RBM-OC-SVRG : A random sampling-based solver for optimal control problems with variance reduction

Byungjoon Lee
The Catholic University of Korea (CUK)

2024.06.28. FAU

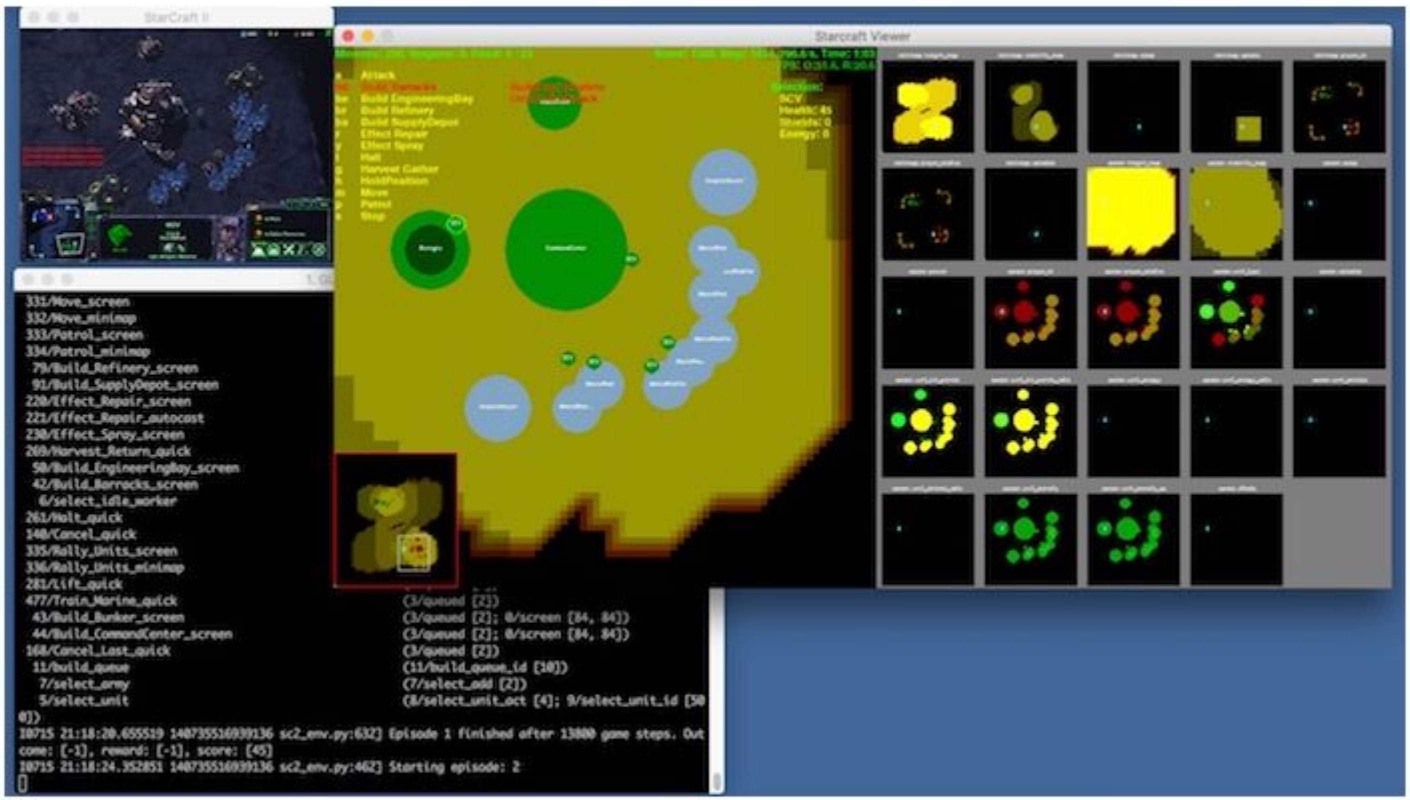
Outline

1. Introduction
2. Optimal control(OC) problem
3. RBM-OC-SVRG
4. Numerical results

1. Introduction

Reinforcement learning (RL)

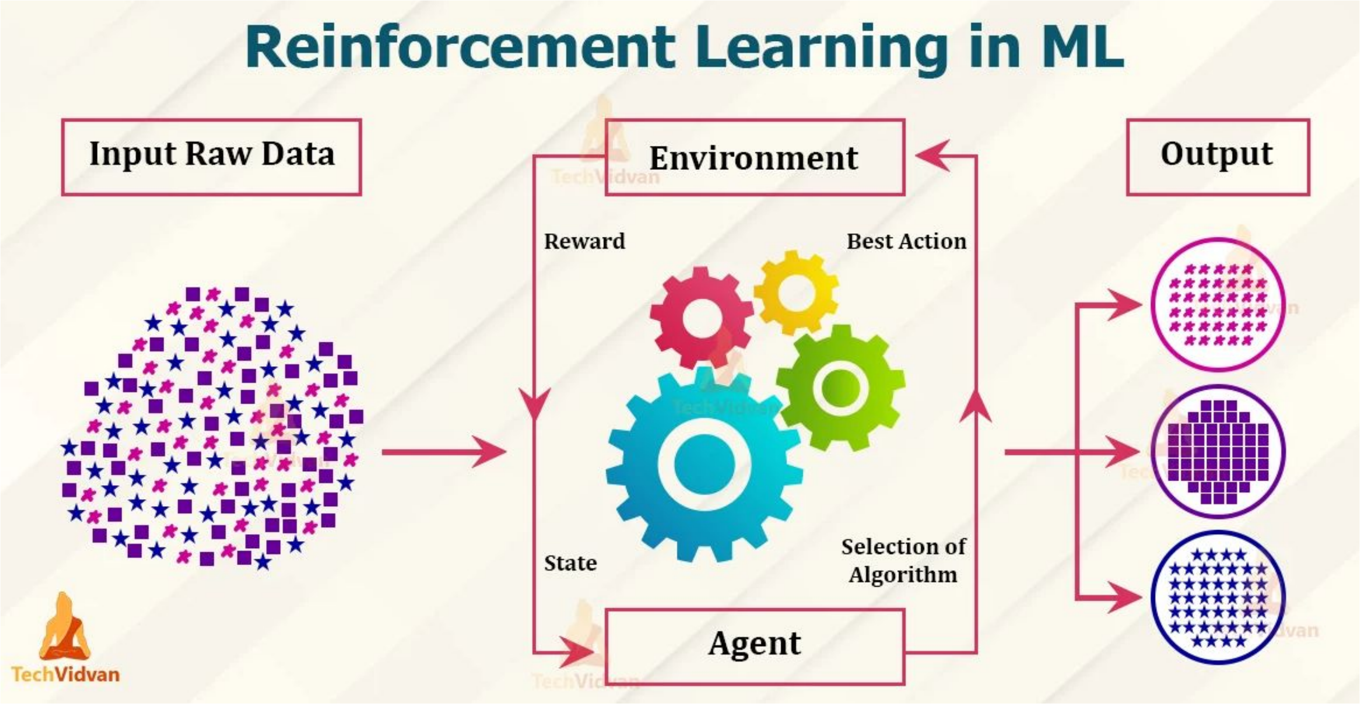
- Starcraft via reinforcement learning [1]



[1] https://bennycheung.github.io/adventures-in-deep-reinforcement-learning

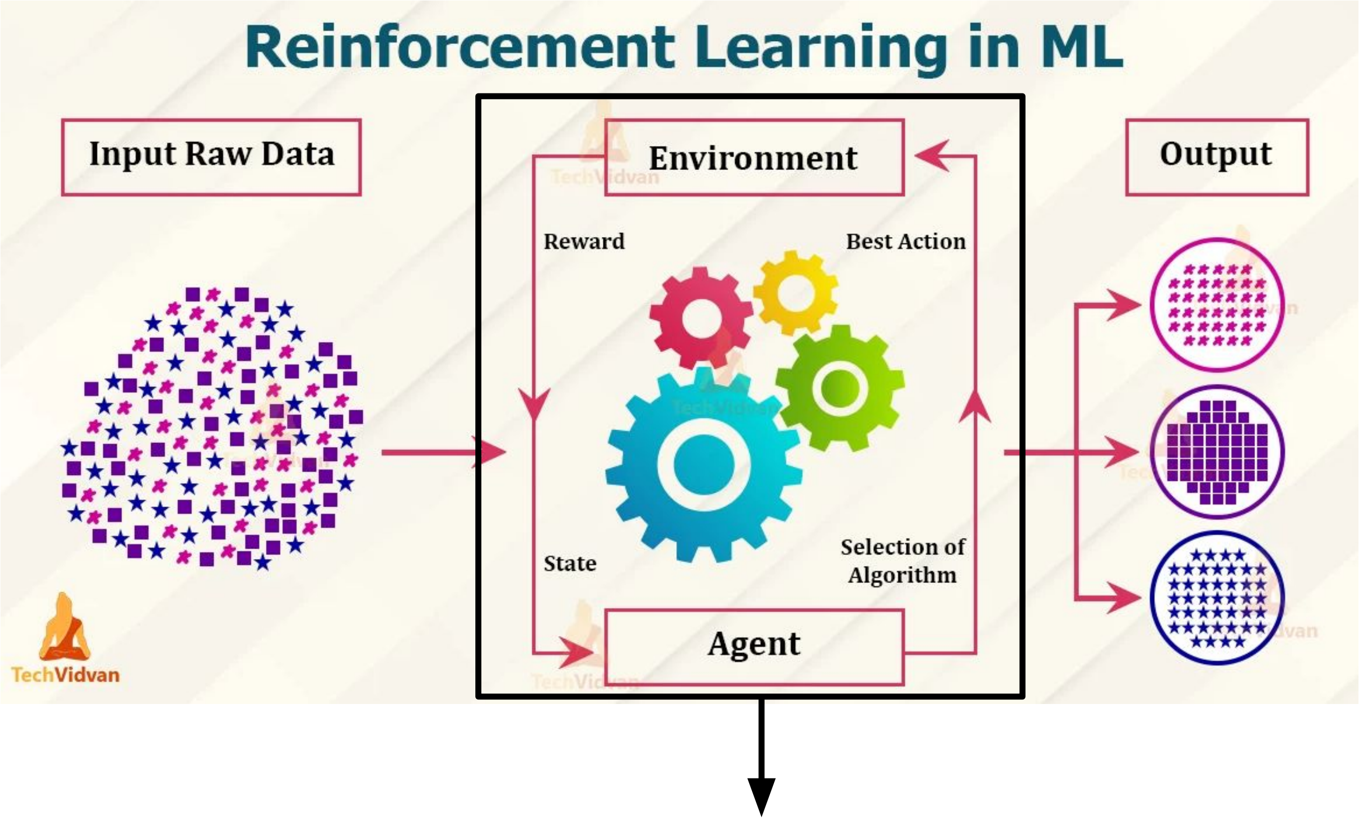
Reinforcement learning (RL)

- Reinforcement learning=find optimal strategies [2]



[2] https://techvidvan.com/tutorials/reinforcement-learning/

RL = Optimal control problem



Optimal control ☺

$$\text{maximize}_w \quad \text{Reward}(x, w)$$
$$\text{subject to} \quad x_{t+1} = f(x_t, w)$$

RL = Large-scale OC

- RL in reality : Tons of states
  - ex) Tetris(4X5) = 2^{20}

- Homework : How to solve large-scale OC

- **Aim of this talk** : Random batch method(RBM) based approach for OC to solve large-scale OC

2. Optimal control problem

Optimal control problem

- Cost functional

$$J\left[u\left(t\right)\right] := \int_{t}^{T} l\left(x\left(s\right), u\left(s\right)\right) ds + g\left(x\left(T\right)\right)$$

- **Optimal control problem**

$$\text{minimize } J\left[u\left(t\right)\right]$$
$$\text{subject to } \begin{cases} \dot{x}\left(s\right) & = f\left(x\left(s\right), u\left(s\right)\right) \\ x\left(t\right) & = x \end{cases}, t \leq s \leq T$$

Pontryagin principle

- Pontryagin principle : Perturbation analysis

$$J_\epsilon\left[u(t)+\epsilon v(t)\right]=\int_t^T l\left(x_\epsilon(s),u(s)+\epsilon v(s)\right)ds+g\left(x(T)\right)$$

$$\dot{x}_\epsilon(s)=f\left(x_\epsilon(s),u(s)+\epsilon v(s)\right)$$

$$z(t):=\frac{d}{d\epsilon}\Big|_{\epsilon=0}x_\epsilon(t)\ ,\ \ z(0)=0$$

$$z'(t)=\frac{\partial f}{\partial x}\left(x(t),u(t)\right)\cdot z(t)+\frac{\partial f}{\partial u}\left(x(t),u(t)\right)\cdot v(t)$$

Pontryagin principle

- Directional derivative of cost function

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} J_\epsilon \left[u(t) + \epsilon v(t)\right] = \int_t^T \left(\frac{\partial l}{\partial x}\left(x(s), u(s)\right) \cdot z(s) + \frac{\partial l}{\partial u}\left(x(s), u(s)\right) \cdot v(s) \right) ds$$

Direction to be determined!

- Aim : Choose the direction $v(s)$ minimizes the cost most!

Adjoint system!

Adjoint system

- Hint : Gradient flow

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} J_\epsilon\left[u(t)+\epsilon v(t)\right] = \int_t^T \square \cdot v(s)\,ds \quad \Rightarrow \quad v(s) = -\square$$

- Adjoint system

$$\begin{cases} p'(t) &= -\frac{\partial f}{\partial x}(x(t), u(t)) \cdot p(t) - \frac{\partial l}{\partial x}(x(t), u(t)) \\ p(T) &= 0 \end{cases}$$

Adjoint system

- Key : Integration-by-part

$$p'(t) = -\frac{\partial f}{\partial x}(x(t),u(t)) \cdot p(t) - \frac{\partial l}{\partial x}(x(t),u(t))$$

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} J_\epsilon[u(t)+\epsilon v(t)] = \int_t^T \left(\frac{\partial l}{\partial x}(x(s),u(s)) \cdot z(s) + \frac{\partial l}{\partial u}(x(s),u(s)) \cdot v(s)\right) ds$$

$$= \int_t^T \left(-p'(s)\, z(s) - \frac{\partial f}{\partial x}(x(s),u(s))\, p(s)\, z(s) \right.$$

$$[-p(s) z$$

$$v(s) = -p(s)\frac{\partial f}{\partial u}(x(s),u(s)) - \frac{\partial l}{\partial u}(x(s),u(s))$$

$$= \int_t^T \left(p(s)\left[z'(s) - \frac{\partial f}{\partial x}(x(s),u(s))\, z(s)\right] + \frac{\partial l}{\partial u}(x(s),u(s))\, v(s)\right) ds$$

$$= \int_t^T \left(\frac{\partial f}{\partial u}(x(s),u(s))\, p(s) + \frac{\partial l}{\partial u}(x(s),u(s))\right) v(s)\, ds$$

$$z'(t) = \frac{\partial f}{\partial x}(x(t),u(t)) \cdot z(t) + \frac{\partial f}{\partial u}(x(t),u(t)) \cdot v(t)$$

Optimal control solver

1. Forward dynamics

$$\begin{cases} \dot{x}(s) &= f(x(s), u(s)) \\ x(t) &= x \end{cases}, t \leq s \leq T$$

2. Backward dynamics (Adjoint system)

$$\begin{cases} p'(t) &= -\frac{\partial f}{\partial x}(x(t), u(t)) \cdot p(t) - \frac{\partial l}{\partial x}(x(t), u(t)) \\ p(T) &= 0 \end{cases}$$

3. Update

$$v(s) = -p(s)\frac{\partial f}{\partial u}(x(s), u(s)) - \frac{\partial l}{\partial u}(x(s), u(s))$$

$$u(s) \leftarrow u(s) + \epsilon v(s)$$

Discrete Optimal control

- Cost functional

$$J\left[u\left(t\right)\right] := \int_t^T l\left(x\left(s\right), u\left(s\right)\right) ds + g\left(x\left(T\right)\right)$$

$$J(u) = h\sum_{k=0}^{N-1} l(x^k, u^k) + g(x^N)$$

- Forward dynamics

$$\begin{cases} \dot{x}\left(s\right) & = f\left(x\left(s\right), u\left(s\right)\right) \\ x\left(t\right) & = x \end{cases}, t \leq s \leq T$$

$$\frac{x^{k+1} - x^k}{h} = f(x^k, u^k), \quad x^0 = x_0, \quad k = 0, 1, \ldots, N-1$$

Discrete Optimal control

- Adjoint system

$$\begin{cases} p'(t) &= -\frac{\partial f}{\partial x}(x(t), u(t)) \cdot p(t) - \frac{\partial l}{\partial x}(x(t), u(t)) \\ p(T) &= 0 \end{cases}$$

$$\frac{p^{k+1} - p^k}{h} = -\nabla_x l(x^k, u^k) - \nabla_x f(x^k, u^k)^\top p^{k+1}, \quad p^N = \nabla_x g(x^N)$$

- Gradient by

$$u_{\ell+1} \leftarrow u_\ell - \alpha \frac{\delta J}{\delta u}(u_\ell)$$

$$\frac{\delta J}{\delta u}(u) = h \sum_{k=0}^{N-1} \left[(\nabla_u f(x^k, u^k))^\top p^{k+1} + \nabla_u l(x^k, u^k) \right]$$

$$v(s) = -p(s) \frac{\partial f}{\partial u}(x(s), u(s)) - \frac{\partial l}{\partial u}(x(s), u(s))$$

3. RBM-OC-SVRG

Random batch method (RBM) [3]

- Random batch method (RBM) : Solve dynamics with **random batch**

$$dX^i = -\nabla V(X^i)\,dt + \frac{1}{N-1}\sum_{j:j\neq i} K(X^i - X^j)\,dt + \sigma\,dB^i,\ i = 1, \ldots, N.$$

Algorithm 1 (RBM-1).

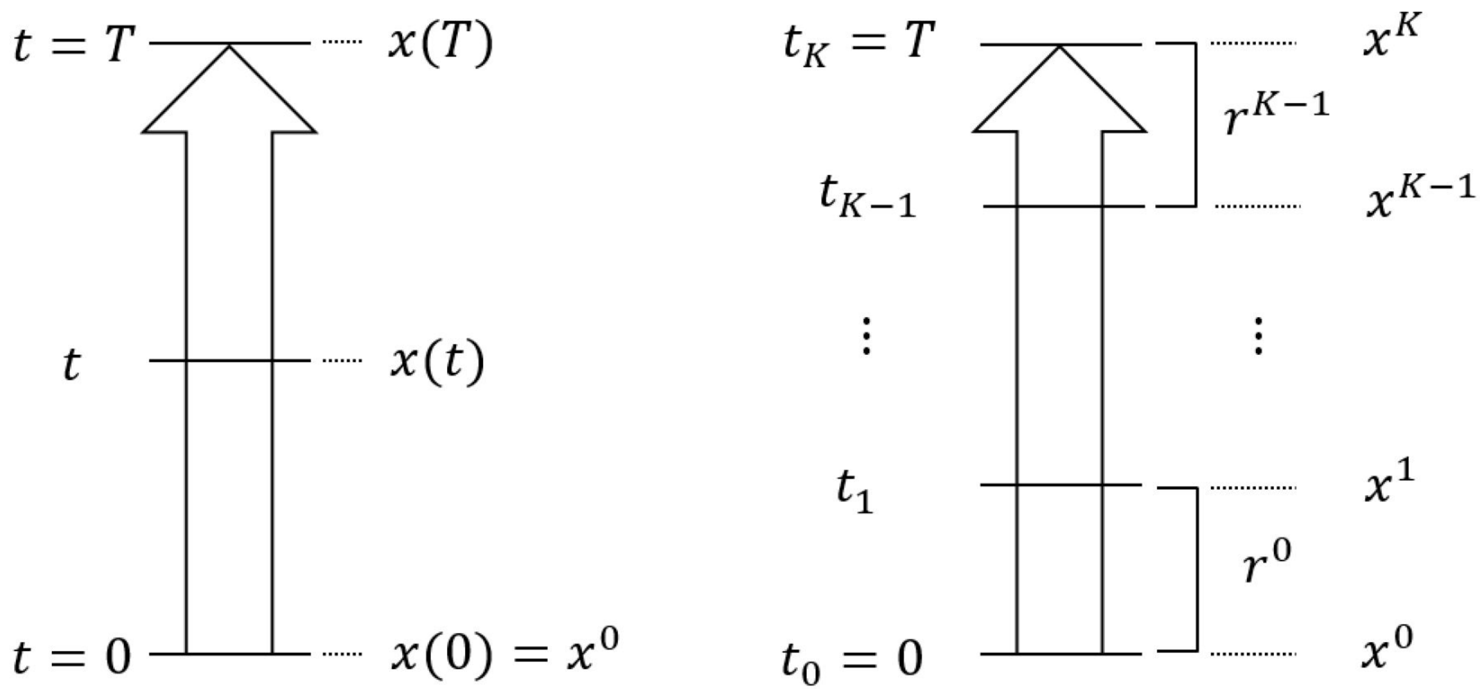
1: **for** m in $1 : [T/\tau]$ **do**
2: Divide $\{1, 2, \ldots, pn\}$ into n batches randomly.
3: **for** each batch $\mathcal{C}_q$ **do**
4: Update X^i's ($i \in \mathcal{C}_q$) by solving the following SDE with $t \in [t_{m-1}, t_m)$.

$$dX^i = -\nabla V(X^i)dt + \frac{1}{p-1}\sum_{j\in\mathcal{C}_q, j\neq i} K(X^i - X^j)dt + \sigma\,dB^i.$$

5: **end for**
6: **end for**

[3] S. Jin, L. Li, J.-G.Liu, Random Batch Methods (RBM) for interacting particle systems, Journal of Computational Physics, 2020

Random batch method (RBM) [3]



Dynamical system

$$x'(t) = f(x), 0 \leq t \leq T$$

Random Batch Method

$$x'(t) = f\left(x, r^k\right), t \in [t_k, t_{k+1}]$$

[3] S. Jin, L. Li, J.-G.Liu, Random Batch Methods (RBM) for interacting particle systems, Journal of Computational Physics, 2020

RBM+Optimal control?

- Optimal control : Forward dynamics + Adjoint system

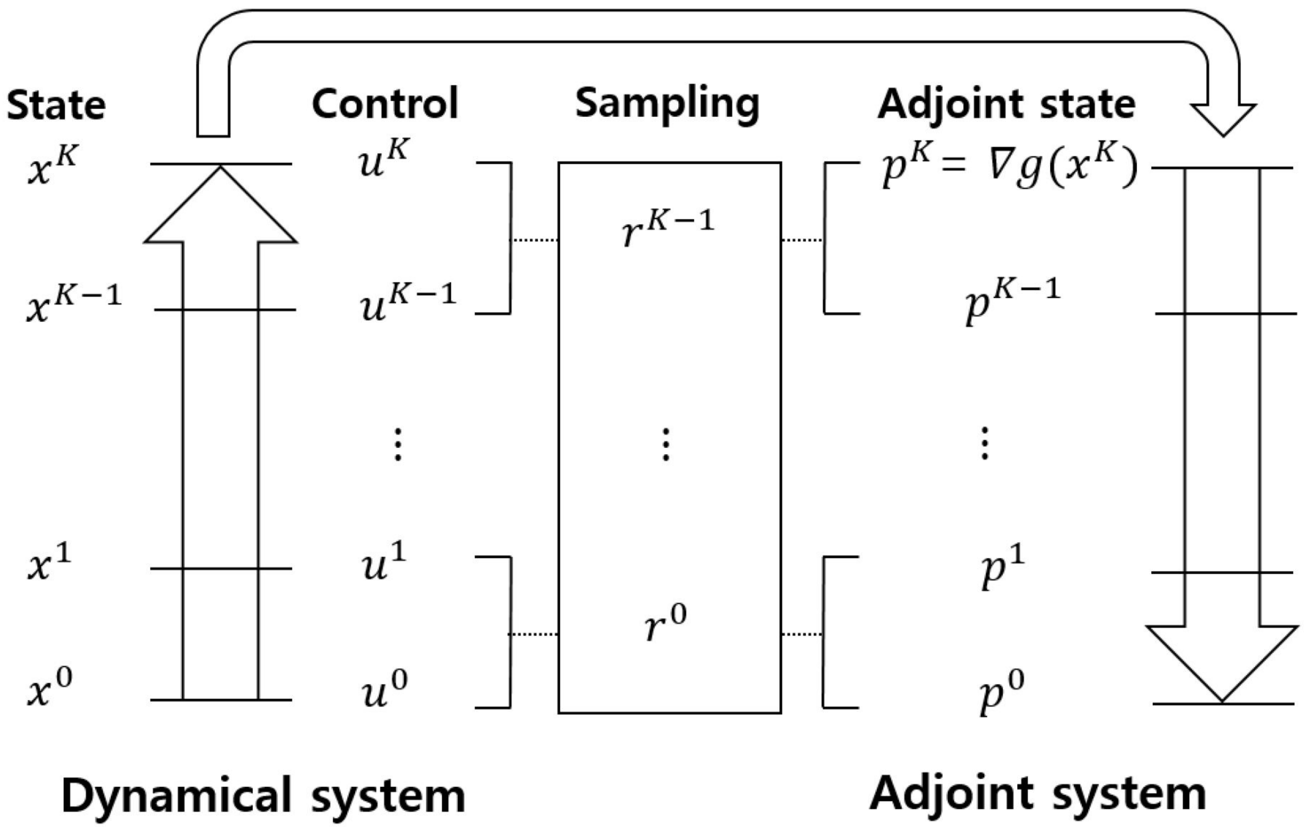
  RBM applicable ☺

- Attempts : RBM+Optimal control with a fixed sampling [4]

$$
\begin{aligned}
&\text{minimize } J_\omega(u,\omega) = h\sum_{k=0}^{N-1} l\left(x_\omega^k, u^k\right) + g\left(x_\omega^N\right), \\
&\text{subject to } \frac{x_\omega^{k+1} - x_\omega^k}{h} = f_\omega\left(x_\omega^k, u^k\right), \; k = 0, 1, \cdots, N-1, \\
&\qquad\qquad\quad x_\omega^0 = x_0,
\end{aligned}
$$

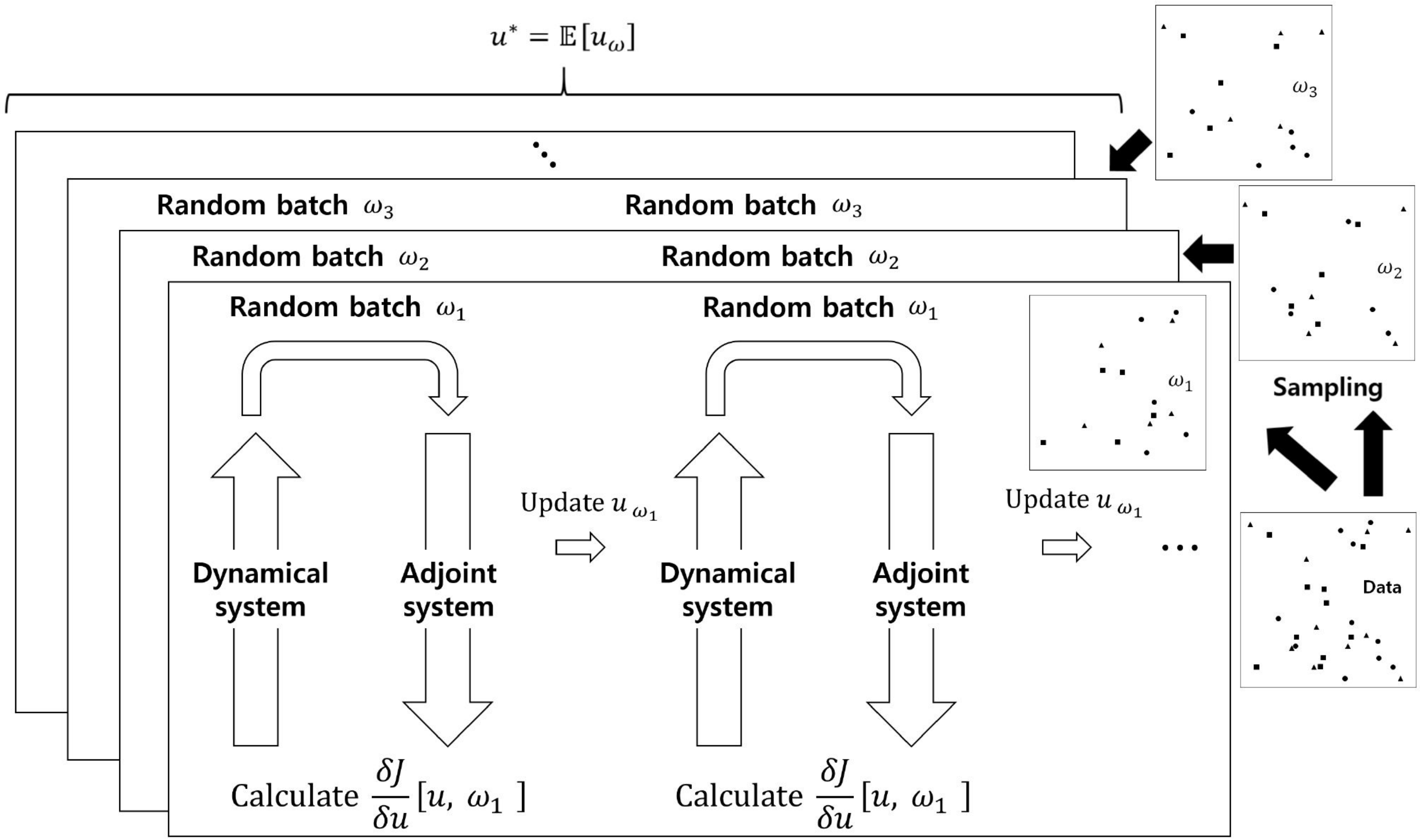
[4] D. Ko, E. Zuazua, Model predictive control with random batch methods for a guiding problem, Mathematical Models and Methods in Applied Sciences, 2021

RBM-OC [4]



[4] D. Ko, E. Zuazua, Model predictive control with random batch methods for a guiding problem, Mathematical Models and Methods in Applied Sciences, 2021

RBM-OC [4]



[4] D. Ko, E. Zuazua, Model predictive control with random batch methods for a guiding problem, Mathematical Models and Methods in Applied Sciences, 2021

RBM-OC [4]

- RBM-OC : Theoretical support when the dynamics are linear

- **Objective** : optimal to the approximated model

$$\mathbb{E}[u^*_\omega] = \mathbb{E}[\arg\min_u J_\omega(u,\omega)] \simeq \arg\min_u J(u) = u^*$$

[4] D. Ko, E. Zuazua, Model predictive control with random batch methods for a guiding problem, Mathematical Models and Methods in Applied Sciences, 2021

RBM-OC in Stochastic way [5]

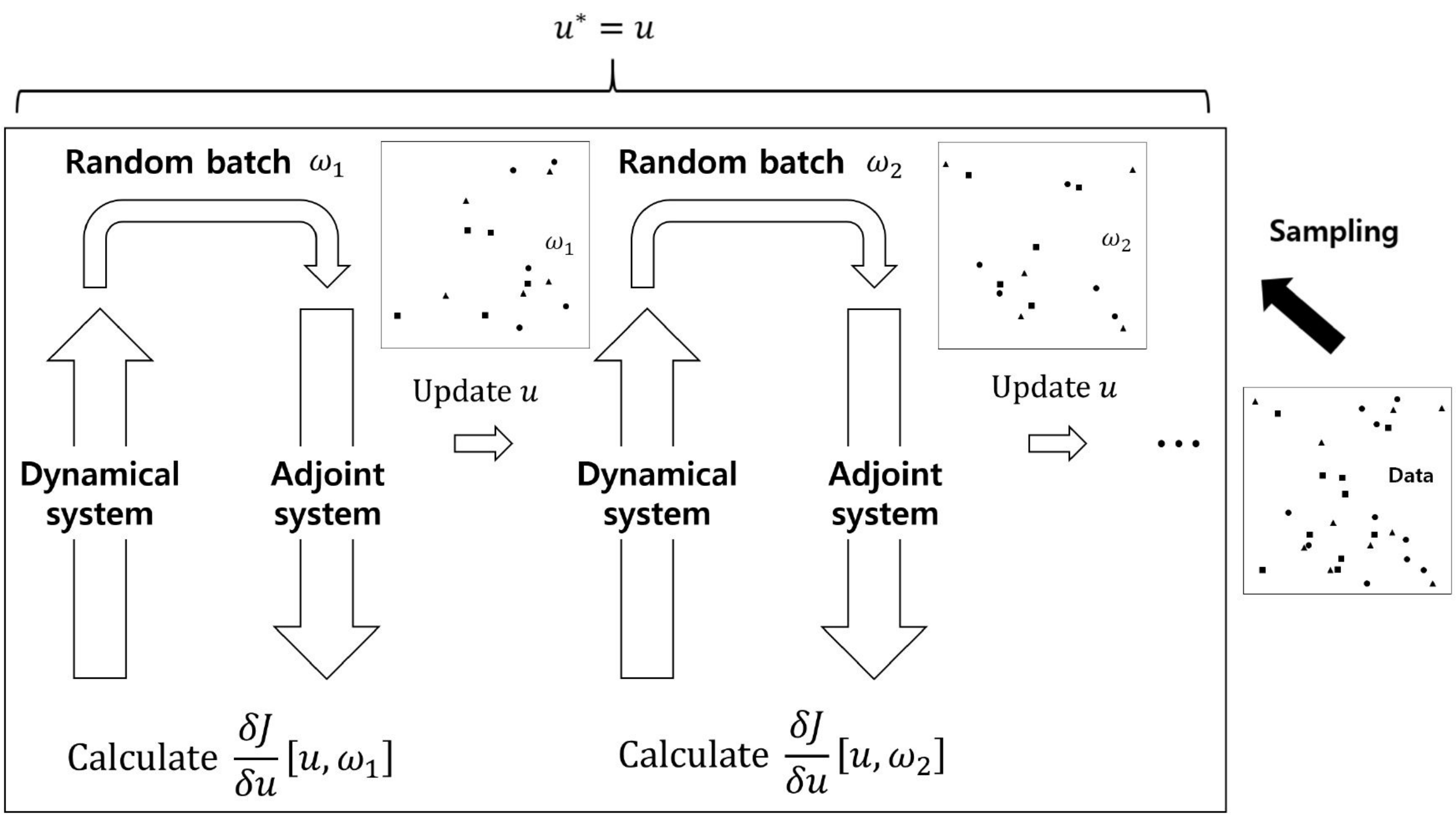
- Lightbulb : Do the **fully random sampling** instead of “fixing” sampling

- Newly proposed method : **RBM+OC+SGD**

$$u^*_{RBM} = \arg\min_u \mathbb{E}[J_\omega(u,\omega)] \simeq \arg\min_u J(u) = u^*$$

[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

RBM-OC in Stochastic way [5]



[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

RBM-OC in Stochastic way [5]

Algorithm 2 RBM-OC-SGD

Input: Initial data x^0, initial control $u_0 = \{u_0^k\}_{k=0}^{K-1}$, step size η

Output: Optimal control $u^* = \{u^{*,k}\}_{k=0}^{K-1}$

Repeat $\ell = 0, 1, \cdots$ **do**

1. Generate the random sampling $\omega_\ell = (r_\ell^0, \ldots, r_\ell^{K-1})$.
2. $\frac{\delta J}{\delta u}[u_\ell, \omega_\ell] \leftarrow \text{ComputeGradient}\left(x^0, u_\ell, \omega_\ell\right)$
3. $u_{\ell+1} \leftarrow u_\ell - \eta \frac{\delta J}{\delta u}[u_\ell, \omega]$

until converges

Algorithm 1 (RBM-1).

for m in $1 : [T/\tau]$ **do**
  Divide $\{1, 2, \ldots, pn\}$ into n batches randomly.
  for each batch $\mathcal{C}_q$ **do**
    Update X^i's ($i \in \mathcal{C}_q$) by solving the following SDE with $t \in [t_{m-1}, t_m)$.

$$dX^i = -\nabla V(X^i)dt + \frac{1}{p-1} \sum_{j \in \mathcal{C}_q, j \neq i} K(X^i - X^j)dt + \sigma dB^i.$$

  end for
end for

[5] J. Kim, D. Ko, C.
problems using ra

SGD+Variance reduction [6]

- **Stochastic variance reduced gradient (SVRG)**

$$\min P(w), \qquad P(w) := \frac{1}{n}\sum^{n} \psi_i(w)$$

Theorem 1. *Consider SVRG in Figure 1 with option II. Assume that all ψ_i are convex and both (5) and (6) hold with $\gamma > 0$. Let $w_* = \arg\min_w P(w)$. Assume that m is sufficiently large so that*

$$\alpha = \frac{1}{\gamma\eta(1-2L\eta)m} + \frac{2L\eta}{1-2L\eta} < 1,$$

then we have geometric convergence in expectation for SVRG:

$$\mathbb{E}\, P(\tilde{w}_s) \leq \mathbb{E}\, P(w_*) + \alpha^s [P(\tilde{w}_0) - P(w_*)]$$

Randomly pick $i_t \in \{1, \ldots, n\}$ and update weight
$w_t = w_{t-1} - \eta(\nabla\psi_{i_t}(w_{t-1}) - \nabla\psi_{i_t}(\tilde{w}) + \tilde{\mu})$
end
option I: set $\tilde{w}_s = w_m$
option II: set $\tilde{w}_s = w_t$ for randomly chosen $t \in \{0, \ldots, m-1\}$
end

[6] R. Johnson, T. Zhang, Accelerating Stochastic Gradient Descent using Predictive Variance Reduction, NeurIPS, 2013

RBM-OC-SGD+Variance reduction

- **RBM-OC-SGD + Variance reduction**

$$\min P(w), \qquad P(w) := \frac{1}{n}\sum_{i=1}^{n} \psi_i(w)$$

$$\text{minimize} \;\; J(u) := \mathbb{E}[J_\omega(u,\omega)]$$

RBM-OC-SVRG [5]

- **RBM-OC-SVRG**

Algorithm 3 RBM-OC-SVRG

Input: Initial data x_0, initial control $\tilde{u}_0 = \{\tilde{u}_0^k\}_{k=0}^{K-1}$, step size η, iteration number m

Output: Optimal control $u^* = \{u^{*,k}\}_{k=0}^{K-1}$

Repeat $\ell = 0, 1, \cdots$ **do**

1. $\tilde{u} \leftarrow \tilde{u}_\ell$
2. Compute $\mathbb{E}\left[\frac{\delta J}{\delta u}[\tilde{u}, \omega]\right]$ over random samplings ω.
3. $u_0 \leftarrow \tilde{u}$
4. Generate the random sampling $\omega_\ell = (r_\ell^0, \ldots, r_\ell^{K-1})$.

   Repeat $t = 0$ **to** $m - 1$ **do**

   5. $\frac{\delta J}{\delta u}[u_t, \omega_\ell] \leftarrow \text{ComputeGradient}(x_0, u_t, \omega_\ell)$
   6. $\frac{\delta J}{\delta u}[\tilde{u}, \omega_\ell] \leftarrow \text{ComputeGradient}(x_0, \tilde{u}, \omega_\ell)$
   7. $u_{t+1} \leftarrow u_t - \eta\left(\frac{\delta J}{\delta u}[u_t, \omega_\ell] - \frac{\delta J}{\delta u}[\tilde{u}, \omega_\ell] + \mathbb{E}\left[\frac{\delta J}{\delta u}[\tilde{u}, \omega]\right]\right)$

   end repeat

8. $\tilde{u}_{\ell+1} \leftarrow u_t$ where t is randomly chosen from $\{0, 1, \cdots, m\}$.

until converges

[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

Several Theoretical supports [5]

- Main theorem (For LQR) : Thm2+Thm3

Theorem 1. *Let s be a fixed stage. Assume that the number of inner iterations m is sufficiently large so that*

$$\alpha = \frac{1}{\gamma\eta(1-2L\eta)m} + \frac{2L\eta}{1-2L\eta} < 1.$$

Then, there exists a positive constant C, which depends on $\widetilde{u}_s$ and u^* such that

$$\mathbb{E}_{\mathrm{VR}}\left[\mathbb{E}[J(\widetilde{u}_s) - J(u^*)]\right] \le C\sqrt{\mathrm{Var}(A_\omega)h} + \alpha^s\left[\overline{J}(\widetilde{u}_0) - \overline{J}(u^*_{RBM})\right]. \tag{13}$$

Here, $\mathbb{E}_{\mathrm{VR}}[\cdot]$ denotes the expectation of choosing $\widetilde{u}_s$ in the variance reduction process, while $\mathbb{E}[\cdot]$ denotes the expectation with respect to the realization $\omega = (r^0, r^1, \ldots, r^{N-1})$.

.

[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

Several Theoretical supports

- Thm2 (From SVRG [6])

Theorem 2. [15] *Let ℓ be a number of outer iterations in RBM-OC-SVRG. Assume that the number of inner iterations m is sufficiently large so that*

$$\alpha := \frac{1}{\gamma\eta(1-2L\eta)m} + \frac{2L\eta}{1-2L\eta} < 1.$$

Then,

$$\mathbb{E}_{\mathrm{VR}}\left[\mathbb{E}\left[\bar{J}[\tilde{u}_\ell] - \bar{J}[u^*_{RBM}]\right]\right] \le \alpha^\ell \left(\bar{J}[\tilde{u}_0] - \bar{J}[u^*_{RBM}]\right).$$

[6] R. Johnson, T. Zhang, Accelerating Stochastic Gradient Descent using Predictive Variance Reduction, NeurIPS, 2013

- Thm3 (From Zuazua [7])

Theorem 3. *For any control sequence $\{u^k\}_{k=0}^{K-1}$,*

$$|\bar{J}[u] - J[u]| = |\mathbb{E}[J[u,\omega]] - J[u]| \le C_2\sqrt{\mathrm{Var}(A_\omega)\Delta t},$$

where $C_2 = C_0\sqrt{C_1}(T\|Q\| + \|P\|)$.

[7] D. Veldman, E. Zuazua, A framework for randomized time-splitting in linear-quadratic optimal control, Numerische Mathematik, 2022

4. Numerical results

Example 1 : Linear case [5]

$$J(u) = \frac{h}{2}\sum_{k=0}^{N-1}(|x^k|^2 + |u^k|^2) + \frac{1}{2}|x^N|^2.$$

$$\frac{x^{k+1} - x^k}{h} = Ax^k + Bu^k$$

$$A = A_1 + A_2 + A_3 + A_4$$

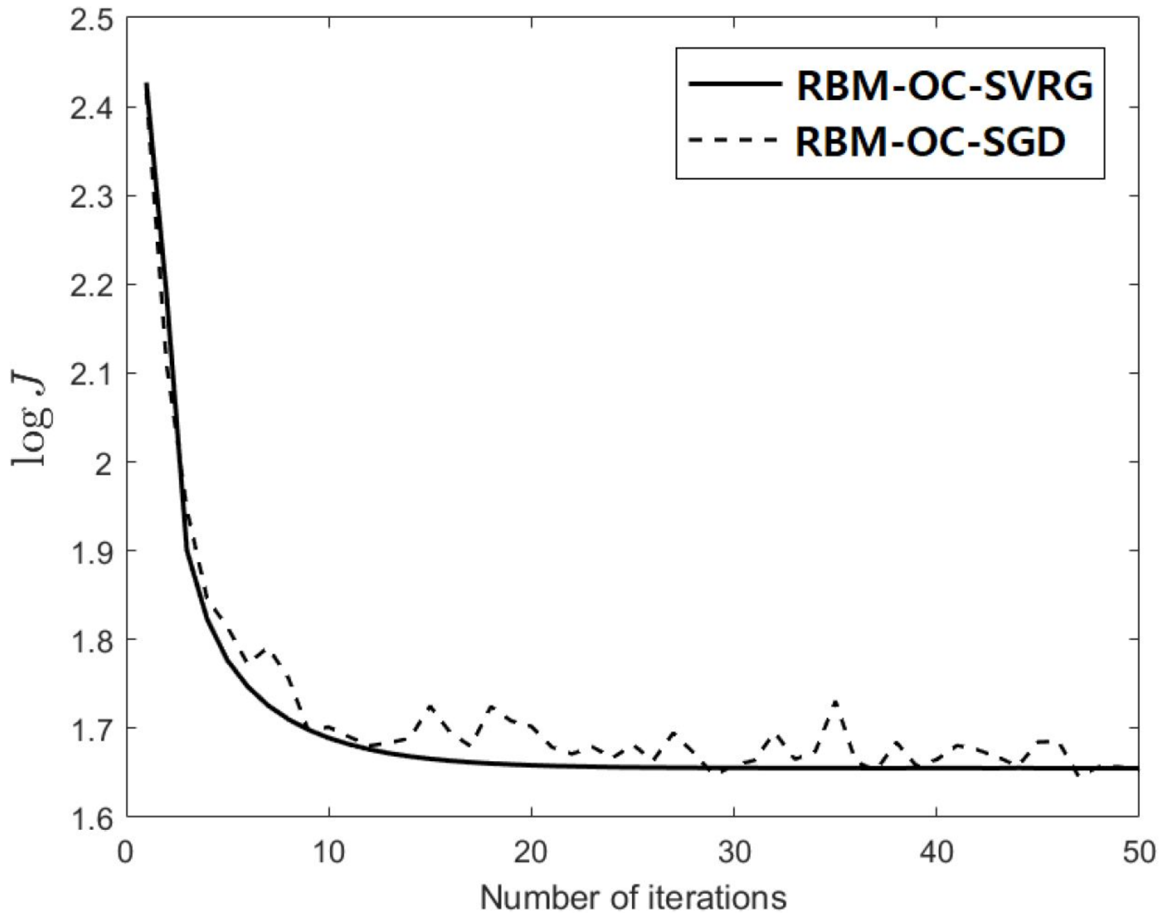
$$A = \begin{bmatrix} -2 & 2 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -2 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 2 & -2 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} A_{11} & 0_{16\times 45} \\ 0_{45\times 16} & 0_{45\times 45} \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0_{15\times 15} & 0_{15\times 16} & 0_{15\times 30} \\ 0_{16\times 15} & A_{22} & 0_{16\times 30} \\ 0_{30\times 15} & 0_{30\times 16} & 0_{30\times 30} \end{bmatrix},$$

$$A_3 = \begin{bmatrix} 0_{30\times 30} & 0_{30\times 16} & 0_{30\times 15} \\ 0_{16\times 30} & A_{33} & 0_{16\times 15} \\ 0_{15\times 30} & 0_{15\times 16} & 0_{15\times 15} \end{bmatrix}, \quad A_4 = \begin{bmatrix} 0_{45\times 45} & 0_{45\times 16} \\ 0_{16\times 45} & A_{44} \end{bmatrix},$$

[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

Example 1 : Linear case [5]



[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

Example 2 : Opinion dynamics [5]

$$J(u) = \frac{h}{2}\sum_{k=0}^{N-1}(|x^k|^2 + |u^k|^2) + \frac{1}{2}|x^N|^2$$

$$\frac{x_i^{k+1} - x_i^k}{h} = \frac{\alpha}{N}\sum_{j=1}^{n}\phi(x_j - x_i) + (Bu^k)_i, \quad B \in \mathbb{R}^{n\times m}, \quad u^k \in \mathbb{R}^m$$

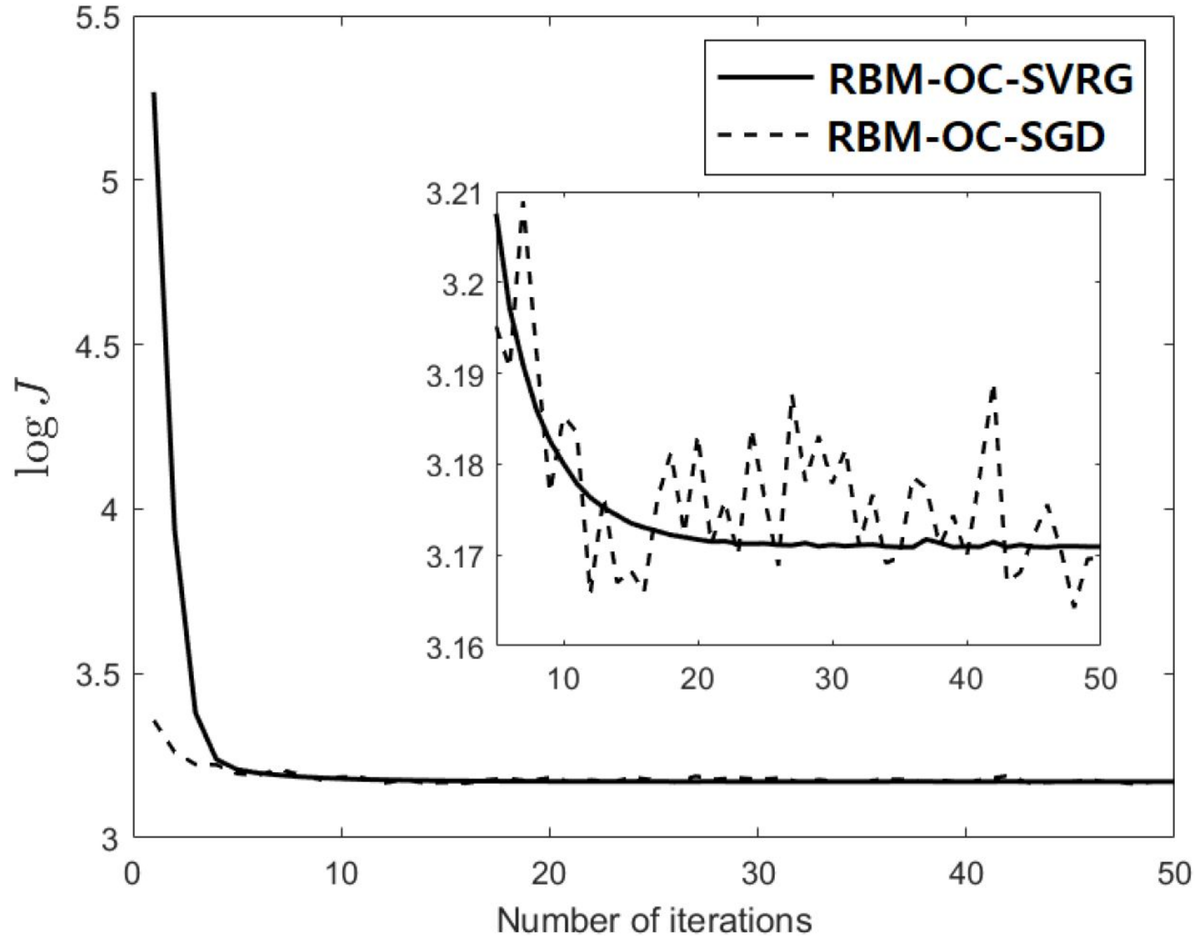
$$A = A_1 + A_2 + A_3 + A_4$$

$$A_1 = \frac{1}{N}\begin{bmatrix} 1_{N/2\times N/2} & 0_{N/2\times N/2} \\ 0_{N/2\times N/2} & 0_{N/2\times N/2} \end{bmatrix}, \quad A_2 = \frac{1}{N}\begin{bmatrix} 0_{N/2\times N/2} & 1_{N/2\times N/2} \\ 0_{N/2\times N/2} & 0_{N/2\times N/2} \end{bmatrix},$$

$$A_3 = \frac{1}{N}\begin{bmatrix} 0_{N/2\times N/2} & 0_{N/2\times N/2} \\ 1_{N/2\times N/2} & 0_{N/2\times N/2} \end{bmatrix}, \quad A_4 = \frac{1}{N}\begin{bmatrix} 0_{N/2\times N/2} & 0_{N/2\times N/2} \\ 0_{N/2\times N/2} & 1_{N/2\times N/2} \end{bmatrix},$$

[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

Example 2 : Opinion dynamics [5]



[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

Example 3 : Guiding problem [5]

[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

Example 3 : Guiding problem [5]

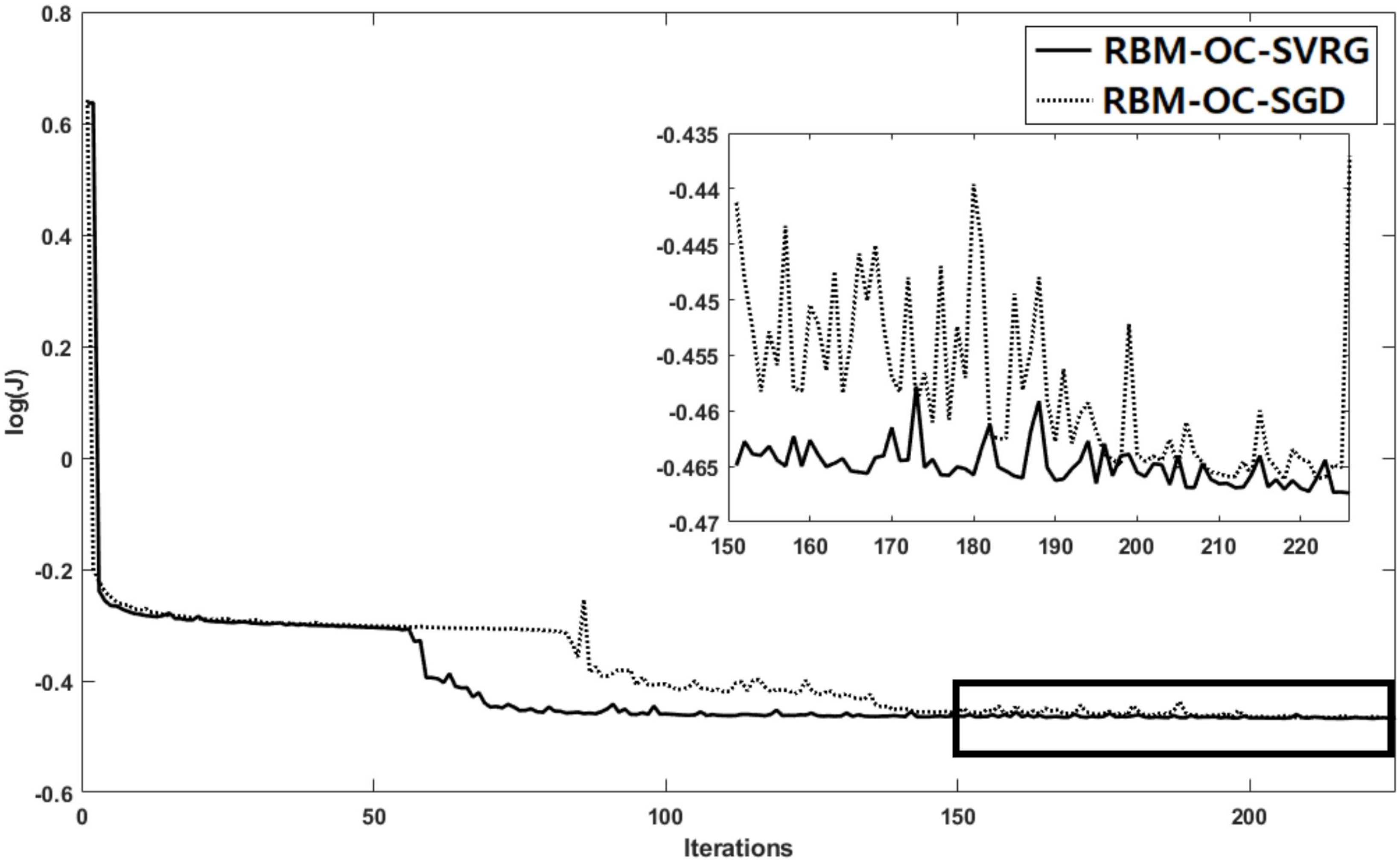
$$J(u) := h \sum_{t=1}^{T} \left[\frac{\alpha_1}{N} \sum_{k=1}^{N} \left| x_k^t - x_f \right|^2 + \frac{\alpha_2}{M} \sum_{j=1}^{M} \left| u_j^t \right|^2 + \frac{\alpha_3}{M} \sum_{j=1}^{M} \left| y_j^t - x_f \right|^2 \right]$$

$$\begin{cases} \frac{x_i^{k+1} - x_i^k}{h} & = v_i^k, \\ \frac{v_i^{k+1} - v_i^k}{h} & = \frac{1}{N-1} \sum_{l=1, l \neq i} a\left(x_l^k - x_i^k\right)\left(v_l^k - v_i^k\right) + \frac{1}{N-1} \sum_{l=1, l \neq i} g\left(x_l^k - x_i^k\right)\left(x_l^k - x_i^k\right) \\ & \quad - \frac{1}{M} \sum_{j=1}^{M} f\left(y_j^k - x_i^k\right)\left(y_j^k - x_i^k\right) \\ \frac{y_j^{k+1} - y_j^k}{h} & = u_j^k \end{cases}$$

$$a(x) := 1, \;\; f(x) := 4 \exp\left(-8 |x|^2\right) \;\text{ and }\; g(x) := \begin{cases} 2\left(1 - \frac{1}{3\sqrt{N}|x|^2}\right) & x \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

Example 3 : Guiding problem [5]



[5] J. Kim, D. Ko, C. Min, B. Lee, Random sampling-based gradient descent method for optimal control problems with variance reduction, preprint, 2024

4. Concluding remark

Optimal control problems

- Random batch method(RBM) based optimal control techniques are introduced.

- Theoretical supports for the linear-quadratic case are presented.

- Future work : Analysis for the case of nonlinear dynamics, Relation with RL

Discussion

Any Question?