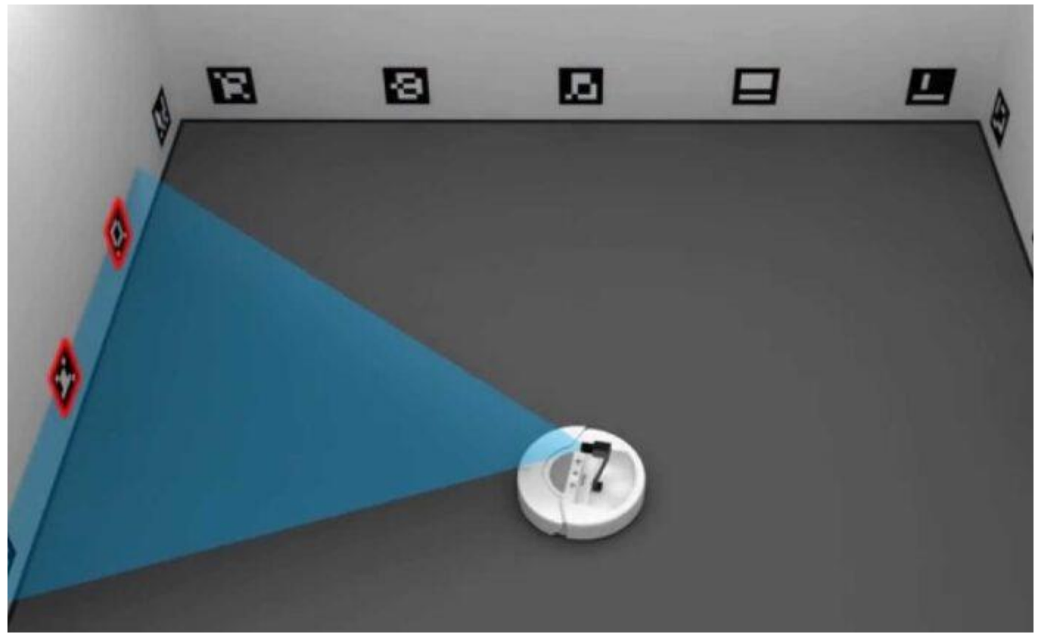
A Slices Perspective for Incremental Nonparametric Inference in High Dimensional State Spaces

Moshe Shienman, Ohad Levy-Or, Michael Kaess and Vadim Indelman

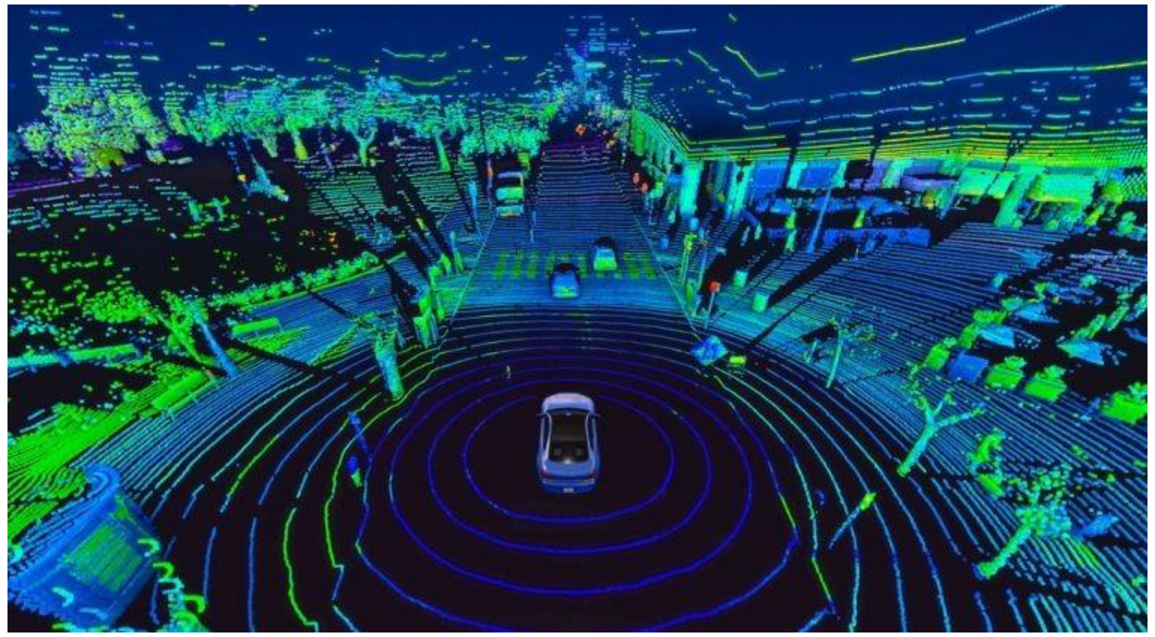
Technion, Israel Institute of Technology

Motivation

Modern autonomous robots perform a wide range of tasks



Localization

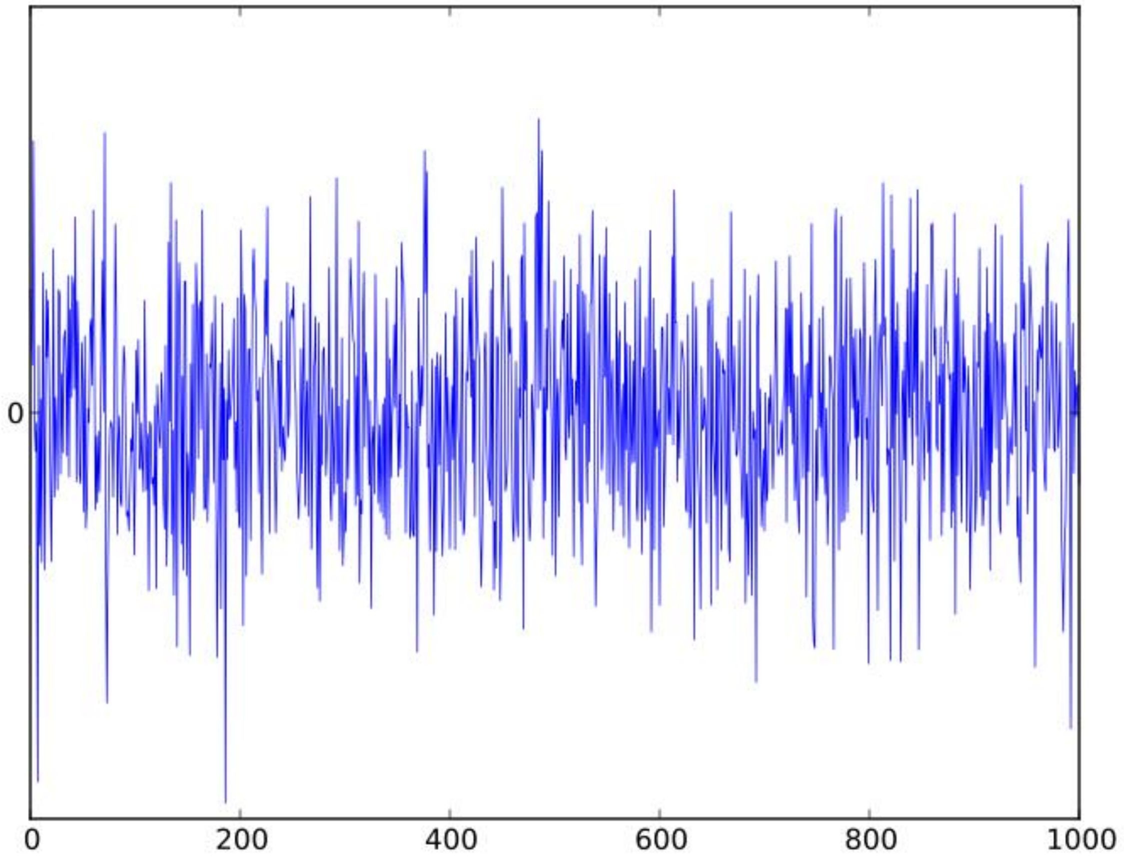


Mapping

Object manipulation

Motivation

Operating under different sources of uncertainty



Noisy measurements

Imprecise actions

Dynamic environments

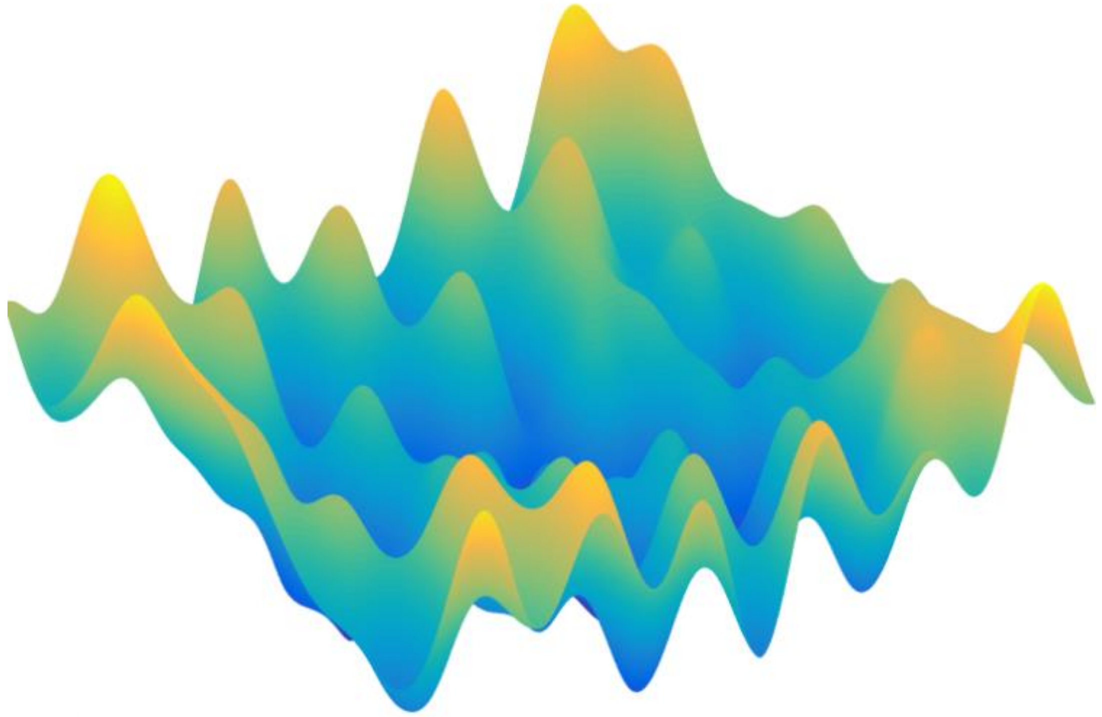
→ Posterior distributions estimates

Motivation

Challenges

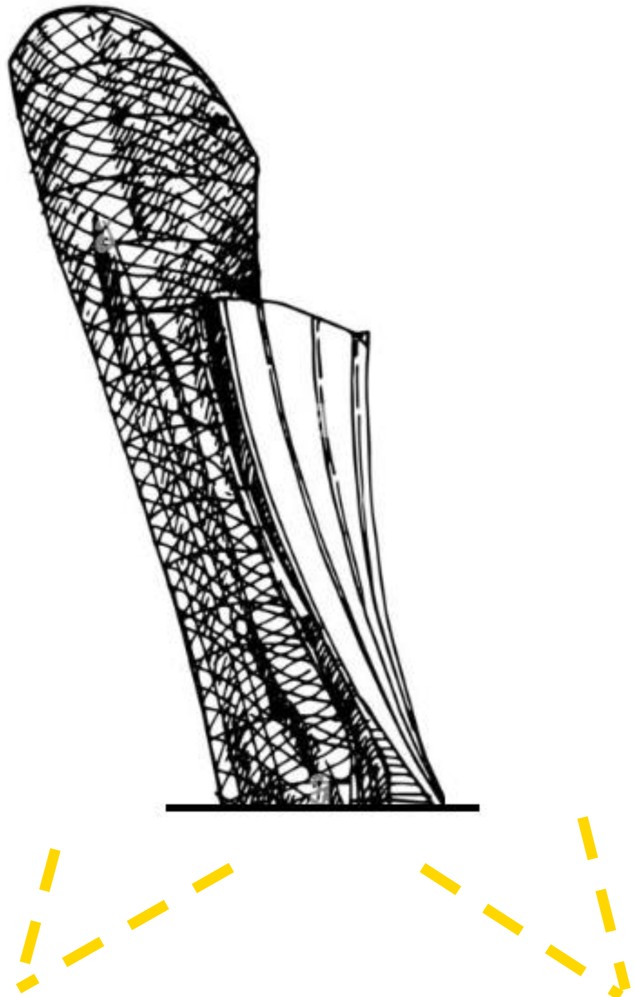
- Handling high-dimensional state spaces and operating in real-time

- In real-world problems the posterior distribution is often non-Gaussian

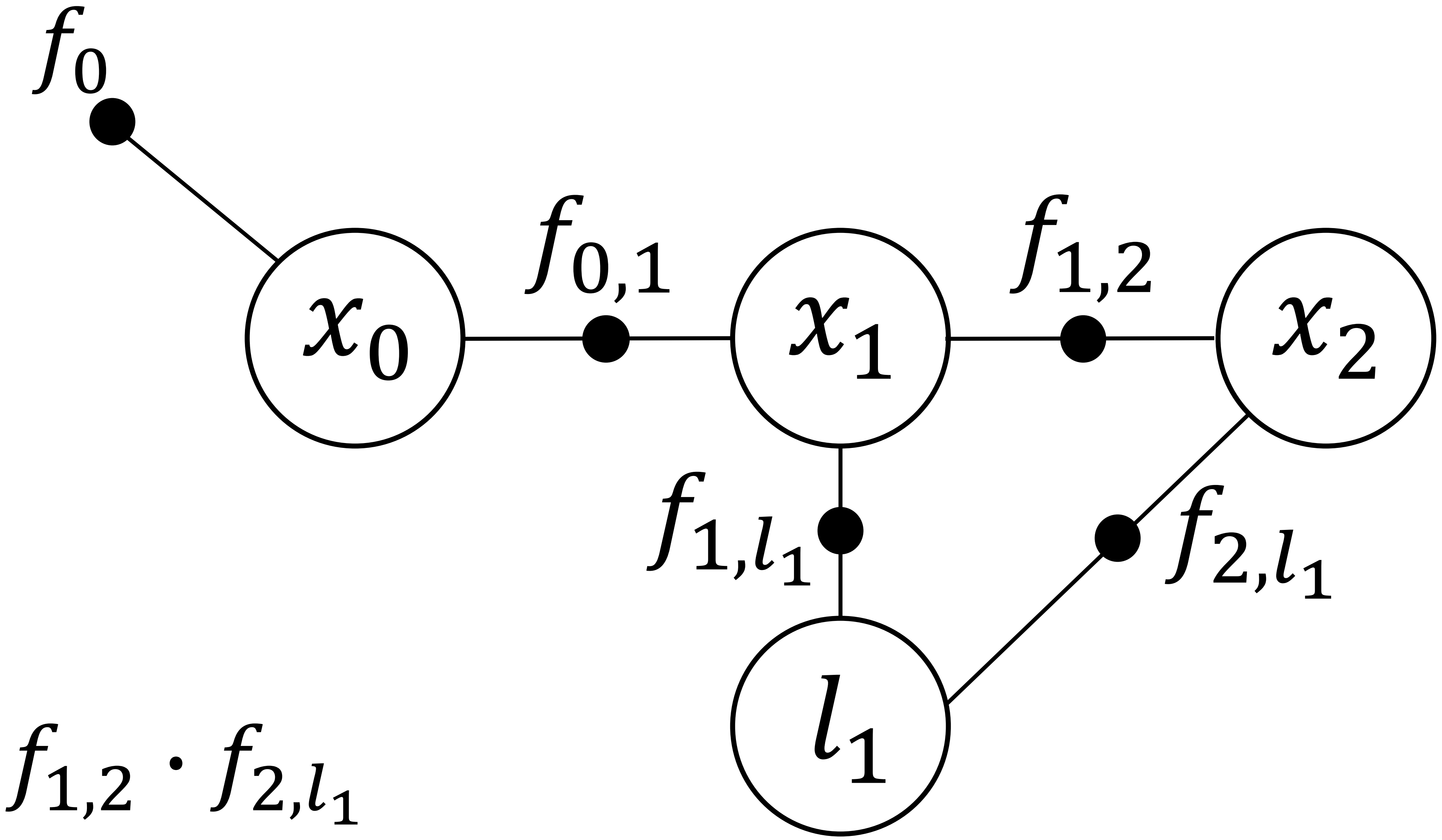


Background

Factor Graph – a probabilistic graphical model

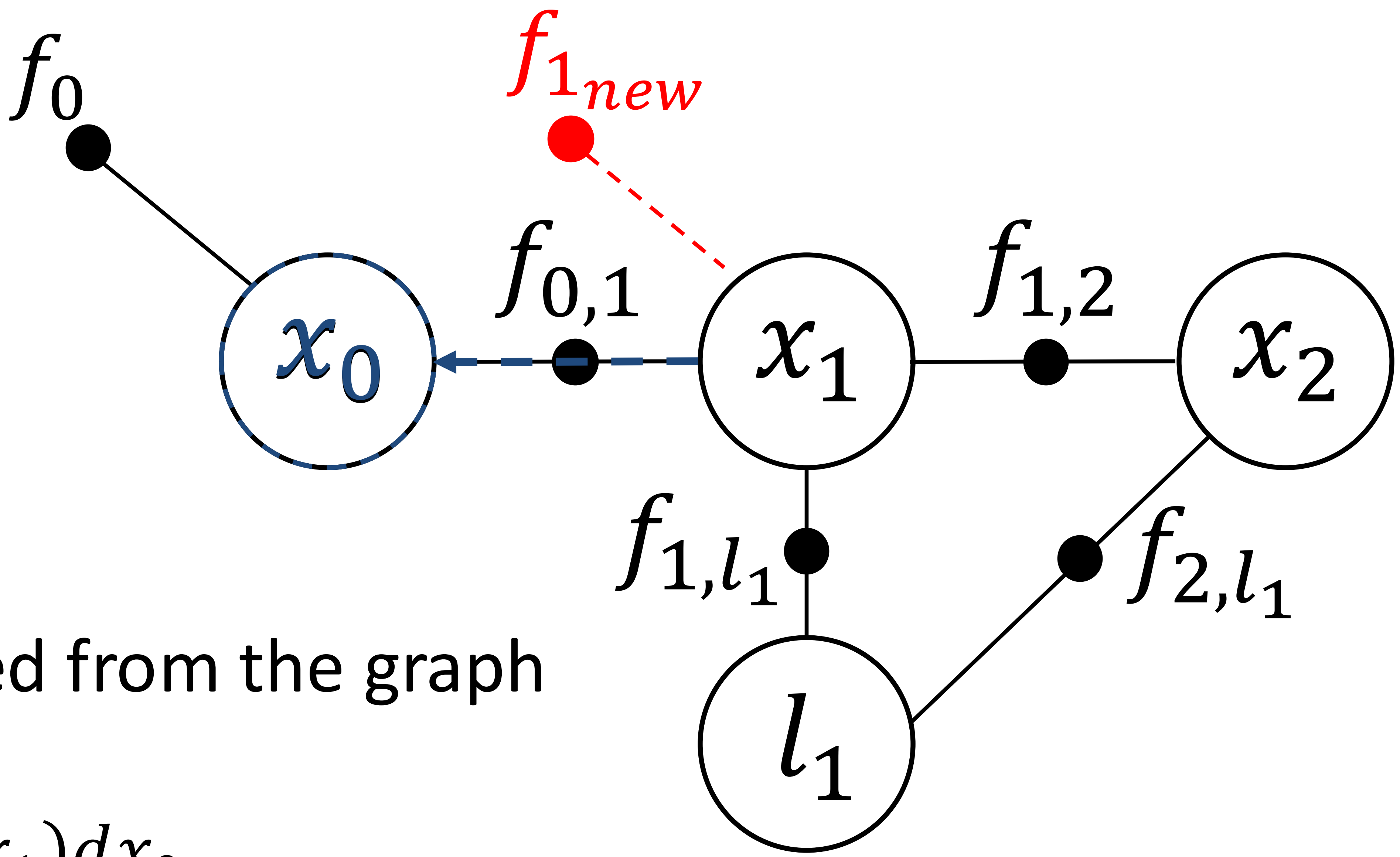


$$p(x_0, x_1, x_2, l_1) \propto f_0 \cdot f_{0,1} \cdot f_{1,l_1} \cdot f_{1,2} \cdot f_{2,l_1}$$



Background

The *forward-backward* algorithm



forward pass:

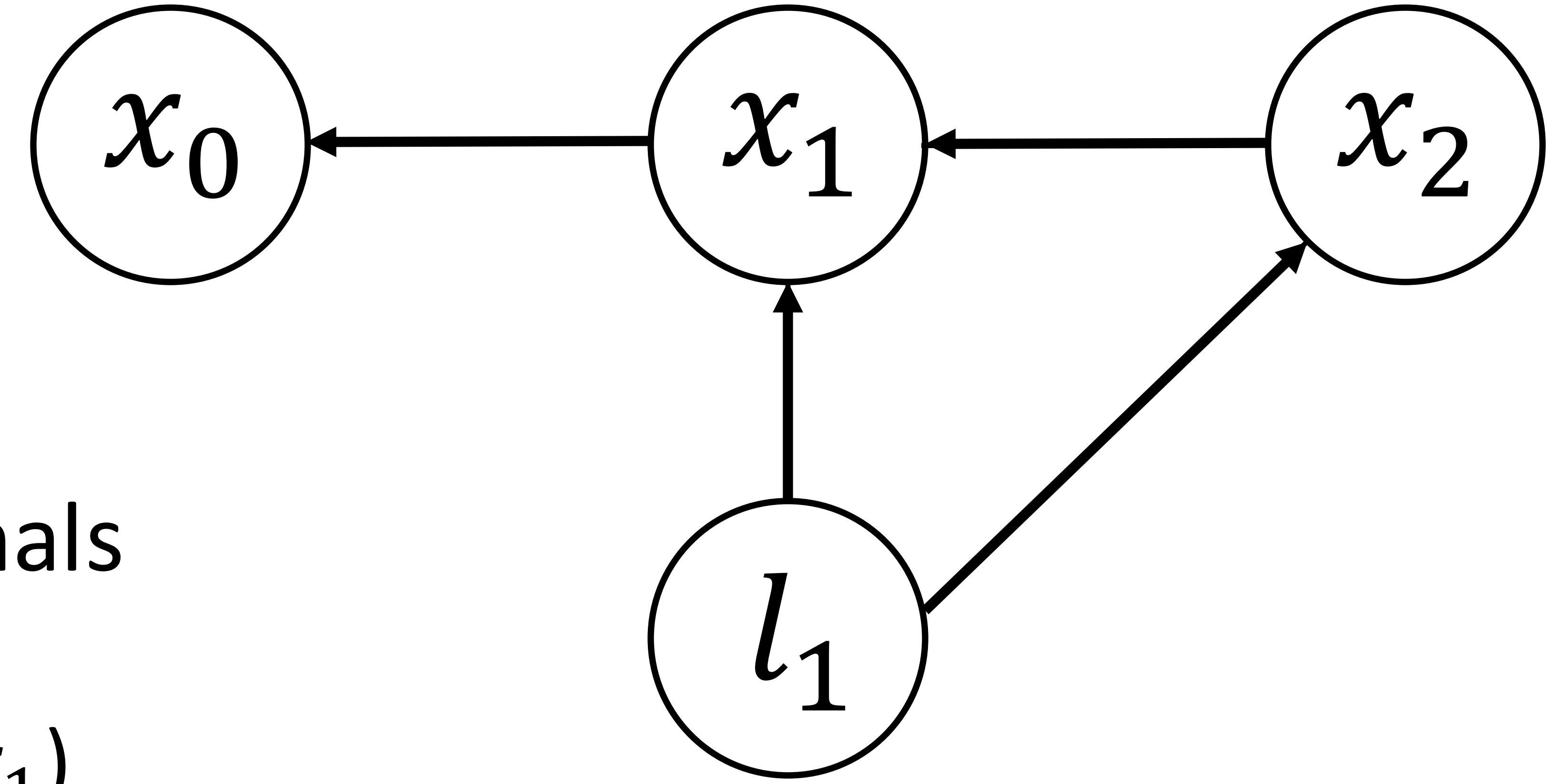
in each step a single variable is eliminated from the graph

$$f_{1_{new}}(x_1) = \underset{x_0 \sim f_0}{\mathbb{E}}[f_{0,1}] = \int_{x_0} f_0(x_0) f_{0,1}(x_0, x_1) dx_0$$

$p(x_0|x_1)$

Background

The *forward-backward* algorithm



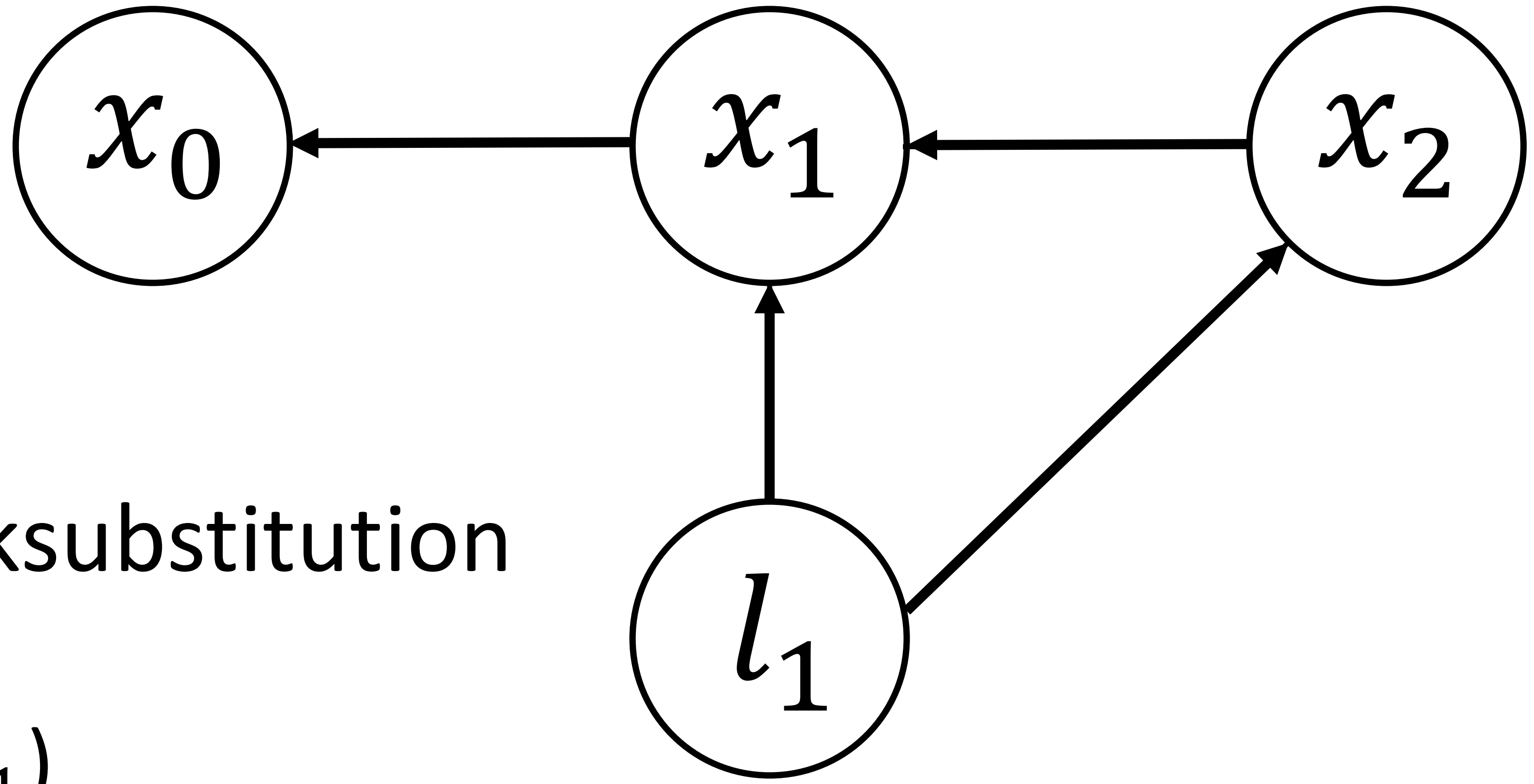
forward pass: when completed
the joint distribution is expressed via conditionals

$$p(x_0, x_1, x_2, l_1) = p(l_1)\, p(x_2|l_1)\, p(x_1|x_2, l_1)\, p(x_0|x_1)$$

Background

The *forward-backward* algorithm

Problem :
No closed-form solution in the general non-parametric case!



backward pass:
marginal distributions are calculated using backsubstitution

$$p(x_0, x_1, x_2, l_1) = p(l_1)\, p(x_2|l_1)\, p(x_1|x_2, l_1)\, p(x_0|x_1)$$

Previous Works

Non-Parametric High Dimensional Settings

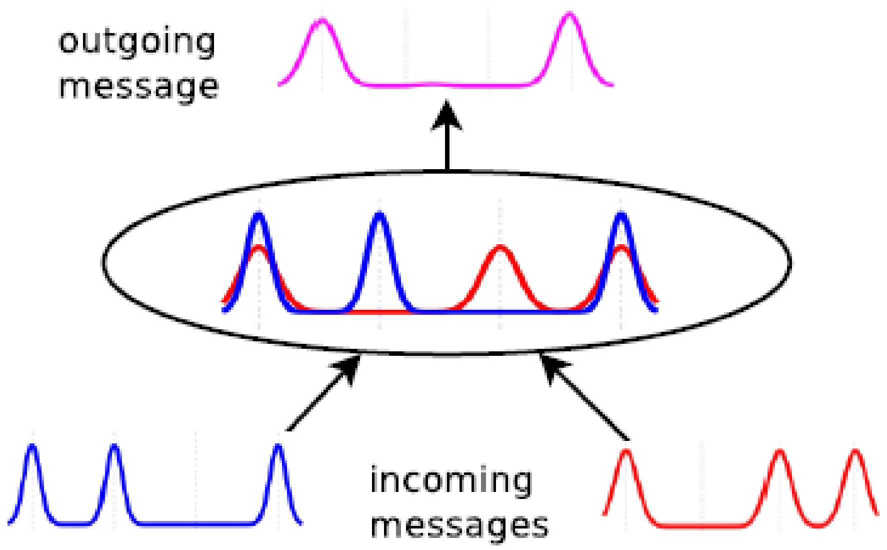


Fig. 1. Illustration of a Bayesian clique operation as part of a larger multi-modal belief propagation on a Bayes tree. Two incoming messages are *combined* with local potentials to produce one outgoing message during the upward pass procedure towards the root. Multi-modality is allowed to exist amongst cliques, rather than selecting a single mode as a maximum-product type algorithm would.

[Fourie et al. IROS 2016]

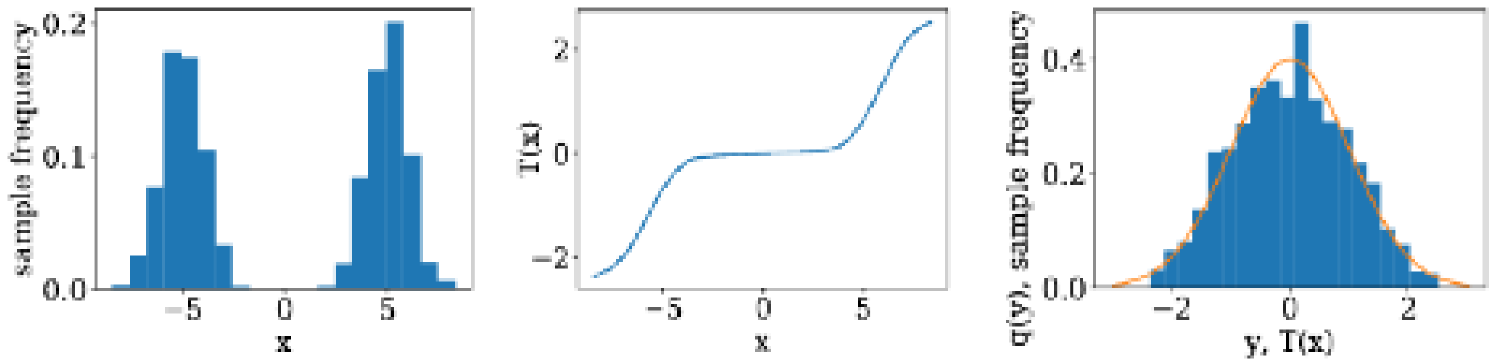
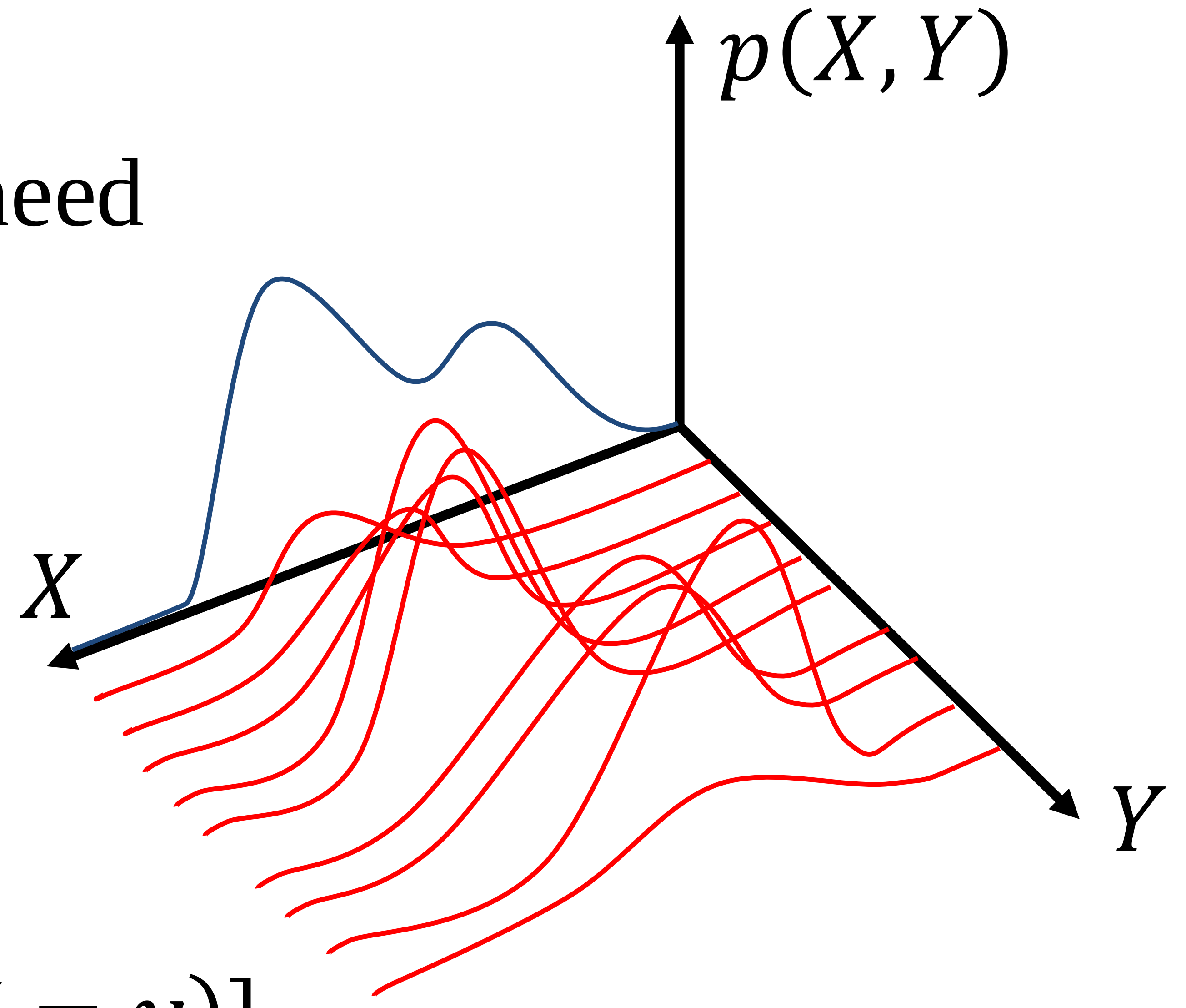


Fig. 2. A one-dimensional example of normalizing flow: histogram of sample x (left), transformation function $T(x)$ (middle), and histogram of transformed samples and reference variable $y \sim N(0, 1)$ (right).

[Huang et al. ICRA 2021]

Key Observation



Distributions can be directly reconstructed without the need for any additional learning techniques or KDE

Example:

joint: $p(X,Y) = p(X|Y)p(Y)$

marginal: $p(X) = \int_Y p(X|Y)p(Y)dY = \underset{y\sim p(Y)}{\mathbb{E}}[p(X|Y=y)]$

estimated marginal: $\hat{p}(X) = \frac{1}{N}\sum_{i=1}^{N} p(X|Y=y^i)$

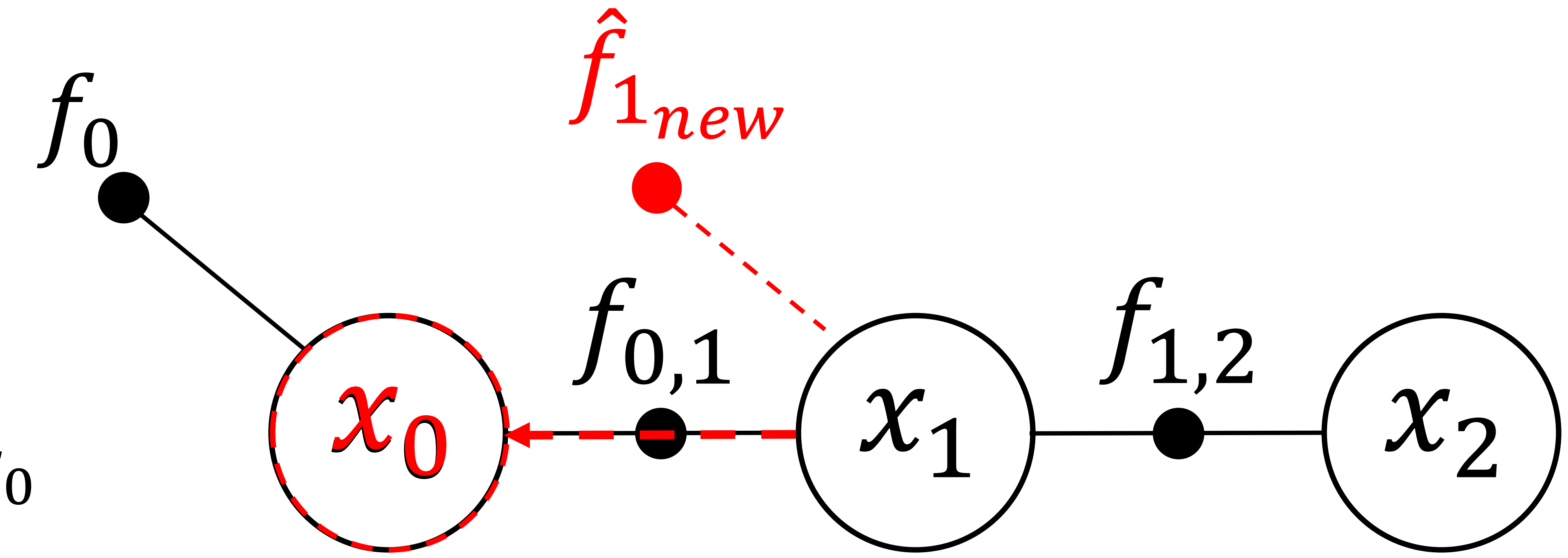
Our Contribution (#1)

Leverage *slices* from high-dimensional surfaces to approximate joint and marginal posterior distributions without any further intermediate reconstructions

forward pass:

$$f_{1_{new}}(x_1) = \eta \int_{x_0} f_0(x_0) f_{0,1}(x_0, x_1) dx_0$$

$$\hat{f}_{1_{new}}(x_1) = \frac{\eta}{N} \sum_{i=1}^{N} f_{0,1}(x_0^i, x_1)$$

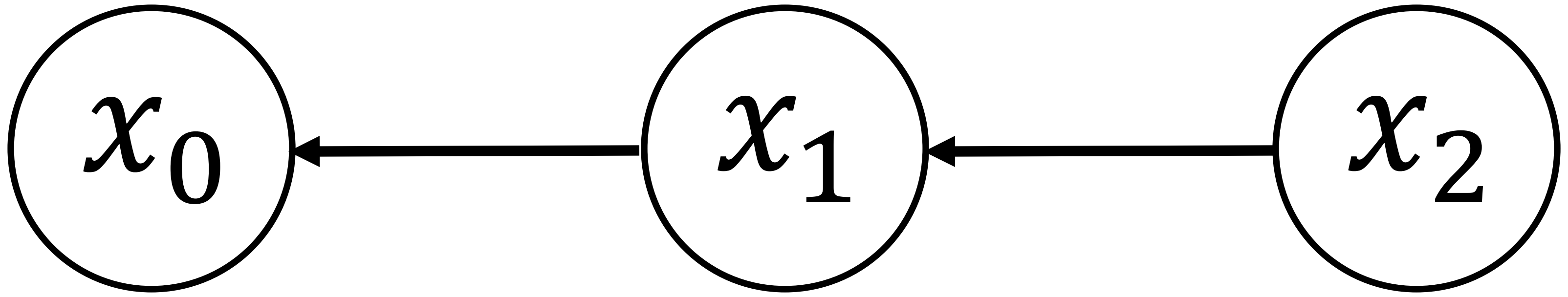


Our Contribution (#1)

Leverage *slices* from high-dimensional surfaces to approximate joint and marginal posterior distributions without any further intermediate reconstructions

backward pass:

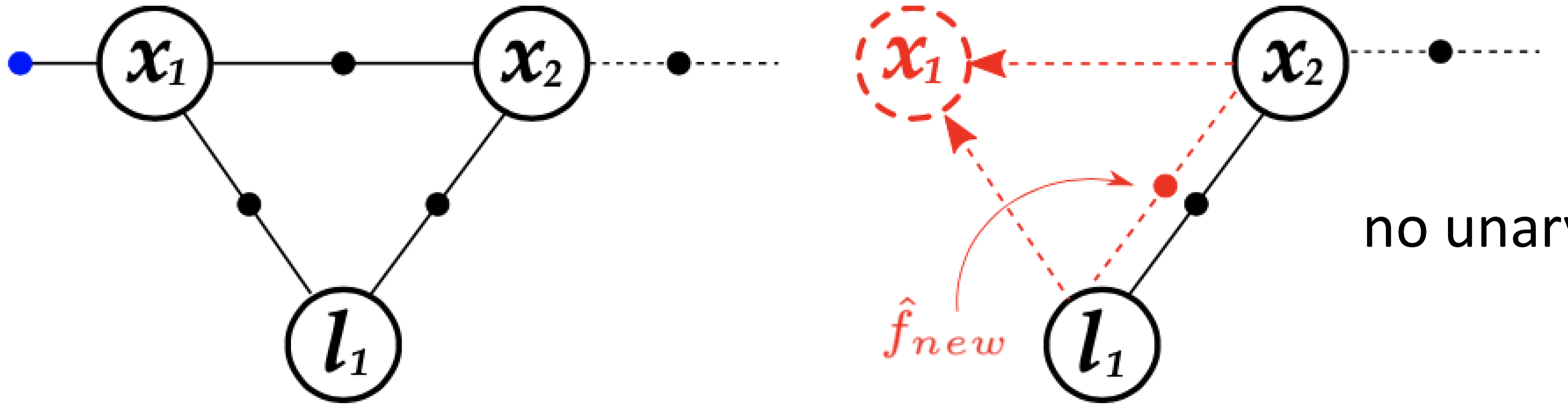
$$\hat{p}(x_1) = \frac{1}{N}\sum_{i=1}^{N} \hat{p}\left(x_1 | x_2^i\right)$$



Our Contribution (#1)

- Our approach supports the elimination of variables even in the **absence** of unary factors (e.g. no GPS)

- For instance, elimination of x_2 or l_1 after eliminating x_1

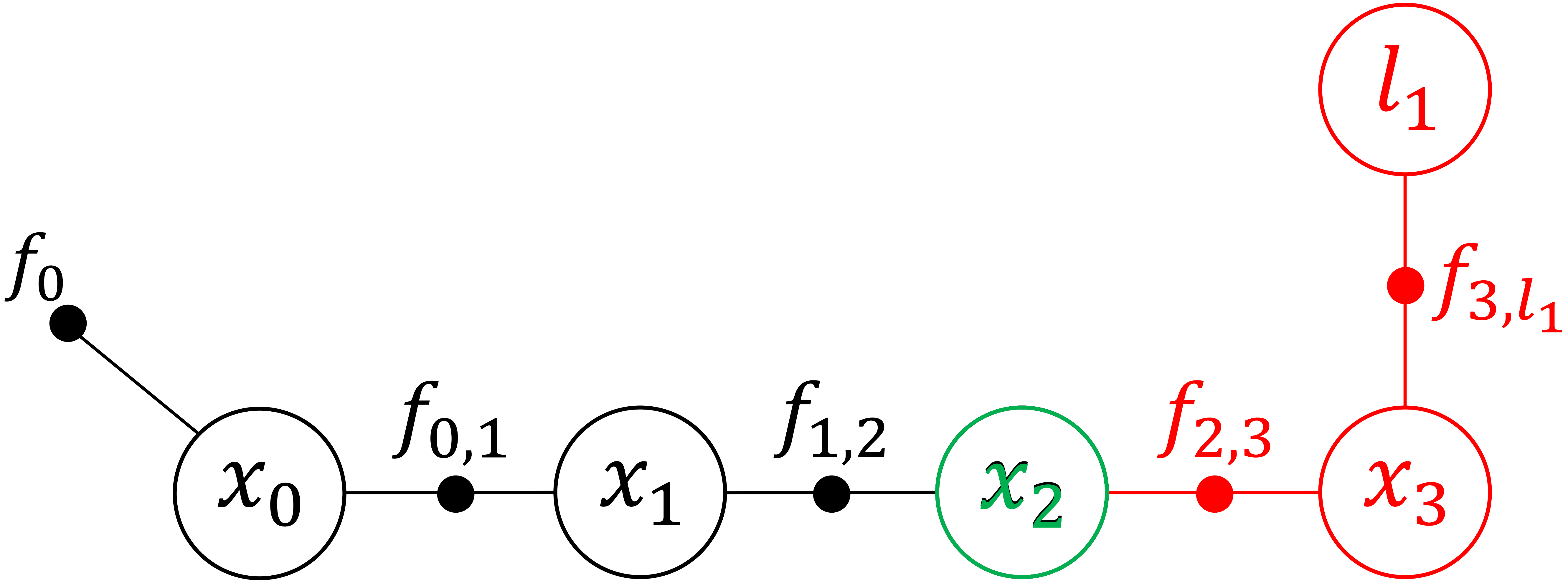


- See Lemma 1 in the paper for details

Lemma 1: Given a factor graph $G = (\mathcal{F}, \Theta, \mathcal{E})$ and an elimination order $\mathcal{O}$, if each eliminated variable $\theta_j \in \Theta$ either has a unary factor connected to it, i.e. $\exists f(\theta_j) \in \mathcal{F}_{j-1}(\theta_j)$ or, $\exists \theta_i \in \Theta$ such that θ_i was previously eliminated and $f(\theta_i, \theta_j) \in \mathcal{F}$, then samples of θ_j can be drawn from one of the factors $\mathcal{F}_{j-1}(\theta_j)$ (see proof in appendix).

Our Contribution (#2)

A novel early stopping heuristic criteria (*backward* pass) in the non-parametric case

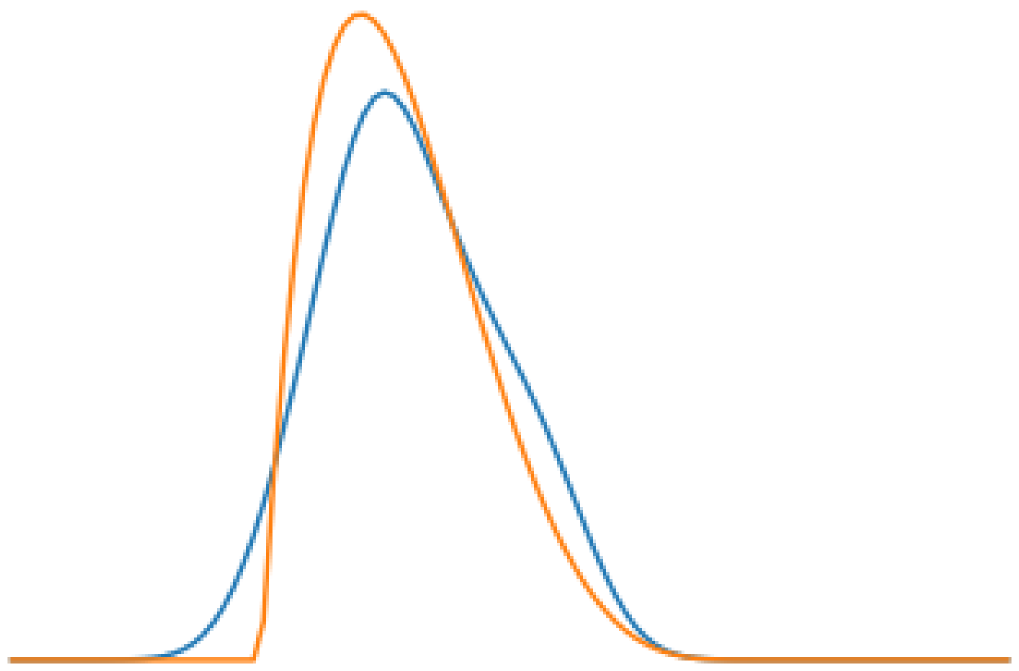


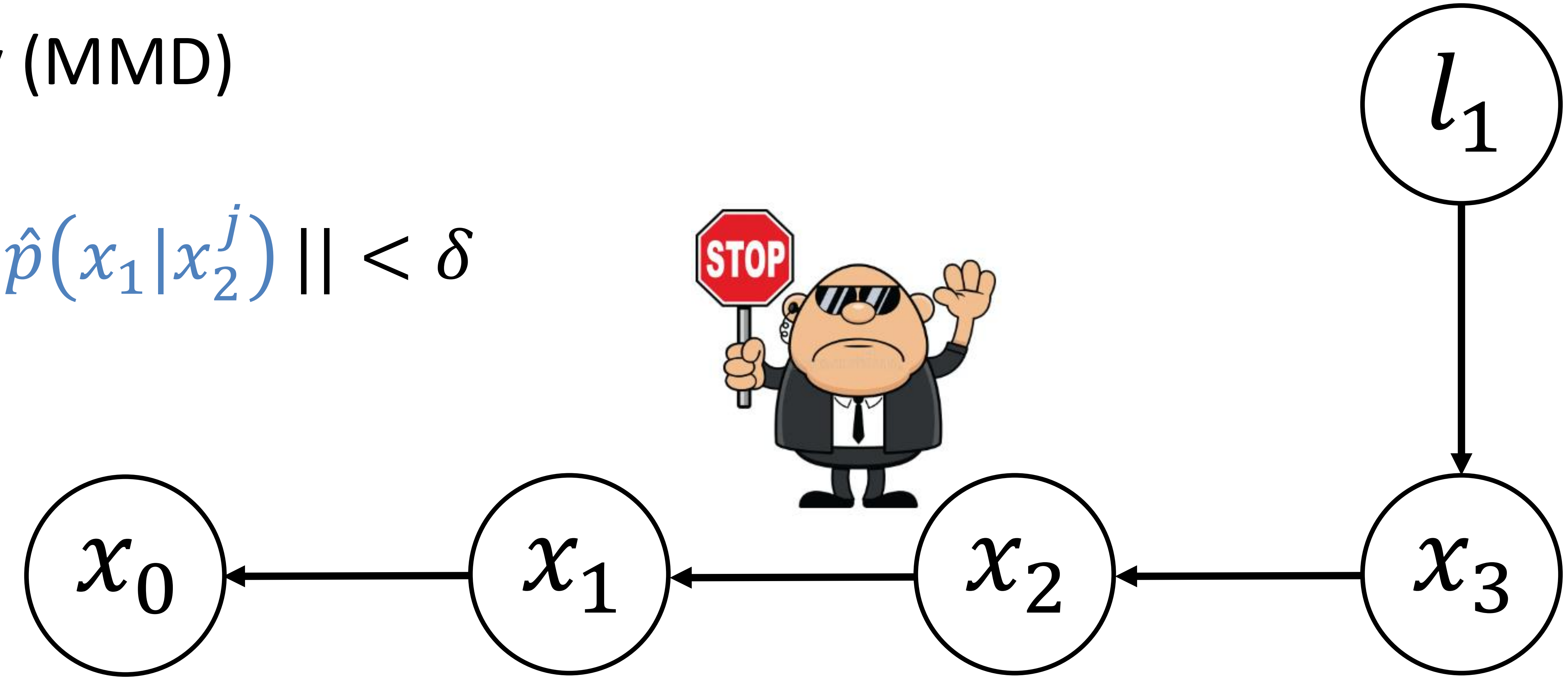
Our Contribution (#2)

A novel early stopping heuristic criteria (*backward* pass) in the non-parametric case

Maximum Mean Discrepancy (MMD)

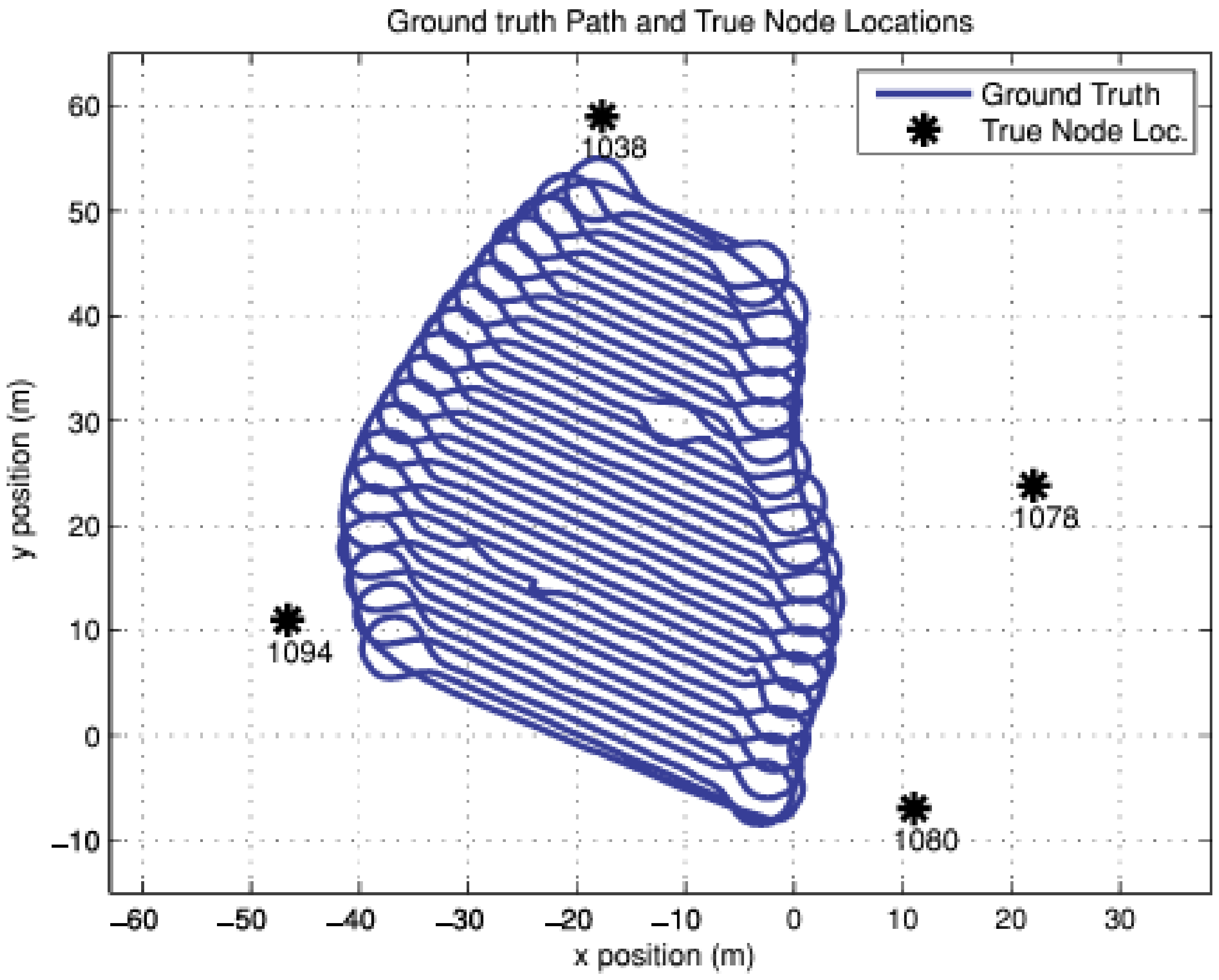
$$|| \frac{1}{N}\sum_{i=1}^{N}\hat{p}\left(x_1|x_2^i\right) - \frac{1}{M}\sum_{j=1}^{M}\hat{p}\left(x_1|x_2^j\right) || < \delta$$

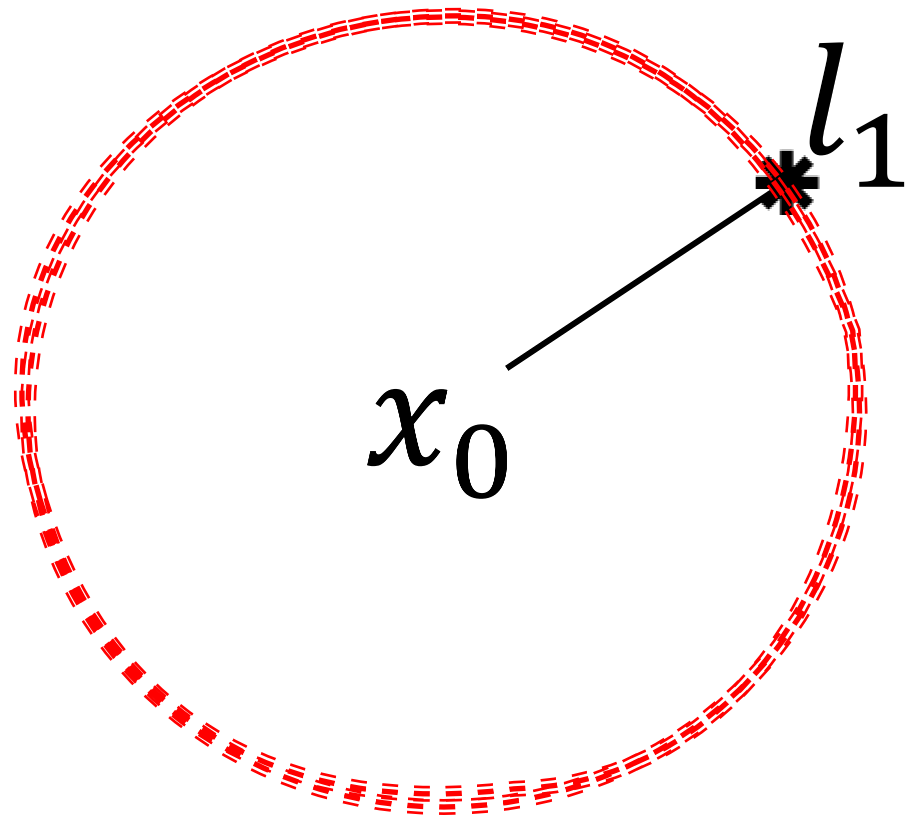




Experimental Results

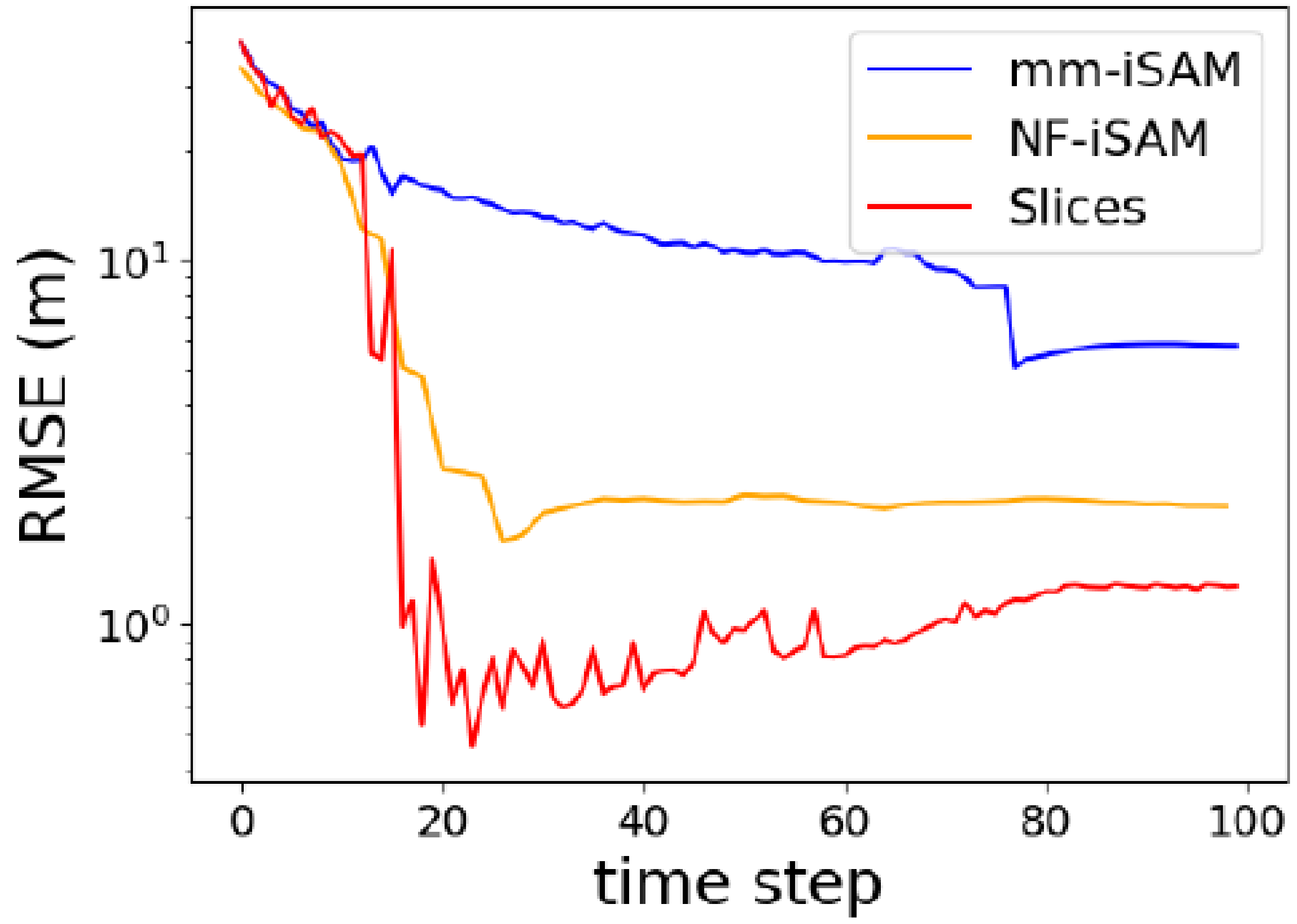
Plaza : a real-world dataset with range measurements [Djugash, et al. JFR 2009]



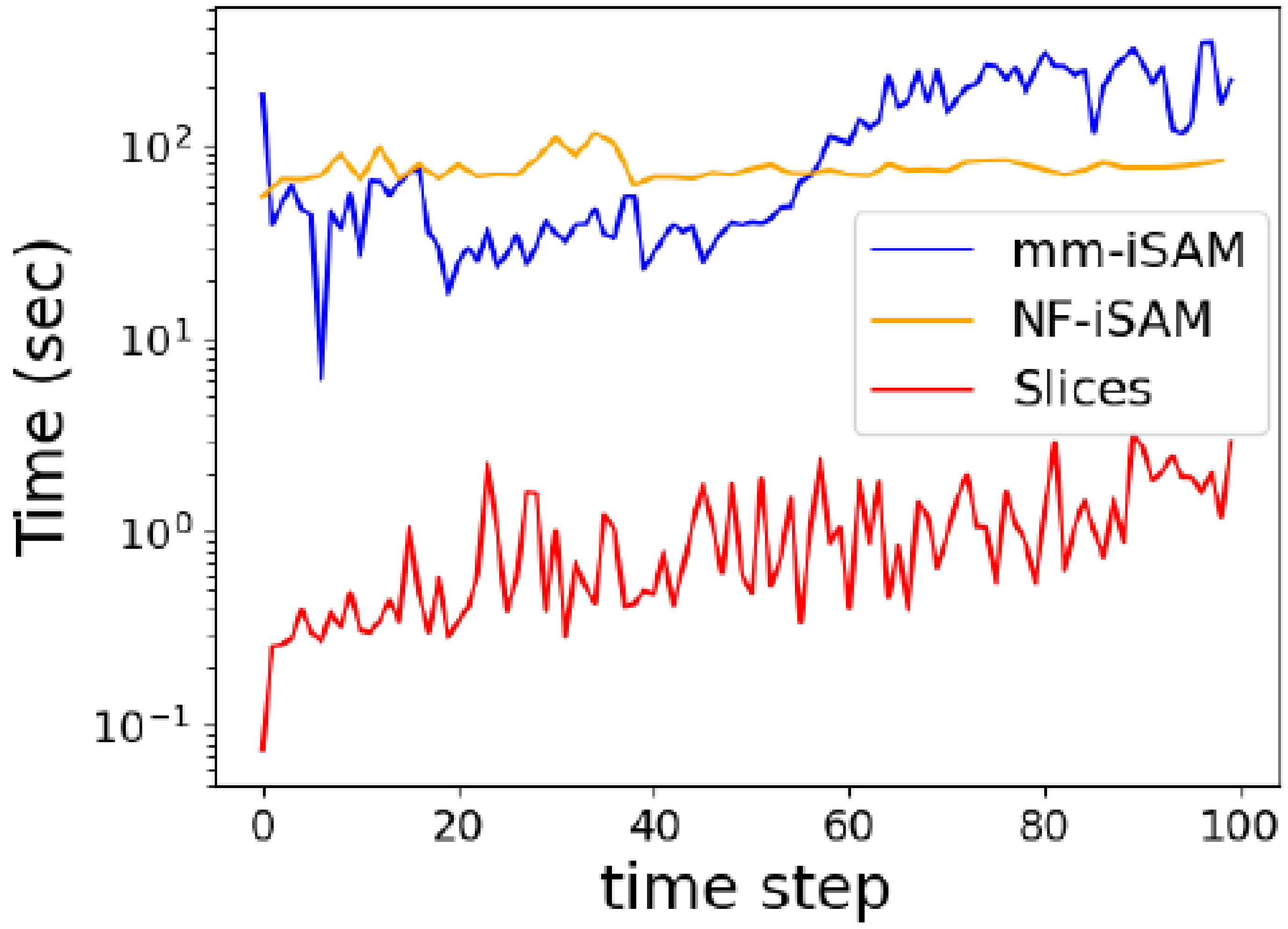


Experimental Results

Plaza : a real-world dataset with range measurements



(a)



(b)

Summary

- Leverage *slices* from high-dimensional surfaces to approximate joint and marginal posterior distributions without any further intermediate reconstructions
- A novel early stopping heuristic criteria to further speed up calculations
- Requires less samples and consistently outperforms state-of-the-art nonparametric inference algorithms in terms of accuracy and computational complexity

Thank you!

IROS '24
ABU DHABI