Safe Non-Stochastic Control of Linear Dynamical Systems

Hongyu Zhou and Vasileios Tzoumas

MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

Drone Delivery

Goal: Deliver a package in a moving vehicle.

Complication: Unpredictable wind and wake disturbances.

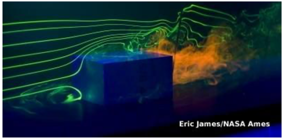
Problem: How can the drone choose collision-free optimal control actions?[1]

[1] Ackerman, IEEE Spectrum' 13

Drone Delivery

Goal: Deliver a package in a moving vehicle.

Complication: Unpredictable wind and wake disturbances.



Problem: How can the drone choose collision-free optimal control actions?[1]

[1] Ackerman, IEEE Spectrum' 13

Inspection and Maintenance

Goal: Inspect and repair facilities using onboard cameras.

Complication: Unpredictable wind disturbances.

Problem: How can the drone choose collision-free optimal control actions?[1]

[1] Seneviratne, Dammika, et al. Acta Imeko '18

Safe Non-Stochastic Control Problems

Target Tracking

Goal: Minimize distance to a target that moves in a cluttered environment.

Complication: Unpredictable wind disturbances.

Problem: How can the drone choose collision-free optimal control actions?[1]

[1] Chen, Liu, Shen, IROS' 16

All above scenarios are optimal control problems with safety constraints

Goal: Find control input to minimize loss subject to system dynamics and safety constraints:

$$\begin{aligned} \min_{\text{control input}} \quad & \text{loss} \qquad \textit{e.g.}\text{, tracking error} \\ \text{subject to} \quad & \text{system dynamics;} \\ & \text{safety constraints.} \end{aligned} \tag{1}$$

All above scenarios are optimal control problems with safety constraints

Goal: Find control input to minimize loss subject to system dynamics and safety constraints:

$$
\begin{aligned}
\min_{u_t} \quad & \sum_{t=1}^{T} c_t\left(x_{t+1}, u_t\right) \\
\text{subject to} \quad & x_{t+1} = A_t x_t + B_t u_t + w_t; \\
& x_t \in \mathcal{S}_t, \; u_t \in \mathcal{U}_t.
\end{aligned}
\tag{1}
$$

known system matrices

constraints

All above scenarios are optimal control problems with safety constraints

Goal: Find control input to minimize loss subject to system dynamics and safety constraints:

$$
\begin{aligned}
\min_{u_t} \quad & \sum_{t=1}^{T} c_t\left(x_{t+1}, u_t\right) \\
\text{subject to} \quad & x_{t+1} = A_t x_t + B_t u_t + w_t; \\
& x_t \in \mathcal{S}_t \triangleq \{x \mid L_{x,t} x \leq l_{x,t}\}, \\
& u_t \in \mathcal{U}_t \triangleq \{u \mid L_{u,t} u \leq l_{u,t}\}.
\end{aligned} \tag{1}
$$

known system matrices

polytopic constraints
for simplicity

All above scenarios are optimal control problems with safety constraints

Goal: Find control input to minimize loss subject to system dynamics and safety constraints:

$$
\begin{aligned}
\min_{u_t} \quad & \sum_{t=1}^{T} c_t\left(x_{t+1}, u_t\right) \\
\text{subject to} \quad & x_{t+1} = A_t x_t + B_t u_t + w_t; \\
& x_t \in \mathcal{S}_t \triangleq \{x \mid L_{x,t} x \leq l_{x,t}\}, \\
& u_t \in \mathcal{U}_t \triangleq \{u \mid L_{u,t} u \leq l_{u,t}\}.
\end{aligned}
\tag{1}
$$

known system matrices

polytopic constraints
for simplicity

Difficulty:

- The noise w_t can be unknown and unstructured, instead of Gaussian.

Limitation of Classical Control Approaches

Classical control approaches can be too optimistic or too pessimistic against unknown and unstructured noise

$\mathcal{H}_\infty$ (Assumes worst-case noise)	(No noise assumptions)	$\mathcal{H}_2$ (Assumes Gaussian noise)
Pessimistic	Ideal	Optimistic

Limitation of Classical Control Approaches

Classical control approaches can be too optimistic or too pessimistic against unknown and unstructured noise

$\mathcal{H}_\infty$	★	$\mathcal{H}_2$
(Assumes worst-case noise)	(No noise assumptions)	(Assumes Gaussian noise)
Pessimistic	Ideal	Optimistic

The ideal approach would give suboptimality guarantees against the optimal method in hindsight for any noise realization.

Suboptimality Metric against Optimal Control Policies in Hindsight

Definition (Dynamic Regret)

Assume a lookahead time horizon of operation T, and loss functions c_t, $t = 1, \ldots, T$. Then, dynamic regret is defined as

$$\text{Regret-NSC}_T^D = \sum_{t=1}^{T} c_t\left(x_{t+1}, u_t\right) - \sum_{t=1}^{T} c_t\left(x_{t+1}^*, u_t^*\right), \tag{2}$$

where x_t^* and u_t^* are the optimal trajectory and control input in hindsight given the noise realization due to $\{u_1, \ldots, u_T\}$.

Remark:

- The dynamic regret is sublinear if $\lim_{T\to\infty} \frac{\text{Regret-NSC}_T^D}{T} \to 0$, which implies $c_t\left(x_{t+1}, u_t\right) - c_t\left(x_{t+1}^*, u_t^*\right) \to 0$ as $T \to \infty$.

Problem

Assume the initial state of the system is safe, i.e., $x_0 \in \mathcal{S}_0$. *At each* $t = 1, \ldots, T$,

- *first a control input* $u_t \in \mathcal{U}_t$ *is chosen;*
- *then, a noise* $w_t \in \mathbb{R}^{d_x}$ *is revealed and the system evolves to state* $x_{t+1} \in \mathcal{S}_{t+1}$*;*
- *the controller suffers a loss* $c_t(x_{t+1}, u_t)$.

The goal is to guarantee $x_{t+1} \in \mathcal{S}_{t+1}$ *and* $u_t \in \mathcal{U}_t$ *for all* t *and that minimize*

$$\text{Regret-NSC}_T^D = \sum_{t=1}^{T} c_t\left(x_{t+1}, u_t\right) - \sum_{t=1}^{T} c_t\left(x_{t+1}^*, u_t^*\right).$$

Linear-Feedback control policy:

$u_t = -K_t x_t - K_t^s x_t$, where K_t is to be designed such that

$$\|K_t\| \leq \kappa, \longrightarrow \text{Compact domain set}$$
$$\|K_t x_t\| \leq \gamma, \longrightarrow \text{Bounded state}$$

given K_t^s that is sequentially stabilizing,[2] and desired $\kappa > 0$ and $\gamma > 0$.

[2] Gradu et al., L4DC '23

Assumptions

Assumption (Bounded Noise)

$w_t \in \mathcal{W} \triangleq \{w \mid \|w\| \leq W\}$ *where* W *is given.*

Remark:

- We assume no stochastic model for the process noise w_t: the noise may even be adversarial, subject to the bound W.

Example:

- Wind and wake disturbances of bounded magnitude, whose evolution may not be governed by a known stochastic model.

Assumptions

Assumption (Convex and Bounded Loss Function with Bounded Gradient)

The loss function $c_t(x_{t+1}, u_t) : \mathbb{R}^{d_x} \times \mathbb{R}^{d_u} \mapsto \mathbb{R}$ is convex in x_{t+1} and u_t. Further, when $\|x\|$ and $\|u\|$ are bounded, then $|c_t(x, u)|$, $\|\nabla_x c_t(x, u)\|$, and $\|\nabla_u c_t(x, u)\|$ are also bounded.

Example:

- Quadratic cost $c_t\left(x_{t+1}, u_t\right) = x_{t+1} Q x_{t+1}^\top + u_t R u_t^\top$.

Remark:

- The above are standard assumptions in the literature of non-stochastic control.[3]

[3] Agarwal et al., ICML '19; Hazan et al., ALT '20; Li et al., AAAI '21; Zhao et al., AISTATS '22; Gradu et al., L4DC '23; Zhou et al., CDC '23; ...

Regret Optimal Control[4]

- selects control inputs over a lookahead horizon;
- guarantees satisfaction of time-varying safety constraints **BUT**:
  - assuming a worst-case noise.

[4] Goel et al., '20; Sabag et al., ACC '21; Goel et al., L4DC '21; Martin et al., L4DC '22; Didier et al., L-CSS '22; Zhou et al., CDC '23; ...

Closest Related Work

Regret Optimal Control[4]

- selects control inputs over a lookahead horizon;
- guarantees satisfaction of time-varying safety constraints **BUT**:
  - assuming a worst-case noise.

Online Learning for Control[5]

- selects control inputs based on past information only;
- consider non-stochastic noise **BUT**:
  - considers no safety constraints or
  - considers time-invariant safety constraints with static regret guarantee.

---

[4] Goel et al., '20; Sabag et al., ACC '21; Goel et al., L4DC '21; Martin et al., L4DC '22; Didier et al., L-CSS '22; Zhou et al., CDC '23; ...

[5] Agarwal et al., ICML '19; Hazan et al., ALT '20; Li et al., AAAI '21; Zhao et al., AISTATS '22; Gradu et al., L4DC '23; Zhou et al., CDC '23; ...

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Initialisation: Time horizon T, step size η, domain set $\mathcal{K}_1$, and $K_1 \in \mathcal{K}_1$.

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Initialisation: Time horizon T, step size η, domain set $\mathcal{K}_1$, and $K_1 \in \mathcal{K}_1$.

At each iteration $t = 1, \ldots, T$:

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Initialisation: Time horizon T, step size η, domain set $\mathcal{K}_1$, and $K_1 \in \mathcal{K}_1$.

At each iteration $t = 1, \ldots, T$:

1. Output $u_t = -K_t x_t$;

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Initialisation: Time horizon T, step size η, domain set $\mathcal{K}_1$, and $K_1 \in \mathcal{K}_1$.

At each iteration $t = 1, \ldots, T$:

1. Output $u_t = -K_t x_t$;
2. Observe state x_{t+1} and noise $w_t = x_{t+1} - A_t x_t - B_t u_t$;

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Initialisation: Time horizon T, step size η, domain set $\mathcal{K}_1$, and $K_1 \in \mathcal{K}_1$.

At each iteration $t = 1, \ldots, T$:

1. Output $u_t = -K_t x_t$;
2. Observe state x_{t+1} and noise $w_t = x_{t+1} - A_t x_t - B_t u_t$;
3. Suffer the loss $c_t(x_{t+1}, u_t)$;

The loss function $c_t\left(x_{t+1}, u_t\right) : \mathbb{R}^{d_x} \times \mathbb{R}^{d_u} \rightarrow \mathbb{R}$ is convex in K_t.

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Initialisation: Time horizon T, step size η, domain set $\mathcal{K}_1$, and $K_1 \in \mathcal{K}_1$.

At each iteration $t = 1, \ldots, T$:

1. Output $u_t = -K_t x_t$;
2. Observe state x_{t+1} and noise $w_t = x_{t+1} - A_t x_t - B_t u_t$;
3. Suffer the loss $c_t(x_{t+1}, u_t)$;
4. Express the loss function in K_t as $f_t(K_t)$;

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Initialisation: Time horizon T, step size η, domain set $\mathcal{K}_1$, and $K_1 \in \mathcal{K}_1$.

At each iteration $t = 1, \ldots, T$:

1. Output $u_t = -K_t x_t$;
2. Observe state x_{t+1} and noise $w_t = x_{t+1} - A_t x_t - B_t u_t$;
3. Suffer the loss $c_t(x_{t+1}, u_t)$;
4. Express the loss function in K_t as $f_t(K_t)$;
5. Obtain gradient $\nabla_K f_t(K_t)$ and update $K'_{t+1} = K_t - \eta \nabla_K f_t(K_t)$;

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Initialisation: Time horizon T, step size η, domain set $\mathcal{K}_1$, and $K_1 \in \mathcal{K}_1$.

At each iteration $t = 1, \ldots, T$:

1. Output $u_t = -K_t x_t$;
2. Observe state x_{t+1} and noise $w_t = x_{t+1} - A_t x_t - B_t u_t$;
3. Suffer the loss $c_t(x_{t+1}, u_t)$;
4. Express the loss function in K_t as $f_t(K_t)$;
5. Obtain gradient $\nabla_K f_t(K_t)$ and update $K'_{t+1} = K_t - \eta \nabla_K f_t(K_t)$;
6. Obtain domain set $\mathcal{K}_{t+1}$ and project $K_{t+1} = \Pi_{\mathcal{K}_{t+1}}(K'_{t+1})$;

Algorithm: Safe Online Gradient Descent (Safe-OGD)

Safe-OGD strictly satisfies time-varying constraints with bounded dynamic regret

Initialisation: Time horizon T, step size η, domain set $\mathcal{K}_1$, and $K_1 \in \mathcal{K}_1$.

At each iteration $t = 1, \ldots, T$:

1. Output $u_t = -K_t x_t$;
2. Observe state x_{t+1} and noise $w_t = x_{t+1} - A_t x_t - B_t u_t$;
3. Suffer the loss $c_t(x_{t+1}, u_t)$;
4. Express the loss function in K_t as $f_t(K_t)$;
5. Obtain gradient $\nabla_K f_t(K_t)$ and update $K'_{t+1} = K_t - \eta \nabla_K f_t(K_t)$;
6. Obtain domain set $\mathcal{K}_{t+1}$ and project $K_{t+1} = \Pi_{\mathcal{K}_{t+1}}(K'_{t+1})$;

We guarantee $x_{t+1} \in \mathcal{S}_{t+1}$ and $u_t \in \mathcal{U}_t$ at each time step t by choosing $K_t \in \mathcal{K}_t$, where

$$\mathcal{K}_t \triangleq \{K \mid -L_{x,t} B_t K x_t \leq l_{x,t} - L_{x,t} A_t x_t - W\|L_{x,t}\|, \\ -L_{u,t} K x_t \leq l_{u,t},\ \|K\| \leq \kappa,\ \|K x_t\| \leq \gamma\}.$$

Dynamic Regret Analysis

Theorem (Dynamic Regret Bound of Safe-OGD)

Safe-OGD with step size $\eta = \mathcal{O}\left(1/\sqrt{T}\right)$ *achieves*

$$\text{Regret}_T^D \leq \mathcal{O}\left(\sqrt{T}\,(1 + C_T + S_T)\right),$$

where

- $C_T \triangleq \sum_{t=2}^{T} \|K_{t-1}^* - K_t^*\|_{\mathrm{F}}$; $\longrightarrow$ Captures how fast K_t^* changes
- $S_T \triangleq \sum_{t=1}^{T} \left\|\Pi_{\mathcal{K}_t}(K_{t+1}') - \Pi_{\mathcal{K}_{t+1}}(K_{t+1}')\right\|_{\mathrm{F}}$. $\longrightarrow$ Captures how fast $\mathcal{K}_t$ changes

Corollary

When $\mathcal{K}_1 = \cdots = \mathcal{K}_T$, *Safe-OGD with step size* $\eta = \mathcal{O}\left(1/\sqrt{T}\right)$ *achieves:*

$$\text{Regret}_T^D \leq \mathcal{O}\left(\sqrt{T}\,(1 + C_T + \cancel{S_T})\right).$$

$S_T = 0$ when $\mathcal{K}_1 = \cdots = \mathcal{K}_T$ since $S_T \triangleq \sum_{t=1}^{T} \left\|\Pi_{\mathcal{K}_t}(K'_{t+1}) - \Pi_{\mathcal{K}_{t+1}}(K'_{t+1})\right\|_{\mathrm{F}}$

Near-Optimality Under Time-Invariant Domain Set

Corollary

When $\mathcal{K}_1 = \cdots = \mathcal{K}_T$, Safe-OGD with step size $\eta = \mathcal{O}\left(1/\sqrt{T}\right)$ achieves:

$$\mathrm{Regret}_T^D \leq \mathcal{O}\left(\sqrt{T}\,(1 + C_T)\right).$$

Comparison with OGD

- OGD with step size $\eta = \mathcal{O}\left(1/\sqrt{T}\right)$ achieves the dynamic regret bound[6]

$$\mathrm{Regret}_T^D \leq \mathcal{O}\left(\sqrt{T}\,(1 + C_T)\right).$$

- The above bound is near-optimal compared to the optimal bound $\Omega\left(\sqrt{T(1+C_T)}\right)$.[7]

[6] Zinkevich, ICML '03
[7] Zhang et al.,, NurIPS '18

Optimality Against Time-Invariant Optimal Controller

Corollary

When $\mathcal{K}_1 = \cdots = \mathcal{K}_T$ *and* $K_1^\star = \cdots = K_T^\star$*, Safe-OGD with step size* $\eta = \mathcal{O}\left(1/\sqrt{T}\right)$ *achieves:*

$$\text{Regret}_T^D \leq \mathcal{O}\left(\sqrt{T}\,(1 + \cancel{C_T})\right).$$

$C_T = 0$ when $K_1^\star = \cdots = K_T^\star$ since $C_T \triangleq \sum_{t=2}^{T} \|K_{t-1}^* - K_t^*\|_{\mathrm{F}}$

Optimality Against Time-Invariant Optimal Controller

Corollary

When $\mathcal{K}_1 = \cdots = \mathcal{K}_T$ and $K_1^\star = \cdots = K_T^\star$, Safe-OGD with step size $\eta = \mathcal{O}\left(1/\sqrt{T}\right)$ achieves:

$$\mathrm{Regret}_T^D \leq \mathcal{O}\left(\sqrt{T}\right).$$

Remark:

- Safe-OGD converges asymptotically to the optimal controller since $\lim_{T\to\infty} \frac{\mathrm{Regret}_T^D}{T} \to 0$.

Example:

- In the Linear–Quadratic–Gaussian control setting,

$$\begin{aligned} x_{t+1} &= A_t x_t + B_t u_t + w_t, \\ y_t &= x_t, \end{aligned} \tag{3}$$

Safe-OGD converges to the optimal linear feedback controller $u_t = -K^* x_t$.

Numerical Evaluation

Setup:

- linear system: $x_{t+1} = Ax_t + Bu_t + w_t$ with $x_t \in \mathbb{R}^6$ and $u_t \in \mathbb{R}^3$;
- $\|w_t\| \leq 0.1$ sampled from various distributions;
- safety constraints: $-\mathbf{1}_{6\times 1} \leq x_t \leq \mathbf{1}_{6\times 1}$, and $[-\pi \ -\pi \ -20]^\top \leq u_t \leq [\pi \ \pi \ 20]^\top$;
- quadratic loss function: $c_t(x_{t+1}, u_t) = \|x_{t+1}\|^2 + \|u_t\|^2$;
- total iteration $T = 500$
- comparison with safe H_2 and H_∞ with lookahead horizon $N = 1,\ 5,\ 10$.[8]

K_t needs to be chosen from time-varying $\mathcal{K}_t$ even the safety constraints are time-invariant, since $\mathcal{K}_t$ depends on time-varying x_t, *i.e.*,

$$\mathcal{K}_t \triangleq \{K \mid -L_{x,t}B_tKx_t \leq l_{x,t} - L_{x,t}A_tx_t - W\|L_{x,t}\|, \\ -L_{u,t}Kx_t \leq l_{u,t},\ \|K\| \leq \kappa,\ \|Kx_t\| \leq \gamma\}.$$

[8] Anderson et al., ARC '19, Martin et al., L4DC '22

Numerical Evaluation

Table: Comparison in terms of cumulative loss.

Noise Distribution	Ours	$N = 1$		$N = 5$		$N = 10$	
		H_2	H_∞	H_2	H_∞	H_2	H_∞
Gaussian	44.05	61.81	93.44	47.96	52.03	30.66	48.69
Uniform	151.49	724.98	1859.61	331.32	323.42	100.21	53.86
Gamma	159.21	811.09	2082.12	372.52	364.26	112.90	60.77
Beta	186.98	836.41	2152.63	386.30	375.73	116.70	62.40
Exponential	126.69	552.73	1421.90	259.82	250.76	79.25	44.35
Weibull	195.71	873.09	2246.31	405.70	392.94	122.63	65.86
Average	142.50	643.35	1642.67	300.60	293.19	93.72	55.99
Standard Deviation	53.92	307.00	814.06	134.16	128.60	34.53	8.43

Lower Loss

Numerical Evaluation

Table: Blue corresponds to best runtime; red corresponds to worse runtime.

Noise Distribution	Ours	$N=1$		$N=5$		$N=10$	
		H_2	H_∞	H_2	H_∞	H_2	H_∞
Average	0.1484	0.3712	0.6429	0.6033	1.3693	1.3854	17.0248
Standard Deviation	0.0342	0.0143	0.0116	0.0282	0.2741	0.0673	0.3691

Faster

Numerical Evaluation

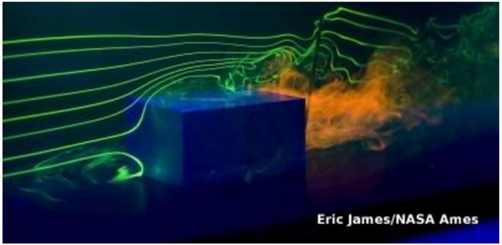
Table: Blue corresponds to best runtime; red corresponds to worse runtime.

Noise Distribution	Ours	$N=1$		$N=5$		$N=10$	
		H_2	H_∞	H_2	H_∞	H_2	H_∞
Average	0.1484	0.3712	0.6429	0.6033	1.3693	1.3854	17.0248
Standard Deviation	0.0342	0.0143	0.0116	0.0282	0.2741	0.0673	0.3691

Best on Average

Online learning for control algorithm that

- guarantees safety despite non-stochastic disturbances, and
- provides dynamic regret performance guarantees under time-varying constraints.



Next steps:

- optimality of the regret bounds;
- recursive feasibility;[8]
- safe non-linear control.[9]

[8] Zhou, and Tzoumas. Safe Non-Stochastic Control of Linear Dynamical Systems, arXiv:2308.12395

[9] Zhou, Song, and Tzoumas. Safe Non-Stochastic Control of Control-Affine Systems: An Online Convex Optimization Approach, IEEE Robotics and Automation Letters (RA-L) '23