A Flexible Tensor Block Coordinate Ascent Scheme for Hypergraph Matching

(Supplementary Material)

Quynh Nguyen Ngoc[1,2], Antoine Gautier[2], Matthias Hein[2]

[1]Max Planck Institute for Informatics, Saarbrücken, Germany
[2]Saarland University, Saarbrücken, Germany

Abstract

This supplementary material contains the proofs of all theorems and some additional experimental results. We keep the numbering of all the theorems as in the paper and repeat the algorithms for the convenience of the reader.

Let $\mathcal{F}^m$ be an m-th order tensor, we recall from the paper the notations (see Definition 3.2)

$$F^m(\mathbf{x}^1,\ldots,\mathbf{x}^m)=\sum_{i_1,\ldots,i_m=1}^{n}\mathcal{F}^m_{i_1\ldots i_m}\mathbf{x}^1_{i_1}\ldots\mathbf{x}^m_{i_m},$$

$S^m(\mathbf{x}) := F^m(\mathbf{x},\ldots,\mathbf{x})$, and the relation

$$S^4(\mathbf{x}) = 4\,F^3(\mathbf{x},\mathbf{x},\mathbf{x})\sum_{i=1}^{n}\mathbf{x}_i = 4\,S^3(\mathbf{x})\sum_{i=1}^{n}\mathbf{x}_i. \quad (1)$$

For simplicity, the superscript in the multilinear form can be omitted when there is no ambiguity. For example, $F(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t})$ can be used interchangeably with $F^4(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t})$ as the number of variables is four, which already implies a fourth order tensor. Similarly, $F(\mathbf{x},\mathbf{y},\cdot,\cdot)$ is equivalent to $F^4(\mathbf{x},\mathbf{y},\cdot,\cdot)$, and so on. In the same way for a third order tensor, $F(\mathbf{x},\mathbf{y},\mathbf{z})$ can be used interchangeably with $F^3(\mathbf{x},\mathbf{y},\mathbf{z})$. We write $F^4(\mathbf{x},\mathbf{x},\mathbf{x},\cdot)$ to denote the vector in $\mathbb{R}^n$ such that

$$F^4(\mathbf{x},\mathbf{x},\mathbf{x},\cdot)_l=\sum_{i,j,k=1}^{n}\mathcal{F}^4_{ijkl}\mathbf{x}_i\mathbf{x}_j\mathbf{x}_k$$

for all $1\leq l\leq n$. We write $F^4(\mathbf{x},\mathbf{x},\cdot,\cdot)$ to denote the matrix in $\mathbb{R}^{n\times n}$ such that

$$F^4(\mathbf{x},\mathbf{x},\cdot,\cdot)_{kl}=\sum_{i,j=1}^{n}\mathcal{F}^4_{ijkl}\mathbf{x}_i\mathbf{x}_j$$

for all $1\leq k,l\leq n$. For an m-th order symmetric tensor it holds for any permutation σ of $\{1,\ldots,m\}$

$$F^m(\mathbf{x}^1,\ldots,\mathbf{x}^m)=F^m(\mathbf{x}^{\sigma(1)},\ldots,\mathbf{x}^{\sigma(x)}).$$

Note that if F^4 is associated to a symmetric tensor, then the position of the dots do not matter. For example, one has $F^4(\mathbf{x},\mathbf{x},\mathbf{x},\cdot) = F^4(\mathbf{x},\mathbf{x},\cdot,\mathbf{x}) = F^4(\mathbf{x},\cdot,\mathbf{x},\mathbf{x}) = F^4(\cdot,\mathbf{x},\mathbf{x},\mathbf{x})$. Similar properties hold for multilinear forms of other orders.

1. Proofs

According to Section 2 in the paper, the third order hypergraph matching problem is formulated as the maximization of the following function:

$$S^3(\mathbf{x})=\sum_{i,j,k=1}^{n}\mathcal{F}^3_{ijk}\,\mathbf{x}_i\,\mathbf{x}_j\,\mathbf{x}_k. \quad (2)$$

where $\mathcal{F}^3$ is a third order symmetric affinity tensor.

Lemma 3.1 *Let S^3 be defined as in Equation (2). If S^3 is not constantly zero, then $S^3:\mathbb{R}^n\to\mathbb{R}$ is not convex.*

Proof: The Hessian HS^3 of S^3 satisfies

$$HS^3(\mathbf{x})_{jk}=6\sum_{i=1}^{n}\mathcal{F}^3_{ijk}\mathbf{x}_i, \qquad \forall 1\leq j,k\leq n, \quad (3)$$

for every $\mathbf{x}\in\mathbb{R}^n$, i.e. $HS^3(\mathbf{x})=6F^3(\mathbf{x},\cdot,\cdot)$. Now, if S^3 is convex, then

$$0\leq\langle\mathbf{y},HS^3(\mathbf{x})\mathbf{y}\rangle=6F^3(\mathbf{x},\mathbf{y},\mathbf{y}) \qquad \forall\mathbf{x},\mathbf{y}\in\mathbb{R}^n. \quad (4)$$

It follows that

$$0\leq F(\mathbf{x},\mathbf{y},\mathbf{y}) \text{ and } 0\leq F(-\mathbf{x},\mathbf{y},\mathbf{y})=-F(\mathbf{x},\mathbf{y},\mathbf{y}) \quad (5)$$

for every $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$. In particular, for $\mathbf{x}=\mathbf{y}$ we get $0=F(\mathbf{x},\mathbf{x},\mathbf{x})=S^3(\mathbf{x})$ for every $\mathbf{x}\in\mathbb{R}^n$. □

The following lemma shows that even if we optimize the multilinear form F^4 instead of S^4 we get ascent in S^4.

□

Lemma 3.2 *Let $\mathcal{F}^4$ be a fourth order symmetric tensor. If $S^4\colon \mathbb{R}^n \to \mathbb{R}$ is convex, then for all $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t} \in \mathbb{R}^n$:*

1. $F^4(\mathbf{x}, \mathbf{x}, \mathbf{y}, \mathbf{y}) \leq \max_{\mathbf{u} \in \{\mathbf{x}, \mathbf{y}\}} F^4(\mathbf{u}, \mathbf{u}, \mathbf{u}, \mathbf{u})$,
2. $F^4(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}) \leq \max_{\mathbf{u} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}\}} F^4(\mathbf{u}, \mathbf{u}, \mathbf{u}, \mathbf{u})$.

Proof: First, we prove a characterization for the convexity of S^4. The gradient and the Hessian of S^4 are given by

$$\nabla S^4(\mathbf{x}) = 4F^4(\mathbf{x}, \mathbf{x}, \mathbf{x}, \cdot), \quad (6)$$
$$HS^4(\mathbf{x}) = 12F^4(\mathbf{x}, \mathbf{x}, \cdot, \cdot), \quad (7)$$

where we used the symmetry of $\mathcal{F}^4$. S^4 is convex if and only if $HS^4(\mathbf{x})$ is positive semi-definite for all $\mathbf{x} \in \mathbb{R}^n$ which is equivalent to

$$\langle \mathbf{y}, HS^4(\mathbf{x})\, \mathbf{y} \rangle = 12F^4(\mathbf{x}, \mathbf{x}, \mathbf{y}, \mathbf{y}) \geq 0, \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \quad (8)$$

1. By (8) and the multilinearity of F, we have

$$\begin{aligned} 0 &\leq F(\mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}, \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}) \quad (9) \\ &= F(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) + F(\mathbf{y}, \mathbf{y}, \mathbf{y}, \mathbf{y}) - 2F(\mathbf{x}, \mathbf{x}, \mathbf{y}, \mathbf{y}) \end{aligned}$$

   for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. It follows that

$$\begin{aligned} 2F(\mathbf{x}, \mathbf{x}, \mathbf{y}, \mathbf{y}) &\leq F(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) + F(\mathbf{y}, \mathbf{y}, \mathbf{y}, \mathbf{y}) \\ &\leq 2 \max_{\mathbf{u} \in \{\mathbf{x}, \mathbf{y}\}} F^4(\mathbf{u}, \mathbf{u}, \mathbf{u}, \mathbf{u}). \quad (10) \end{aligned}$$

   for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

2. Similarly, for all $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t} \in \mathbb{R}^n$, we have

$$\begin{aligned} 0 &\leq F(\mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}, \mathbf{z}-\mathbf{t}, \mathbf{z}-\mathbf{t}) \\ &= F(\mathbf{x}, \mathbf{x}, \mathbf{z}, \mathbf{z}) + F(\mathbf{x}, \mathbf{x}, \mathbf{t}, \mathbf{t}) + F(\mathbf{y}, \mathbf{y}, \mathbf{z}, \mathbf{z}) \\ &\quad + F(\mathbf{y}, \mathbf{y}, \mathbf{t}, \mathbf{t}) + 2F(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{z}) \quad (11) \\ &\quad + 2F(\mathbf{x}, \mathbf{y}, \mathbf{t}, \mathbf{t}) - 2F(\mathbf{x}, \mathbf{x}, \mathbf{z}, \mathbf{t}) \\ &\quad - 2F(\mathbf{y}, \mathbf{y}, \mathbf{z}, \mathbf{t}) - 4F(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}). \end{aligned}$$

   Switching the variables from $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t})$ to $(\mathbf{z}, \mathbf{t}, \mathbf{x}, \mathbf{y})$ and applying the same inequality, we get

$$\begin{aligned} 0 &\leq F(\mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}, \mathbf{z}+\mathbf{t}, \mathbf{z}+\mathbf{t}) \\ &= F(\mathbf{x}, \mathbf{x}, \mathbf{z}, \mathbf{z}) + F(\mathbf{x}, \mathbf{x}, \mathbf{t}, \mathbf{t}) + F(\mathbf{y}, \mathbf{y}, \mathbf{z}, \mathbf{z}) \\ &\quad + F(\mathbf{y}, \mathbf{y}, \mathbf{t}, \mathbf{t}) - 2F(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{z}) \quad (12) \\ &\quad - 2F(\mathbf{x}, \mathbf{y}, \mathbf{t}, \mathbf{t}) + 2F(\mathbf{x}, \mathbf{x}, \mathbf{z}, \mathbf{t}) \\ &\quad + 2F(\mathbf{y}, \mathbf{y}, \mathbf{z}, \mathbf{t}) - 4F(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}). \end{aligned}$$

   Summing up inequalities (11) and (12) we obtain:

$$\begin{aligned} 4F(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}) &\leq F(\mathbf{x}, \mathbf{x}, \mathbf{z}, \mathbf{z}) + F(\mathbf{x}, \mathbf{x}, \mathbf{t}, \mathbf{t}) \quad (13) \\ &\quad + F(\mathbf{y}, \mathbf{y}, \mathbf{z}, \mathbf{z}) + F(\mathbf{y}, \mathbf{y}, \mathbf{t}, \mathbf{t}). \end{aligned}$$

   Finally, applying the first result finishes the proof.

The following theorem shows that the optimization of the multilinear form $F^4(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t})$ and the score function $S^4(\mathbf{x}) = F^4(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x})$ are equivalent in the sense that there exists a globally optimal solution of the first problem which is also a globally optimal solution of the second problem.

Theorem 3.3 *Let $\mathcal{F}^4$ be a fourth order symmetric tensor and suppose that $S^4\colon \mathbb{R}^n \to \mathbb{R}$ is convex. Then it holds for any compact constraint set $D \subset \mathbb{R}^n$,*

$$\begin{aligned} \max_{\mathbf{x} \in D} F^4(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) &= \max_{\mathbf{x}, \mathbf{y} \in D} F^4(\mathbf{x}, \mathbf{x}, \mathbf{y}, \mathbf{y}) \quad (14) \\ &= \max_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t} \in D} F^4(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}). \end{aligned}$$

Proof: For any compact set $D \subset \mathbb{R}^n$ it holds:

$$\begin{aligned} \max_{\mathbf{x} \in D} F^4(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) &\leq \max_{\mathbf{x}, \mathbf{y} \in D} F^4(\mathbf{x}, \mathbf{x}, \mathbf{y}, \mathbf{y}) \quad (15) \\ &\leq \max_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t} \in D} F^4(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}). \end{aligned}$$

However, the second inequality in Lemma 3.2 shows

$$F^4(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}) \leq \max_{\mathbf{u} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}\}} F^4(\mathbf{u}, \mathbf{u}, \mathbf{u}, \mathbf{u}) \quad (16)$$

which leads to

$$\max_{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t} \in D} F^4(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}) \leq \max_{\mathbf{x} \in D} F^4(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}), \quad (17)$$

and the proof is done. □

The following proposition shows that we can always make S^4 convex.

Proposition 3.4 *Let $\mathcal{F}^4$ be a fourth order symmetric tensor. Then for any $\alpha \geq 3\left\|\mathcal{F}^4\right\|_2$, where $\left\|\mathcal{F}^4\right\|_2 := \sqrt{\sum_{i,j,k,l=1}^{n} (\mathcal{F}^4_{ijkl})^2}$, the function*

$$S^4_\alpha(\mathbf{x}) := S^4(\mathbf{x}) + \alpha \left\|\mathbf{x}\right\|_2^4 \quad (18)$$

is convex on $\mathbb{R}^n$, and for any $\mathbf{x} \in M$,

$$S^4_\alpha(\mathbf{x}) = S^4(\mathbf{x}) + \alpha\, n_1^2. \quad (19)$$

Proof: The gradient and the Hessian of S^4_α can be computed as:

$$\nabla S^4_\alpha(\mathbf{x}) = 4F^4(\mathbf{x}, \mathbf{x}, \mathbf{x}, \cdot) + 4\alpha \left\|\mathbf{x}\right\|_2^2 \mathbf{x}, \quad (20)$$
$$HS^4_\alpha(\mathbf{x}) = 12F^4(\mathbf{x}, \mathbf{x}, \cdot, \cdot) + 8\alpha \mathbf{x}\mathbf{x}^T + 4\alpha \left\|\mathbf{x}\right\|_2^2 I, \quad (21)$$

where I is the identity matrix. S^4_α is convex if and only if

$$\langle \mathbf{y}, HS^4_\alpha(\mathbf{x})\mathbf{y}\rangle \geq 0 \qquad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \tag{22}$$

This is equivalent to

$$12F^4(\mathbf{x},\mathbf{x},\mathbf{y},\mathbf{y}) + 8\alpha \langle \mathbf{x},\mathbf{y}\rangle^2 + 4\alpha \|\mathbf{x}\|_2^2 \|\mathbf{y}\|_2^2 \geq 0 \tag{23}$$

for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. By the Cauchy-Schwarz inequality we have

$$\begin{aligned} \left|F^4(\mathbf{x},\mathbf{x},\mathbf{y},\mathbf{y})\right| &= \left|\sum_{i,j,k,l=1}^n \mathcal{F}^4_{ijkl}\mathbf{x}_i\mathbf{x}_j\mathbf{y}_k\mathbf{y}_l\right| \\ &\leq \sqrt{\sum_{i,j,k,l=1}^n \left(\mathcal{F}^4_{ijkl}\right)^2 \sum_{i,j,k,l=1}^n \mathbf{x}_i^2\mathbf{x}_j^2\mathbf{y}_k^2\mathbf{y}_l^2} \\ &= \left\|\mathcal{F}^4\right\|_2 \|\mathbf{x}\|_2^2 \|\mathbf{y}\|_2^2, \end{aligned} \tag{24}$$

for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. It follows that for $\alpha \geq 3\left\|\mathcal{F}^4\right\|_2$, inequality (23) is true for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Finally, the second statement of the proposition follows from the fact that $\|\mathbf{x}\|_2^2 = n_1$ for any $\mathbf{x} \in M$. □

The following proposition shows that there exists a multilinear form associated to $\|\mathbf{x}\|_2^4$.

Proposition 3.5 *The symmetric tensor $\mathcal{G}^4 \in \mathbb{R}^{n\times n\times n\times n}$ with corresponding symmetric multilinear form defined as*

$$G^4(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t}) = \frac{\langle \mathbf{x},\mathbf{y}\rangle\langle \mathbf{z},\mathbf{t}\rangle + \langle \mathbf{x},\mathbf{z}\rangle\langle \mathbf{y},\mathbf{t}\rangle + \langle \mathbf{x},\mathbf{t}\rangle\langle \mathbf{y},\mathbf{z}\rangle}{3}$$

satisfies $G^4(\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x}) = \|\mathbf{x}\|_2^4$.

Proof: The proof requires two parts:

1. There exists a fourth order symmetric tensor $\mathcal{G}^4$ such that its multilinear form is given as above.

2. The multilinear form associated to $\mathcal{G}^4$ satisfies $G^4(\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x}) = \|\mathbf{x}\|_2^4$.

Let $\mathcal{G}^4 \in \mathbb{R}^{n\times n\times n\times n}$ be defined as follow:

$$\mathcal{G}^4_{ijkl} = \begin{cases} 1 & \text{if } i=j=k=l, \\ 1/3 & \text{if } i=j\neq k=l, \\ 1/3 & \text{if } i=k\neq j=l, \\ 1/3 & \text{if } i=l\neq j=k, \\ 0 & \text{else}, \end{cases} \qquad \forall 1 \leq i,j,k,l \leq n.$$

The multilinear form associated to $\mathcal{G}$ is then computed as

$$\begin{aligned} G^4(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t}) &= \sum_{i,j,k,l=1}^n \mathcal{G}^4_{ijkl}\mathbf{x}_i\mathbf{y}_j\mathbf{z}_k\mathbf{t}_l \\ &= \sum_{i=1}^n \mathbf{x}_i\mathbf{y}_i\mathbf{z}_i\mathbf{t}_i + \frac{1}{3}\sum_{\substack{i,j=1\\ i\neq j}}^n \mathbf{x}_i\mathbf{y}_i\mathbf{z}_j\mathbf{t}_j \\ &\quad + \frac{1}{3}\sum_{\substack{i,j=1\\ i\neq j}}^n \mathbf{x}_i\mathbf{y}_j\mathbf{z}_i\mathbf{t}_j + \frac{1}{3}\sum_{\substack{i,j=1\\ i\neq j}}^n \mathbf{x}_i\mathbf{y}_j\mathbf{z}_j\mathbf{t}_i \\ &= \frac{1}{3}\Big(\langle \mathbf{x},\mathbf{y}\rangle\langle \mathbf{z},\mathbf{t}\rangle + \langle \mathbf{x},\mathbf{z}\rangle\langle \mathbf{y},\mathbf{t}\rangle + \langle \mathbf{x},\mathbf{t}\rangle\langle \mathbf{y},\mathbf{z}\rangle\Big) \end{aligned}$$

As a result, we have $G^4(\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x}) = \|\mathbf{x}\|_2^4$. □

Our first algorithm optimizes for $\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t} \in M$ the function:

$$F^4_\alpha(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t}) = F^4(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t}) + \alpha\, G^4(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{t}) \tag{25}$$

and our second algorithm optimizes for $\mathbf{x},\mathbf{y} \in M$ the function:

$$F^4_\alpha(\mathbf{x},\mathbf{x},\mathbf{y},\mathbf{y}) = F^4(\mathbf{x},\mathbf{x},\mathbf{y},\mathbf{y}) + \alpha\, G^4(\mathbf{x},\mathbf{x},\mathbf{y},\mathbf{y}) \tag{26}$$

Theorem 4.1 *Let $\mathcal{F}^4$ be a fourth order symmetric tensor. Then the following holds for Algorithm 1:*

1. *The sequence $F_\alpha(\mathbf{x}^k,\mathbf{y}^k,\mathbf{z}^k,\mathbf{t}^k)$ for $k = 1,2,\ldots$ is strictly monotonically increasing or terminates.*
2. *The sequence of scores $S^4(\mathbf{u}^m)$ for $m = 1,2,\ldots$ is strictly monotonically increasing or terminates. For every m, $\mathbf{u}^m \in M$ is a valid assignment matrix.*
3. *The sequence of original third order scores $S^3(\mathbf{u}^m)$ for $m = 1,2,\ldots$ is strictly monotonically increasing or terminates.*
4. *The algorithm terminates after a finite number of iterations.*

Proof: From the definition of steps $1) - 4)$ in Algorithm 1, we get

$$\begin{aligned} F_{\alpha,k} &:= F_\alpha(\mathbf{x}^k,\mathbf{y}^k,\mathbf{z}^k,\mathbf{t}^k) \\ &\leq F_\alpha(\tilde{\mathbf{x}}^{k+1},\mathbf{y}^k,\mathbf{z}^k,\mathbf{t}^k) \\ &\leq F_\alpha(\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{y}}^{k+1},\tilde{\mathbf{z}}^k,\mathbf{t}^k) \\ &\leq F_\alpha(\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{y}}^{k+1},\tilde{\mathbf{z}}^{k+1},\mathbf{t}^k) \\ &\leq F_\alpha(\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{y}}^{k+1},\tilde{\mathbf{z}}^{k+1},\tilde{\mathbf{t}}^{k+1}) =: \tilde{F}_{\alpha,k+1}. \end{aligned} \tag{27}$$

Either $\tilde{F}_{\alpha,k+1} > F_{\alpha,k}$ in which case

$$\mathbf{x}^{k+1} = \tilde{\mathbf{x}}^{k+1}, \mathbf{y}^{k+1} = \tilde{\mathbf{y}}^{k+1}, \mathbf{z}^{k+1} = \tilde{\mathbf{z}}^{k+1}, \mathbf{t}^{k+1} = \tilde{\mathbf{t}}^{k+1}$$

Algorithm 1: BCAGM

Input: Lifted affinity tensor $\mathcal{F}^4$, $\alpha = 3\left\|\mathcal{F}^4\right\|_2$, $(\mathbf{x}^0, \mathbf{y}^0, \mathbf{z}^0, \mathbf{t}^0) \in M \times M \times M \times M$, $k = 0$, $m = 0$
Output: $\mathbf{x}^* \in M$
Repeat

1. $\tilde{\mathbf{x}}^{k+1} = \arg\max_{\mathbf{x}\in M} F_\alpha(\mathbf{x}, \mathbf{y}^k, \mathbf{z}^k, \mathbf{t}^k)$
2. $\tilde{\mathbf{y}}^{k+1} = \arg\max_{\mathbf{y}\in M} F_\alpha(\tilde{\mathbf{x}}^{k+1}, \mathbf{y}, \mathbf{z}^k, \mathbf{t}^k)$
3. $\tilde{\mathbf{z}}^{k+1} = \arg\max_{\mathbf{z}\in M} F_\alpha(\tilde{\mathbf{x}}^{k+1}, \tilde{\mathbf{y}}^{k+1}, \mathbf{z}, \mathbf{t}^k)$
4. $\tilde{\mathbf{t}}^{k+1} = \arg\max_{\mathbf{t}\in M} F_\alpha(\tilde{\mathbf{x}}^{k+1}, \tilde{\mathbf{y}}^{k+1}, \tilde{\mathbf{z}}^{k+1}, \mathbf{t})$
5. **if** $F_\alpha(\tilde{\mathbf{x}}^{k+1}, \tilde{\mathbf{y}}^{k+1}, \tilde{\mathbf{z}}^{k+1}, \tilde{\mathbf{t}}^{k+1}) = F_\alpha(\mathbf{x}^k, \mathbf{y}^k, \mathbf{z}^k, \mathbf{t}^k)$ **then**
   - $\mathbf{u}^{m+1} = \arg\max_{\mathbf{u}\in\{\tilde{\mathbf{x}}^{k+1}, \tilde{\mathbf{y}}^{k+1}, \tilde{\mathbf{z}}^{k+1}, \tilde{\mathbf{t}}^{k+1}\}} F_\alpha(\mathbf{u}, \mathbf{u}, \mathbf{u}, \mathbf{u})$
   - **if** $F_\alpha(\mathbf{u}^{m+1}, \mathbf{u}^{m+1}, \mathbf{u}^{m+1}, \mathbf{u}^{m+1}) = F_\alpha(\tilde{\mathbf{x}}^{k+1}, \tilde{\mathbf{y}}^{k+1}, \tilde{\mathbf{z}}^{k+1}, \tilde{\mathbf{t}}^{k+1})$ **then return**
   - $\mathbf{x}^{k+1} = \mathbf{y}^{k+1} = \mathbf{z}^{k+1} = \mathbf{t}^{k+1} = \mathbf{u}^{m+1}$
   - $m = m + 1$

   else $\mathbf{x}^{k+1} = \tilde{\mathbf{x}}^{k+1}$, $\mathbf{y}^{k+1} = \tilde{\mathbf{y}}^{k+1}$, $\mathbf{z}^{k+1} = \tilde{\mathbf{z}}^{k+1}$, $\mathbf{t}^{k+1} = \tilde{\mathbf{t}}^{k+1}$

   end
6. $k = k + 1$

and

$$F_\alpha(\mathbf{x}^{k+1}, \mathbf{y}^{k+1}, \mathbf{z}^{k+1}, \mathbf{t}^{k+1}) > F_\alpha(\mathbf{x}^k, \mathbf{y}^k, \mathbf{z}^k, \mathbf{t}^k),$$

or $\tilde{F}_{\alpha,k+1} = F_{\alpha,k}$ and the algorithm enters step 5. Since, by Proposition 3.4, S^4_α is convex for the chosen value of α, applying Lemma 3.2 we get

$$\begin{aligned}\tilde{F}_{\alpha,k+1} &\leq \max_{\mathbf{v}\in\{\tilde{\mathbf{x}}^{k+1}, \tilde{\mathbf{y}}^{k+1}, \tilde{\mathbf{z}}^{k+1}, \tilde{\mathbf{t}}^{k+1}\}} F_\alpha(\mathbf{v}, \mathbf{v}, \mathbf{v}, \mathbf{v}) \\ &= F_\alpha(\mathbf{u}^{m+1}, \mathbf{u}^{m+1}, \mathbf{u}^{m+1}, \mathbf{u}^{m+1}) \\ &= S^4_\alpha(\mathbf{u}^{m+1}).\end{aligned} \tag{28}$$

If the inequality is an equality, then the termination condition of the algorithm is met. Otherwise, we get

$$\begin{aligned}\tilde{F}_{\alpha,k+1} &< F_\alpha(\mathbf{u}^{m+1}, \mathbf{u}^{m+1}, \mathbf{u}^{m+1}, \mathbf{u}^{m+1}) \\ &= S^4_\alpha(\mathbf{u}^{m+1}) = F_\alpha(\mathbf{x}^{k+1}, \mathbf{y}^{k+1}, \mathbf{z}^{k+1}, \mathbf{t}^{k+1}).\end{aligned} \tag{29}$$

This proves the first statement of the theorem.

From

$$\begin{aligned}S^4_\alpha(\mathbf{u}^{m+1}) &= F_\alpha(\mathbf{u}^{m+1}, \mathbf{u}^{m+1}, \mathbf{u}^{m+1}, \mathbf{u}^{m+1}) \\ &= F_\alpha(\mathbf{x}^{k+1}, \mathbf{y}^{k+1}, \mathbf{z}^{k+1}, \mathbf{t}^{k+1})\end{aligned} \tag{30}$$

it follows that $S^4_\alpha(\mathbf{u}^m), m = 1, 2, \ldots$ is a subsequence of $F_\alpha(\mathbf{x}^k, \mathbf{y}^k, \mathbf{z}^k, \mathbf{t}^k), k = 1, 2, \ldots$ and thus it holds either $S^4_\alpha(\mathbf{u}^m) = S^4_\alpha(\mathbf{u}^{m+1})$ in which case the algorithm terminates or $S^4_\alpha(\mathbf{u}^m) < S^4_\alpha(\mathbf{u}^{m+1})$. However, by Equation (19), the additional term which has been added to S^4 to get a convex function is constant on M, that is

$$S^4_\alpha(\mathbf{x}) = S^4(\mathbf{x}) + \alpha n_1^2 \qquad \forall \mathbf{x} \in M.$$

It follows that either $S^4(\mathbf{u}^m) = S^4(\mathbf{u}^{m+1})$ and the algorithm terminates or $S^4(\mathbf{u}^m) < S^4(\mathbf{u}^{m+1})$ which proves the second part of the theorem.

By Equation (1), we have $S^4(\mathbf{x}) = 4\,n_1\,S^3(\mathbf{x})$ for any $\mathbf{x} \in M$. Thus the statements made for S^4 directly translate into corresponding statements for the original third order score S^3. This proves the penultimate statement.

Finally, the algorithm yields a strictly monotonically increasing sequence $S^4(\mathbf{u}^m), m = 1, 2, \ldots$ or it terminates. Since there is only a finite number of possible assignment matrices, the sequence has to terminate after a finite number of steps. □

We prove below Theorem 4.2 for Algorithm 2 as presented in the paper. Assume Ψ is a sub-algorithm which delivers monotonic ascent for the quadratic assignment problem (QAP)

$$\max_{\mathbf{x}\in M} \langle \mathbf{x}, A\mathbf{x}\rangle = \max_{\mathbf{x}\in M} \sum_{i,j=1}^{n} A_{ij}\mathbf{x}_i\mathbf{x}_j, \tag{31}$$

Theorem 4.2 *Let $\mathcal{F}^4$ be a fourth order symmetric tensor and let Ψ be an algorithm for the QAP which yields monotonic ascent. Then the following holds for Algorithm 2:*

1. *The sequence $F_\alpha(\mathbf{x}^k, \mathbf{x}^k, \mathbf{y}^k, \mathbf{y}^k)$ for $k = 1, 2, \ldots$ is strictly monotonically increasing or terminates.*
2. *The sequence of scores $S^4(\mathbf{u}^m)$ for $m = 1, 2, \ldots$ is strictly monotonically increasing or terminates. For every m, $\mathbf{u}^m \in M$ is a valid assignment matrix.*
3. *The sequence of original third order scores $S^3(\mathbf{u}^m)$ for $m = 1, 2, \ldots$ is strictly monotonically increasing or terminates.*
4. *The algorithm terminates after a finite number of iterations.*

Proof: From the definition of steps $1) - 2)$ in Algorithm 2, we get

$$\begin{aligned}F_{\alpha,k} &:= F_\alpha(\mathbf{x}^k, \mathbf{x}^k, \mathbf{y}^k, \mathbf{y}^k) \\ &\leq F_\alpha(\tilde{\mathbf{x}}^{k+1}, \tilde{\mathbf{x}}^{k+1}, \mathbf{y}^k, \mathbf{y}^k) \\ &\leq F_\alpha(\tilde{\mathbf{x}}^{k+1}, \tilde{\mathbf{x}}^{k+1}, \tilde{\mathbf{y}}^{k+1}, \tilde{\mathbf{y}}^{k+1}) =: \tilde{F}_{\alpha,k+1}.\end{aligned} \tag{32}$$

Algorithm 2: BCAGM-Ψ

Input: Lifted affinity tensor $\mathcal{F}^4$, $\alpha = 3\left\|\mathcal{F}^4\right\|_2$,
$(\mathbf{x}^0, \mathbf{y}^0) \in M \times M$, $k = 0$, $m = 0$,
$\mathbf{z} = \Psi(A, \mathbf{x}^k)$ is an algorithm for the QAP in (31) which provides monotonic ascent, that is
$\langle \mathbf{z}, A\mathbf{z}\rangle \geq \langle \mathbf{x}^k, A\mathbf{x}^k\rangle$
Output: $\mathbf{x}^* \in M$
Repeat

1. $\tilde{\mathbf{x}}^{k+1} = \Psi\big(F_\alpha(\cdot,\cdot,\mathbf{y}^k,\mathbf{y}^k),\, \mathbf{x}^k\big)$
2. $\tilde{\mathbf{y}}^{k+1} = \Psi\big(F_\alpha(\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{x}}^{k+1},\cdot,\cdot), \mathbf{y}^k\big)$
3. **if** $F_\alpha(\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{y}}^{k+1},\tilde{\mathbf{y}}^{k+1}) = F_\alpha(\mathbf{x}^k,\mathbf{x}^k,\mathbf{y}^k,\mathbf{y}^k)$ **then**
   - $\mathbf{u}^{m+1} = \underset{\mathbf{u}\in\{\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{y}}^{k+1}\}}{\arg\max}\ F_\alpha(\mathbf{u},\mathbf{u},\mathbf{u},\mathbf{u})$
   - **if** $F_\alpha(\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{y}}^{k+1},\tilde{\mathbf{y}}^{k+1}) = F_\alpha(\mathbf{u}^{m+1},\mathbf{u}^{m+1},\mathbf{u}^{m+1},\mathbf{u}^{m+1})$ **then return**
   - $\mathbf{x}^{k+1} = \mathbf{y}^{k+1} = \mathbf{u}^{m+1}$
   - $m = m+1$

   else $\mathbf{x}^{k+1} = \tilde{\mathbf{x}}^{k+1}$, $\mathbf{y}^{k+1} = \tilde{\mathbf{y}}^{k+1}$.
   end
4. $k = k+1$

Either $\tilde{F}_{\alpha,k+1} > F_{\alpha,k}$ in which case

$$\mathbf{x}^{k+1} = \tilde{\mathbf{x}}^{k+1}, \qquad \mathbf{y}^{k+1} = \tilde{\mathbf{y}}^{k+1}$$

and

$$F_\alpha(\mathbf{x}^{k+1},\mathbf{x}^{k+1},\mathbf{y}^{k+1},\mathbf{y}^{k+1}) > F_\alpha(\mathbf{x}^k,\mathbf{x}^k,\mathbf{y}^k,\mathbf{y}^k),$$

or $\tilde{F}_{\alpha,k+1} = F_{\alpha,k}$ and the algorithm enters step 3. Since, by Proposition 3.4, S^4_α is convex for the chosen value of α, applying Lemma 3.2 we get

$$\begin{aligned}\tilde{F}_{\alpha,k+1} &\leq \max_{\mathbf{v}\in\{\tilde{\mathbf{x}}^{k+1},\tilde{\mathbf{y}}^{k+1}\}} F_\alpha(\mathbf{v},\mathbf{v},\mathbf{v},\mathbf{v}) \\ &= F_\alpha(\mathbf{u}^{m+1},\mathbf{u}^{m+1},\mathbf{u}^{m+1},\mathbf{u}^{m+1}) \qquad (33) \\ &= S^4_\alpha(\mathbf{u}^{m+1}).\end{aligned}$$

If the inequality is an equality, then the termination condition of the algorithm is met. Otherwise, we get

$$\begin{aligned}\tilde{F}_{\alpha,k+1} &< F_\alpha(\mathbf{u}^{m+1},\mathbf{u}^{m+1},\mathbf{u}^{m+1},\mathbf{u}^{m+1}) \qquad (34) \\ &= S^4_\alpha(\mathbf{u}^{m+1}) = F_\alpha(\mathbf{x}^{k+1},\mathbf{x}^{k+1},\mathbf{y}^{k+1},\mathbf{y}^{k+1}).\end{aligned}$$

This proves the first statement of the theorem.

From

$$\begin{aligned}S^4_\alpha(\mathbf{u}^{m+1}) &= F_\alpha(\mathbf{u}^{m+1},\mathbf{u}^{m+1},\mathbf{u}^{m+1},\mathbf{u}^{m+1}) \\ &= F_\alpha(\mathbf{x}^{k+1},\mathbf{x}^{k+1},\mathbf{y}^{k+1},\mathbf{y}^{k+1}) \qquad (35)\end{aligned}$$

it follows that $S^4_\alpha(\mathbf{u}^m), m = 1,2,\ldots$ is a subsequence of $F_\alpha(\mathbf{x}^k,\mathbf{x}^k,\mathbf{y}^k,\mathbf{y}^k), k = 1,2,\ldots$ and thus it holds either $S^4_\alpha(\mathbf{u}^m) = S^4_\alpha(\mathbf{u}^{m+1})$ in which case the algorithm terminates or $S^4_\alpha(\mathbf{u}^m) < S^4_\alpha(\mathbf{u}^{m+1})$. However, by Equation (19), the additional term which has been added to S^4 to get a convex function is constant on M, that is

$$S^4_\alpha(\mathbf{x}) = S^4(\mathbf{x}) + \alpha n_1^2 \qquad \forall \mathbf{x} \in M.$$

It follows that either $S^4(\mathbf{u}^m) = S^4(\mathbf{u}^{m+1})$ and the algorithm terminates or $S^4(\mathbf{u}^m) < S^4(\mathbf{u}^{m+1})$ which proves the second part of the theorem.

By Equation (1), we have $S^4(\mathbf{x}) = 4\,n_1\,S^3(\mathbf{x})$ for any $\mathbf{x} \in M$. Thus the statements made for S^4 directly translate into corresponding statements for the original third order score S^3. This proves the penultimate statement.

Finally, the algorithm yields a strictly monotonically increasing sequence $S^4(\mathbf{u}^m), m = 1,2,\ldots$ or it terminates. Since there is only a finite number of possible assignment matrices, the sequence has to terminate after a finite number of steps. □

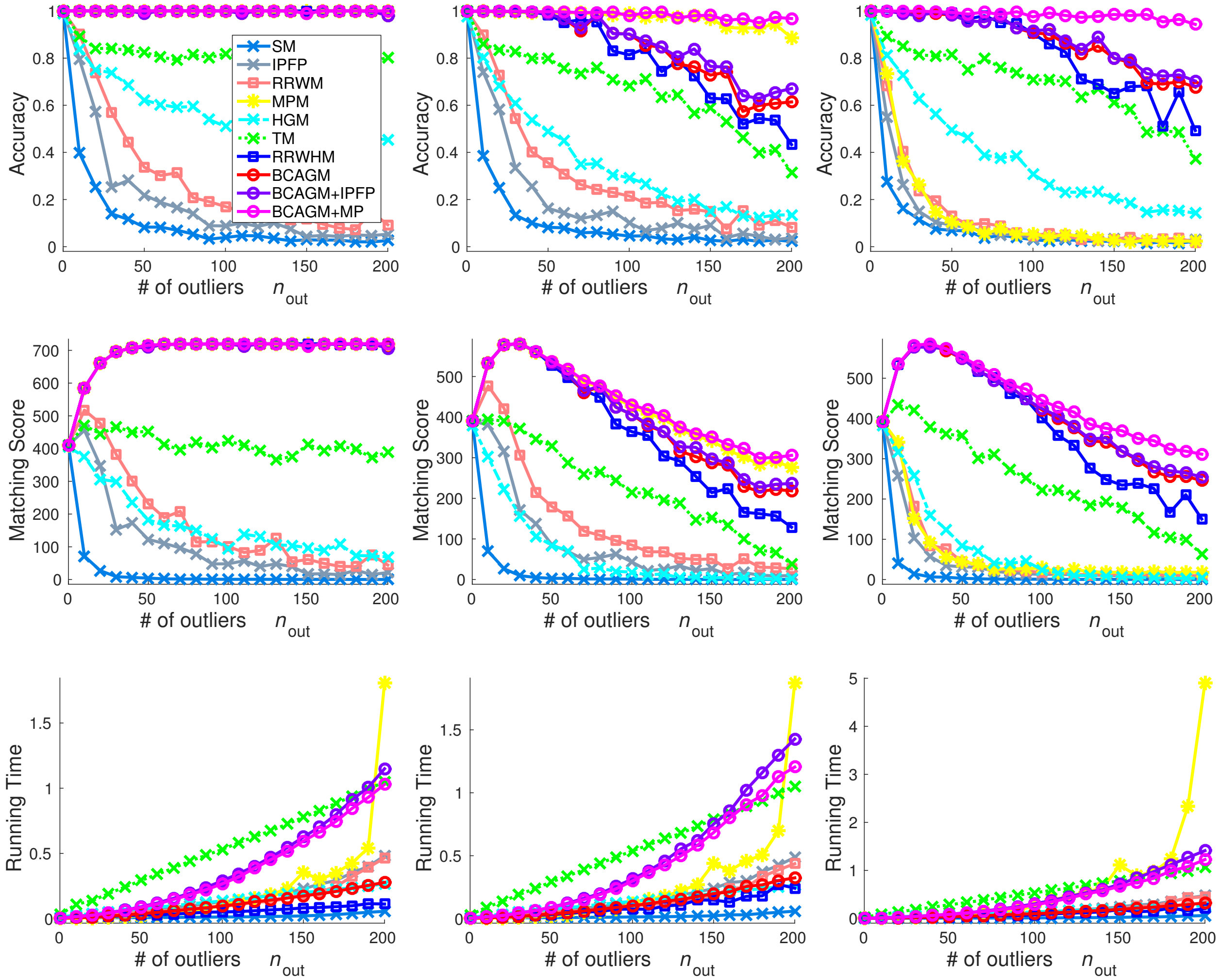
2. Additional Experiments

We provide below additional experimental results on the synthetic and CMU House dataset. The running time of all the algorithms is reported along with accuracy and matching score.

2.1. Synthetic Dataset

In this experiment, further results on the outlier and deformation settings are provided. The same has been done for the experiments in the paper. Figure 1 shows additional results in the outlier setting, where the number of outliers was varied from 0 to 200. The number of inliers was fixed to 10 while σ and $scale$ were set to different values. It is interesting to see that when there is no deformation and scaling, our algorithms together with RRWHM [3] and MPM [1] achieve an almost perfect result. However, our algorithms outperform all other higher order approaches when deformation and scaling are slightly present. Compared to second order methods, our algorithms can take advantage of higher order features, therefore, achieve superior performance when transformations such as scaling are present.

Figure 2 shows further results in the deformation setting, where the number of inliers was set to 30 and 40 accordingly, and no other form of noise was used. As we show the runtime for all the experiments, the result for 20 inliers points is repeated from the paper as well. One can observe from Figure 2 that our algorithms always stay competitive with other state-of-the-art higher order methods, in particular RRWHM [3], even when the deformation is significant.



(a) $n_{in} = 10, \sigma = 0, scale = 1$ (b) $n_{in} = 10, \sigma = 0.01, scale = 1.01$ (c) $n_{in} = 10, \sigma = 0.01, scale = 1.1$

Figure 1: Matching point sets in $\mathbb{R}^2$ (outliers test): The top row shows average accuracy while the middle row shows the average matching score and the bottom row shows average running time. The number of outliers was varied from 0 to 200 with the interval 10. (a) Increasing number of outliers without deformation and scaling. (b) Increasing number of outliers with slight deformation and small scaling. (c) Increasing number of outliers with slight deformation and large scaling. (Best viewed in color.)

2.2. CMU House Dataset

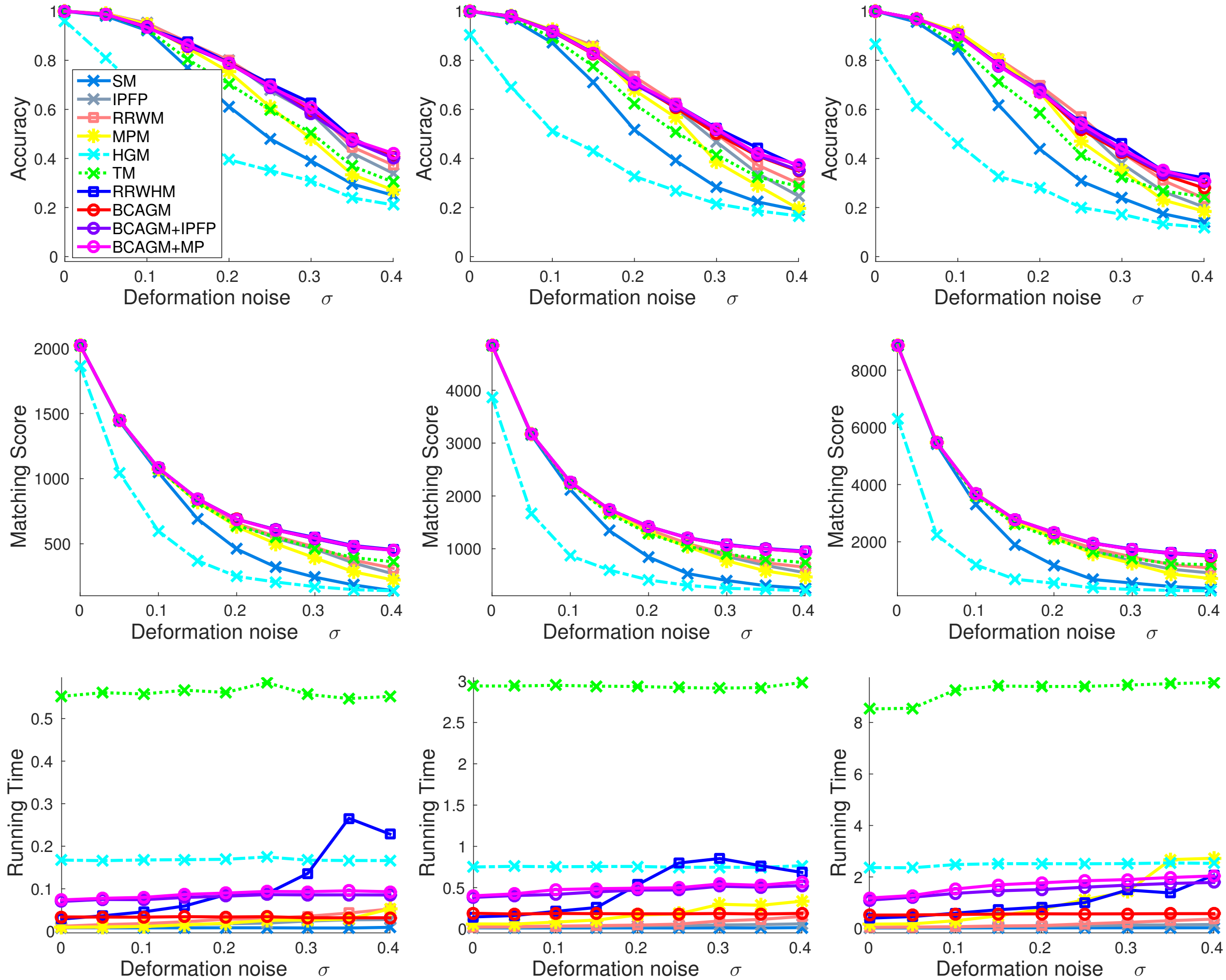
This section presents further results on the CMU house dataset. In particular, we evaluate GM algorithms on two tasks where we match 15 points to 30 points and 25 points to 30 points in two corresponding images. For each task, we match all the possible image pairs and compute the average result for each baseline as done in the paper. The results in Figure 3 show that our algorithms achieve competitive or better results than other methods for all the baselines.

2.3. Car/Motorbike Dataset

Figure 4 shows the running time of higher order methods on the Car and Motorbike dataset.

2.4. Runtime

From all the runtime plots, one observes that our algorithms have competitive running time compared to the other methods. In particular, BCAGM is always among the methods with lowest running time. Compared to other third order approaches, Figures 2, 3, 4 show that our algorithms are much faster than TM [2] and HGM [4], while staying quite



(a) $n_{in} = 20$ (b) $n_{in} = 30$ (c) $n_{in} = 40$

Figure 2: Matching point sets in $\mathbb{R}^2$ (deformation test): The top row shows average accuracy while the middle row shows average objective score and the bottom row shows average running time. (a) 20 inliers (b) 30 inliers (c) 40 inliers (Best viewed in color.)

competitive with RRWHM.

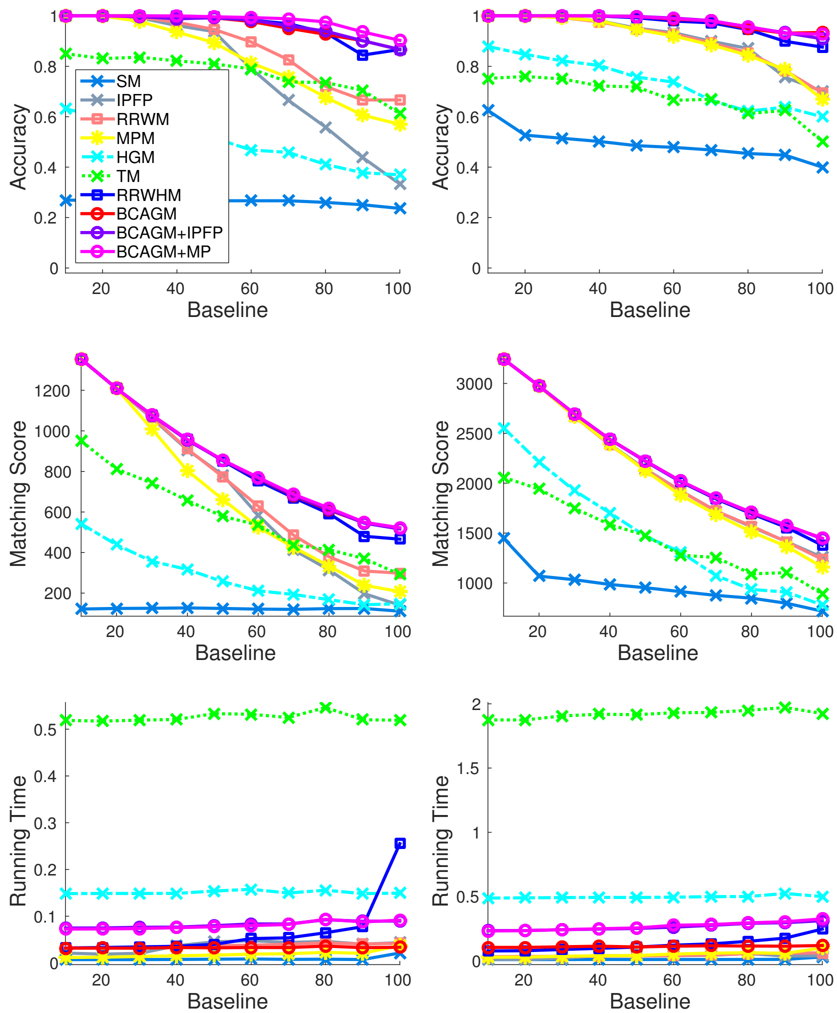
References

[1] M. Cho, J. Sun, O. Duchenne, and J. Ponce. Finding matches in a haystack: A max-pooling strategy for graph matching in the presence of outliers. In *CVPR*, 2014. 5

[2] O. Duchenne, F. Bach, I. Kweon, and J. Ponce. A tensor-based algorithm for high-order graph matching. *PAMI*, 33:2383–2395, 2011. 6

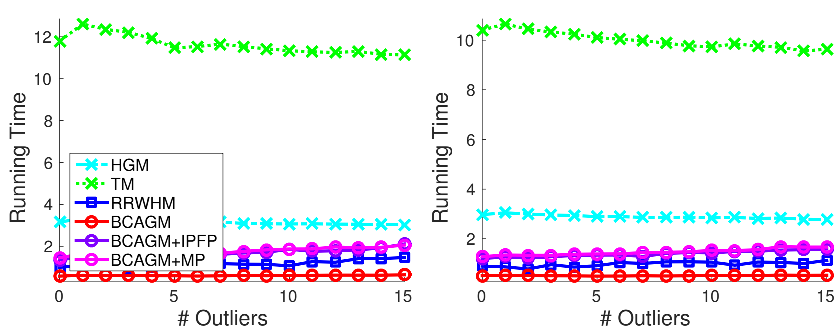
[3] J. Lee, M. Cho, and K. M. Lee. Hyper-graph matching via reweighted random walks. In *CVPR*, 2011. 5

[4] R. Zass and A. Shashua. Probabilistic graph and hypergraph matching. In *CVPR*, 2008. 6



(a) 15 pts vs 30 pts (b) 25 pts vs 30 pts

Figure 3: CMU house dataset: The top row shows average matching accuracy while the middle row shows average objective score and the bottom row shows average running time. (Best viewed in color.)



(a) Car dataset (b) Motorbike dataset

Figure 4: Car and Motorbike dataset: Running time of the third order methods. The accuracy and matching score can be found in the paper. (Best viewed in color.)