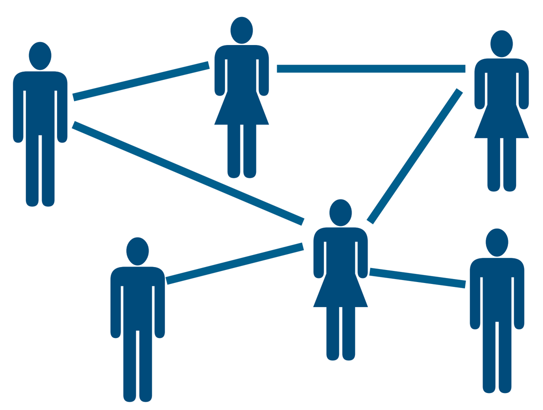
Scaling MPE Inference for Constrained Continuous Markov Random Fields with Consensus Optimization

Stephen H. Bach, Matthias Broecheler, Lise Getoor, and Dianne P. O'Leary

Introduction

- **Probabilistic graphical models** are powerful tools, but joint inference can become **expensive** for large problems
- We **show how to scale up inference** in an **expressive** class of Markov random fields defined over **continuous variables**
- Our method **scales linearly in practice** and can solve problems in 2 minutes that take a commercial interior-point method over 1 hour

Example: Opinion Modeling



Model **combines** influences of **predispositions** and **social network** on strengths of voters' preferences for parties

Preference(, /) ≥ Predisposition(, /)

Link(,) → [Preference(, /) ≥ Preference (, /)]

Preference(,) + Preference(,) ≤ 1

Hinge-loss CCMRFs

- Class of **Markov random fields** defined over **continuous variables** in [0,1] using **hinge-loss potentials** and **linear constraints**
- Density is **log-concave** and defined over a **convex domain**
- Previously used for collective classification, link prediction, entity resolution, trust analysis in social networks, group detection in social networks, multi-relational clustering, and more

Definition

$$f(\mathbf{X}) = \frac{1}{Z(\Lambda)} \exp\left[-\sum_{j=1}^{m} \Lambda_j \phi_j(\mathbf{X})\right]$$

$$\phi_j(\mathbf{X}) = [\max\{\ell_j(\mathbf{X}), 0\}]^{p_j}, \quad p_j \in \{1, 2\}$$

$$\tilde{\mathbf{D}} \equiv \{\mathbf{X} \in [0,1]^n \mid \mathbf{X} \text{ satisfies all constraints }\}$$

The MPE Problem

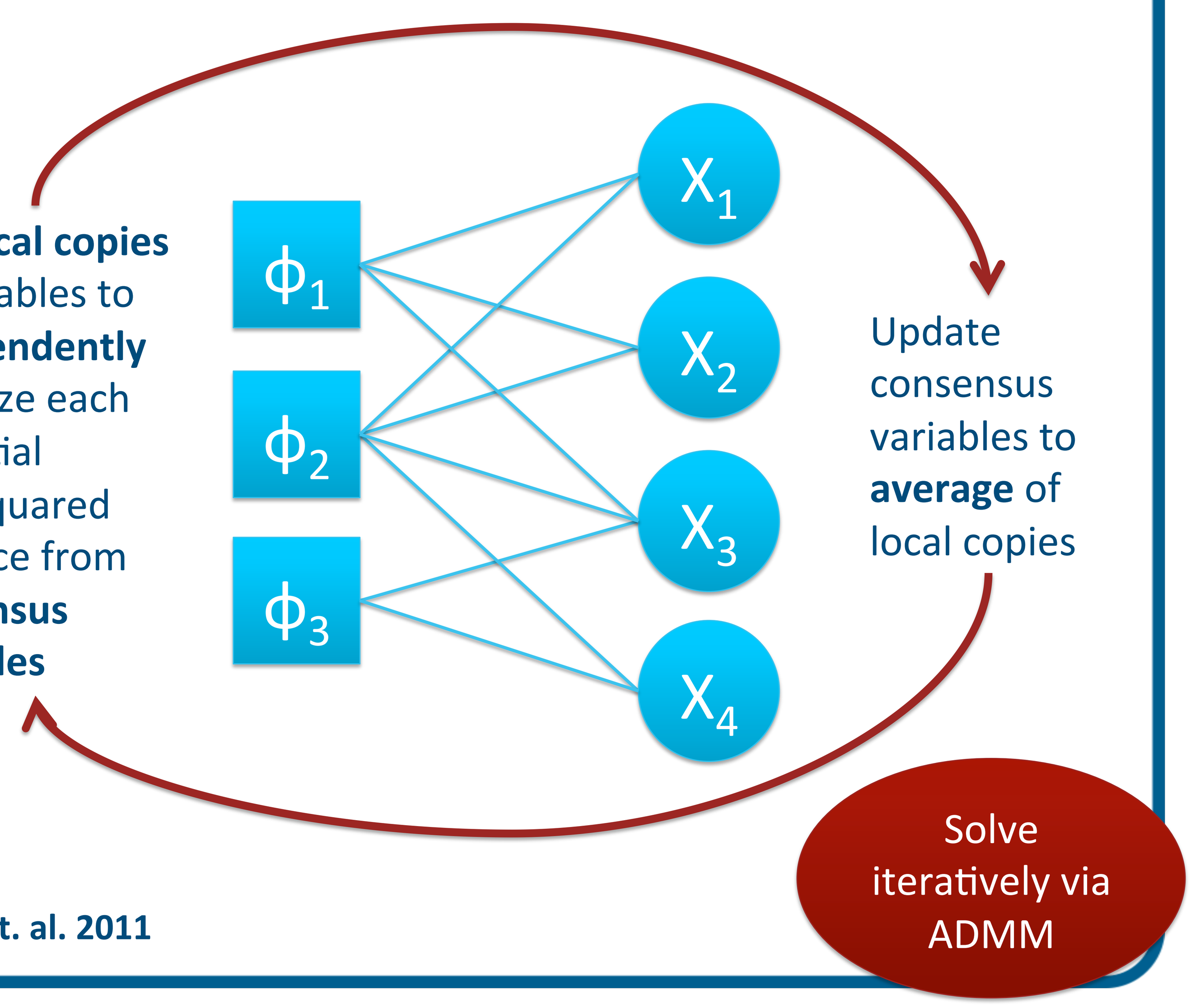
Task is to find a **most-probable explanation** (MPE), an assignment to the random variables that maximizes the probability density:

Objective

$$\arg\max f(\mathbf{X}) \equiv \arg\min_{\mathbf{X} \in \tilde{\mathbf{D}}} \sum_{j=1}^{m} \Lambda_j \phi_j(\mathbf{X})$$

Consensus Optimization

Use **local copies** of variables to **independently** optimize each potential **plus** squared distance from **consensus variables**



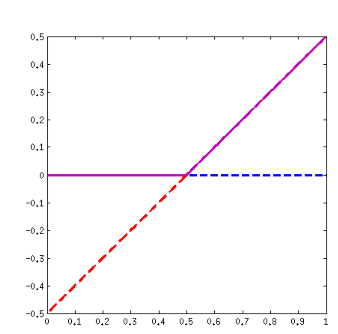
Update consensus variables to **average** of local copies

Boyd, et. al. 2011

Solving Subproblems

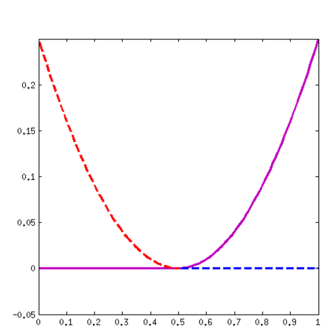
Algorithm: Optimize using modified potentials below. Select an optimizer with same score in both original and modified potentials.

CO-Linear for p = 1



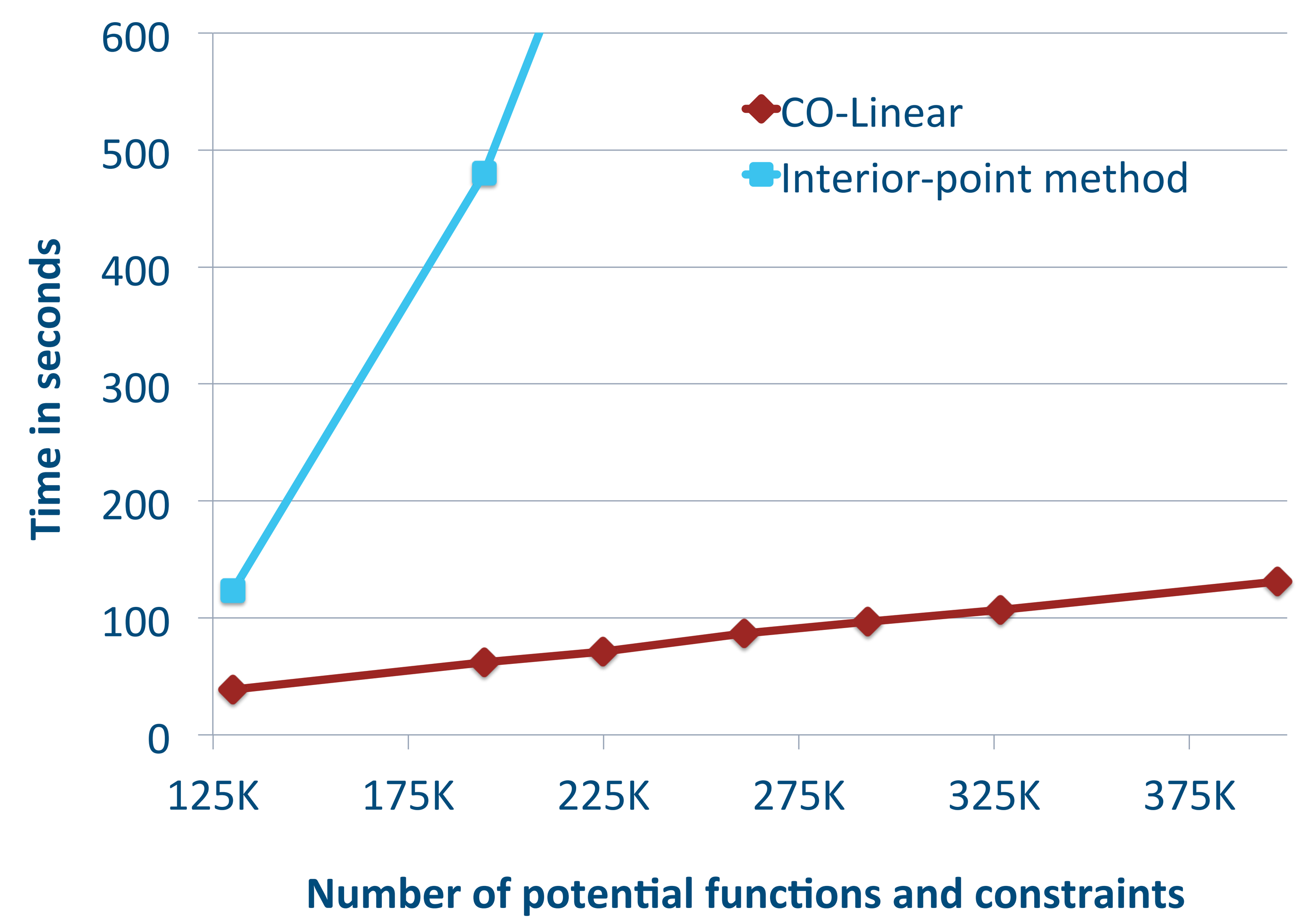
1. without hyperplane
2. with hyperplane
3. where hyperplane is 0

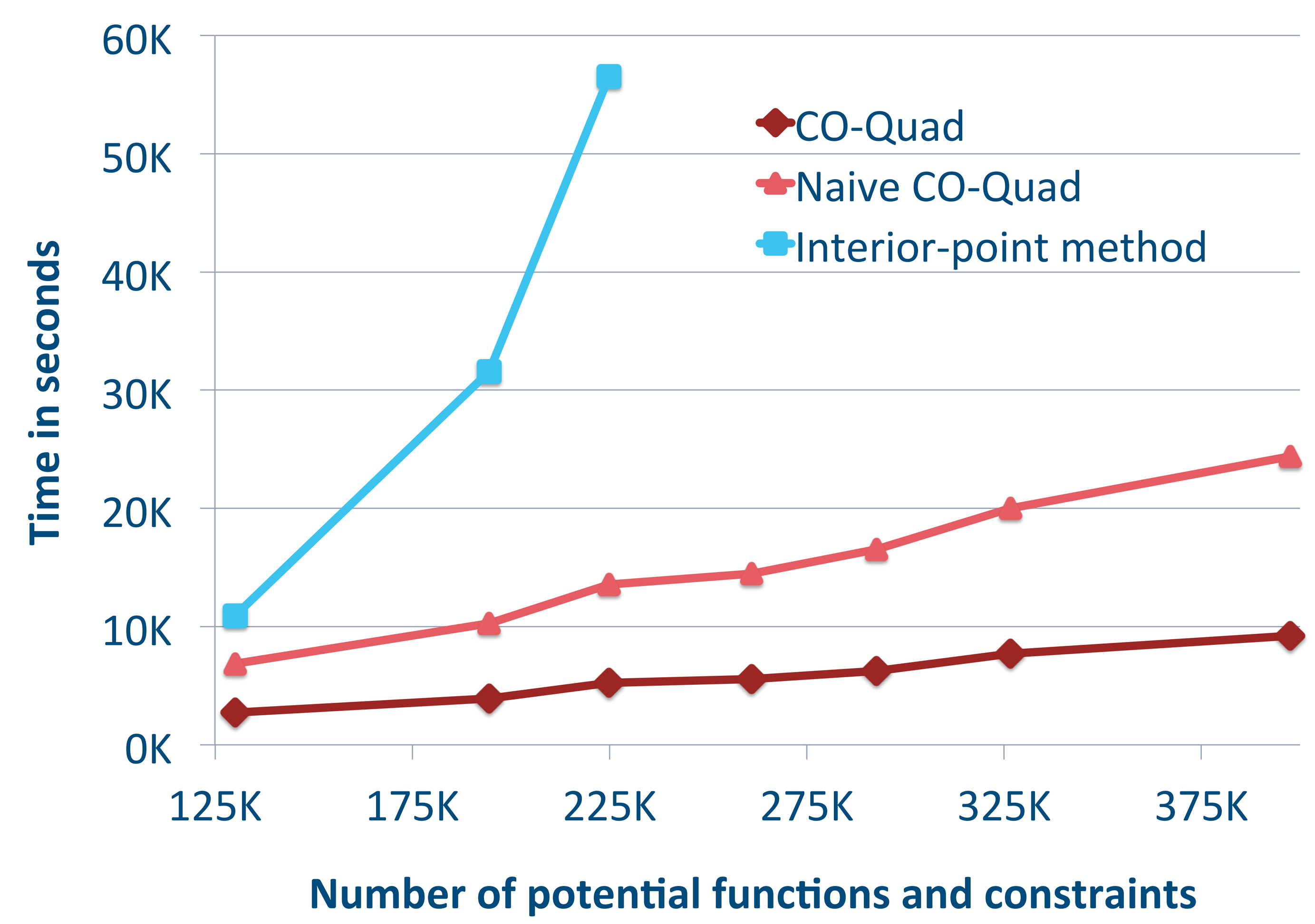
CO-Quad for p = 2



1. without squared hyperplane
2. with squared hyperplane

Experimental Results





Apache-licensed implementation with declarative-language interface:

http://psl.umiacs.umd.edu

This material is based upon work supported by the National Science Foundation under Grant No. 0937094, the Department of Energy under Grant No. DESC0002218, and the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center (DoI/NBC) contract number D12PC00337.