Asynchronous and Parallel Distributed Pose Graph Optimization

Yulun Tian[1] Alec Koppel[2] Amrit S. Bedi[2] Jonathan P. How[1]

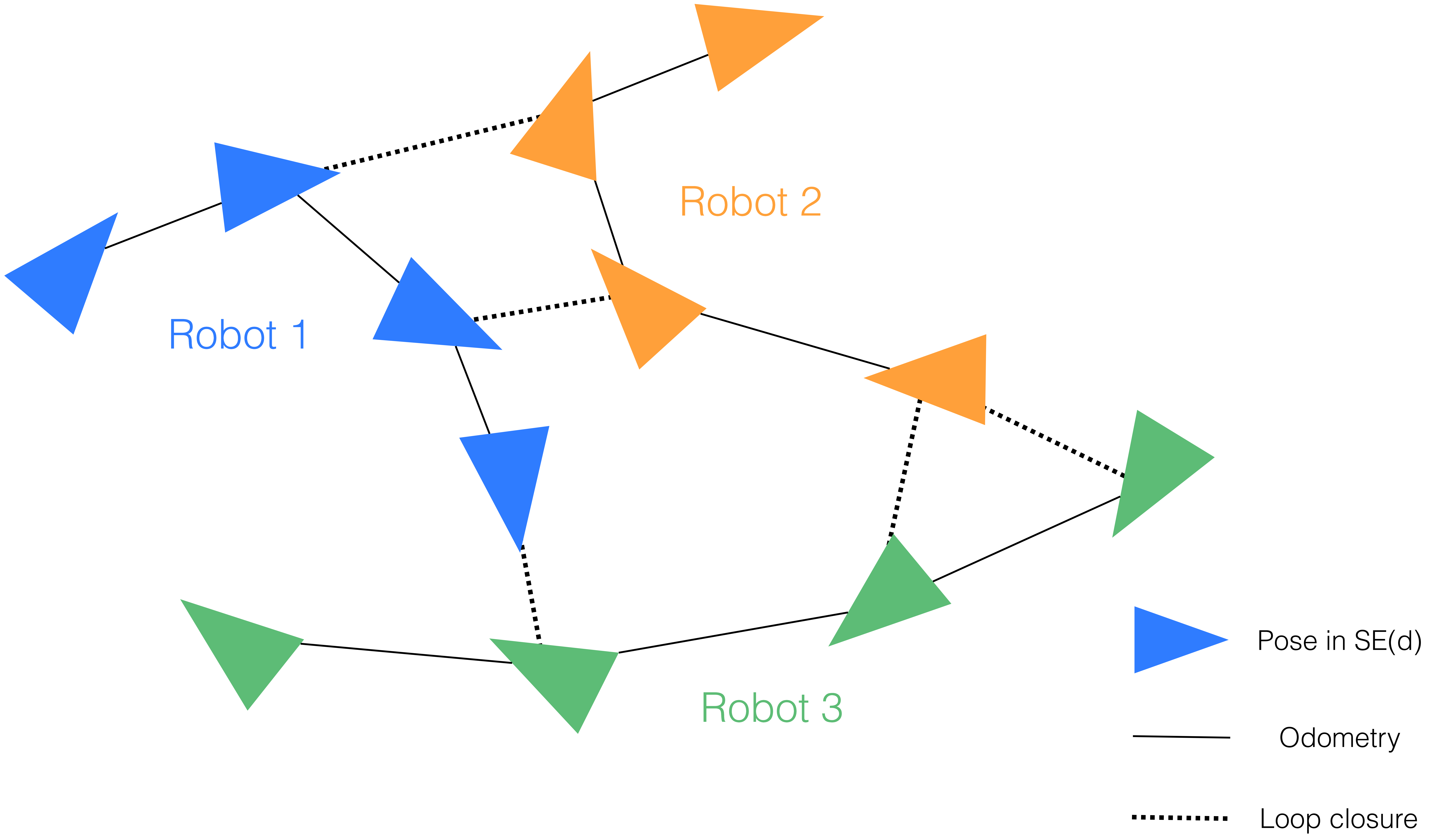
Massachusetts Institute of Technology

MIT AEROASTRO

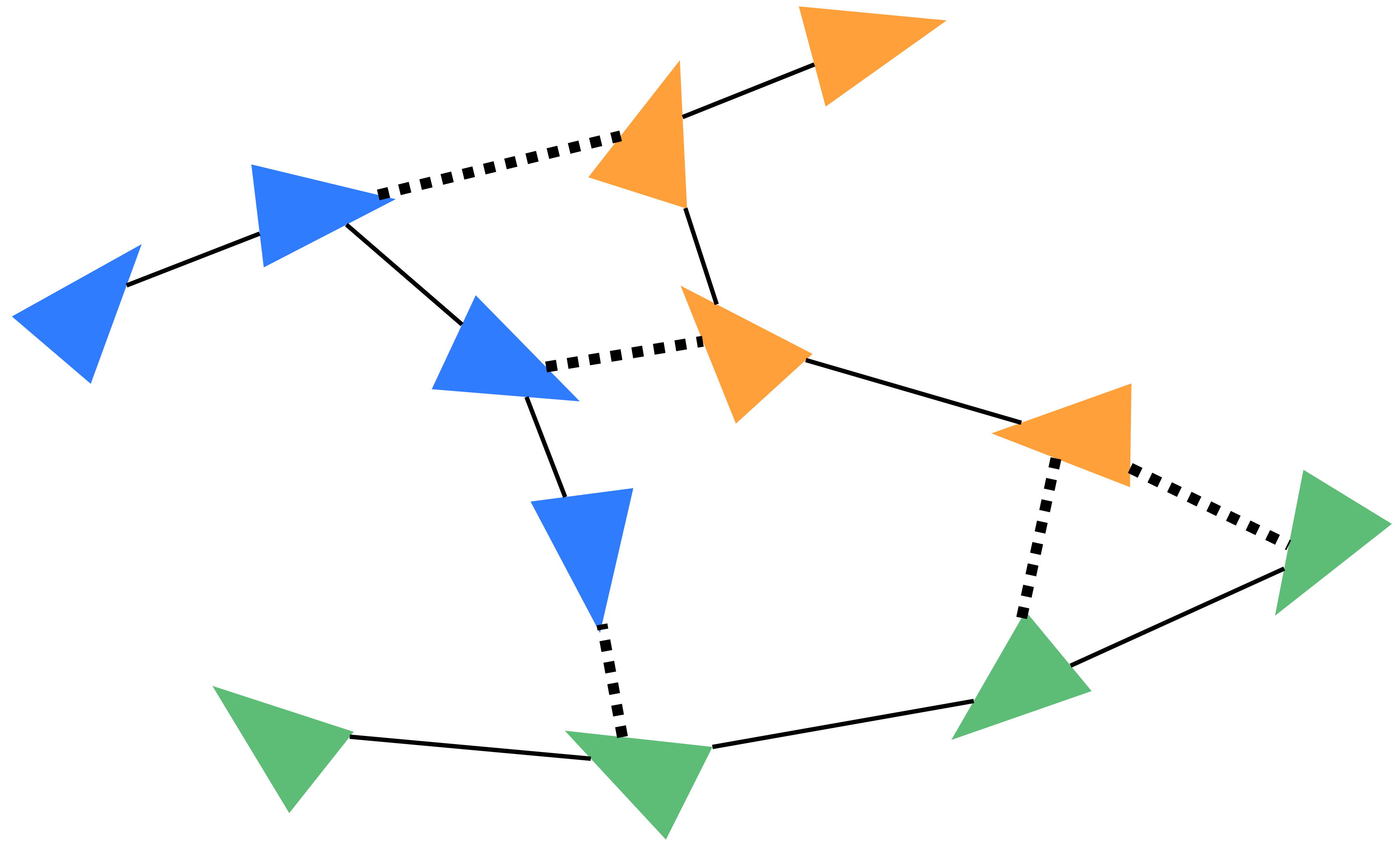
DEVCOM ARMY RESEARCH LABORATORY

1. Department of Aeronautics and Astronautics, MIT
2. U.S. Army Research Laboratory

Distributed Pose Graph Optimization (PGO)



Distributed Pose Graph Optimization (PGO)



Pose graph $G = (V, E)$

Maximum Likelihood Estimation (MLE):

$$\min \sum_{(i_\tau, j_s) \in E} w_R \left\| R_{j_s} - R_{i_\tau} \widetilde{R}^{i_\tau}_{j_s} \right\|_F^2 + w_t \left\| t_{j_s} - t_{i_\tau} - R_{i_\tau} \widetilde{t}^{i_\tau}_{j_s} \right\|_2^2,$$

$$\text{s.t.} \quad R_{i_\tau} \in \mathrm{SO}(d),\ t_{i_\tau} \in \mathbb{R}^d,\ \forall i \in \mathcal{R},\ \forall \tau.$$

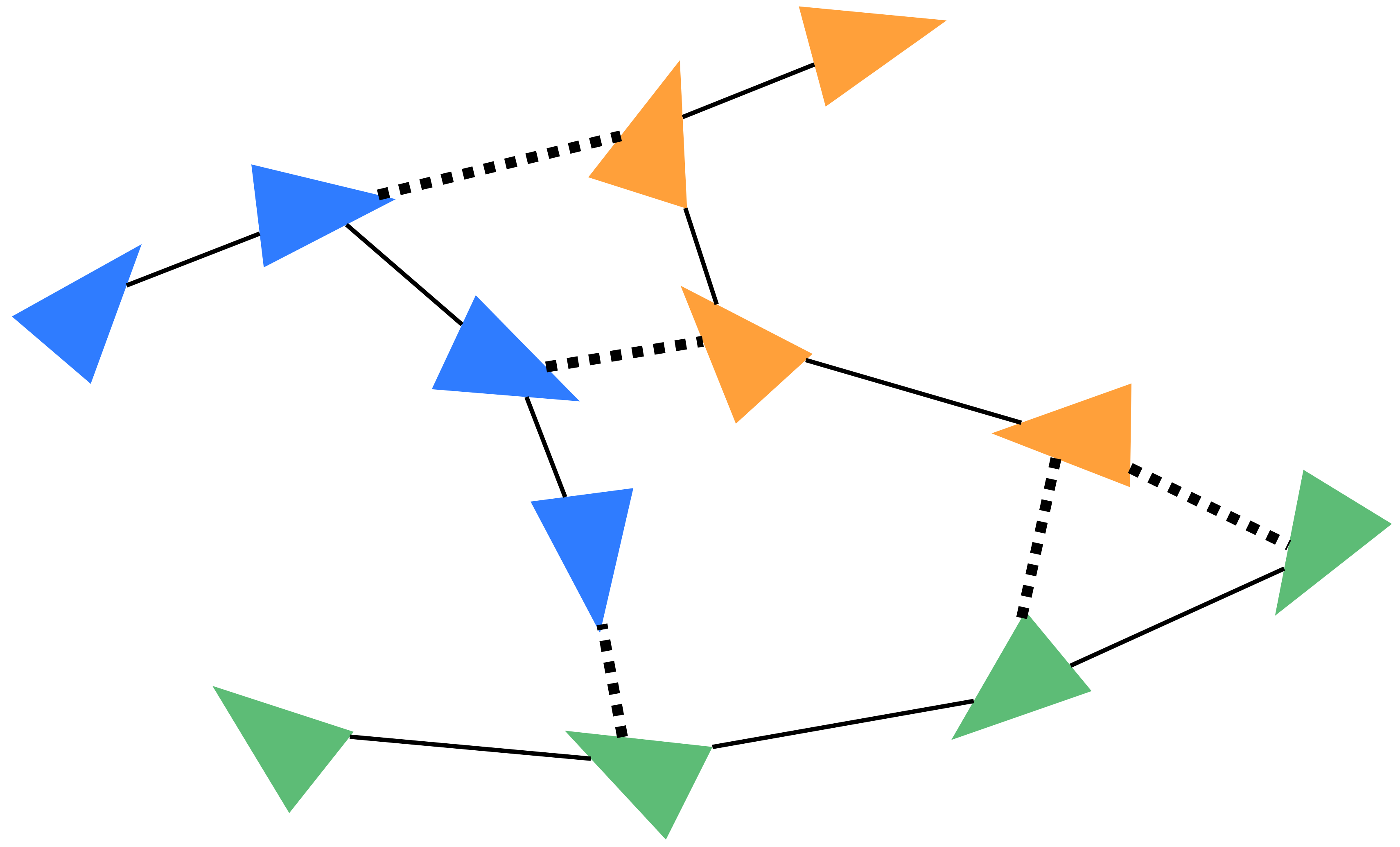
Variables:

- R_{i_τ}: rotation of robot $i \in \mathcal{R}$ at time τ
- t_{i_τ}: translation of robot $i \in \mathcal{R}$ at time τ

Measurements:

- $\widetilde{R}^{i_\tau}_{j_s}$: noisy relative rotation from (i, τ) to (j, s)
- $\widetilde{t}^{i_\tau}_{j_s}$: noisy relative translation from (i, τ) to (j, s)

Distributed Pose Graph Optimization (PGO)



Pose graph $G = (V, E)$

Maximum Likelihood Estimation (MLE):

$$\min \sum_{(i_\tau, j_s) \in E} w_R \left\| R_{j_s} - R_{i_\tau} \widetilde{R}^{i_\tau}_{j_s} \right\|_{\mathrm{F}}^2 + w_t \left\| t_{j_s} - t_{i_\tau} - R_{i_\tau} \widetilde{t}^{i_\tau}_{j_s} \right\|_2^2,$$
$$\text{s.t.} \quad R_{i_\tau} \in \mathrm{SO}(d),\ t_{i_\tau} \in \mathbb{R}^d,\ \forall i \in \mathcal{R},\ \forall \tau.$$

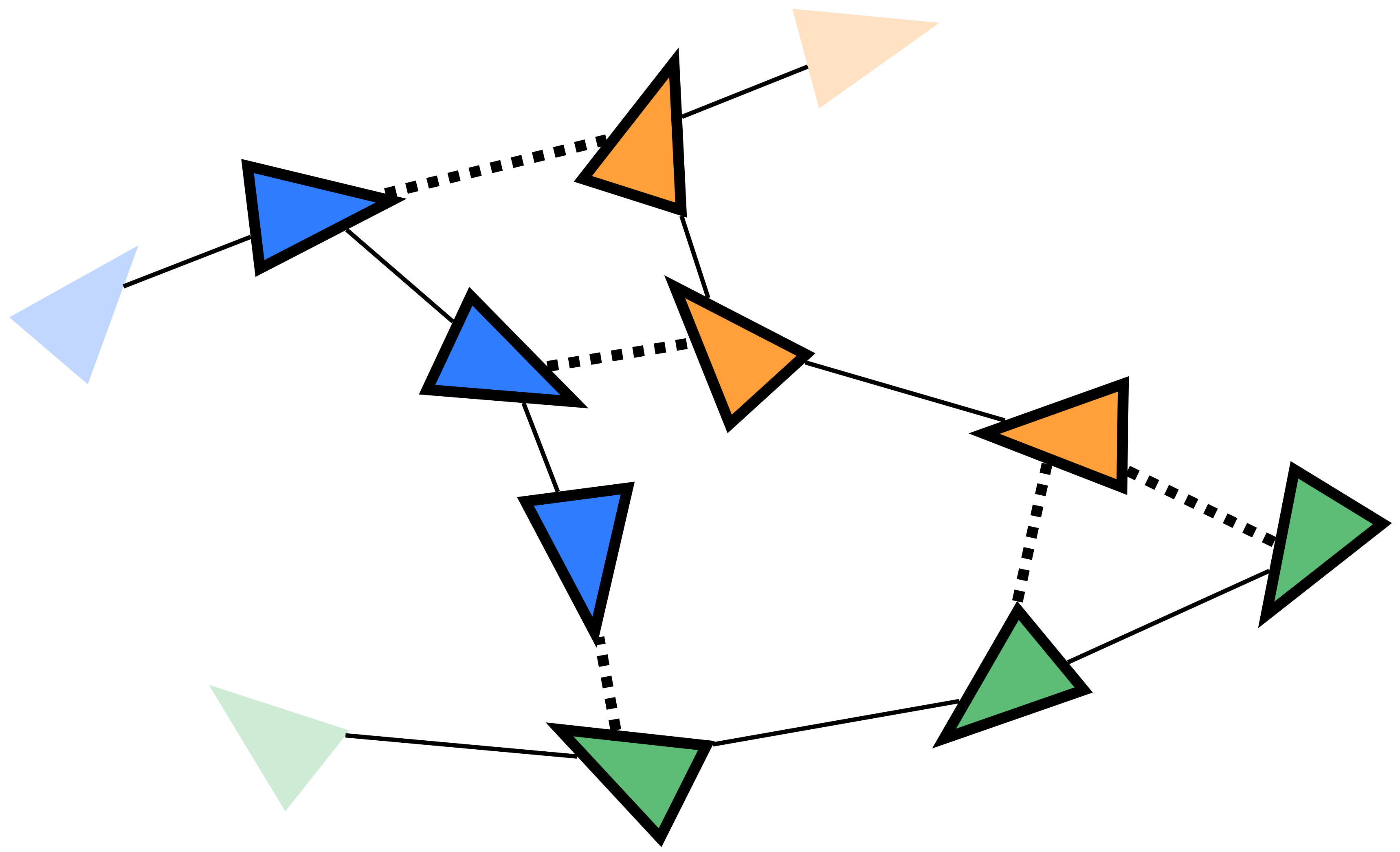
Rank-r Relaxation ($r \geq d$):

$$\min \sum_{(i_\tau, j_s) \in E} w_R \left\| Y_{j_s} - Y_{i_\tau} \widetilde{R}^{i_\tau}_{j_s} \right\|_{\mathrm{F}}^2 + w_t \left\| p_{j_s} - p_{i_\tau} - Y_{i_\tau} \widetilde{t}^{i_\tau}_{j_s} \right\|_2^2,$$
$$\text{s.t.} \quad Y_{i_\tau} \in \mathrm{St}(d, r),\ p_{i_\tau} \in \mathbb{R}^r,\ \forall i \in \mathcal{R},\ \forall \tau.$$

Useful for **global** optimization: Burer-Monteiro approach for SDP relaxation
[Rosen et al. 2016, Briales et al. 2017, Tian et al. 2019]

The role of communication

- **Communication is essential for distributed PGO**
- **Existing methods operate by exchanging "public" poses**
  - Riemannian gradient descent [Tron et al. 2009, Knuth and Barooah 2012]
  - Alternating direction method of multipliers (ADMM) [Choudhary et al. 2015]
  - Distributed Gauss-Seidel [Choudhary et al. 2017]
  - Block-coordinate descent on product manifold [Tian et al. 2019]
  - Generalized proximal method [Fan and Murphey 2019, 2020]
- **However, all existing methods are limited by synchronization:**
  - Robots synchronize to use up-to-date public poses from each other
  - Induce penalty on communication and overall execution time
  - Increase complexity of implementation

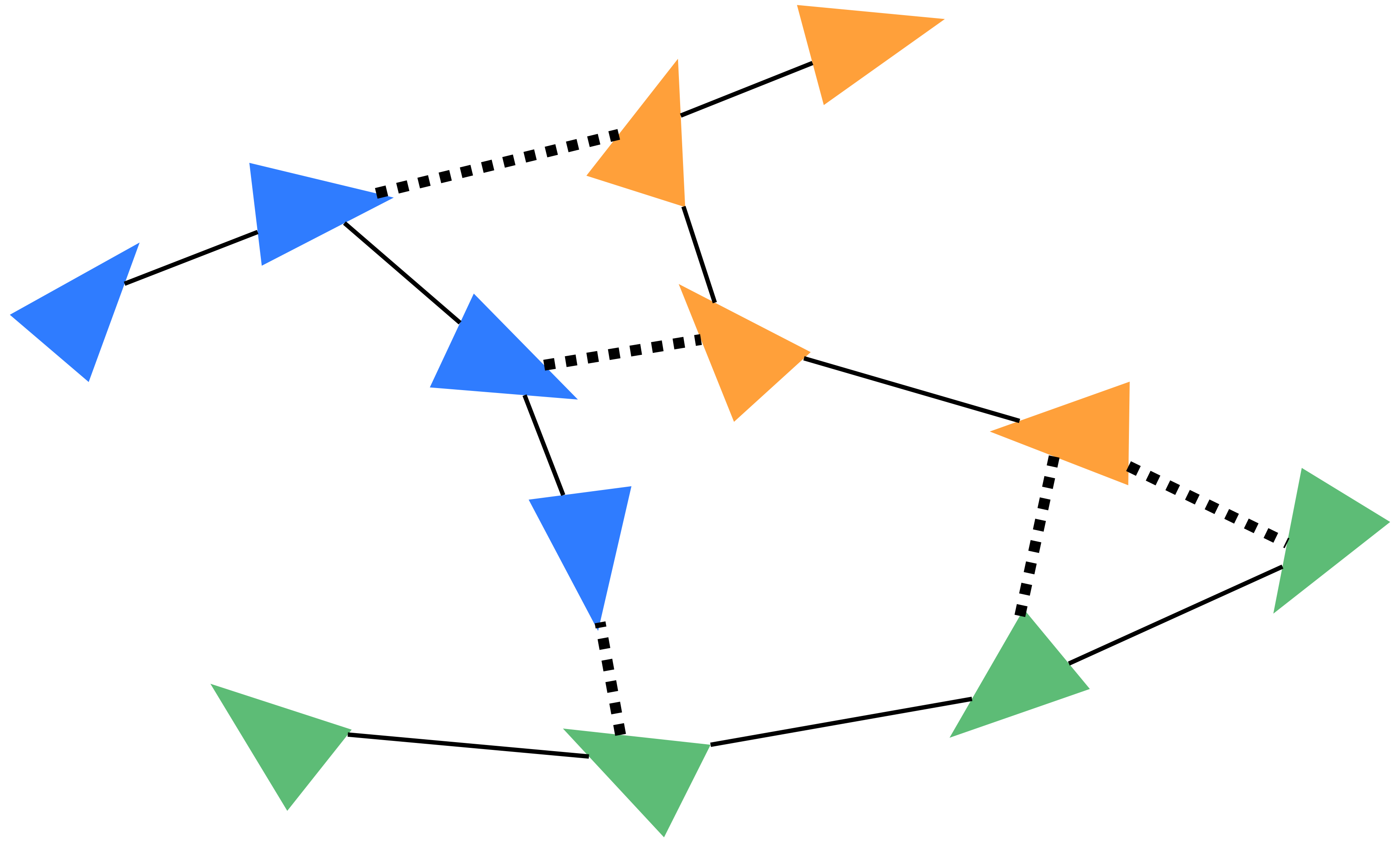


Contributions

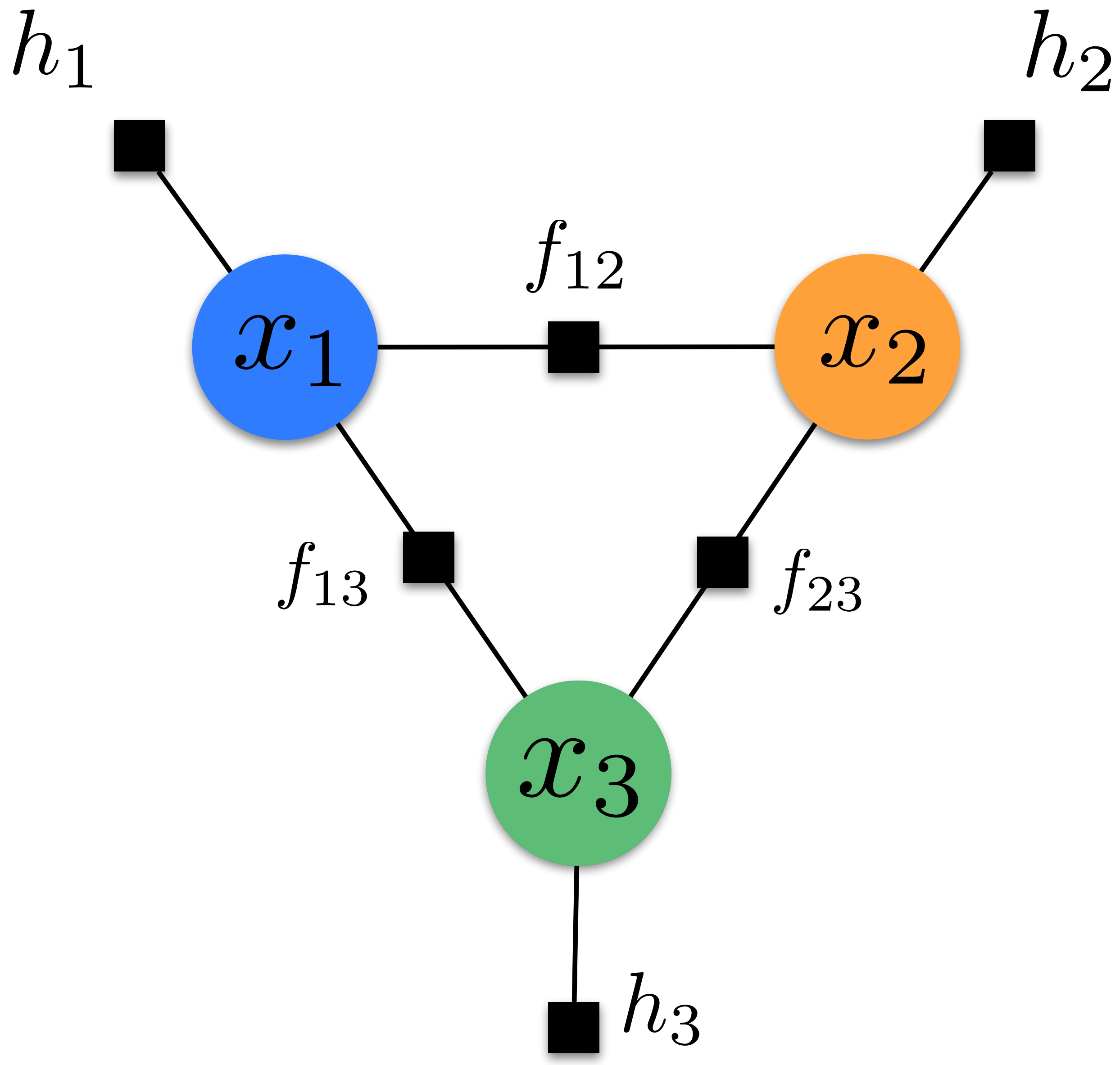
- Inspired by asynchronous algorithms in distributed optimization
  - Asynchronous coordinate descent [Bertsekas and Tsitsilis 1989]
  - Asynchronous stochastic optimization [Niu et al. 2011, Liu et al. 2015]
  - Asynchronous non-convex optimization [Lian et al. 2015]

- **Contributions:**
  - **ASAPP — asynchronous stochastic parallel pose graph optimization**
  - **The first asynchronous and provably convergent algorithm for distributed PGO and its rank-restricted relaxations**
  - **Sufficient conditions for convergence with respect to *maximum delay* and *problem sparsity***

Forming the robot-level problem



Pose graph $G = (V, E)$

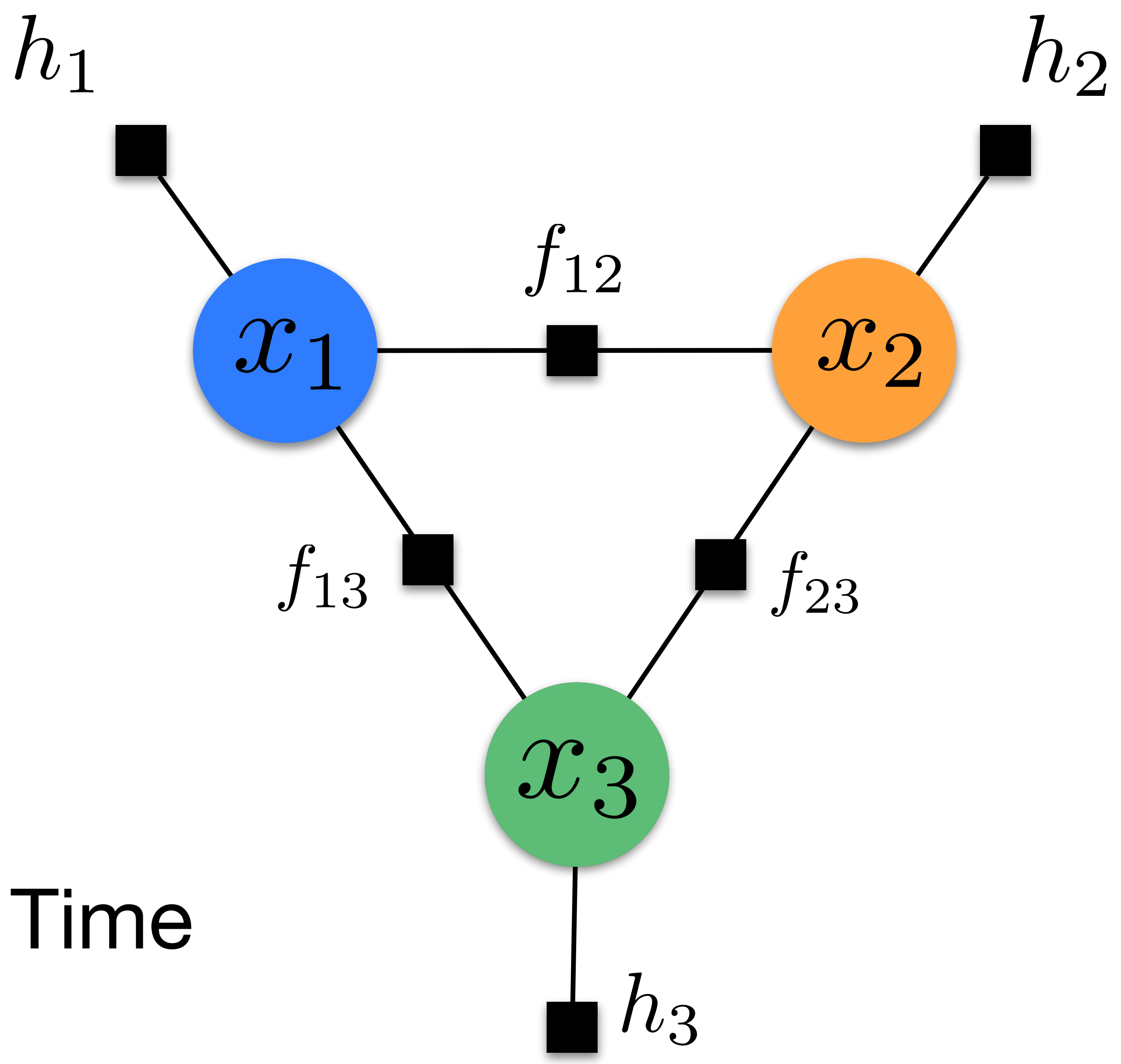


Robot-level graph $G_{\mathcal{R}} = (\mathcal{R}, E_{\mathcal{R}})$
(factor graph form)

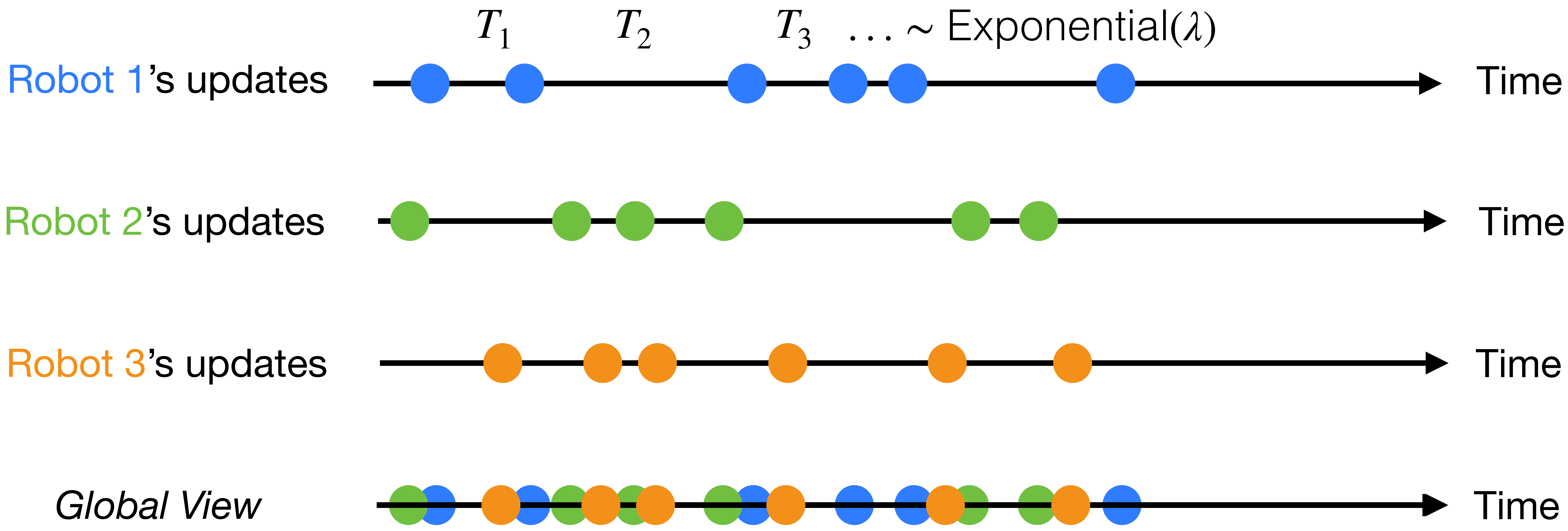
Robot-level Optimization Problem:

$$\min \sum_{(i,j)\in E_{\mathcal{R}}} f_{ij}(x_i, x_j) + \sum_{i\in\mathcal{R}} h_i(x_i),$$
$$\text{s.t.} \quad x_i \in \mathcal{M}_i, \ \forall i \in \mathcal{R}.$$

ASAPP: Global View



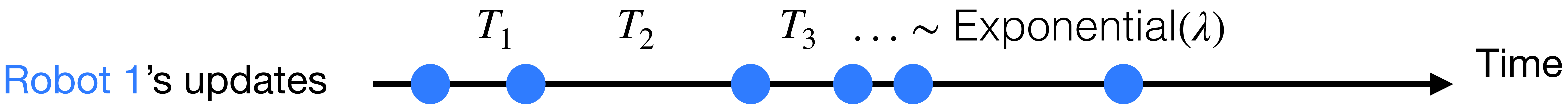
- Each robot shares public poses with neighbors periodically
- Each robot performs **asynchronous** local updates using Poisson clocks*



Key idea: Merging n independent Poisson processes leads to **uniform sampling** in the global view.

S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, “Randomized gossip algorithms,” IEEE Transactions on Information Theory, 2006.

ASAPP: Local Updates



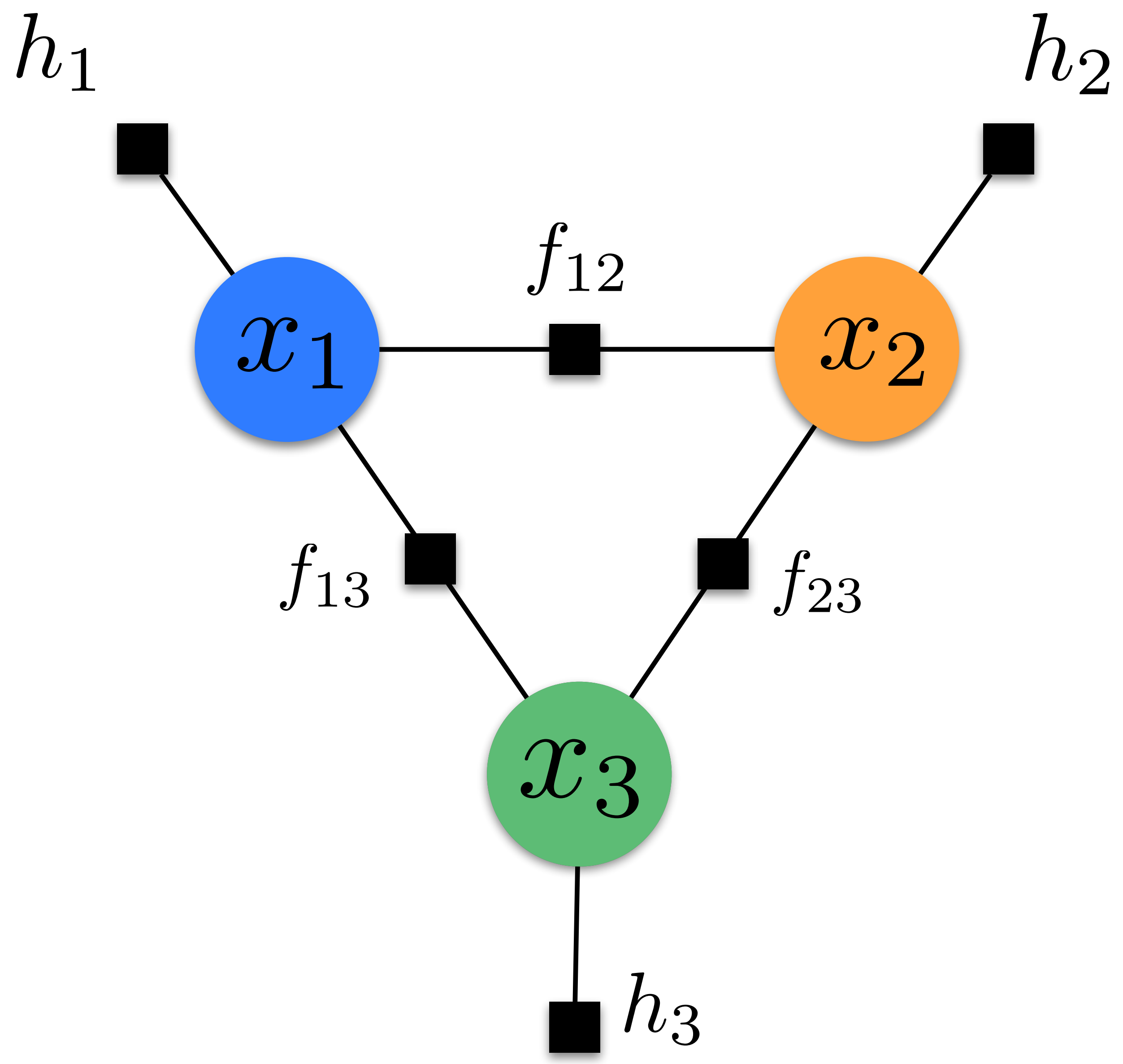
- Forms local cost function using **delayed** information from other robots:

Delayed variables from neighbors

$$g_i(x_i) = h_i(x_i) + \sum_{j \in \mathrm{N}(i)} f_{ij}(x_i, \hat{x}_j).$$

- Apply **asynchronous** Riemannian gradient descent*

$$x_i \leftarrow \mathrm{Retr}_{\hat{x}_i}(-\gamma \mathrm{grad} g_i(\hat{x}_i)).$$



- **Question**: select stepsize γ to guarantee convergence?

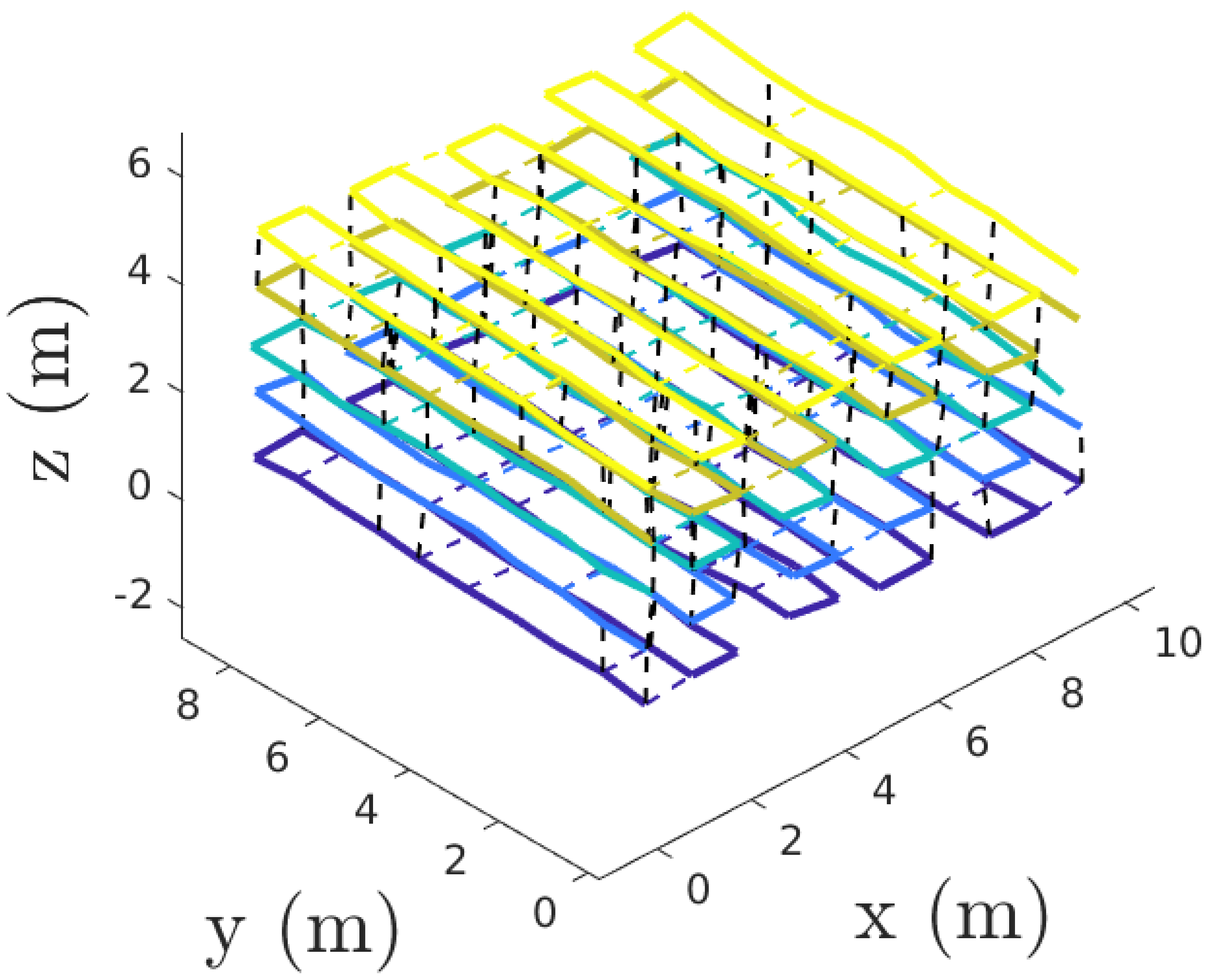
*Can be extended to use preconditioner.

Global convergence

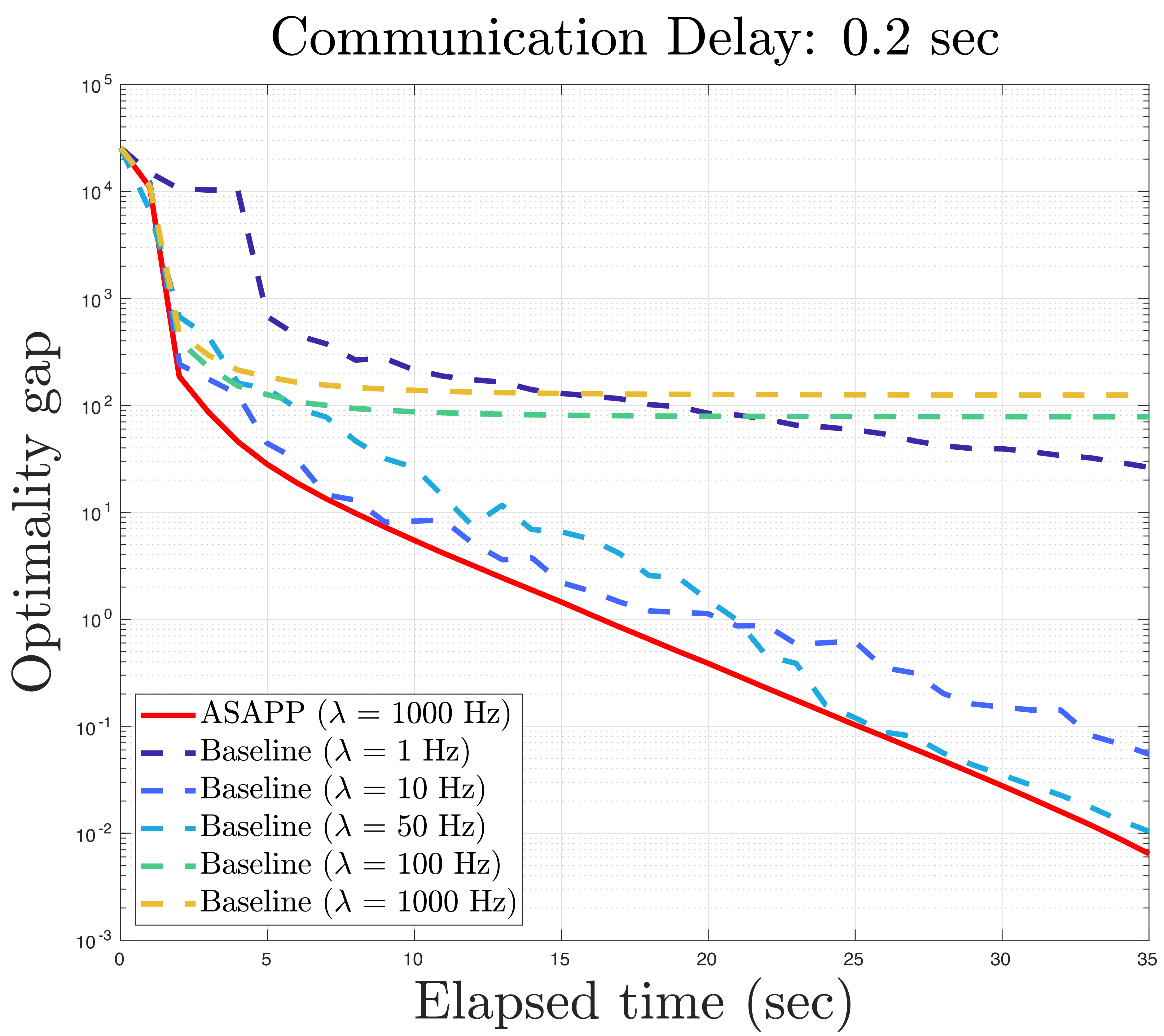
Theorem 1 (Global convergence of ASAPP)**.** *Let $f^\star$ be the global minimum. There exists an upper bound on the stepsize $\bar{\gamma}$ such that if* $0 < \gamma \leq \bar{\gamma}$, ASAPP *converges to a first-order critical point with the following rate,*

$$\min_{k \in \{0,\ldots,K-1\}} \mathbb{E}\left[\left\| \operatorname{grad} f(x^k) \right\|^2 \right] \leq \frac{2n(f(x^0) - f^\star)}{\gamma K}.$$

Experimental Results

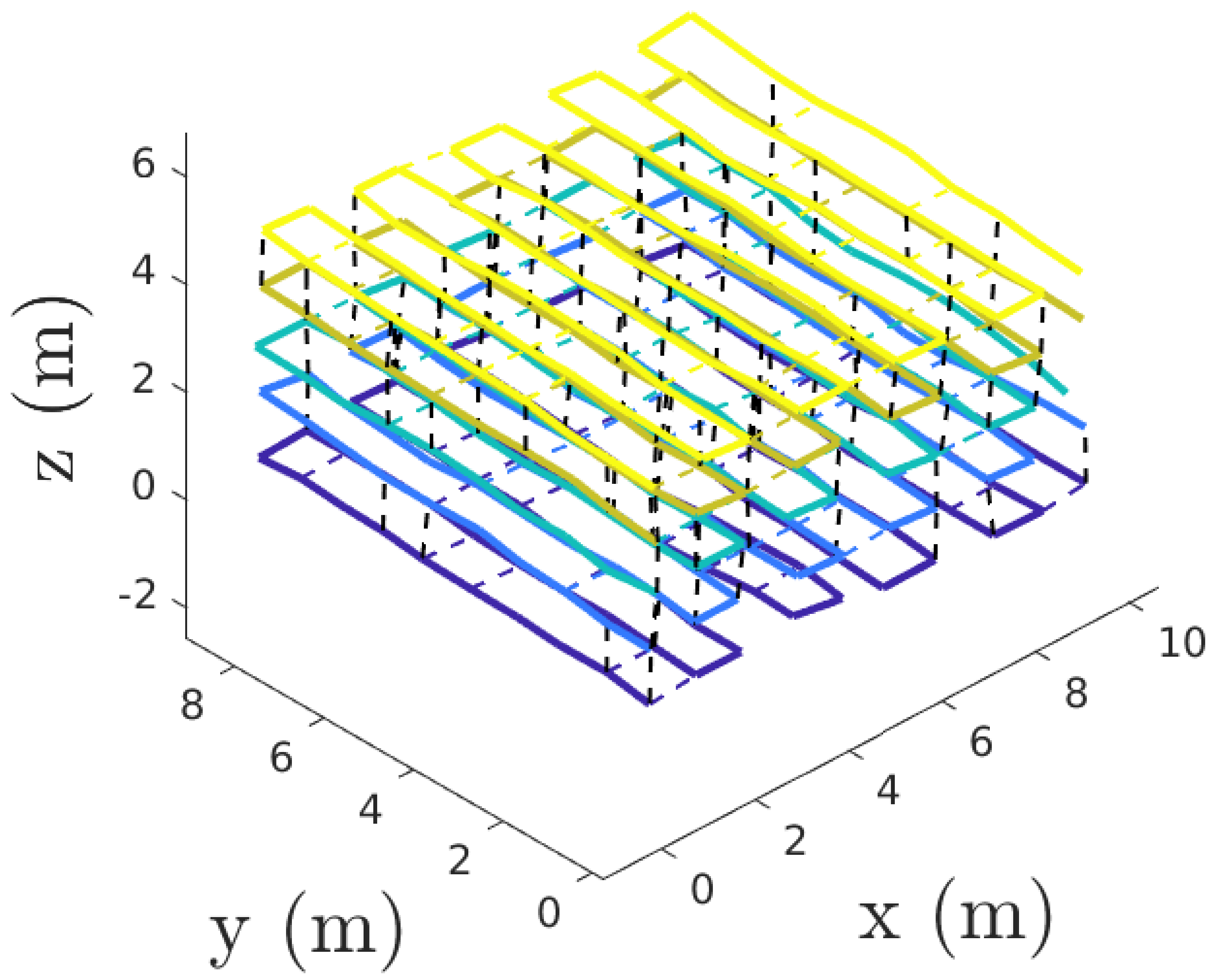


3D Simulation

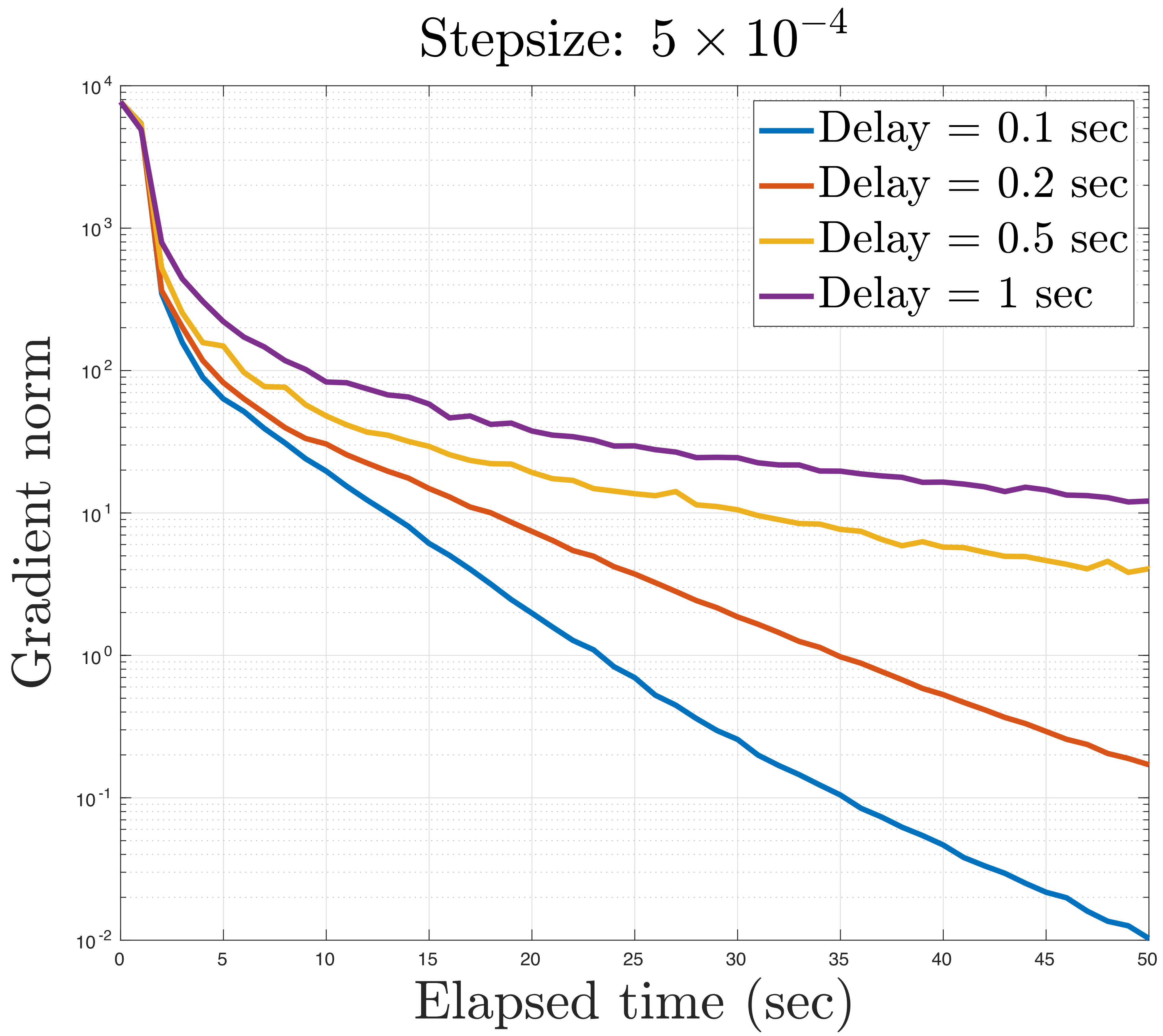


Optimality Gap

Experimental Results



3D Simulation



Performance under Varying Delay

Experimental Results

Dataset	# Poses	# Edges	Cost value f				Additional statistics for ASAPP			
			Opt. [9]	Initial	DGS [1]	ASAPP [ours]	Grad. Norm	Rot. RMSE [chordal]	Trans. RMSE [m]	Stepsize
CSAIL (2D)	1045	1171	31.47	36.10	31.54	**31.51**	0.36	0.0017	0.001	1.0
Intel (2D)	1228	1483	393.7	1205.9	394.6	**393.7**	0.002	2×10^{-6}	2×10^{-6}	1.0
Manhattan (2D)	3500	5453	193.9	1187.2	242.7	**227.7**	5.42	0.07	0.02	0.03
Garage (3D)	1661	6275	1.267	4.477	**1.277**	1.309	0.04	0.014	0.01	0.05
Sphere (3D)	2500	4949	1687.0	3075.7	1743.1	**1711.7**	2.73	0.02	0.01	0.23
Torus (3D)	5000	9048	24227	25812	24305	**24240**	8.38	0.017	0.001	1.0
Cubicle (3D)	5750	16869	717.1	916.2	**720.8**	734.5	12.67	0.030	0.001	0.065

TABLE I: Evaluation on benchmark PGO datasets. Each dataset is divided into trajectories of 5 robots. We run ASAPP for 60 s under a fixed communication delay of 0.1 s. For reference, we also run DGS [1] for 600 synchronous iterations. We compare the final cost values of the two approaches, and highlight the better solution in bold. For ASAPP, we also report the used stepsize, achieved gradient norm, and rotation and translation root mean squared errors (RMSE) with respect to the global minimizer, computed by Cartan-Sync [9].

Conclusion

- We present ASAPP, the first **asynchronous** algorithm for distributed pose graph optimization
- Sufficient conditions for **convergence** w.r.t. worst-case delay, sparsity, and problem smoothness
- **Resilience** against a wide range of communication delay

Massachusetts Institute of Technology

MIT AEROASTRO

DEVCOM ARMY RESEARCH LABORATORY