Consensus Maximisation Using Influences of Monotone Boolean Functions

Ruwan Tennakoon[1], David Suter[2], Erchuan Zhang[2], Tat-Jun Chin[3] and Alireza Bab-Hadiashar[1]
[1]RMIT University, Australia. [2]Edith Cowen University, Australia. [3]University of Adelaide, Australia.

CVPR VIRTUAL JUNE 19-25

Abstract

In this paper we:

1) Provide a precise (and abstract) definition of the maximum consensus (MaxCon) problem in terms of finding the maximum upper zero of a Monotone Boolean Function (MBF) defined over the Boolean Cube.
2) Link the concept of influences in a MBF to the concept of outlier in MaxCon and provide theoretical analysis to show that influences of points belonging to the largest structure in data would be smallest under "ideal" conditions.
3) Derive a greedy algorithm that searches for the maximum upper zero of a MBF to efficiently solve the MaxCon problem.

Problem Definition

Maximum Consensus (MaxCon): Find the largest subset of data that fits the model within some tolerance level.

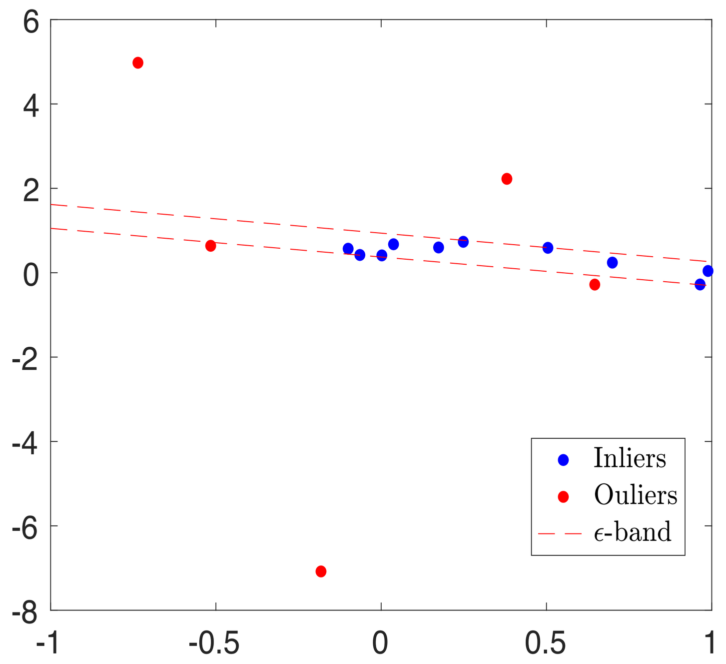


Fig 1: Simple 2D line fitting example with 15 points.

$$\max_{\theta,\ \mathcal{I}\subseteq\mathcal{D}} |\mathcal{I}|$$
$$\text{subject to } \ r_{\mathbf{p}_i}(\theta) \le \epsilon \quad \forall \mathbf{p}_i \in \mathcal{I}$$

$r_{\mathbf{p}_i}(\theta)$: Distance of $\mathbf{p}_i$ from model θ

Feasible subset: All data points belonging to that subset fits its model within a tolerance level (ϵ). We call this "being within an epsilon-band of a model".

Each subset can be represented as a length-n bit vector ($n = 5$):

$$\text{subset } \{p_2, p_4, p_5\} \rightarrow \text{Bit-vector '01011'}$$

Insight 1: Relationship between MaxCon and MBF

Which of the subsets are feasible/infeasible, is an evaluation of a **Boolean function over the n-dimensional Boolean Cube**.

The Boolean function associated with a MaxCon problem is **Monotone:**

- Adding to a subset can only move the function towards infeasibility.
- Deleting points from a subset can move only towards feasible.

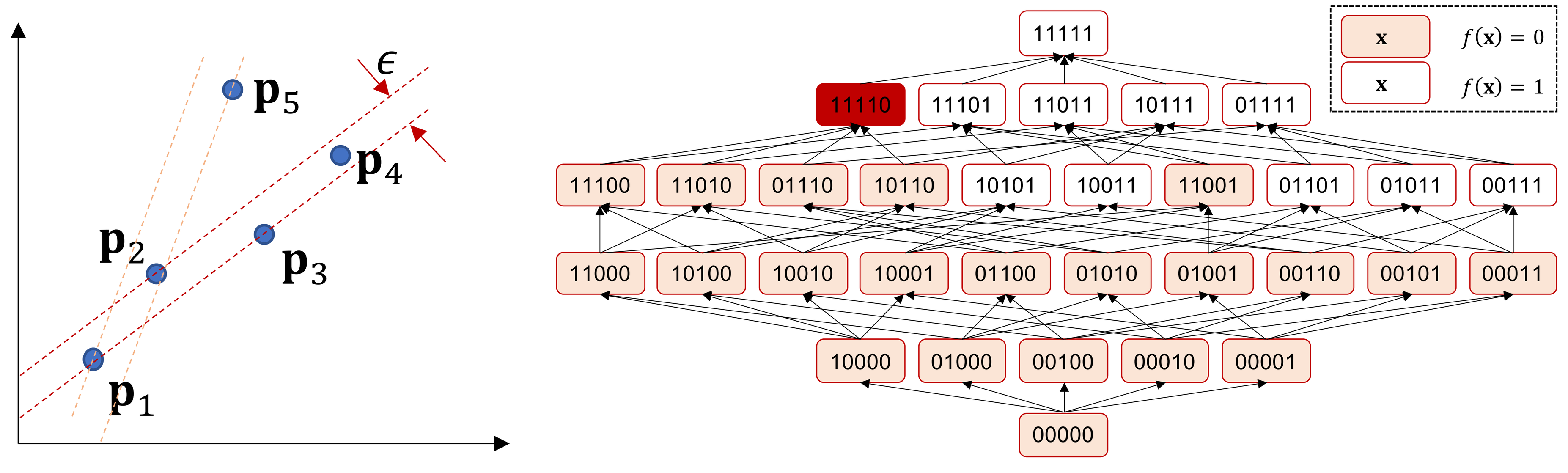


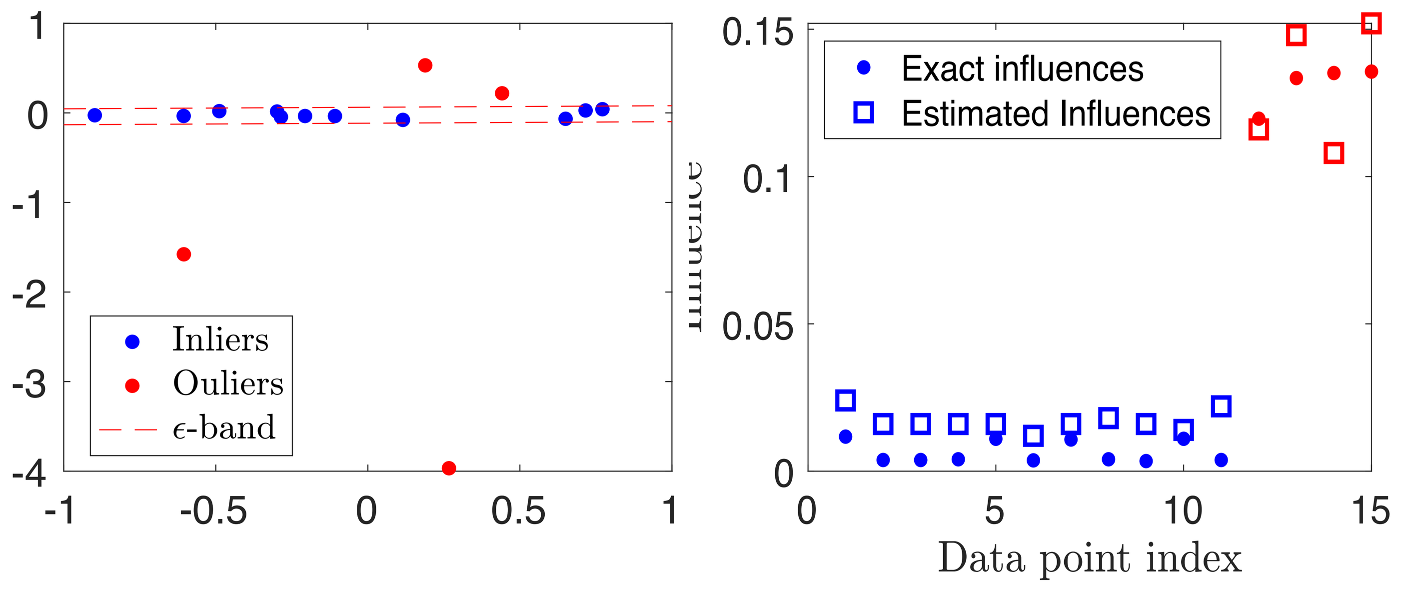
Fig 3: Simple 2D Example with 5 points and the associated Boolean Cube and MBF - represented in Hasse diagram format.

Solving MaxCon is equivalent to finding the maximum upper zero of an MBF

Insight 2: Influence of MBF & Outliers in MaxCon

Influence: The probability that the i^{th} coordinate "affects" the output of the function

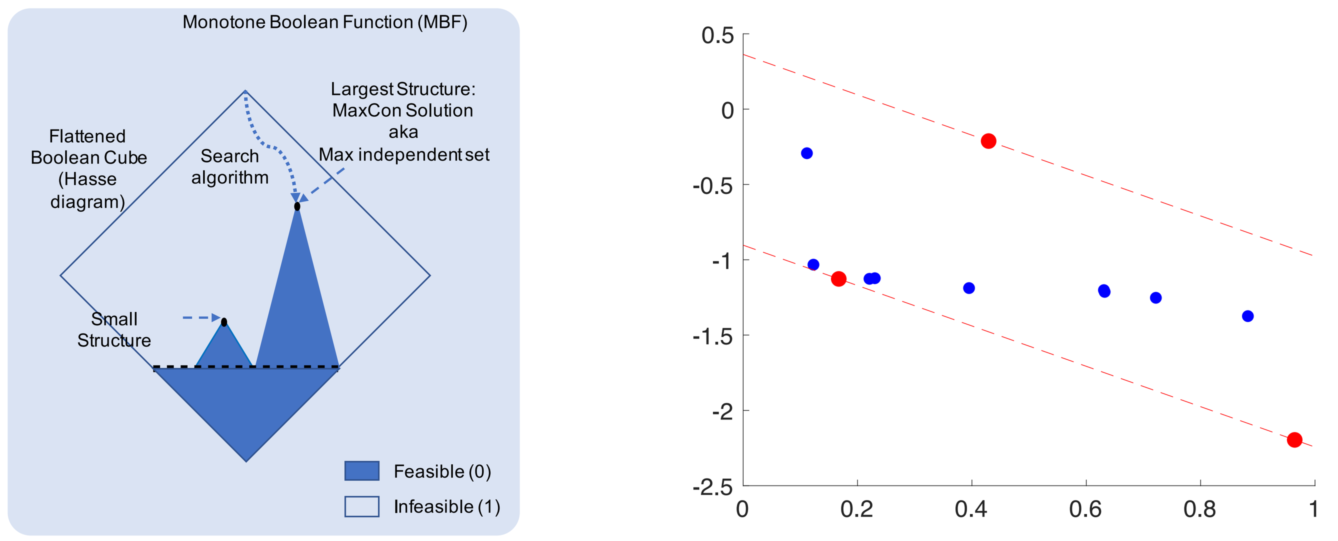
$$\mathrm{Inf}_i[f] = \Pr_{\mathbf{x}\sim\{0,1\}^n}\left[f(\mathbf{x}) \neq f\left(\mathbf{x}^{\oplus i}\right)\right]$$



Influence of an inlier data point is likely to be smaller than the influence of an outlier data point

Theoretical analysis under "ideal" conditions presented in paper.

A greedy algorithm to solve MaxCon (MBF-MaxCon)

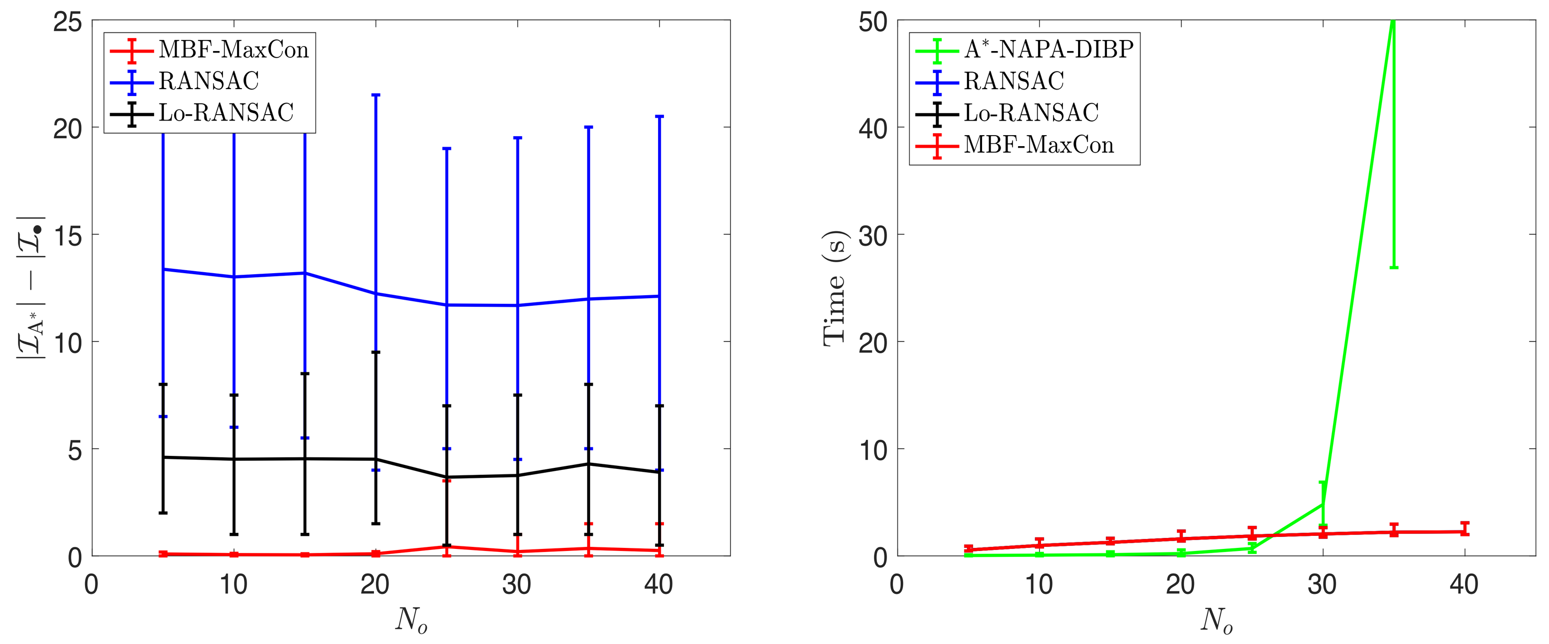


MBF-MaxCon: starts with the set of all data points and then gradually remove one data point at a time (data point with the largest influence) until the remaining set of points is within the ϵ band.

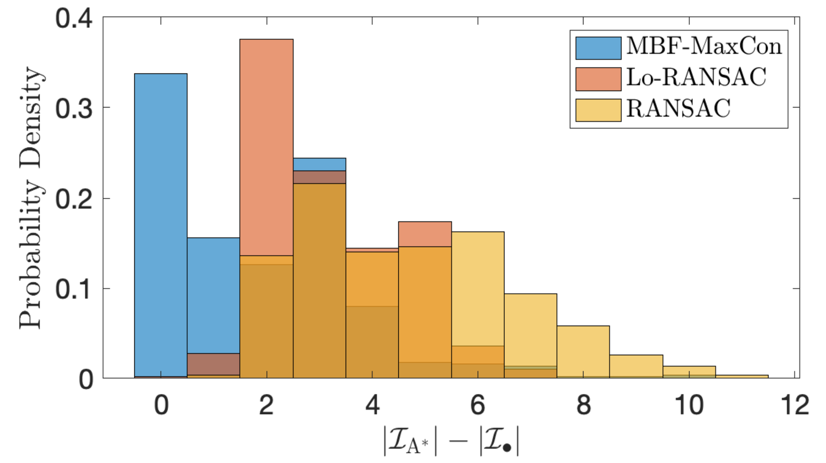
Use of basis to make the algorithm efficient & local update to guarantee convergence to local optimal (upper-zero of MBF).

Results

8D robust Linier regression with synthetic data:



Linearized fundamental matrix estimation:



Results on KITTI Odometry dataset: we generally get much closer to A* performance - including often finding the optimal, which RANSAC or Lo-RANSAC rarely do in this experiment.