Travesti : TRaffic Volume Estimation by Space-Time Inference

Jean-Marc Lasgouttes Cyril Furtlehner Anne Auger

INRIA, teams Imara and TAO

PEPS meeting, october 8-9 2009

The Travesti project

Partners INRIA (TAO and Imara teams), Mines de Paris (and initially Laboratoire d'Économetrie du Transport, Lyon)

Main goal to provide a macroscopic description of non-linear network-based interactions

Application traffic reconstruction and prediction

Proposed strategy

- identification of relevant degrees of freedom and their statistical interactions
- identifications of relevant macroscopic observables
- design of a Markov random field to encode the dependencies between the degrees of freedom
- design of a (stochastic) dynamical system to encode the dynamics of the macroscopic observables
- setup a real-time decoding algorithm based on message-passing algorithms

The traffic reconstruction problem

Goal reconstruct and predict road traffic on smaller roads where having magnetic loops is not feasible

Solution have a fleet of vehicles that are equipped with GPS units and can send data to the system

Available data FCD (Floating Car Data) are small data packets sent at regular intervals and containing time, car position, speed, . . .

Expected result be able to represent a map (a la Sytadin) of current traffic conditions in a city, and to predict travel times between points

Reconstruction/prediction scheme

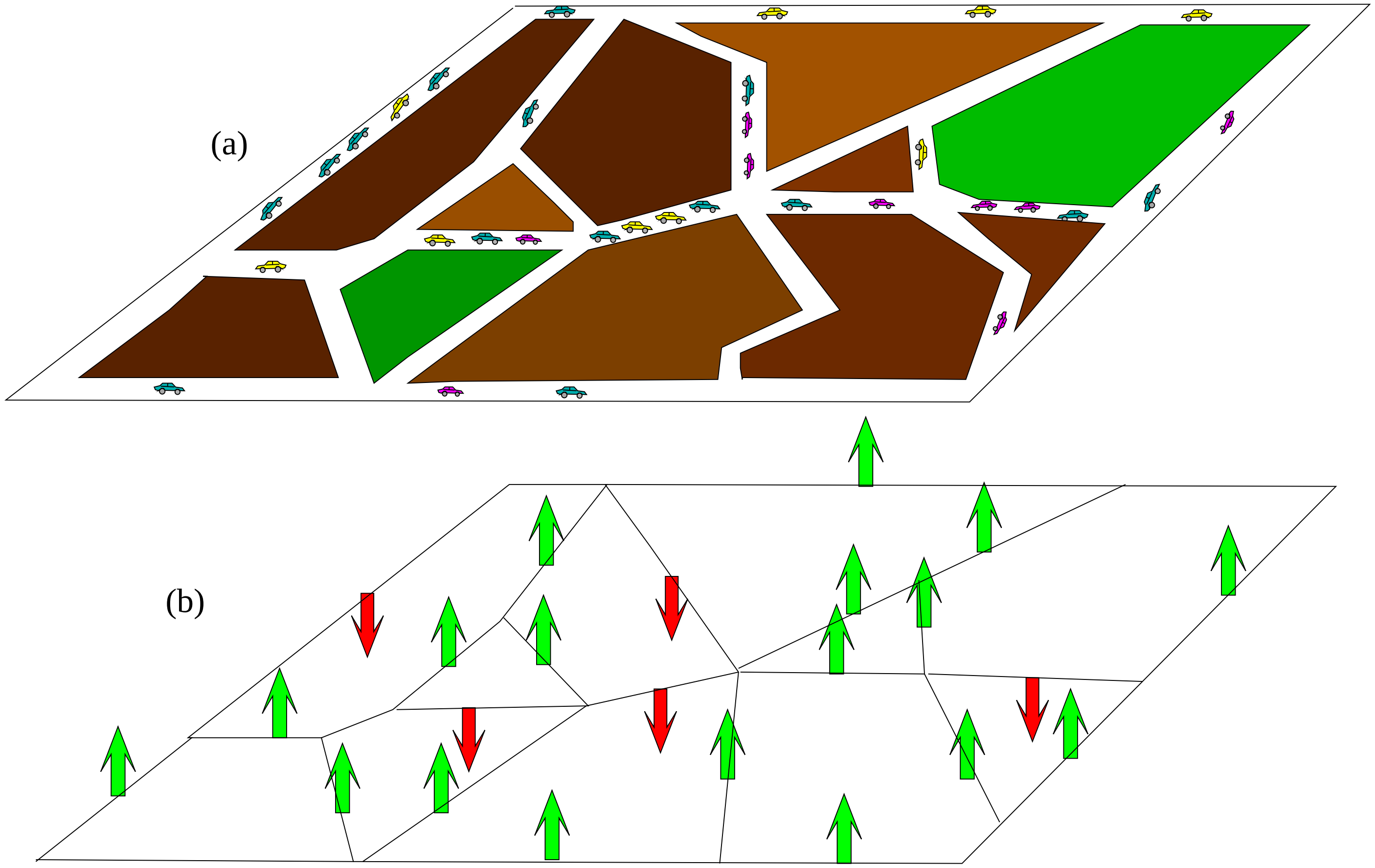
Data assume a large quantity of FCD sent by probe vehicles

Proposed Method

- use the FCD to build a historical database as a list of empirical correlations between consecutive segments of the traffic network.
- Encode these correlations into a graphical model which vertices are couples (segment,date)
- compute the congestion probabilities for each segment conditionally to the real-time data with Belief Propagation algorithm
- Prediction and reconstruction constitute are the two faces of the same problem in this viewpoint.

Space time graph information are propagated in space and time (whence the strange project name)

Traffic network representation



Description of the data

Microscopic description space-time graph

- discretization $i = (\ell, t) \in \mathcal{V}$ represents road segment ℓ at time slot t (a few minutes)
- attached local variable $x_i \in E$ a descriptor of the corresponding segment
  Example : $E = \{0, 1\}$, 0 for congested and 1 for fluid
- information of interest
  - historical data : $\hat{p}(x_i)$, $\hat{p}(x_i, x_j)$ statistical interaction, where $j = (\ell', t+1)$
  - reconstruction and prediction : $p(x_i|\mathcal{V}^*)$, $\mathcal{V}^* =$ set of observed variables.

Macroscopic description

- principal components (linear analysis)
- spatial and temporal patterns (non-linear analysis)
- hierarchical structures
- traffic indexes associated to regions or components
- weather (rain. . .)

The Model

Two components

- a Markov random field on which BP can compute marginals at the microscopic level
- a dynamical system to encode the transition between spatial patterns

Factor graph bipartite graph with hierarchical levels that encapsulates the model

- $\mathcal{V}$ (variables) : vertices corresponding to variables
  - low level (road segments)
  - high level (macroscopic observables)
- $\mathcal{F}$ (factors) : vertices that encode the dependency between variables
  - low level statistical interactions
  - between low level variables and high level indicators
  - between high level indicators for the dynamic of the system

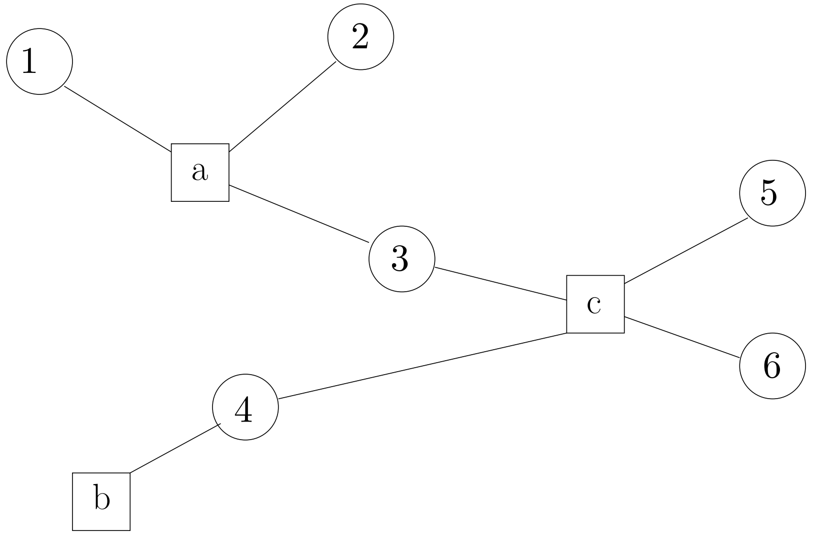
Notation for BP

Joint distribution with $\mathbf{x} = \{x_1, x_2 \ldots, x_N\}$ and $\mathbf{x}_a = \{x_i, i \in a\}$, $a \in \mathcal{F}$ is a set of variable indices

$$p(\mathbf{x}) = \prod_{a \in \mathcal{F}} \psi_a(\mathbf{x}_a) \prod_{i \in \mathcal{V}} \phi_i(x_i)$$

Example

$$p(x_1, \ldots, x_6) = \frac{1}{Z} \psi_a(x_1, x_2, x_3) \psi_b(x_4) \psi_c(x_3, x_4, x_{,5}, x_6)$$



Factor nodes

$$a = \{1, 2, 3\}$$

$$b = \{4\}$$

$$c = \{3, 4, 5, 6\}$$

Belief Propagation on a factor graph

Message update rules

$$m_{a\to i}(x_i) \leftarrow \sum_{\mathbf{x}_{a\setminus i}} \psi_a(\mathbf{x}_a) \prod_{j\in a\setminus i} n_{j\to a}(x_j)$$

where $$n_{i\to a}(x_i) \stackrel{\text{def}}{=} \phi_i(x_i) \prod_{b\ni i, b\neq a} m_{b\to i}(x_i)$$

Resulting beliefs

$$b_i(x_i) \stackrel{\text{def}}{=} \frac{1}{Z_i}\phi_i(x_i) \prod_{a\ni i} m_{a\to i}(x_i) \qquad b_a(\mathbf{x}_a) \stackrel{\text{def}}{=} \frac{1}{Z_a}\psi_a(\mathbf{x}_a) \prod_{i\in a} n_{i\to a}(x_i)$$

In general, messages are normalized so that $\sum_{x_i} m_{a\to i}(x_i) = 1$

The inference problem

The data

- set of variables $x_i \in \{0, \dots, q\},\ i \in \mathcal{V}$
- subsets of variables $a \subset \mathcal{V}$ for which past observations give us marginal laws $\hat{p}_a(\mathbf{x}_a)$
- a distribution law on x of the form

$$p(\mathbf{x}) = \prod_{a \in \mathcal{F}} \psi_a(\mathbf{x}_a) \prod_{i \in \mathcal{V}} \phi_i(x_i)$$

The problem based on a partial observation $\mathbf{x}^* = \{x_i, i \in \mathcal{V}^*\}$ for a subset $\mathcal{V}^*$ of $\mathcal{V}$, what prediction can be made on the complementary set of $\mathcal{V}^*$?

Sub questions

- how to encode historical observations $\hat{p}$?
- how to decode the information in real time, i.e. compute $p(x_i|\mathbf{x}^*)$ for $i \notin \mathcal{V}^*$?

Model calibration : using the "wrong" model

Idea setting the model from the marginals is very expensive, therefore look for ϕ and ψ such that the beliefs match the historical marginals

$$b_a(\mathbf{x}_a) = \hat{p}_a(\mathbf{x}_a)$$

Fixed point property for any fixed point $\mathbf{b}$

$$p(\mathbf{x}) = \prod_{i\in\mathcal{V}} \phi_i(x_i) \prod_{a\in\mathcal{F}} \psi_a(\mathbf{x}_a) = \prod_{i\in\mathcal{V}} b_i(x_i) \prod_{a\in\mathcal{F}} \frac{b_a(\mathbf{x}_a)}{\prod_{i\in a} b_i(x_i)}.$$

Canonical choice for the functions ϕ and ψ

$$\hat{\phi}_i(x_i) \stackrel{\text{def}}{=} \hat{p}_i(x_i), \quad \hat{\psi}_a(\mathbf{x}_a) \stackrel{\text{def}}{=} \frac{\hat{p}_a(\mathbf{x}_a)}{\prod_{i\in a} \hat{p}_i(x_i)}$$ going along with $m_{a\to i}(x_i) \equiv 1$

Model calibration (2)

Property All other choices produce the same fixed points and the same convergence properties

Therefore, setting *one* fixed point determines all the other fixed points of the system

The hope is that these other fixed points are related to the local minima of the free energy

Rewriting of BP

$$m_{a\to i}(x_i) \leftarrow \sum_{\mathbf{x}_{a\setminus i}} \hat{p}_{ai}(\mathbf{x}_a|x_i) \prod_{\substack{j\in a\\ j\neq i}} n_{j\to a}(x_j)$$

where $$n_{j\to a}(x_j) \overset{\text{def}}{=} \prod_{\substack{a'\ni j\\ a'\neq a}} m_{a'\to j}(x_j)$$

Real-time decoding

Conditioning to compute marginals conditional to variables in $\mathcal{V}^*$, one has to consider $\mathsf{P}(x_i|\mathbf{x}^*) = \mathsf{P}(x_i, \mathbf{x}^*)/\,\mathsf{P}(\mathbf{x}^*)$, which is obtained nullifying some messages

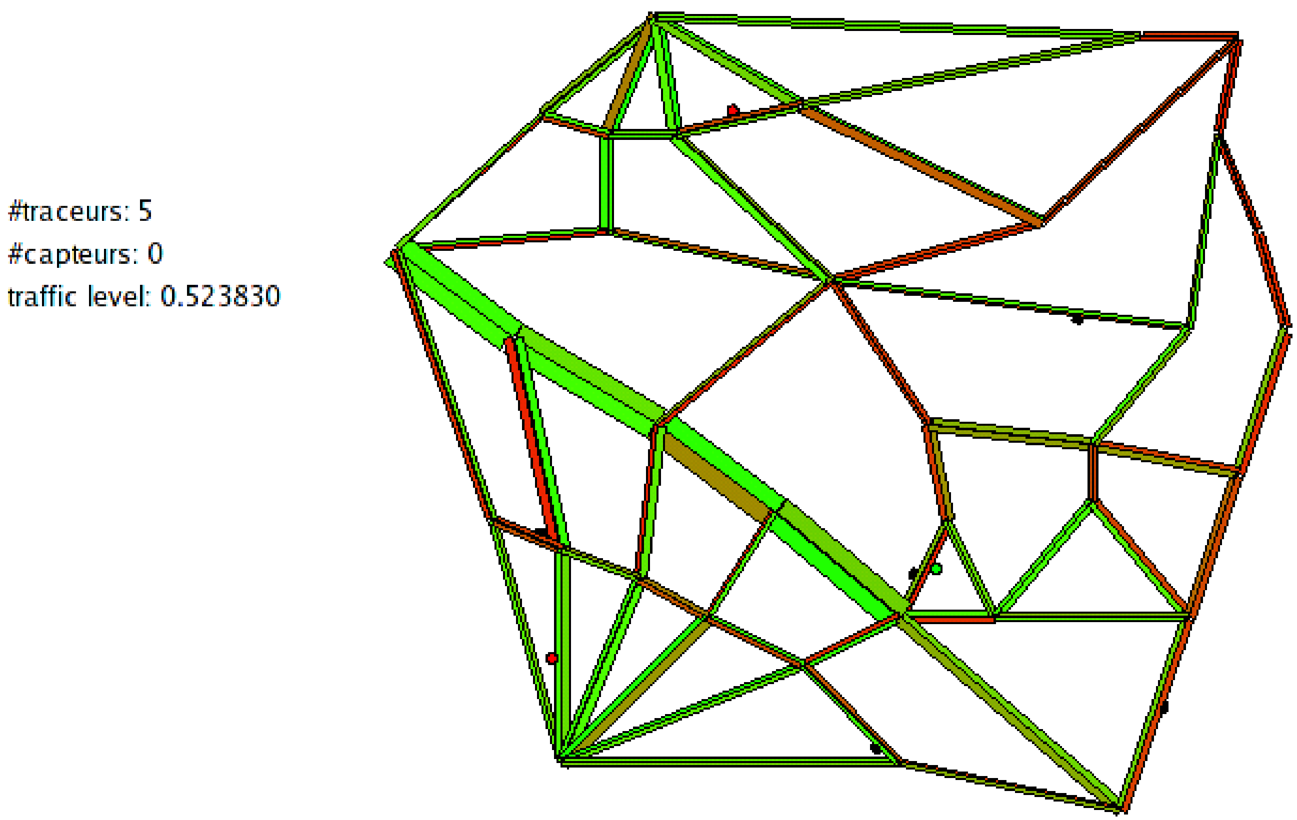
$$n_{i\to a}(x_i) = \prod_{a'\ni i, a'\neq a} m_{b\to i}(x_i)\mathbb{1}_{\{x_i = x_i^*\}} \qquad \text{when } i \in \mathcal{V}^*$$

Uncertain data assuming that we only know $p_i^*(x_i)$

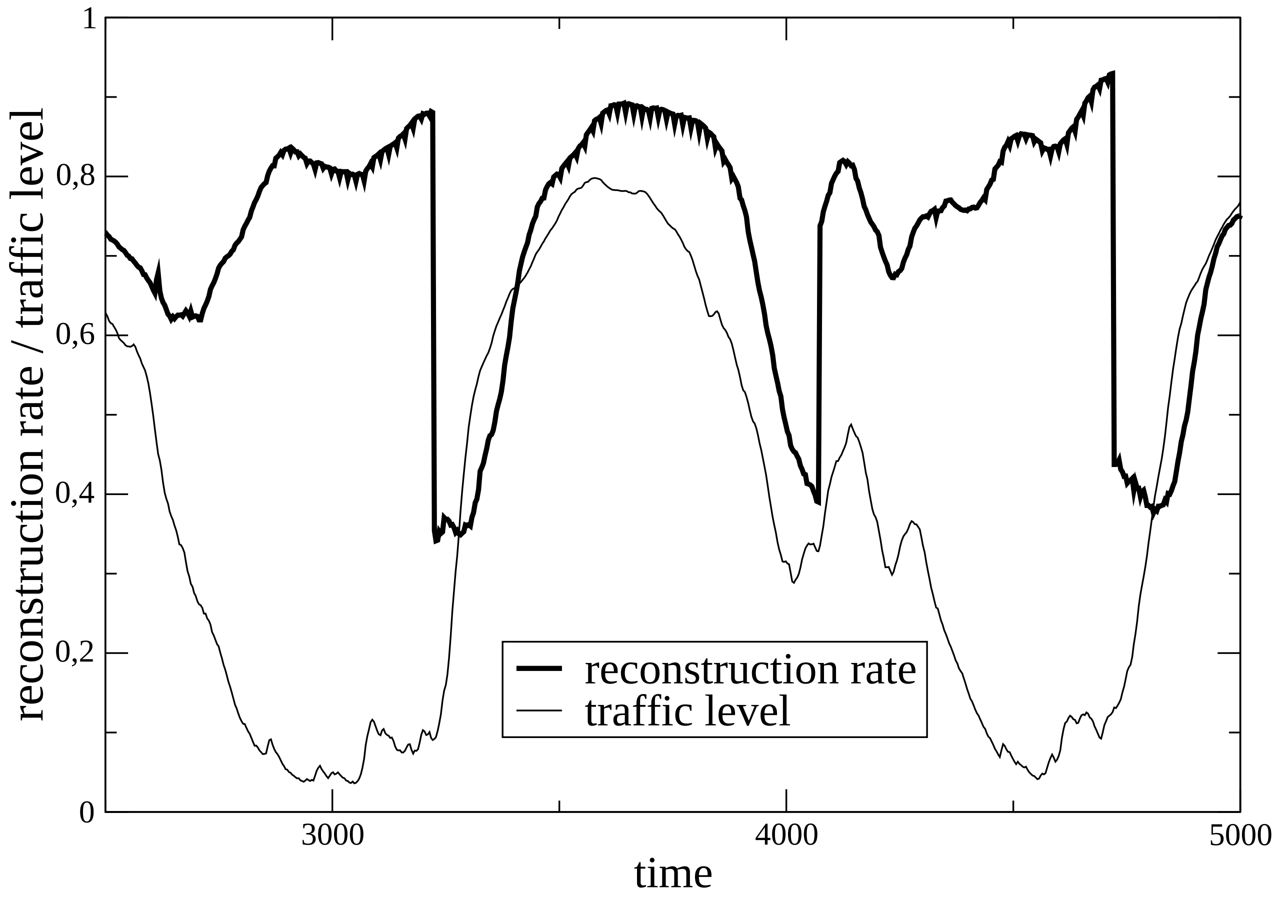
$$n_{i\to a}(x_i) = \prod_{a'\ni i, a'\neq a} m_{b\to i}(x_i)\frac{p_i^*(x_i)}{\hat{p}_i(x_i)} \qquad \text{when } i \in \mathcal{V}^*$$

Numerical experiments

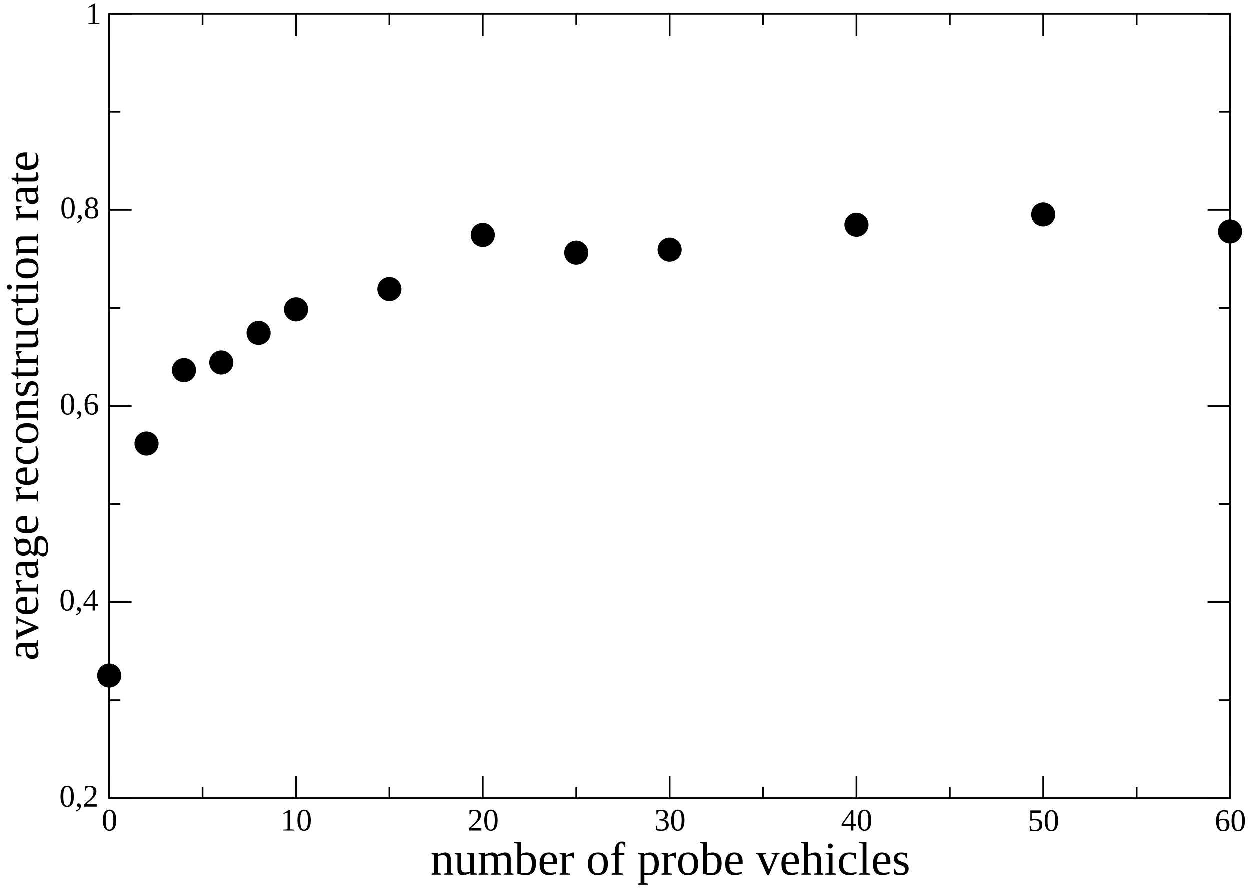
We use the toy network



Reconstruction as a function of time (10 vehicles)



Reconstruction as a function of vehicles



Encoding many patterns

Case study a probabilistic mixture of product form with C components

$$p(\mathbf{x}) \stackrel{\text{def}}{=} \frac{1}{C}\sum_{c=1}^{C}\prod_{i\in\mathcal{V}} p_i^c(x_i),$$

One parameter model α

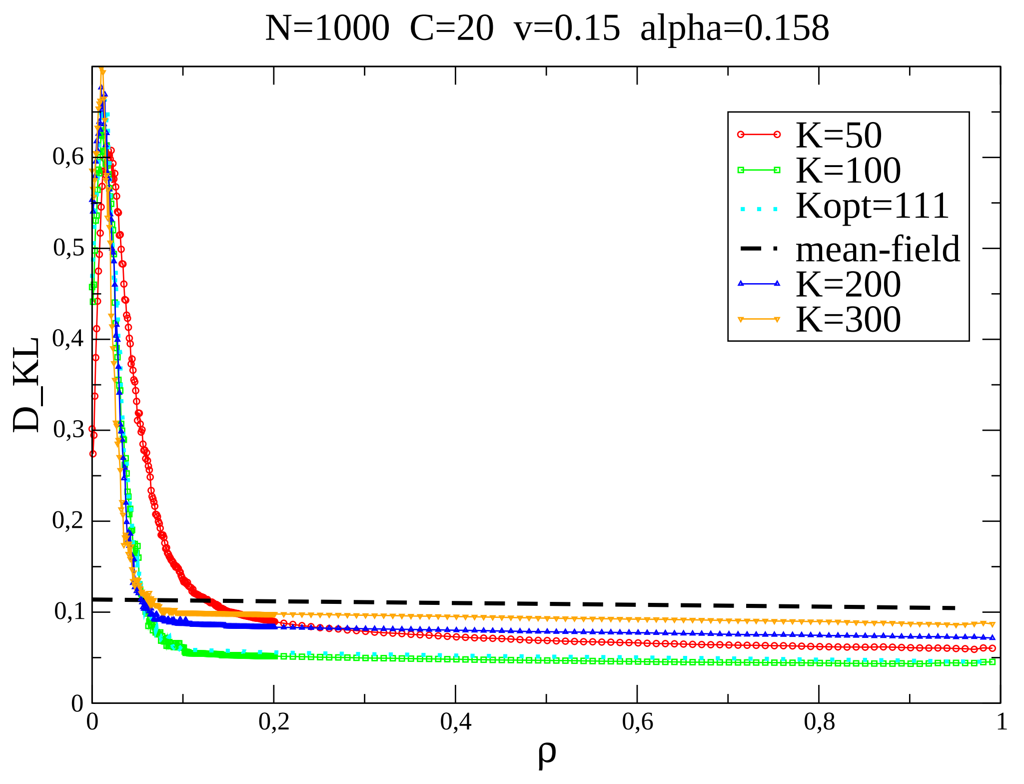
$$\phi_i(x_i) = \hat{p}_i(x_i), \quad \psi_a(\mathbf{x}_a) = \Bigl(\frac{\hat{p}_a(\mathbf{x}_a)}{\prod_{i\in a}\hat{p}_i(x_i)}\Bigr)^{\alpha}$$

where $\hat{p}$ are empirical marginals taken from the hidden mixture P.

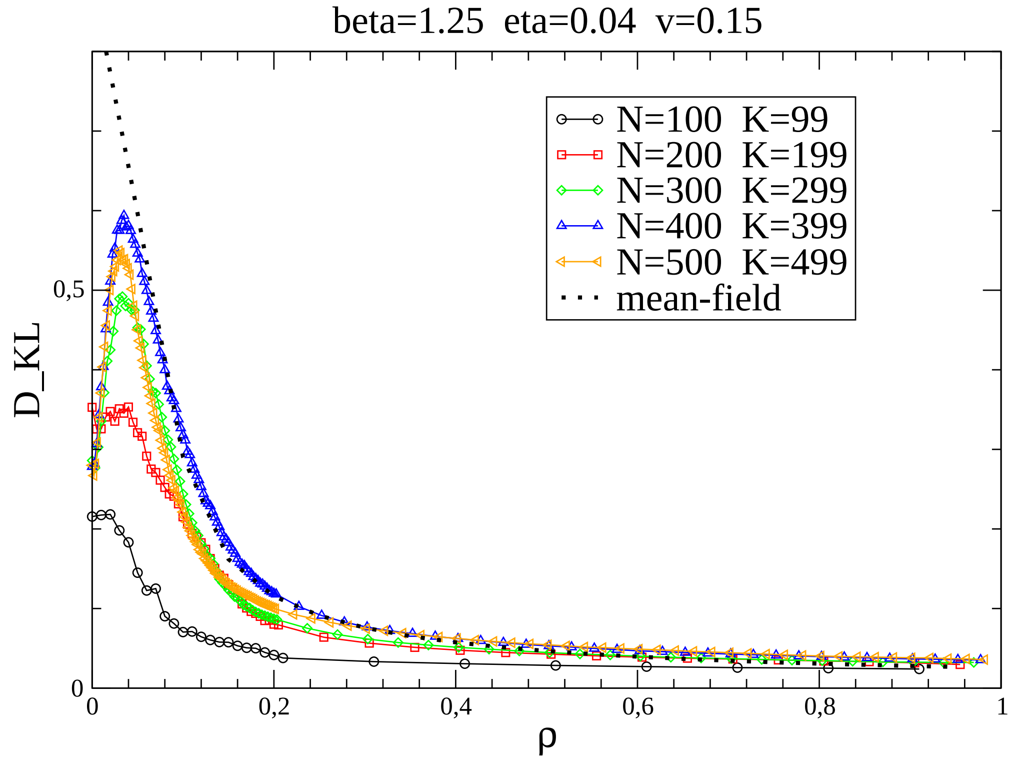
Multi-parameter model

- Sort and cluster the links of the graphs w.r.t. the correlation level
- One parameter α for each class
- Optimization with CMAES

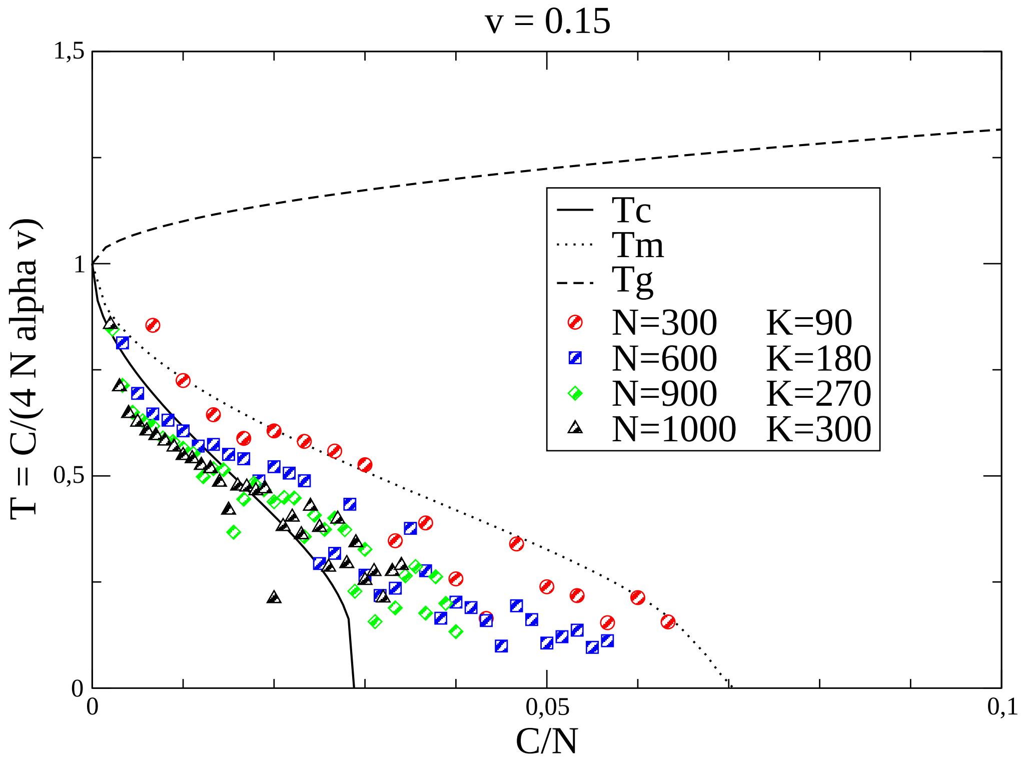
Results : Kullback-Leibler error



Results : more components



Phase Diagram



Open problems

Properties of BP

- is there a form of BP where the stability properties of the fixed points match the minima/maxima of the Bethe free energy ?
- when fitting the model to one Bethe free energy stationary point, is it possible to qualify of the approximation of the other points ?

Improving convergence

- what is the correct way of pruning the graph to improve the convergence of the model ?
- alternatively, how can we choose weights to the Bethe free energy (cf. Fractional BP, tree-reweighted BP) ?