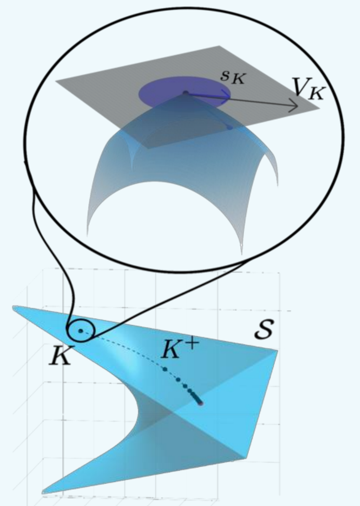
Shahriar Talebi (Harvard University)

Harvard John A. Paulson School of Engineering and Applied Sciences

DYNAMICS AI

01 - Geometry and Policies

❖ How to utilize **the geometry of policies** for learning efficiently under constraints? [CDC '22, TAC '24]



- **Riemannian geometry** of stabilizing policies for LTI systems

$$\langle V_K, W_K\rangle_K \triangleq \mathrm{tr}[V_K^\intercal W_K Y_K], \quad Y_K = \mathrm{Lyap}(A_K, \Sigma)$$

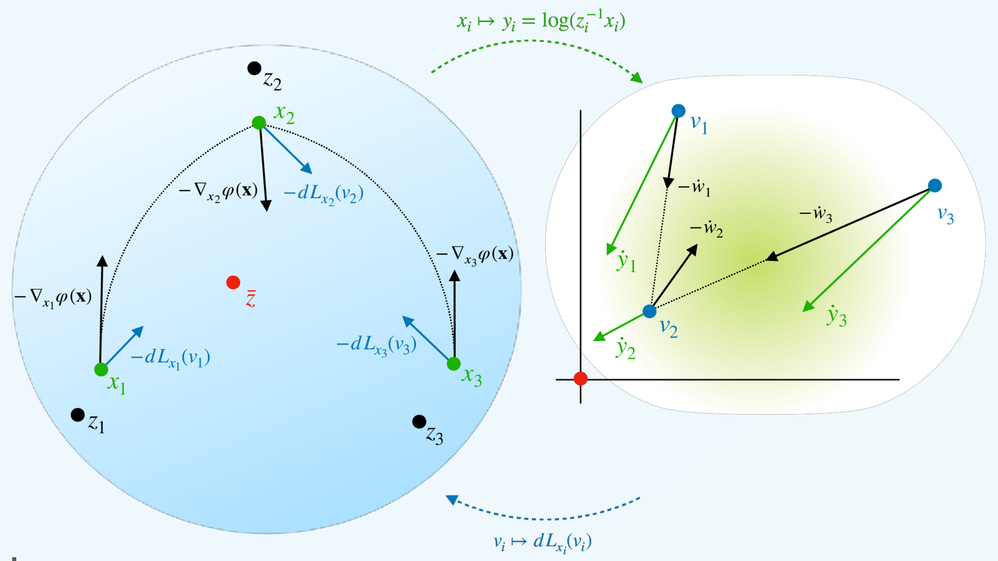
- Riemannian Gradient/Newton Descent for constrained LQR
- RGD / RND: $\begin{cases} V_K = -\mathrm{grad}_K f \quad \text{or} \quad \mathrm{Hess}\, f[V_K] = -\mathrm{grad}_K f \\ K^+ = \mathcal{R}_K(s_K V_K), \quad s_K^{-1} \propto \|V_K\|\,\lambda_{\max}(Y_K)/\lambda_{\min}(\Sigma) \end{cases}$

❖ **Geometric consensus** to RCM on bi-invariant Lie groups: [CDC '23, CCTA '23]

$$\mathbf{f}(x) := \tfrac{1}{2}\textstyle\sum_{i=1}^N \mathrm{d}(x_i, z_i)^2$$

$$\varphi(x) := \textstyle\sum_{\{i,j\}\in E} \mathrm{d}(x_i, x_j)^2$$

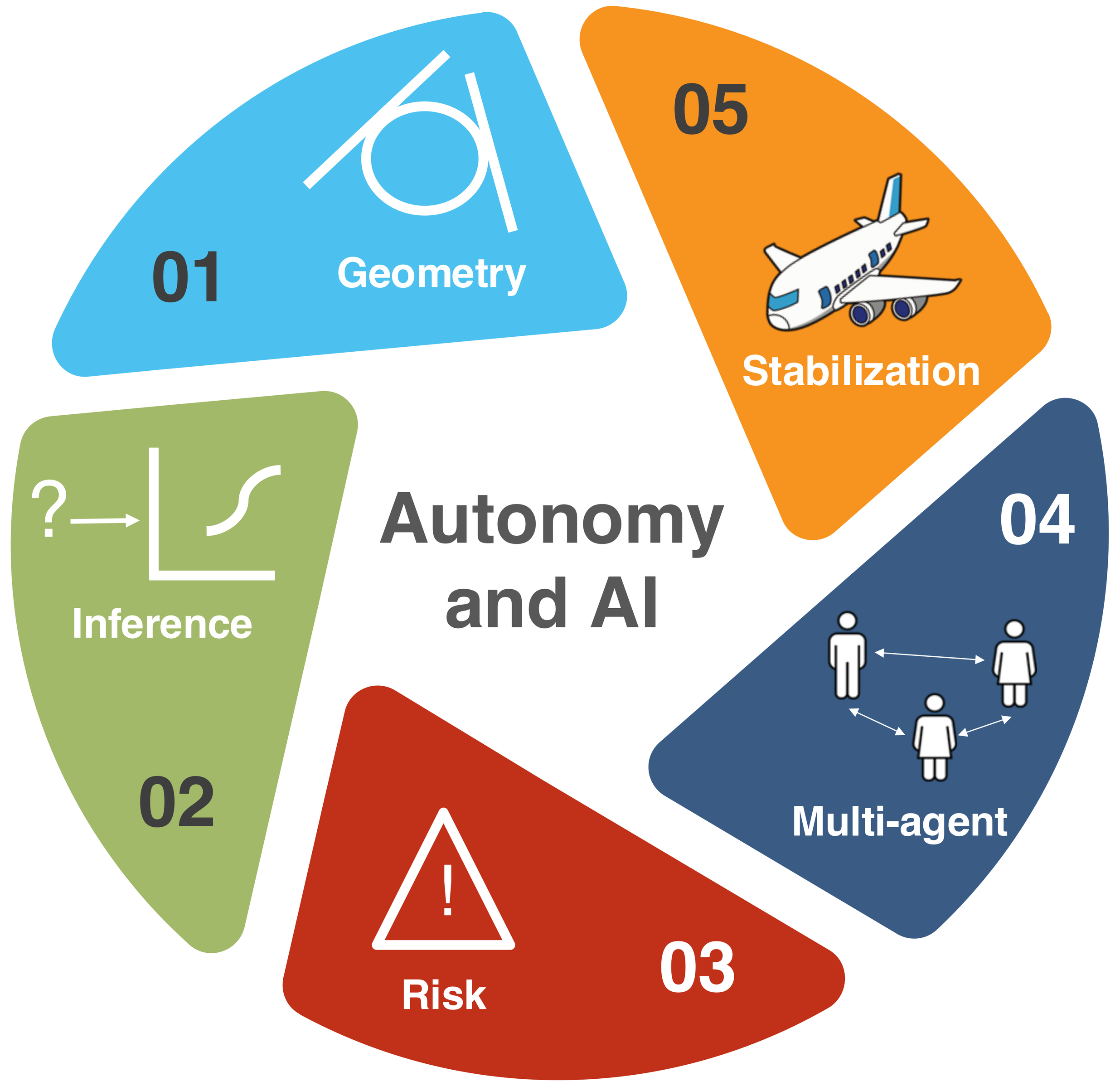
- RCM: $\begin{cases} \dot{x} = -\nabla\varphi(x) - dL_x v, \\ \dot{w} = \mathbf{L}v \\ v = -w + dL_{x^{-1}}\nabla f(x) \end{cases}$



↠ Generalization to robust, decentralized policies

↠ Interactions with the geometry of stable manifolds

* Talebi & Mesbahi, M. (2022). Riemannian constrained policy optimization via geometric stability certificates. 61st IEEE Conf on Decision and Control (**CDC**), 1472–1478
* Talebi & Mesbahi, M. (2024). Policy optimization over submanifolds for linearly constrained feedback synthesis. IEEE Transactions on Automatic Control (**TAC**), 69(5), 3024–3039
* Kraisler, Talebi, & Mesbahi, (2023). Distributed consensus on manifolds using the Riemannian center of mass. 2023 IEEE Conf on Control Technology and Applications (**CCTA**), 130–135
* Kraisler, Talebi, & Mesbahi, (2023). Consensus on Lie groups for the Riemannian center of mass. 62nd IEEE Conf on Decision and Control (**CDC**), 4461–4466



02 - Learning and Inference

❖ How to efficiently infer from partially observable systems with **unknown noise** or **high-dim measurements**? [ACC '23, NeurIPS '23]

- Optimal MSE estimation **without noise covariances** Q and R:

$$\min_L J_T^{\mathrm{est}}(L) := \mathrm{E}\|y(T) - \hat{y}_L(T)\|^2$$

- Filtering policy optimization via **Estimation-Control duality**:

$$L_{\mathrm{Kalman}}(A, H, Q, R) \equiv K_{\mathrm{LQR}}(A^\intercal, H^\intercal, Q, R, H^\intercal H)^\intercal$$

$$L^+ = L - s_L \nabla_L \widehat{J}_T^{\mathrm{est}}(L)$$

❖ SysID from high-dimensional observation (applications in meta-learning): [ICML '24]

- **High-dim SysID**: $\begin{cases} \hat{\Phi}_C \leftarrow \text{Col-approx}(\{y_t\}) \\ \hat{A},\ \hat{B},\ \hat{\Phi}_C^\intercal C \leftarrow \text{Ho-Kalman}(\{u_t, \hat{\Phi}_C y_t\}) \\ \hat{C} \leftarrow \hat{\Phi}_C \hat{\Phi}_C^\intercal C \end{cases}$

- Complexity $\mathcal{O}(d_y \cdot \mathrm{poly}(d_x, d_u)/\epsilon^2) \xrightarrow{\text{reduces}} \mathcal{O}([d_y + \mathrm{poly}(d_x, d_u)]/\epsilon^2)$ "optimal"

↠ Time-varying, partially unknown settings

↠ Optimal SysID through nonlinear high-dim observations

* Talebi, Taghvaei, & Mesbahi, (2023). Data-driven optimal filtering for linear systems with unknown noise covariances. Adv in Neural Inform Processing Systems (**NeurIPS**), 36, 69546–85
* Talebi, Taghvaei, & Mesbahi, (2023). Duality-based stochastic policy optimization for estimation with unknown noise covariances. 2023 American Control Conference (**ACC**), 622–627
* Zhang, Talebi, & Li, (2024). Learning low-dim latent dynamics from high-dim observations. 41st Int Conf on Machine Learning (**ICML**), 235, 59851–96

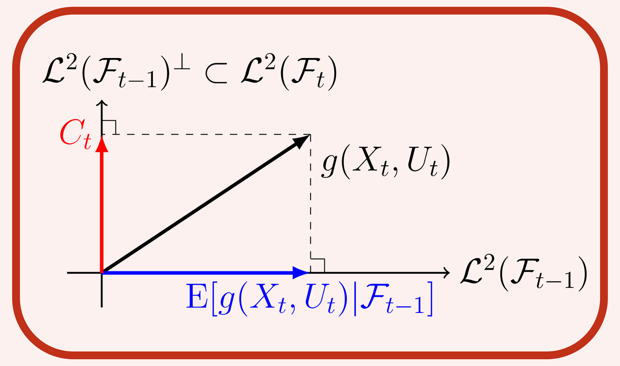
03 – Ergodic-risk Control

❖ How to manage long-term cumulative risks under **heavy-tailed noise**, while maintaining **stability** and **performance**? [ACC '25(sub), SICON '25(sub)]

- **Ergodic-risk** criteria:

$$C_t := g(X_t, U_t) - \mathrm{E}[g(X_t, U_t)|\mathcal{F}_{t-1}]$$

$$\tfrac{1}{\sqrt{t}}\textstyle\sum_{s=0}^t C_s \xrightarrow{d} C_\infty$$



- **Theorem** (using ergodic theory):
  a. The closed-loop system is $(\|\cdot\|^4 + 1)$-uniformly ergodic if
     1) the controller is stabilizing, and
     2) noise has finite fourth moment.
  b. The Functional CLT holds.
- Ergodic-risk COCP:

$$\min\ J_{\mathrm{LQR}}(U)$$
$$\text{s.t.}\quad X_{t+1} = AX_t + BU_t + HW_{t+1}, \quad \{U_t\}\ \text{admissible}$$
$$\varrho(C_\infty) \le \varrho_0, \quad \varrho\ \text{risk measure}$$

- **(Strong duality)** Primal-dual PO balancing performance and risk

↠ **Distributional policy optimization** for stochastically dominant policies

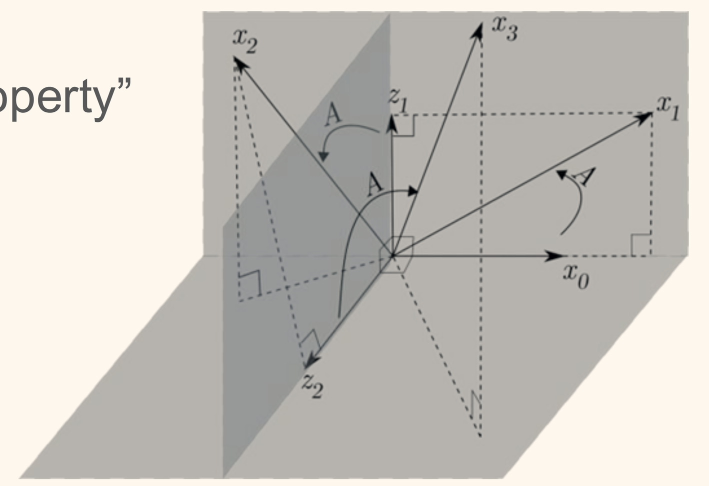
$$g(X_t, U_t) \succeq_{SD} g(X_t^{\mathrm{baseline}}, U_t^{\mathrm{baseline}})$$

* Talebi, & Li, (2024). Uniform Ergodicity and Ergodic-Risk Constrained Policy Optimization. arXiv Preprint ArXiv:2409.10767 (SIAM **SICON** submitted)
* Talebi, & Li, (2024). Ergodic-risk constrained policy optimization: The linear quadratic case. 2025 American Control Conference (**ACC** submitted)

05 - Online Stabilization

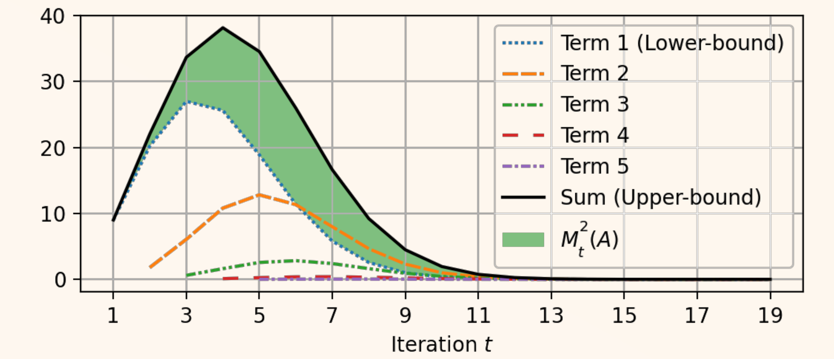
❖ Can we achieve real-time **online stabilization** and control for unstable systems? [CDC '20, TAC '22]

- (A, B) is **Regularizable** if $\mathrm{Proj}_{\mathcal{R}(B)^\perp} A$ is stable,
  "a **geometric criterion** characterizing finite-time stabilization property"



- **Theorem:** The certainty-equivalent minimum-energy control
  a. stabilizes (A, B) $\iff$ (A, B) is regularizable,
  b. generates informative data
- Performance guarantees: $\dfrac{\|x_{t+1}\|}{\|x_0\|} \propto M_t(A)$
- The "instability number of order t":

$$M_t(A) \triangleq \sup_{\{v_1,\ldots,v_t\}\in\mathcal{O}_t^n} \|Av_1\|\,\|Av_2\| \cdots \|Av_t\| \propto \left[\frac{\sigma_1(A)}{t}\right]^t$$



↠ Output-feedback and distributed stabilization

↠ Geometric classification of control-affine systems

* Talebi, Alemzadeh, Rahimi, & Mesbahi, (2020). Online regulation of unstable linear systems from a single trajectory. IEEE Conference on Decision and Control (**CDC**), 4784–4789
* Talebi, Alemzadeh, Rahimi, & Mesbahi, (2022). On Regularizability and its application to online control of unstable LTI systems. IEEE Trans on Automatic Control (**TAC**), 67(12), 6413–6428

04 - Learning in Multi-agent Systems

❖ How to utilize **algebraic structures** of interconnections for MARL? [CDC '21, TAC '24]

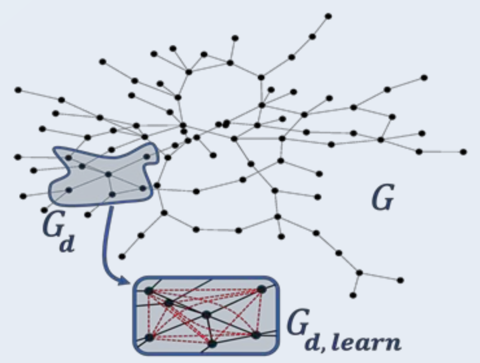
- The invariance of the **Patterned Monoid** under Lyapunov eq:

$$\mathrm{PM}(r \times n, \mathbb{R}) \triangleq \{\mathbf{N}_r \in \mathrm{GL}(rn, \mathbb{R}) \mid A \in \mathrm{GL}(n, \mathbb{R}) \cap \mathbb{S}^n,\ B \in \mathbb{S}^n\} \qquad \mathbf{N}_r = \begin{pmatrix} A & B & \cdots & B \\ B & A & \ddots & \vdots \\ \vdots & \ddots & \ddots & B \\ B & \cdots & B & A \end{pmatrix}$$

Theorem: if A is stable and $P = A^\intercal P A + Q$, then

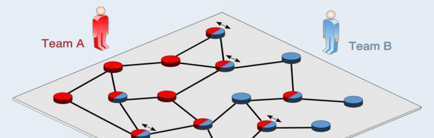
$$P \in \mathrm{PM}(r \times n, \mathbb{R}) \iff Q \in \mathrm{PM}(r \times n, \mathbb{R})$$

- Policy optimization via sub-network data of homogeneous agents
  1. PO on $G_{d,\,learn}$ using the invariance of $\mathrm{PM}(r \times n, \mathbb{R})$
  2. Propagate policies to G using robust control techniques
  3. Guaranteed near-optimal performance



❖ Sub-network **distributed games** [CDC '19]

$$\begin{cases} \min_{x_A} f_A(x_A, x_B) = \sum_i f_{A,i}(x_A, x_B) \\ \min_{x_B} f_B(x_A, x_B) = \sum_i f_{B,i}(x_A, x_B) \end{cases}$$



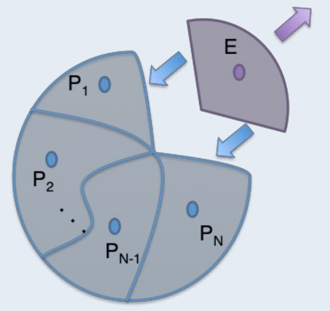
- "Team-based dual averaging":

$$\ell \in \{A, B\}: \begin{cases} z_{\ell,i}(t+1) &= \sum_{j\in\mathcal{N}_{\ell,i}} P_{ij} z_{\ell,j}(t) + g_{\ell,i}(t) \\ x_{\ell,i}(t+1) &= \Pi_{\mathcal{X}_\ell}^{\psi}\big(-z_{\ell,i}(t+1), \alpha(t)\big) \end{cases}$$

- No-regret guarantee $\mathcal{R}_{\ell,i}(T) = \mathcal{O}(\sqrt{T}\log(T)/\sqrt{1-\sigma_2(P)})$

❖ **Noninferior-Nash** Strategies [CDC '19, ECC '18]

- LQ game $\{\mathrm{P_i}\}$ vs E:
  1. Non-coop Nash
  2. Team Nash
  3. Greedy Nash



- Finite time-capture if $\sum_i Q_{\mathrm{P_i}} > -Q_{\mathrm{E}}$:

$$t \ge \mathcal{O}(-\log\epsilon) \implies \|x_t\| \le \epsilon$$

↠ MARL in heterogeneous, non-cooperative settings

* Talebi, Alemzadeh, & Mesbahi, (2024). Data-driven structured policy iteration for homogeneous distributed systems. IEEE Trans on Automatic Control (**TAC**), 69(9), 5979–5994
* Talebi, Alemzadeh, & Mesbahi, (2021). Distributed model-free policy iteration for networks of homogeneous systems. 60th IEEE Conf on Decision and Control (**CDC**), 6970–6975
* Talebi, Alemzadeh, Ratliff, & Mesbahi, (2019). Distributed learning in network games: A dual averaging approach. 58th IEEE Conf on Decision and Control (**CDC**), 5544–5549
* Talebi, Simaan, & Qu, (2019). Decision-making in complex dynamical systems of systems with one opposing subsystem. 18th European Control Conf (**ECC**), 2789–2795
* Talebi, Simaan, & Qu, (2019). Cooperative design of systems of systems against attack on one subsystem. 58th Conf on Decision and Control (**CDC**), 7313–7318