, U_t)\mathrm{d}t + \ell_f(X_T)$$

Dynamic constraint: $$\frac{\mathrm{d}X_t}{\mathrm{d}t} = a(X_t, U_t), \quad X_0 = x_0$$

- $X_t \in \mathbb{R}^n$ is the state
- $U_t \in \mathbb{R}^m$ is the control input

Solution methodologies:

1. Dynamical programming $\longrightarrow$ HJB eq.
2. First-order optimality condition $\longrightarrow$ Maximum principle

D. Liberzon, Calculus of variations and optimal control theory: a concise introduction. Princeton university press; 2011

Pontryagin's Maximum principle

- Hamiltonian function

$$H(x,u,y) := \ell(x,u) + y^\top a(x,u)$$

- Hamiltonian sys.

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \partial_y H(X_t,U_t,Y_t), \quad X_0 = x_0$$
$$-\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \partial_x H(X_t,U_t,Y_t), \quad Y_T = \partial_x \ell_f(X_T)$$

- Optimality condition

$$U_t \in \arg\min_u H(X_t,u,Y_t), \quad \forall t \in [0,T]$$

D. Liberzon, Calculus of variations and optimal control theory: a concise introduction. Princeton university press; 2011

Pontryagin's Maximum principle

- Hamiltonian function

$$H(x,u,y) := \ell(x,u) + y^\top a(x,u)$$

- Hamiltonian sys.

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \partial_y H(X_t, U_t, Y_t), \quad X_0 = x_0$$
$$-\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \partial_x H(X_t, U_t, Y_t), \quad Y_T = \partial_x \ell_f(X_T)$$

- Optimality condition

$$U_t \in \arg\min_u H(X_t, u, Y_t), \quad \forall t \in [0,T]$$

D. Liberzon, Calculus of variations and optimal control theory: a concise introduction. Princeton university press; 2011

Pontryagin's Maximum principle

- Hamiltonian function

$$H(x,u,y) := \ell(x,u) + y^\top a(x,u)$$

- Hamiltonian sys.

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \partial_y H(X_t, U_t, Y_t), \quad X_0 = x_0$$
$$-\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \partial_x H(X_t, U_t, Y_t), \quad Y_T = \partial_x \ell_f(X_T)$$

- Optimality condition

$$U_t \in \arg\min_u H(X_t, u, Y_t), \quad \forall t \in [0,T]$$

D. Liberzon, Calculus of variations and optimal control theory: a concise introduction. Princeton university press; 2011

Iterative approach to find the optimal control

Gradient descent:

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Gradient formula:

$$\nabla J(U)\Big|_t = \partial_u H(X_t, U_t, Y_t)$$

where (X_t, Y_t) are obtained from

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \partial_y H(X_t, U_t, Y_t), \quad X_0 = x_0$$

$$-\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \partial_x H(X_t, U_t, Y_t), \quad Y_T = \partial_x \ell_f(X_T)$$

Question: How to extend it to the stochastic setting?

Iterative approach to find the optimal control

Gradient descent:

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Gradient formula:

$$\nabla J(U)\Big|_t = \partial_u H(X_t, U_t, Y_t)$$

where (X_t, Y_t) are obtained from

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \partial_y H(X_t, U_t, Y_t), \quad X_0 = x_0$$

$$-\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \partial_x H(X_t, U_t, Y_t), \quad Y_T = \partial_x \ell_f(X_T)$$

Question: How to extend it to the stochastic setting?

Iterative approach to find the optimal control

Gradient descent:

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Gradient formula:

$$\nabla J(U)\Big|_t = \partial_u H(X_t, U_t, Y_t)$$

where (X_t, Y_t) are obtained from

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \partial_y H(X_t, U_t, Y_t), \quad X_0 = x_0$$

$$-\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \partial_x H(X_t, U_t, Y_t), \quad Y_T = \partial_x \ell_f(X_T)$$

Question: How to extend it to the stochastic setting?

- Maximum principle in the deterministic setting
- Maximum principle in the stochastic setting
- Numerical solution of FBSDE $\rightarrow$ time-reversal formulation

- Maximum principle in the deterministic setting
- Maximum principle in the stochastic setting
- Numerical solution of FBSDE $\rightarrow$ time-reversal formulation

Optimal control problem
Stochastic setting

Objective function: $$\min_{U_t \in \mathcal{F}_t} J(U) = \mathbb{E}\left[\int_0^T \ell(X_t, U_t)\mathrm{d}t + \ell_f(X_T)\right]$$

Dynamic constraint: $$\mathrm{d}X_t = a(X_t, U_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 = x_0$$

- $W_t \in \mathbb{R}^m$ is standard Wiener process
- $\mathcal{F}_t$ is the filtration generated by W_t

Solution methodologies:

1. Dynamical programming $\longrightarrow$ HJB eq.
2. First-order optimality condition $\longrightarrow$ Stochastic Maximum principle

J. Yong and X. Zhou, Stochastic controls: Hamiltonian systems and HJB equations. Springer Science & Business Media, 1999, vol. 43.

Optimal control problem
Stochastic setting

Objective function: $$\min_{U_t \in \mathcal{F}_t} J(U) = \mathbb{E}\left[\int_0^T \ell(X_t, U_t)\mathrm{d}t + \ell_f(X_T)\right]$$

Dynamic constraint: $$\mathrm{d}X_t = a(X_t, U_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 = x_0$$

- $W_t \in \mathbb{R}^m$ is standard Wiener process
- $\mathcal{F}_t$ is the filtration generated by W_t

Solution methodologies:

1. Dynamical programming $\longrightarrow$ HJB eq.
2. First-order optimality condition $\longrightarrow$ Stochastic Maximum principle

J. Yong and X. Zhou, Stochastic controls: Hamiltonian systems and HJB equations. Springer Science & Business Media, 1999, vol. 43.

- Hamiltonian function

$$H(x,u,y,z) := \ell(x,u) + y^\top a(x,u) + \mathrm{tr}(z\sigma(x))$$

- Hamiltonian sys.

$$\begin{aligned} \mathrm{d}X_t &= \partial_y H(X_t,U_t,Y_t,Z_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 = x_0 \\ -\mathrm{d}Y_t &= \partial_x H(X_t,U_t,Y_t,Z_t)\mathrm{d}t - Z_t\mathrm{d}W_t, \quad Y_T = \partial_x \ell_f(X_T) \end{aligned}$$

- Optimality condition

$$U_t \in \arg\min_u H(X_t,u,Y_t,Z_t), \quad \forall t \in [0,T]$$

J. Yong and X. Zhou, Stochastic controls: Hamiltonian systems and HJB equations. Springer Science & Business Media, 1999, vol. 43.

- Hamiltonian function

$$H(x,u,y,z) := \ell(x,u) + y^\top a(x,u) + \mathrm{tr}(z\sigma(x))$$

- Hamiltonian sys.

$$\begin{aligned} \mathrm{d}X_t &= \partial_y H(X_t,U_t,Y_t,Z_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 = x_0 \\ -\mathrm{d}Y_t &= \partial_x H(X_t,U_t,Y_t,Z_t)\mathrm{d}t - Z_t\mathrm{d}W_t, \quad Y_T = \partial_x \ell_f(X_T) \end{aligned}$$

- Optimality condition

$$U_t \in \arg\min_u H(X_t,u,Y_t,Z_t), \quad \forall t \in [0,T]$$

J. Yong and X. Zhou, Stochastic controls: Hamiltonian systems and HJB equations. Springer Science & Business Media, 1999, vol. 43.

- Hamiltonian function

$$H(x,u,y,z) := \ell(x,u) + y^\top a(x,u) + \mathsf{tr}(z\sigma(x))$$

- Hamiltonian sys.

$$\begin{aligned} \mathrm{d}X_t &= \partial_y H(X_t,U_t,Y_t,Z_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 = x_0 \\ -\mathrm{d}Y_t &= \partial_x H(X_t,U_t,Y_t,Z_t)\mathrm{d}t - Z_t\mathrm{d}W_t, \quad Y_T = \partial_x \ell_f(X_T) \end{aligned}$$

- Optimality condition

$$U_t \in \arg\min_u H(X_t,u,Y_t,Z_t), \quad \forall t \in [0,T]$$

J. Yong and X. Zhou, Stochastic controls: Hamiltonian systems and HJB equations. Springer Science & Business Media, 1999, vol. 43.

Iterative approach to find the optimal control

Gradient descent:

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Gradient formula:

$$\nabla J(U)\Big|_t = \partial_u H(X_t, U_t, Y_t, Z_t)$$

where (X_t, Y_t, Z_t) are obtained from

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \partial_y H(X_t, U_t, Y_t) + \sigma(X_t)\mathrm{d}W_t, \quad X_0 = x_0$$

$$-\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \partial_x H(X_t, U_t, Y_t) - Z_t \mathrm{d}W_t, \quad Y_T = \partial_x \ell_f(X_T)$$

Question: How to numerically find (Y_t, Z_t)?

Iterative approach to find the optimal control

Gradient descent:

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Gradient formula:

$$\nabla J(U)\Big|_t = \partial_u H(X_t, U_t, Y_t, Z_t)$$

where (X_t, Y_t, Z_t) are obtained from

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \partial_y H(X_t, U_t, Y_t) + \sigma(X_t)\mathrm{d}W_t, \quad X_0 = x_0$$

$$-\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \partial_x H(X_t, U_t, Y_t) - Z_t \mathrm{d}W_t, \quad Y_T = \partial_x \ell_f(X_T)$$

Question: How to numerically find (Y_t, Z_t)?

Iterative approach to find the optimal control

Gradient descent:

$$U^{k+1} = U^k - \eta \nabla J(U^k)$$

Gradient formula:

$$\nabla J(U)\Big|_t = \partial_u H(X_t, U_t, Y_t, Z_t)$$

where (X_t, Y_t, Z_t) are obtained from

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \partial_y H(X_t, U_t, Y_t) + \sigma(X_t)\mathrm{d}W_t, \quad X_0 = x_0$$

$$-\frac{\mathrm{d}Y_t}{\mathrm{d}t} = \partial_x H(X_t, U_t, Y_t) - Z_t \mathrm{d}W_t, \quad Y_T = \partial_x \ell_f(X_T)$$

Question: How to numerically find (Y_t, Z_t)?

- Maximum principle in the deterministic setting
- Maximum principle in the stochastic setting
- Numerical solution of FBSDE $\rightarrow$ time-reversal formulation

Presentation overview

Outline

- Maximum principle in the deterministic setting
- Maximum principle in the stochastic setting
- Numerical solution of FBSDE $\rightarrow$ time-reversal formulation

Numerical solution of the FBSDE

FBSDE system:

$$\mathrm{d}X_t = a(t, X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 \sim P_0$$

$$-\mathrm{d}Y_t = g(t, X_t, Y_t, Z_t)\mathrm{d}t - Z_t\mathrm{d}W_t, \quad Y_T = g_f(X_T)$$

Existing approaches:

- PDE-based approach: [Peng 91, Ma et. al. 94, ...]
- Conditional expectation based approach: [Zhang 2004, Exarchos & Theodorou 2016]

This paper: A time-reversal approach to solve the FBSDE

Numerical solution of the FBSDE

FBSDE system:

$$\mathrm{d}X_t = a(t, X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 \sim P_0$$

$$-\mathrm{d}Y_t = g(t, X_t, Y_t, Z_t)\mathrm{d}t - Z_t\mathrm{d}W_t, \quad Y_T = g_f(X_T)$$

Existing approaches:

- PDE-based approach: [Peng 91, Ma et. al. 94, ...]

$$(Y_t, Z_t) = (\phi(t, X_t), \sigma(X_t)^\top \partial_x \phi(t, X_t))$$

where $\phi : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ solves a PDE

- Conditional expectation-based approach: [Zhang 2004, Exarchos & Theodorou 2016]

$$Y_s = \mathbb{E}\left[Y_t + \int_s^t g(\tau, X_\tau, Y_\tau, Z_\tau)\mathrm{d}\tau \,\middle|\, \mathcal{F}_s\right]$$

This paper: A time-reversal approach to solve the FBSDE

Numerical solution of the FBSDE

FBSDE system:

$$\mathrm{d}X_t = a(t, X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 \sim P_0$$

$$-\mathrm{d}Y_t = g(t, X_t, Y_t, Z_t)\mathrm{d}t - Z_t\mathrm{d}W_t, \quad Y_T = g_f(X_T)$$

Existing approaches:

- PDE-based approach: [Peng 91, Ma et. al. 94, . . .]

$$(Y_t, Z_t) = (\phi(t, X_t), \sigma(X_t)^\top \partial_x \phi(t, X_t))$$

where $\phi : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ solves a PDE

- Conditional expectation-based approach: [Zhang 2004, Exarchos & Theodorou 2016]

$$Y_s = \mathbb{E}\left[Y_t + \int_s^t g(\tau, X_\tau, Y_\tau, Z_\tau)\mathrm{d}\tau \,\middle|\, \mathcal{F}_s\right]$$

This paper: A time-reversal approach to solve the FBSDE

Numerical solution of the FBSDE

FBSDE system:

$$\mathrm{d}X_t = a(t, X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 \sim P_0$$

$$-\mathrm{d}Y_t = g(t, X_t, Y_t, Z_t)\mathrm{d}t - Z_t\mathrm{d}W_t, \quad Y_T = g_f(X_T)$$

Existing approaches:

- PDE-based approach: [Peng 91, Ma et. al. 94, ...]

$$(Y_t, Z_t) = (\phi(t, X_t), \sigma(X_t)^\top \partial_x \phi(t, X_t))$$

where $\phi : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ solves a PDE

- Conditional expectation-based approach: [Zhang 2004, Exarchos & Theodorou 2016]

$$Y_s = \mathbb{E}\left[Y_t + \int_s^t g(\tau, X_\tau, Y_\tau, Z_\tau)\mathrm{d}\tau \,\middle|\, \mathcal{F}_s\right]$$

This paper: A time-reversal approach to solve the FBSDE

Time-reversal of diffusions

- Let X and $\overleftarrow{X}$ be the solutions to

$$\mathrm{d}X_t = a(t, X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 \sim P_0$$

$$\mathrm{d}\overleftarrow{X}_t = a(t, \overleftarrow{X}_t)\mathrm{d}t + \sigma(\overleftarrow{X}_t)\overleftarrow{\mathrm{d}}W_t + b(t, \overleftarrow{X}_t)\mathrm{d}t, \quad \overleftarrow{X}_T \sim P_T$$

where P_t is the probability law of X_t and

$$b(t,x) = \frac{1}{P_t(x)} \partial_x(\sigma(x)\sigma(x)^\top P_t(x))$$

Theorem [Cattiaux, et. al., 2021]

If $D(x) := \sigma(x)\sigma(x)^\top$ is positive-definite everywhere, and a finite entropy condition holds, then

$$\{X_t; 0 \leq t \leq T\} \stackrel{\mathrm{d}}{=} \{\overleftarrow{X}_t; 0 \leq t \leq T\}$$

B. D. Anderson, "Reverse-time diffusion equation models," Stochastic Processes and their Applications, vol. 12, no. 3, pp. 313–326, 1982
U. G. Haussmann and E. Pardoux, "Time reversal of diffusions," The Annals of Probability, pp. 1188–1205, 1986
P. Cattiaux, G. Conforti, I. Gentil, and C. Leonard,"Timereversalof diffusion processes under a finite entropy condition", 2021

Time-reversal of diffusions

- Let X and $\overleftarrow{X}$ be the solutions to

$$\mathrm{d}X_t = a(t, X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 \sim P_0$$

$$\mathrm{d}\overleftarrow{X}_t = a(t, \overleftarrow{X}_t)\mathrm{d}t + \sigma(\overleftarrow{X}_t)\overleftarrow{\mathrm{d}}W_t + b(t, \overleftarrow{X}_t)\mathrm{d}t, \quad \overleftarrow{X}_T \sim P_T$$

where P_t is the probability law of X_t and

$$b(t, x) = \frac{1}{P_t(x)}\partial_x(\sigma(x)\sigma(x)^\top P_t(x))$$

Theorem [Cattiaux, et. al., 2021]

If $D(x) := \sigma(x)\sigma(x)^\top$ is positive-definite everywhere, and a finite entropy condition holds, then

$$\{X_t; 0 \le t \le T\} \overset{\mathrm{d}}{=} \{\overleftarrow{X}_t; 0 \le t \le T\}$$

B. D. Anderson, "Reverse-time diffusion equation models," Stochastic Processes and their Applications, vol. 12, no. 3, pp. 313–326, 1982
U. G. Haussmann and E. Pardoux, "Time reversal of diffusions," The Annals of Probability, pp. 1188–1205, 1986
P. Cattiaux, G. Conforti, I. Gentil, and C. Leonard,"Timereversalof diffusion processes under a finite entropy condition", 2021

Time-reversal of diffusions

- Let X and $\overleftarrow{X}$ be the solutions to

$$\mathrm{d}X_t = a(t, X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \quad X_0 \sim P_0$$

$$\mathrm{d}\overleftarrow{X}_t = a(t, \overleftarrow{X}_t)\mathrm{d}t + \sigma(\overleftarrow{X}_t)\overleftarrow{\mathrm{d}}W_t + b(t, \overleftarrow{X}_t)\mathrm{d}t, \quad \overleftarrow{X}_T \sim P_T$$

where P_t is the probability law of X_t and

$$b(t, x) = \frac{1}{P_t(x)}\partial_x(\sigma(x)\sigma(x)^\top P_t(x))$$

Theorem [Cattiaux, et. al., 2021]

If $D(x) := \sigma(x)\sigma(x)^\top$ is positive-definite everywhere, and a finite entropy condition holds, then

$$\{X_t; 0 \leq t \leq T\} \overset{\mathsf{d}}{=} \{\overleftarrow{X}_t; 0 \leq t \leq T\}$$

B. D. Anderson, "Reverse-time diffusion equation models," Stochastic Processes and their Applications, vol. 12, no. 3, pp. 313–326, 1982
U. G. Haussmann and E. Pardoux, "Time reversal of diffusions," The Annals of Probability, pp. 1188–1205, 1986
P. Cattiaux, G. Conforti, I. Gentil, and C. Leonard,"Timereversalof diffusion processes under a finite entropy condition", 2021

Time-reversal formulation of the FBSDE

- Let Y and $\overleftarrow{Y}$ be the solutions to

$$-\mathrm{d}Y_t = g(t, X_t, Y_t, Z_t)\mathrm{d}t - Z_t \mathrm{d}W_t, \quad Y_T = g_f(X_T)$$

$$-\mathrm{d}\overleftarrow{Y}_t = g(t, \overleftarrow{X}_t, \overleftarrow{Y}_t, \overleftarrow{Z}_t)\mathrm{d}t - \overleftarrow{Z}_t \overleftarrow{\mathrm{d}} W_t + c(t, \overleftarrow{X}_t)\mathrm{d}t, \quad \overleftarrow{Y}_T = g_f(\overleftarrow{X}_T)$$

where

$$c(t,x) = \mathrm{tr}(D(x)\partial_{xx}\phi(t,x)) - \partial_x \phi(t,x)^\top b(t,x)$$

$$\phi(t,x) = \mathbb{E}[\overleftarrow{Y}_t | \overleftarrow{X}_t = x], \quad \overleftarrow{Z}_t = \sigma(\overleftarrow{X}_t)^\top \partial_x \phi(t, \overleftarrow{X}_t)$$

Theorem

If, additionally, the PDE $\partial_t \phi + \mathcal{L}\phi + g(t, x, \phi, \sigma(x)^\top \partial_x \phi) = 0$ has a smooth solution, then

$$\{(X_t, Y_t, Z_t); 0 \le t \le T\} \stackrel{\mathrm{d}}{=} \{(\overleftarrow{X}_t, \overleftarrow{Y}_t, \overleftarrow{Z}_t); 0 \le t \le T\}$$

Time-reversal formulation of the FBSDE

- Let Y and $\overleftarrow{Y}$ be the solutions to

$$-\mathrm{d}Y_t = g(t, X_t, Y_t, Z_t)\mathrm{d}t - Z_t \mathrm{d}W_t, \quad Y_T = g_f(X_T)$$

$$-\mathrm{d}\overleftarrow{Y}_t = g(t, \overleftarrow{X}_t, \overleftarrow{Y}_t, \overleftarrow{Z}_t)\mathrm{d}t - \overleftarrow{Z}_t \overleftarrow{\mathrm{d}}W_t + c(t, \overleftarrow{X}_t)\mathrm{d}t, \quad \overleftarrow{Y}_T = g_f(\overleftarrow{X}_T)$$

where

$$c(t,x) = \mathsf{tr}(D(x)\partial_{xx}\phi(t,x)) - \partial_x\phi(t,x)^\top b(t,x)$$

$$\phi(t,x) = \mathbb{E}[\overleftarrow{Y}_t | \overleftarrow{X}_t = x], \quad \overleftarrow{Z}_t = \sigma(\overleftarrow{X}_t)^\top \partial_x \phi(t, \overleftarrow{X}_t)$$

Theorem

If, additionally, the PDE $\partial_t\phi + \mathcal{L}\phi + g(t, x, \phi, \sigma(x)^\top \partial_x \phi) = 0$ has a smooth solution, then

$$\{(X_t, Y_t, Z_t); 0 \le t \le T\} \stackrel{\mathrm{d}}{=} \{(\overleftarrow{X}_t, \overleftarrow{Y}_t, \overleftarrow{Z}_t); 0 \le t \le T\}$$

Time-reversal formulation of the FBSDE

- Let Y and $\overleftarrow{Y}$ be the solutions to

$$-\mathrm{d}Y_t = g(t, X_t, Y_t, Z_t)\mathrm{d}t - Z_t \mathrm{d}W_t, \quad Y_T = g_f(X_T)$$

$$-\mathrm{d}\overleftarrow{Y}_t = g(t, \overleftarrow{X}_t, \overleftarrow{Y}_t, \overleftarrow{Z}_t)\mathrm{d}t - \overleftarrow{Z}_t \overleftarrow{\mathrm{d}}W_t + c(t, \overleftarrow{X}_t)\mathrm{d}t, \quad \overleftarrow{Y}_T = g_f(\overleftarrow{X}_T)$$

where

$$c(t,x) = \mathsf{tr}(D(x)\partial_{xx}\phi(t,x)) - \partial_x\phi(t,x)^\top b(t,x)$$

$$\phi(t,x) = \mathbb{E}[\overleftarrow{Y}_t | \overleftarrow{X}_t = x], \quad \overleftarrow{Z}_t = \sigma(\overleftarrow{X}_t)^\top \partial_x \phi(t, \overleftarrow{X}_t)$$

Theorem

If, additionally, the PDE $\partial_t\phi + \mathcal{L}\phi + g(t, x, \phi, \sigma(x)^\top \partial_x\phi) = 0$ has a smooth solution, then

$$\{(X_t, Y_t, Z_t); 0 \le t \le T\} \overset{\mathsf{d}}{=} \{(\overleftarrow{X}_t, \overleftarrow{Y}_t, \overleftarrow{Z}_t); 0 \le t \le T\}$$

Iterative approach to find the optimal control

1. Forward Monte-Carlo simulations and score function approximation:

$$\mathrm{d}X_t^i = a(X_t^i, U_t^i)\mathrm{d}t + \sigma(X_t^i)\mathrm{d}W_t^i, \quad X_0^i \sim P_0$$

$$\min_b \mathbb{E}\left[\frac{1}{2}\|b(t,X_t)\|^2 + \mathsf{Tr}(D(X_t)\partial_x b(t,X_t))\right]$$

2. Simulation of time-reversed FBSDE

$$\mathrm{d}\overleftarrow{X}_t^i = a(\overleftarrow{X}_t^i, \overleftarrow{U}_t^i)\mathrm{d}t + \sigma(\overleftarrow{X}_t^i)\overleftarrow{\mathrm{d}}W_t^i + b(t, X_t^i)\mathrm{d}t, \quad \overleftarrow{X}_T^i \sim P_T$$

$$-\mathrm{d}\overleftarrow{Y}_t^i = \partial_x H(\overleftarrow{X}_t^i, \overleftarrow{U}_t^i, \overleftarrow{Y}_t^i, \overleftarrow{Z}_t^i)\mathrm{d}t - \overleftarrow{Z}_t^i \overleftarrow{\mathrm{d}}W_t^i + c(t, \overleftarrow{X}_t^i)\mathrm{d}t, \quad \overleftarrow{Y}_T^i = g_f(\overleftarrow{X}_T^i)$$

$$\min_\phi \mathbb{E}\left[\|\overleftarrow{Y}_t - \phi(t, \overleftarrow{X}_t)\|^2\right], \quad \overleftarrow{Z}_t^i = \sigma(\overleftarrow{X}_t^i)^\top \partial_x \phi(t, \overleftarrow{X}_t^i)$$

3. Control update

$$U_t^i \leftarrow U_t^i - \eta \partial_u H(X_t^i, U_t^i, Y_t^i, Z_t^i)$$
$$\overleftarrow{U}_t^i \leftarrow \overleftarrow{U}_t^i - \eta \partial_u H(\overleftarrow{X}_t^i, \overleftarrow{U}_t^i, \overleftarrow{Y}_t^i, \overleftarrow{Z}_t^i)$$

Iterative approach to find the optimal control

1. Forward Monte-Carlo simulations and score function approximation:

$$\mathrm{d}X_t^i = a(X_t^i, U_t^i)\mathrm{d}t + \sigma(X_t^i)\mathrm{d}W_t^i, \quad X_0^i \sim P_0$$

$$\min_b \mathbb{E}\left[\frac{1}{2}\|b(t,X_t)\|^2 + \mathsf{Tr}(D(X_t)\partial_x b(t,X_t))\right]$$

2. Simulation of time-reversed FBSDE

$$\mathrm{d}\overleftarrow{X}_t^i = a(\overleftarrow{X}_t^i, \overleftarrow{U}_t^i)\mathrm{d}t + \sigma(\overleftarrow{X}_t^i)\overleftarrow{\mathrm{d}}W_t^i + b(t, X_t^i)\mathrm{d}t, \quad \overleftarrow{X}_T^i \sim P_T$$

$$-\mathrm{d}\overleftarrow{Y}_t^i = \partial_x H(\overleftarrow{X}_t^i, \overleftarrow{U}_t^i, \overleftarrow{Y}_t^i, \overleftarrow{Z}_t^i)\mathrm{d}t - \overleftarrow{Z}_t^i\overleftarrow{\mathrm{d}}W_t^i + c(t, \overleftarrow{X}_t^i)\mathrm{d}t, \quad \overleftarrow{Y}_T^i = g_f(\overleftarrow{X}_T^i)$$

$$\min_\phi \mathbb{E}\left[\|\overleftarrow{Y}_t - \phi(t, \overleftarrow{X}_t)\|^2\right], \quad \overleftarrow{Z}_t^i = \sigma(\overleftarrow{X}_t^i)^\top \partial_x\phi(t, \overleftarrow{X}_t^i)$$

3. Control update

$$U_t^i \leftarrow U_t^i - \eta\partial_u H(X_t^i, U_t^i, Y_t^i, Z_t^i)$$
$$\overleftarrow{U}_t^i \leftarrow \overleftarrow{U}_t^i - \eta\partial_u H(\overleftarrow{X}_t^i, \overleftarrow{U}_t^i, \overleftarrow{Y}_t^i, \overleftarrow{Z}_t^i)$$

Iterative approach to find the optimal control

1. Forward Monte-Carlo simulations and score function approximation:

$$\mathrm{d}X_t^i = a(X_t^i, U_t^i)\mathrm{d}t + \sigma(X_t^i)\mathrm{d}W_t^i, \quad X_0^i \sim P_0$$

$$\min_b \mathbb{E}\left[\frac{1}{2}\|b(t, X_t)\|^2 + \mathsf{Tr}(D(X_t)\partial_x b(t, X_t))\right]$$

2. Simulation of time-reversed FBSDE

$$\mathrm{d}\overleftarrow{X}_t^i = a(\overleftarrow{X}_t^i, \overleftarrow{U}_t^i)\mathrm{d}t + \sigma(\overleftarrow{X}_t^i)\overleftarrow{\mathrm{d}}W_t^i + b(t, X_t^i)\mathrm{d}t, \quad \overleftarrow{X}_T^i \sim P_T$$

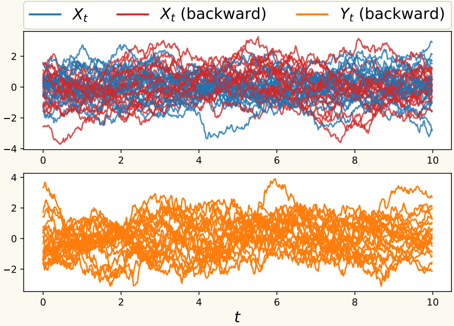
$$-\mathrm{d}\overleftarrow{Y}_t^i = \partial_x H(\overleftarrow{X}_t^i, \overleftarrow{U}_t^i, \overleftarrow{Y}_t^i, \overleftarrow{Z}_t^i)\mathrm{d}t - \overleftarrow{Z}_t^i\overleftarrow{\mathrm{d}}W_t^i + c(t, \overleftarrow{X}_t^i)\mathrm{d}t, \quad \overleftarrow{Y}_T^i = g_f(\overleftarrow{X}_T^i)$$

$$\min_\phi \mathbb{E}\left[\|\overleftarrow{Y}_t - \phi(t, \overleftarrow{X}_t)\|^2\right], \quad \overleftarrow{Z}_t^i = \sigma(\overleftarrow{X}_t^i)^\top \partial_x \phi(t, \overleftarrow{X}_t^i)$$

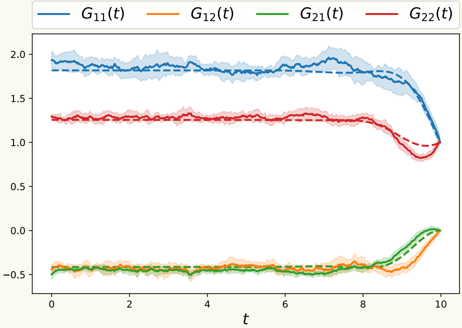
3. Control update

$$U_t^i \leftarrow U_t^i - \eta\partial_u H(X_t^i, U_t^i, Y_t^i, Z_t^i)$$
$$\overleftarrow{U}_t^i \leftarrow \overleftarrow{U}_t^i - \eta\partial_u H(\overleftarrow{X}_t^i, \overleftarrow{U}_t^i, \overleftarrow{Y}_t^i, \overleftarrow{Z}_t^i)$$

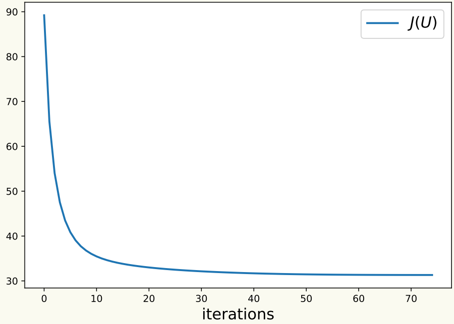
- Trajectories
- Accuracy in estimating $\phi(t,x) = G_t x$
- Convergence of the cost



- Trajectories
- Accuracy in estimating $\phi(t, x) = G_t x$
- Convergence of the cost



- Trajectories
- Accuracy in estimating $\phi(t,x) = G_t x$
- Convergence of the cost



Concluding remarks

- Convergence analysis $\rightarrow$ simpler compared to optimization on policy space
- Significant increase in accuracy compared to existing approach $\rightarrow$ submitted to ACC
- Extension to nonlinear setting $\rightarrow$ use of neural networks to represent b and ϕ
- Incorporating additional constraints
- Application to mean-field control/control of prob. dist.