Supplementary Materials
Maximum Consensus by Weighted Influences of Monotone Boolean Functions

Erchuan Zhang* David Suter*,‡ Ruwan Tennakoon** Tat-Jun Chin†
Alireza Bab-Hadiashar** Giang Truong* Syed Zulqarnain Gilani*,‡
*Centre for AI & ML, Edith Cowan University, Perth Australia.
‡ Nutrition and Health Innovation Research Institute, Edith Cowan University, Perth Australia.
**RMIT University, Melbourne Australia. †University of Adelaide, Adelaide Australia.
{erchuan.zhang, d.suter, h.truong, s.gilani}@ecu.edu.au, {ruwan.tennakoon,abh}@rmit.edu.au
tat-jun.chin@adelaide.edu.au

This supplementary material provides (section 1) proofs of the theoretical results presented in the main paper, some observations of the implications of the expressions for uniform Hamming weighted influences (section 2), algorithmic details of a scheme for employing Bernoulli weighted influences to approximately solve MaxCon (section 3), and some further experimental results (sections 4, 5 6).

1. Proofs of theoretical results

Theorem 2.1. *If $f : \{0,1\}^n \to \{0,1\}$ is a monotone Boolean function, then* $\mathrm{Inf}_i^q[f] = -\frac{1}{\sqrt{q(1-q)}}\hat{f}^q(\{i\})$.

Proof. The *i-th derivative operator* D_i maps a Boolean function f to a function $D_i f$ defined by [1]

$$D_i f(\boldsymbol{b}) := f(\boldsymbol{b}^{i\mapsto 1}) - f(\boldsymbol{b}^{i\mapsto 0}), \tag{1}$$

where $\boldsymbol{b}^{i\mapsto a} = (b_1, \cdots, b_{i-1}, a, b_{i+1}, \cdots, b_n)$. Since $f : \{0,1\}^n \to \{0,1\}$ is Boolean-valued, we have

$$D_i f(\boldsymbol{b}) = \begin{cases} \pm 1, & \text{if } f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i}), \\ 0, & \text{if } f(\boldsymbol{b}) = f(\boldsymbol{b}^{\oplus i}), \end{cases} \tag{2}$$

which means

$$\mathrm{Inf}_i^q[f] = \langle D_i f, D_i f\rangle_q. \tag{3}$$

By definition,

$$D_i \chi_S^q(\boldsymbol{b}) = \begin{cases} -\frac{1}{\sqrt{q(1-q)}}\chi_{S\setminus\{i\}}^q(\boldsymbol{b}), & \text{if } i \in S, \\ 0, & \text{if } i \notin S. \end{cases}$$

Then

$$\begin{aligned} D_i f &= D_i\left(\sum_{S\subseteq[n]} \hat{f}^q(S)\chi_S^q\right) \\ &= -\frac{1}{\sqrt{q(1-q)}}\sum_{S: i\in S}\hat{f}^q(S)\chi_{S\setminus\{i\}}^q. \end{aligned}$$

If f is monotone,

$$\mathrm{Inf}_i^q[f] = \langle D_i f, \chi_\emptyset^q\rangle = -\frac{1}{\sqrt{q(1-q)}}\hat{f}^q(\{i\}).$$

□

Theorem 3.1 (Ideal Single Structure)**.** *If $n_0 = 1$, namely, f is ideal with respect to a single maximum upper zero $\boldsymbol{b}^{k_1}$, then, for $i \in S^1_{\boldsymbol{b}^{k_1}}$ (inliers),*

$$\mathrm{Inf}_i^q[f] = (C_p^{n-1} - C_p^{k_1-1})q^p(1-q)^{n-p-1},$$

and for $i \in S^0_{\boldsymbol{b}^{k_1}}$ (outliers),

$$\mathrm{Inf}_i^q[f] = C_p^{n-1}q^p(1-q)^{n-p-1} + \sum_{l=p+1}^{k_1} C_l^{k_1} q^l(1-q)^{n-l-1},$$

which implies $\mathrm{Inf}_{i_1}^q[f] - \mathrm{Inf}_{i_2}^q[f] = C_p^{k_1-1}q^p(1-q)^{n-p-1} + \sum_{l=p+1}^{k_1} C_l^{k_1}q^l(1-q)^{n-l-1} > 0$*, where $i_1 \in S^0_{\boldsymbol{b}^{k_1}}$, $i_2 \in S^1_{\boldsymbol{b}^{k_1}}$.*

Proof. For any $i \in S^1_{\boldsymbol{b}^{k_1}}$, if $\boldsymbol{b} \in L_{\leqslant p-1}$ or $L_{\geqslant p+2}$, then $f(\boldsymbol{b}) \equiv f(\boldsymbol{b}^{\oplus i})$; if $\boldsymbol{b} \in L_p$, then $f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i})$ holds for $(C_p^{n-1} - C_p^{k_1-1})$-possible vertices $\boldsymbol{b}$; if $\boldsymbol{b} \in L_{p+1}$, then $f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i})$ holds also for $(C_p^{n-1} - C_p^{k_1-1})$-possible vertices $\boldsymbol{b}$. Therefore, we have

$$\begin{aligned} \mathrm{Inf}_i^q[f] &= (C_p^{n-1} - C_p^{k_1-1})q^p(1-q)^{n-p} \\ &+ (C_p^{n-1} - C_p^{k_1-1})q^{p+1}(1-q)^{n-p-1} \\ &= (C_p^{n-1} - C_p^{k_1-1})q^p(1-q)^{n-p-1}. \end{aligned}$$

For any $i \in S^0_{\boldsymbol{b}^{k_1}}$, if $\boldsymbol{b} \in L_{\leqslant p-1}$ or $L_{\geqslant k+2}$, then $f(\boldsymbol{b}) \equiv f(\boldsymbol{b}^{\oplus i})$; if $\boldsymbol{b} \in L_p$, then $f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i})$ holds for C_p^{n-1}-possible vertices $\boldsymbol{b}$; if $\boldsymbol{b} \in L_{p+1}$, then $f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i})$ holds for $(C_p^{n-1} + C_{p+1}^{k_1})$-possible vertices $\boldsymbol{b}$; if $\boldsymbol{b} \in L_l$ $(p+2 \leqslant l \leqslant k_1)$, then $f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i})$ holds for $(C_{l-1}^{k_1} + C_l^{k_1})$-possible vertices $\boldsymbol{b}$; if $\boldsymbol{b} \in L_{k_1+1}$, then $f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i})$ holds for $C_{k_1}^{k_1}$-possible vertices $\boldsymbol{b}$. Therefore, we have

$$\begin{aligned}\mathrm{Inf}_i^q[f] &= C_p^{n-1} q^p (1-q)^{n-p} + (C_p^{n-1} + C_{p+1}^{k_1}) q^{p+1} \times \\ &(1-q)^{n-p-1} + \sum_{l=p+2}^{k_1} (C_{l-1}^{k_1} + C_l^{k_1}) q^l (1-q)^{n-l} \\ &+ C_{k_1}^{k_1} q^{k_1+1} (1-q)^{n-k_1-1} \\ &= C_p^{n-1} q^p (1-q)^{n-p-1} + \sum_{l=p+1}^{k_1} C_l^{k_1} q^l (1-q)^{n-l-1}.\end{aligned}$$

□

Theorem 3.2 (Ideal Multi-Structure)**.** *Suppose $n_0 > 1$, i.e., f is ideal with several upper zeros, then, $\forall\, i \in \cap_{l=1}^{n_0} S^{i_l}_{\boldsymbol{b}^{k_i}}$ (if non-empty),*

$$\begin{aligned}\mathrm{Inf}_i^q[f] \;=\; &(C_p^{n-1} - \sum_{i_s=1} C_p^{k_s-1}) q^p (1-q)^{n-p-1} \\ &+ \sum_{i_s=0} \sum_{l=p+1}^{k_s} C_l^{k_s} q^l (1-q)^{n-l-1}.\end{aligned} \tag{4}$$

Proof. We prove this theorem by induction on n_0. By Theorem 3.1, (4) is true for $n_0 = 1$. Suppose (4) holds for $n_0 - 1$, namely, $\forall\, \hat{i} \in \cap_{l=1}^{n_0-1} S^{i_l}_{\boldsymbol{b}^{k_i}}$,

$$\begin{aligned}\mathrm{Inf}_{\hat{i}}^q[f] = &C_p^{n-1} q^p (1-q)^{n-p-1} \\ &+ \sum_{\substack{i_s=0 \\ 1 \leqslant s \leqslant n_0-1}} \sum_{l=p+1}^{k_s} C_l^{k_s} q^l (1-q)^{n-l-1} \\ &- \sum_{\substack{i_s=1 \\ 1 \leqslant s \leqslant n_0-1}} C_p^{k_s-1} q^p (1-q)^{n-p-1}.\end{aligned}$$

Now we only have to prove that, $\forall\, i \in \cap_{l=1}^{n_0} S^{i_l}_{\boldsymbol{b}^{k_i}}$,

$$\begin{aligned}\mathrm{Inf}_i^q[f] = &\mathrm{Inf}_{\hat{i}}^q[f] \\ &+ \begin{cases} -C_p^{k_{n_0}-1} q^p (1-q)^{n-p-1}, & i_{n_0} = 1, \\ \sum_{l=p+1}^{k_{n_0}} C_l^{k_{n_0}} q^l (1-q)^{n-l-1}, & i_{n_0} = 0. \end{cases}\end{aligned}$$

When adding one more upper zero $\boldsymbol{b}^{k_{n_0}}$, for any $i \in \cap_{l=1}^{n_0-1} S^{i_l}_{\boldsymbol{b}^{k_i}} \cap S^1_{\boldsymbol{b}^{k_{n_0}}}$, i-boundary edges will decrease by $C_p^{k_{n_0}-1}$ at level p and $p+1$. That is, if $\boldsymbol{b} \in L_p$ or L_{p+1}, then $f(\boldsymbol{b}) = f(\boldsymbol{b}^{\oplus i})$ holds for $C_p^{k_{n_0}-1}$ possible vertices $\boldsymbol{b}$. Then, the decrease amount for $\mathrm{Inf}_{\hat{i}}^q[f]$ is

$$\begin{aligned}&C_p^{k_{n_0}-1} q^p (1-q)^{n-p} + C_p^{k_{n_0}-1} q^{p+1} (1-q)^{n-p-1} \\ &= C_p^{k_{n_0}-1} q^p (1-q)^{n-p-1}.\end{aligned}$$

For any $i \in \cap_{l=1}^{n_0-1} S^{i_l}_{\boldsymbol{b}^{k_i}} \cap S^0_{\boldsymbol{b}^{k_{n_0}}}$, i-boundary edges will increase. In details, if $\boldsymbol{b} \in L_{k_{n_0}+1}$, then $f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i})$ holds for $C_{k_{n_0}}^{k_{n_0}} = 1$ possible vertex $\boldsymbol{b}$, if $\boldsymbol{b} \in L_l$ $(p+2 \leqslant l \leqslant k_{n_0})$, then $f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i})$ holds for $(C_l^{k_{n_0}} + C_{l-1}^{k_{n_0}})$-possible vertices $\boldsymbol{b}$, if $\boldsymbol{b} \in L_{p+1}$, the possible vertices $\boldsymbol{b}$ have $C_{p+1}^{k_{n_0}}$. Then, the increase amount for $\mathrm{Inf}_{\hat{i}}^q[f]$ is

$$\begin{aligned}&C_{k_{n_0}}^{k_{n_0}} q^{k_{n_0}+1} (1-q)^{n-k_{n_0}-1} \\ &+ \sum_{l=p+2}^{k_{n_0}} (C_l^{k_{n_0}} + C_{l-1}^{k_{n_0}}) q^l (1-q)^{n-l} \\ &+ C_{p+1}^{k_{n_0}} q^{p+1} (1-q)^{n-p-2} \\ &= \sum_{l=p+1}^{k_{n_0}} C_l^{k_{n_0}} q^l (1-q)^{n-l-1},\end{aligned}$$

which complete this proof. □

Theorem 3.3. *If f is non-ideal, then the weighted influence $\tilde{f}^q(S^{\boldsymbol{i}})$ is given by*

$$\begin{aligned}&(C_p^{n-1} - \sum_{\substack{i_s=1 \\ 1 \leqslant s \leqslant n_0}} C_p^{k_s-1} + \sum_{\substack{i_{n_0+s}=0 \\ 1 \leqslant s \leqslant m_0}} C_p^{\alpha_s-1}) q^p (1-q)^{n-p-1} \\ &+ (\sum_{\substack{i_s=0 \\ 1 \leqslant s \leqslant n_0}} \sum_{l=p+1}^{k_s} C_l^{k_s} - \sum_{\substack{i_{n_0+s}=1 \\ 1 \leqslant s \leqslant m_0}} \sum_{l=p+1}^{\alpha_s} C_l^{\alpha_s}) q^l (1-q)^{n-l-1}.\end{aligned} \tag{5}$$

where $\tilde{f}^q(S^\bullet)$ doesn't exist if $S^\bullet = \emptyset$.

To better understand Theorem 3.3, let us consider the simplest non-ideal case where $n_0 = 2$ and $m_0 = 1$.

Theorem 3.4. *The existing influences are represented as*

$$\begin{aligned}\tilde{f}^q(S^{(110)}) &= \left(C_p^{n-1} - C_p^{k_1-1} - C_p^{k_2-1} + C_p^{\alpha_1-1}\right) \times \\ &q^p (1-q)^{n-p-1}, \\ \tilde{f}^q(S^{(101)}) &= \left(C_p^{n-1} - C_p^{k_1-1}\right) q^p (1-q)^{n-p-1} \\ &+ \left(\sum_{l=p+1}^{k_2} C_l^{k_2} - \sum_{l=p+1}^{\alpha_1} C_l^{\alpha_1}\right) q^l (1-q)^{n-l-1}, \\ \tilde{f}^q(S^{(011)}) &= \left(C_p^{n-1} - C_p^{k_2-1}\right) q^p (1-q)^{n-p-1} \\ &+ \left(\sum_{l=p+1}^{k_1} C_l^{k_1} - \sum_{l=p+1}^{\alpha_1} C_l^{\alpha_1}\right) q^l (1-q)^{n-l-1}, \\ \tilde{f}^q(S^{(001)}) &= C_p^{n-1} q^p (1-q)^{n-p-1} \\ &+ \left(\sum_{l=p+1}^{k_1} C_l^{k_1} + \sum_{l=p+1}^{k_1} C_l^{k_2} - \sum_{l=p+1}^{\alpha_1} C_l^{\alpha_1}\right) \times \\ &q^l (1-q)^{n-l-1},\end{aligned}$$

which implies

$$\tilde{f}^q(S^{(101)}) < \tilde{f}^q(S^{(001)}), \;\; \tilde{f}^q(S^{(011)}) < \tilde{f}^q(S^{(001)}).$$

Proof. By Theorem 3.2, we only have to prove

$$\tilde{f}^q(S^{(\bullet\bullet i_3)}) = \tilde{f}^q(S^{(\bullet\bullet)}) + \begin{cases} C_p^{\alpha_1-1} q^p (1-q)^{n-p-1}, & i_3 = 0, \\ -\sum_{l=p+1}^{\alpha_1} C_l^{\alpha_1} q^l (1-q)^{n-l-1}, & i_3 = 1. \end{cases}$$

$\forall\, i \in S^{(\bullet\bullet 0)}$, the vertices $\boldsymbol{b} \in L_p$ or L_{p+1} for $f(\boldsymbol{b}) \neq f(\boldsymbol{b}^{\oplus i})$ have increased $C_p^{\alpha_1-1}$ because of the ovelap sub-cube $B_{\boldsymbol{b}^{\alpha_1}}$, then

$$\begin{aligned} \tilde{f}^q(S^{(\bullet\bullet 0)}) =& \tilde{f}^q(S_{\bullet\bullet}) + C_p^{\alpha_1-1} q^p (1-q)^{n-p} \\ &+ C_p^{\alpha_1-1} q^{p+1} (1-q)^{n-p-1} \\ =& \tilde{f}^q(S^{(\bullet\bullet)}) + C_p^{\alpha_1-1} q^p (1-q)^{n-p-1}. \end{aligned}$$

$\forall\, i \in S^{(\bullet\bullet 1)}$, if $\boldsymbol{b} \in L_{\alpha_1+1}$, then i-boundary edges decrease by $C_{\alpha_1}^{\alpha_1}$, if $\boldsymbol{b} \in L_l$ $(p+2 \leqslant l \leqslant \alpha_1)$, then i-boundary edges decrease by $C_l^{\alpha_1} + C_{l-1}^{\alpha_1}$, if $\boldsymbol{b} \in L_{p+1}$, the decrease amount is $C_{p+1}^{\alpha_1}$. Therefore,

$$\begin{aligned} \tilde{f}^q(S^{(\bullet\bullet 1)}) =& \tilde{f}^q(S^{(\bullet\bullet)}) - (C_{\alpha_1}^{\alpha_1} q^{\alpha_1+1} (1-q)^{n-\alpha_1-1} \\ &+ \sum_{l=p+2}^{\alpha_1} (C_l^{\alpha_1} + C_{l-1}^{\alpha_1}) q^l (1-q)^{n-l} \\ &+ C_{p+1}^{\alpha_1} q^{p+1} (1-q)^{n-p-1}) \\ =& \tilde{f}^q(S^{(\bullet\bullet)}) - \sum_{l=p+1}^{\alpha_1} C_l^{\alpha_1} q^l (1-q)^{n-l-1}, \end{aligned}$$

which complete the proof. □

By Theorem 3.2 and Theorem 3.4, Theorem 3.3 can be proved by induction on m_0.

2. Uniform Hamming measure Influence

Consider the ideal single structure case define in Theorem 3.1. Firstly, it is easy to see the for the ideal single structure, any algorithm that starts with a feasible set of size large than the combinatorial dimension, and then greedily adds points if the larger set remains feasible, will obtain the MaxCon solution. So it is an easy problem with obvious solution strategies. But if we did decide to use influences we can note that from equation (3.1) the inlier influences came from counting feasible/infeasible transitions between levels p and $p+1$ only. Thus for sampling uniform Hamming level above level $p+1$, there are no feasibility/infeasibility transitions caused by inliers. In other words, the Uniform Hamming influence measure, at that level or above, will be exactly zero for inliers. Since the influences of outliers, for the same measure will be non-zero, this seems to promise a remarkably efficient sampling strategy - one could eliminate outliers at the first sample that revealed a count for the associated influence - without the need to continue with the full estimation process. Of course, practically, this is too good to hold for real data - it is very brittle to our strict assumptions here. Nonetheless, it does hint at the usefulness of a less brittle strategy of early termination of the counting process once a count reaches some degree of statistically significantly higher than the rest, rather than a set number of samples always being used: a future research topic.

Now consider the ideal multiple structure setting. From equation (4), the influence accrued by being inlier to some structure (first term) is only accrued between level p and $p+1$. The subtraction is due to feasibility transitions “that didn’t occur” because the subset with added inlier remains within the same structure). So once again, sampling above level p will not “see” those counts. But since inliers to one structure are outliers to another (we assumed no significant overlap): hence the influence of inliers to any structure will not be zero - different to the single structure case, as all structures are outliers to (all) other structures and thus accrue influence from the second term in the equation. It is also easy to see that so long as the level is “not too high” (above the largest structure) the Hamming sampled influences will be an appropriate guide (influence of inliers of a larger structure will be smaller). (In that second term the largest structure is excluded from the sum over $i_s = 0$, when calculating influence for that structure.)

In more detail, since a point is a member of at most one structure (we forbid overlaps in the definition of “ideal”), we observe that an inlier to any structure is associated with only one term in the subtraction and

Analysis of the non-ideal case is complicated (hugely) by the complex combinatorics of possible overlaps. Nonetheless, for structures with little overlap with any other (we would argue the majority of structures in situations of interest) the “perurbation” from the ideal case calculations will be minimal. For situations with very large overlap in structures, one could alternatively view these as minor variants of one and the same structure (simply including a few extra points and losing one or two) and thus - with respect to the overlaps involving the largest structure, these could be considered as minor sub-optimal variants and essentially recovering one of the slightly smaller variants, compared with the actual optimal, is something of likely minor practical consequence. Of course, we realise that such observations are far short of conclusive argument and we make no claims of otherwise.

3. Algorithms based on Bernoulli weighted influences

Algorithm 1 is essentially similar to that presented in [2], where $p+1$ is the combinatorial dimension of the prescribed model, the function f is evaluated as

$$f(\mathcal{I}) := \mathbb{I}(\min_{\boldsymbol{\theta}} \max_{\boldsymbol{x}_i \in \mathcal{I}} r(\boldsymbol{x}_i, \boldsymbol{\theta}) \leqslant \varepsilon) \tag{6}$$

with $\mathbb{I}(\cdot)$ the indicator function. The key difference is how we evaluate the estimated weighted influences $\widetilde{\mathrm{Inf}}_i^q[f]$.

Algorithm 1: Consensus maximisation using weighted influences (WI)

Input: Dataset $\mathcal{X} = \{\boldsymbol{x}_i\}_{i=1}^n$, probability $q \in (0,1)$, sample size h, threshold $\varepsilon > 0$.

Output: Inliers set $\mathcal{I}^*$

Initialization: $\mathcal{I} \leftarrow \mathbf{1}_{1\times n}$.

while $|\mathcal{I}| > p$ **do**

  Solve the minmax problem

$$\min_{\boldsymbol{\theta}} \max_{\boldsymbol{x}_i \in \mathcal{I}} r(\boldsymbol{x}_i, \boldsymbol{\theta}),$$

  to get a basis $\mathcal{B}$.

  Evaluate the estimated weighted influences $\widetilde{\mathrm{Inf}}_i^q[f]$ for $i \in \mathcal{B}$ by

$$\widetilde{\mathrm{Inf}}_i^q[f] = -\frac{1}{h\sqrt{q(1-q)}} \sum_{j=1}^{h} f(\boldsymbol{b}_j) q_-^{b_{j,i}} q_+^{1-b_{j,i}} \mu_q(\boldsymbol{b}_j).$$

  $\mathcal{I} \leftarrow \mathcal{I} \setminus \arg\max_i \{\widetilde{\mathrm{Inf}}_i^q[f] \mid i \in \mathcal{B}\}$.

  if $f(\mathcal{I}) = 0$ **then**

    $\mathcal{I}^* \leftarrow \mathcal{I}$.

    Break.

  end if

  Conduct Algorithm 2 for local expansion to add possible missing inliers.

end while

return $\mathcal{I}^*$

Algorithm 2: Local expansion step

Input: Dataset $\mathcal{X} = \{\boldsymbol{x}_i\}_{i=1}^n$, threshold $\varepsilon > 0$, initial solution $\mathcal{I}$.

Output: Inliers set $\mathcal{I}$

Candidates $\leftarrow \mathcal{X} \setminus \mathcal{I}$.

for i *in* Candidates **do**

  $\mathcal{I} \leftarrow \mathcal{I} \cup \{i\}$.

  if $f(\mathcal{I}) = 1$ **then**

    $\mathcal{I} \leftarrow \mathcal{I} \setminus \{i\}$.

  end if

end for

return $\mathcal{I}$

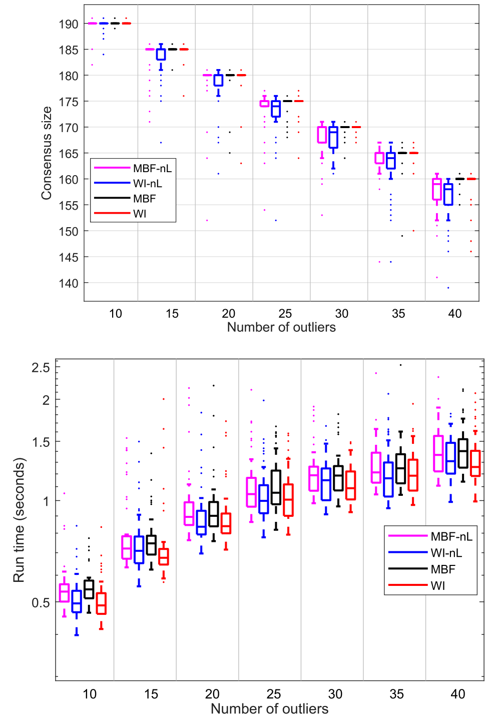

Figure 1. Comparison results of 8 dimensional linear regression with synthetic data: (Top) consensus size and (Bottom) runtime. All experiments were repeated 50 times.

4. Comparison of the effect of local expansion in MBF and WI

In this section, we will compare the effect of local expansion in MBF and WI using the example of 8-dimensional linear regression. The experiment setting is the same as Subsection 4.1 in the main paper. We denote MBF and WI without local expansion by MBF-nL and WI-nL, respectively.

From Figure 1, we find that the number of inliers returned by the proposed method without local expansion WI-nL is less than that of MBF-nL, however, with the help of local expansion, both WI and MBF can find the same number of inliers. More importantly, our method WI (WI-nL) is generally faster than MBF (MBF-nL), especially in the presence of higher number of outliers.

5. Further results on linearised fundamental matrix estimation

This section further examines the performance of our proposed method on linearised fundamental matrix estimation on the KITTI dataset that is used in the main paper. In this experiment, we choose the confidence $\rho = 0.99$ for the standard stopping criteria in both RANSAC and Lo-

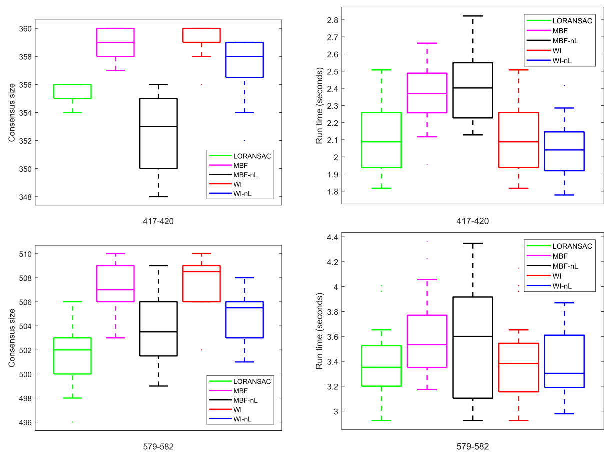

Figure 2. Distributions of consensus size (left column) and run time (right column) of linearised fundamental matrix estimation on KITTI image pairs $417-420$ (top row) and $579-582$ (bottom row).

RANSAC. The average consensus size and runtime including their variance over 20 repeated runs are shown in Table 1.

From Table 1, we can find that: (1) With respect to the standard stopping criteria, RANSAC-p and Lo-RANSAC-p are much faster than the proposed method, however, they sacrifice the consensus size a lot, especially on the image pair 738-742; (2) With the help of local expansion, both MBF and WI improve the returned consensus size from MBF-nL and WI-nL with a small amount of extra time budget. Moreover, without local expansion, WI-nL is slightly better than MBF-nL in terms of returned consensus size and runtime, on average.

To further compare the performance of Lo-RANSAC, MBF/MBF-nL and WI/WI-nL, we plot the distributions of consensus size and runtime on the image pair $417-420$ and $579-582$ in Figure 2. From which, we can see that although WI and MBF can get similar average consensus size, WI has a higher probability to achieve better results with less time budget. Obviously, the more iterations of RANSAC and Lo-RANSAC use, the higher consensus size they return. However, our method as well as MBF can increase results by sampling more vertices in the Boolean cube to get more accurate (weighted) influences.

A fair and safe conclusion is that on some datasets, WI and MBF (including their variants) perform better than RANSAC and Lo-RANSAC with some prescribed time (generally longer than the rule of thumb prescriptions for termination of those algorithms: thus when one is prepared to spend extra computation for better results, WI and MBF may be alternatives worth considering). More importantly, WI is able to achieve similar consensus size with less time budget, which means WI is an effective alternative of MBF.

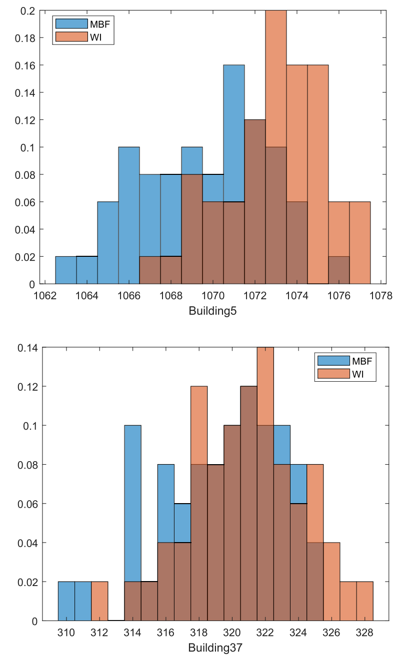

Figure 3. Distributions of consensus size returned by MBF and WI.

6. Further results on linearised homography estimation

In this section, we compare the distributions of consensus size returned by MBF and WI, which is shown in Figure 3. It can be seen that WI can achieve better results with high probability.

References

[1] Ryan O'Donnell. *Analysis of boolean functions*. Cambridge university press, 2014. 1

[2] Ruwan Tennakoon, David Suter, Erchuan Zhang, Tat-Jun Chin, and Alireza Bab-Hadiashar. Consensus maximisation using influences of monotone boolean functions. In *CVPR*, pages 2866–2875, 2021. 3

Table 1. Results for linearised fundamental matrix estimation. RANSAC-p and Lo-RANSAC-p refer to RANSAC and Lo-RANSAC with standard stopping criteria of the confidence $\rho = 0.99$, respectively. MBF-nL and WI-nL refer to implementing MBF and WI without local expansion steps, respectively. All experiments were repeated over 20 random runs.

		104-108	**198-201**	**417-420**	**579-582**	**738-742**
RANSAC-p	Consensus	252.05 (238,266)	276.00 (267,282)	341.20 (317,351)	474.20 (453,496)	411.60 (401,425)
	Time (s)	0.01 (0.01,0.01)	0.01 (0.01,0.01)	0.01 (0.01,0.01)	0.01 (0.01,0.01)	0.01 (0.01,0.01)
Lo-RANSAC-p	Consensus	264.15 (255,269)	281.75 (279,285)	354.05 (352,356)	492.25 (480,500)	423.15 (413,435)
	Time (s)	0.04 (0.01,0.07)	0.04 (0.01,0.07)	0.08 (0.03,0.13)	0.14 (0.08,0.27)	0.13 (0.05,0.26)
MBF-nL	Consensus	261.80 (253,268)	285.60 (281,288)	352.75 (348,356)	503.75 (499,509)	441.40 (434,445)
	Time (s)	3.03 (2.56,3.55)	1.84 (1.65,2.29)	2.41 (2.13,2.82)	3.57 (2.93,4.35)	3.25 (2.82,4.10)
WI-nL	Consensus	269.50 (266,272)	287.40 (282,289)	357.40 (352,359)	504.75 (501,508)	441.90 (439,444)
	Time (s)	2.53 (2.27,2.90)	1.67 (1.56,2.04)	2.07 (1.78,2.69)	3.40 (2.98,4.05)	3.02 (2.79,3.41)