Theoretical Proofs

Xingyu Cao, Yipeng Liu *Senior Member, IEEE*, Jiani Liu, Ce Zhu *Fellow, IEEE*

Abstract

This document contains the proofs of Definition 4, Proposition 1 and Corollary 1 in the main body.

I. Derivation of FHCS for General Tensors

Given $\mathcal{T} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \cdots \circ \mathbf{u}^{(N)} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, if we assign $l = \sum_{n=1}^{N} (i_n - 1) \prod_{j=1}^{n-1} I_j + 1$, then $\text{vec}(\mathcal{T})_l = \mathbf{u}^{(1)}_{i_1} \cdots \mathbf{u}^{(N)}_{i_N}$. We have

$$\begin{aligned}
&\sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} \text{vec}(\mathcal{T})_l \mathbf{s}_1(i_1) \cdots \mathbf{s}_N(i_N) \mathrm{w}^{\mathbf{h}_1(i_1)+\cdots+\mathbf{h}_N(i_N)-N} \\
&= \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} \mathbf{u}^{(1)}_{i_1} \cdots \mathbf{u}^{(N)}_{i_N} \mathbf{s}_1(i_1) \cdots \mathbf{s}_N(i_N) \mathrm{w}^{\mathbf{h}_1(i_1)+\cdots+\mathbf{h}_N(i_N)-N} \\
&= \sum_{i_1} \mathbf{u}^{(1)}_{i_1} \mathbf{s}_1(i_1) \mathrm{w}^{\mathbf{h}_1(i_1)-1} \cdots \sum_{i_N} \mathbf{u}^{(N)}_{i_N} \mathbf{s}_N(i_N) \mathrm{w}^{\mathbf{h}_N(i_N)-1} \\
&= \mathcal{P}_{\text{CS}_1(\mathbf{u}^{(1)})}(\mathrm{w}) \cdots \mathcal{P}_{\text{CS}_N(\mathbf{u}^{(N)})}(\mathrm{w}),
\end{aligned} \tag{1}$$

where $\mathcal{P}_{\text{CS}_n(\mathbf{u}^{(n)})}(\mathrm{w})$ is the polynomial form of $\text{CS}_n(\mathbf{u}^{(n)})$ for $n \in [N]$. Let $\mathbf{s}_{N+1}(l) = \mathbf{s}_1(i_1) \cdots \mathbf{s}_N(i_N)$, $\mathbf{h}_{N+1}(l) = \mathbf{h}_1(i_1) + \cdots + \mathbf{h}_N(i_N) - N + 1$, we have

$$\begin{aligned}
&\sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} \text{vec}(\mathcal{T})_l \mathbf{s}_1(i_1) \cdots \mathbf{s}_N(i_N) \mathrm{w}^{\mathbf{h}_1(i_1)+\cdots+\mathbf{h}_N(i_N)-N} \\
&= \sum_{l} \text{vec}(\mathcal{T})_l \mathbf{s}_{N+1}(l) \mathrm{w}^{\mathbf{h}_{N+1}(l)-1} \\
&= \mathcal{P}_{\text{CS}(\text{vec}(\mathcal{T});\mathbf{h}_{N+1},\mathbf{s}_{N+1})}(\mathrm{w}) \\
&:= \mathcal{P}_{\text{FCS}(\mathcal{T};\{\mathbf{h}_n,\mathbf{s}_n\}_{n=1}^{N})}(\mathrm{w}).
\end{aligned} \tag{2}$$

Therefore, we have

$$\mathcal{P}_{\text{FCS}(\mathcal{T};\{\mathbf{h}_n,\mathbf{s}_n\}_{n=1}^{N})}(\mathrm{w}) = \mathcal{P}_{\text{CS}_1(\mathbf{u}^{(1)})}(\mathrm{w}) \cdots \mathcal{P}_{\text{CS}_N(\mathbf{u}^{(N)})}(\mathrm{w}). \tag{3}$$

Due to the fact that polynomial multiplication equals the convolution of their coefficients, we have

$$\begin{aligned}
\text{FCS}(\mathcal{T}; \{\mathbf{h}_n, \mathbf{s}_n\}_{n=1}^{N}) &= \text{CS}_1(\mathbf{u}^{(1)}) \circledast \cdots \circledast \text{CS}_N(\mathbf{u}^{(N)}) \\
&= \mathbb{F}^{-1}(\mathbb{F}(\text{CS}_1(\mathbf{u}^{(1)})) * \cdots * \mathbb{F}(\text{CS}_N(\mathbf{u}^{(N)}))),
\end{aligned} \tag{4}$$

which completes the proof.

II. Proofs for Proposition 1 and Corollary 1

A. Proof for Proposition 1

First we prove that for any 3-order tensors $\mathcal{M}$, $\mathcal{N}$ with the same size, $\langle \text{FCS}(\mathcal{M}), \text{FCS}(\mathcal{N}) \rangle$ is a consistent estimator of $\langle \mathcal{M}, \mathcal{N} \rangle$. To this end we prove it is unbiased with bounded variance.

1) Unbiasedness:

$$\begin{aligned}
\langle \text{FCS}(\mathcal{M}), \text{FCS}(\mathcal{N}) \rangle &= \sum_{t} \text{FCS}(\mathcal{M})_t \text{FCS}(\mathcal{N})_t \\
&= \sum_{t} \Big(\sum_{\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=t} \mathbf{s}_1(i_1)\mathbf{s}_2(i_2)\mathbf{s}_3(i_3)\mathcal{M}_{i_1,i_2,i_3} \Big) \Big(\sum_{\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)=t} \mathbf{s}_1(j_1)\mathbf{s}_2(j_2) + \mathbf{s}_3(j_3)\mathcal{N}_{j_1,j_2,j_3} \Big),
\end{aligned} \tag{5}$$

$$\mathrm{E}_{\mathbf{h},\mathbf{s}} \langle \text{FCS}(\mathcal{M}), \text{FCS}(\mathcal{N}) \rangle = \sum_{\substack{i_1,i_2,i_3 \\ j_1,j_2,j_3}} \mathrm{E}\left[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3), \mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3))\right] \mathrm{E}[\mathbf{s}_1(i_1)\mathbf{s}_2(i_2)\mathbf{s}_3(i_3)\mathbf{s}_1(j_1)\mathbf{s}_2(j_2)\mathbf{s}_3(j_3)] \mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{j_1,j_2,j_3}, \tag{6}$$

where E denotes expectation. Given $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$, are 2-wise independent, we have

$$\mathrm{E}[\mathbf{s}_1(i_1)\mathbf{s}_2(i_2)\mathbf{s}_3(i_3)\mathbf{s}_1(j_1)\mathbf{s}_2(j_2)\mathbf{s}_3(j_3)] = \begin{cases} 1, \; i_1 = j_1, i_2 = j_2, i_3 = j_3 \\ 0, \text{ otherwise.} \end{cases} \tag{7}$$

Obviously $\mathrm{E}[\delta(\mathbf{h}_1(i_1) + \mathbf{h}_2(i_2) + \mathbf{h}_3(i_3), \mathbf{h}_1(j_1) + \mathbf{h}_2(j_2) + \mathbf{s}_3(j_3))]|_{\substack{i_1=j_1\\ i_2=j_2\\ i_3=j_3}} = 1$. Hence

$$\mathrm{E}_{\mathbf{h},\mathbf{s}} \langle \mathrm{FCS}(\mathcal{M}), \mathrm{FCS}(\mathcal{N}) \rangle = \sum_{i_1,i_2,i_3} \mathcal{M}_{i_1,i_2,i_3} \mathcal{N}_{i_1,i_2,i_3} = \langle \mathcal{M}, \mathcal{N} \rangle . \tag{8}$$

2) Variance boundedness:

$$\begin{aligned}\langle \mathrm{FCS}(\mathcal{M}), \mathrm{FCS}(\mathcal{N}) \rangle^2 = \sum_{\substack{i_1,i_2,i_3\\ j_1,j_2,j_3\\ i'_1,i'_2,i'_3\\ j'_1,j'_2,j'_3}} & \delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3), \mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3))\delta(\mathbf{h}_1(i'_1)+\mathbf{h}_2(i'_2)+\mathbf{h}_3(i'_3), \mathbf{h}_1(j'_1)+\mathbf{h}_2(j'_2)+\mathbf{h}_3(j'_3)) \\ & \mathbf{s}_1(i_1)\mathbf{s}_2(i_2)\mathbf{s}_3(i_3)\mathbf{s}_1(j_1)\mathbf{s}_2(j_2)\mathbf{s}_3(j_3)\mathbf{s}_1(i'_1)\mathbf{s}_2(i'_2)\mathbf{s}_3(i'_3)\mathbf{s}_1(j'_1)\mathbf{s}_2(j'_2)\mathbf{s}_3(j'_3)\mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{j_1,j_2,j_3}\mathcal{M}_{i'_1,i'_2,i'_3}\mathcal{N}_{j'_1,j'_2,j'_3}.\end{aligned} \tag{9}$$

Denote $\mathfrak{G}_1 = \{i_1, j_1, i'_1, j'_1\}, \mathfrak{G}_2 = \{i_2, j_2, i'_2, j'_2\}, \mathfrak{G}_3 = \{i_3, j_3, i'_3, j'_3\}$. We say $\mathfrak{G}_k$ has *best match* if $i_k = j_k, i'_k = j'_k$ for $k \in [3]$. then (9) can be grouped into following cases:

Case 1 $\mathfrak{G}_1, \mathfrak{G}_2, \mathfrak{G}_3$ all have best matches, then

$$\mathrm{S}_1 = \sum_{\substack{i_1,i_2,i_3\\ i'_1,i'_2,i'_3}} \mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{i_1,i_2,i_3}\mathcal{M}_{i'_1,i'_2,i'_3}\mathcal{N}_{i'_1,i'_2,i'_3} = \langle \mathcal{M}, \mathcal{N} \rangle^2 . \tag{10}$$

Case 2 One out of $\mathfrak{G}_1, \mathfrak{G}_2, \mathfrak{G}_3$ does not have best match, take $\mathfrak{G}_1$ for example. Suppose $i_1 = i'_1 \neq j_1 = j'_1$, then

$$\begin{aligned}\mathrm{S}_2 &= \binom{3}{1}\binom{2}{1} \sum_{\substack{i_1,i_2,i_3\\ j_1,i'_2,i'_3}} \mathrm{E}[\delta(\mathbf{h}_1(i_1), \mathbf{h}_1(j_1))^2]\mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{j_1,i_2,i_3}\mathcal{M}_{i_1,i'_2,i'_3}\mathcal{N}_{j_1,i'_2,i'_3} \\ &= \frac{6}{J} \sum_{\substack{i_1,i_2,i_3\\ j_1,i'_2,i'_3}} \mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{j_1,i_2,i_3}\mathcal{M}_{i_1,i'_2,i'_3}\mathcal{N}_{j_1,i'_2,i'_3} \\ &= \frac{6}{J} \sum_{i_1,j_1} \langle \mathcal{M}(\mathbf{e}_{i_1}, \mathbf{I}, \mathbf{I}), \mathcal{N}(\mathbf{e}_{j_1}, \mathbf{I}, \mathbf{I}) \rangle^2 \\ &\leq \frac{6}{J} \sum_{i_1,j_1} \|\mathcal{M}(\mathbf{e}_{i_1}, \mathbf{I}, \mathbf{I})\|_{\mathrm{F}}^2 \|\mathcal{N}(\mathbf{e}_{j_1}, \mathbf{I}, \mathbf{I})\|_{\mathrm{F}}^2 \\ &= \frac{6}{J} \|\mathcal{M}\|_{\mathrm{F}}^2 \|\mathcal{N}\|_{\mathrm{F}}^2\end{aligned} \tag{11}$$

holds due to the Cauchy-Schwartz inequality.

Case 3 One out of $\mathfrak{G}_1, \mathfrak{G}_2, \mathfrak{G}_3$ has best match, take $\mathfrak{G}_1$ for example. Suppose $i_2 = i'_2 \neq j_2 = j'_2, i_3 = i'_3 \neq j_3 = j'_3$, then

$$\mathrm{S}_3 = \binom{3}{1}\binom{2}{1}\binom{2}{1} \sum_{\substack{i_1,i_2,i_3\\ j_2,j_3,i'_1}} \mathrm{E}[\delta(\mathbf{h}_2(i_2) + \mathbf{h}_3(i_3), \mathbf{h}_2(j_2) + \mathbf{h}_3(i_3))^2]\mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{i_1,j_2,j_3}\mathcal{M}_{i'_1,i_2,i_3}\mathcal{N}_{i'_1,j_2,j_3} \tag{12}$$

From 2-wise independence of $\mathbf{h}_2$ and $\mathbf{h}_3$, we have

$$\mathrm{P}[\mathbf{h}_2(i_2) + \mathbf{h}_3(i_3) = t] = \begin{cases} \displaystyle\sum_{k=0}^{t} \mathrm{P}[\mathbf{h}_2(i_2) = k, \mathbf{h}_3(i_3) = t - k] = \frac{t+1}{J^2}, & t \leq J - 1 \\ \displaystyle\sum_{k=t-J+1}^{J-1} \mathrm{P}[\mathbf{h}_2(i_2) = k, \mathbf{h}_3(i_3) = t - k] = \frac{2J - 1 - t}{J^2}, & t \geq J \end{cases}$$

Obviously

$$\begin{aligned}\mathrm{E}[\delta(\mathbf{h}_2(i_2) + \mathbf{h}_3(i_3), \mathbf{h}_2(j_2) + \mathbf{h}_3(i_3))^2] &= \mathrm{E}[\delta(\mathbf{h}_2(i_2) + \mathbf{h}_3(i_3), \mathbf{h}_2(j_2) + \mathbf{h}_3(i_3))] = \mathrm{P}[\mathbf{h}_2(i_2) + \mathbf{h}_3(i_3) = \mathbf{h}_2(j_2) + \mathbf{h}_3(i_3)] \\ &= (\frac{1}{J^2})^2 + (\frac{2}{J^2})^2 + \cdots + (\frac{J}{J^2})^2 + (\frac{J-1}{J^2})^2 + \cdots + (\frac{1}{J^2})^2 = \frac{2J^2+1}{3J^3}.\end{aligned} \tag{13}$$

Therefore

$$\begin{aligned}
\mathrm{S}_3 &= \frac{8J^2+4}{J^3} \sum_{\substack{i_1,i_2,i_3\\ j_2,j_3,i_1'}} \mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{i_1,j_2,j_3}\mathcal{M}_{i_1',i_2,i_3}\mathcal{N}_{i_1',j_2,j_3} \\
&= \frac{8J^2+4}{J^3} \sum_{\substack{i_2,i_3\\ j_2,j_3}} \langle \mathcal{M}(\mathbf{I},\mathbf{e}_{i_2},\mathbf{e}_{i_3}), \mathcal{N}(\mathbf{I},\mathbf{e}_{j_2},\mathbf{e}_{j_3})\rangle^2 \\
&\leq \frac{8J^2+4}{J^3} \sum_{\substack{i_2,i_3\\ j_2,j_3}} \|\mathcal{M}(\mathbf{I},\mathbf{e}_{i_2},\mathbf{e}_{i_3})\|_{\mathrm{F}}^2 \|\mathcal{N}(\mathbf{I},\mathbf{e}_{j_2},\mathbf{e}_{j_3})\|_{\mathrm{F}}^2 \\
&= \frac{8J^2+4}{J^3} \|\mathcal{M}\|_{\mathrm{F}}^2 \|\mathcal{N}\|_{\mathrm{F}}^2.
\end{aligned} \tag{14}$$

Case 4 None of $\mathfrak{G}_1, \mathfrak{G}_2, \mathfrak{G}_3$ has best matches. Suppose $i_1 = i_1' \neq j_1 = j_1', i_2 = i_2' \neq j_2 = j_2', i_3 = i_3' \neq j_3 = j_3'$, then

$$\mathrm{S}_4 = \binom{2}{1}\binom{2}{1}\binom{2}{1} \sum_{\substack{i_1,i_2,i_3\\ j_1,j_2,j_3}} \mathrm{E}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3), \mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3))]\mathcal{M}_{i_1,i_2,i_3}^2\mathcal{N}_{j_1,j_2,j_3}^2. \tag{15}$$

Denote $\mathbf{h}_4(i_2,i_3) = \mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)$. From **Case 3**:

$$\mathrm{P}[\mathbf{h}_4(i_2,i_3)=t] = \begin{cases} \dfrac{t+1}{J^2}, t \leq J-1 \\ \dfrac{2J-1-t}{J^2}, t \geq J \end{cases}$$

Since $\mathrm{P}[\mathbf{h}_1(i_1)=t] = \frac{1}{J}$ for $t = 0, \cdots, J-1$, we have

$$\begin{aligned}
&\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=t] = \mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_4(i_2,i_3)=t] \\
&= \begin{cases} \displaystyle\sum_{k=0}^{t} \mathrm{P}[\mathbf{h}_4(i_2,i_3)=k, \mathbf{h}_1(i_1)=t-k] = \frac{(t+1)(t+2)}{2J^3}, & 0 \leq t \leq J-1 \\ \displaystyle\sum_{k=t-J+1}^{t} \mathrm{P}[\mathbf{h}_4(i_2,i_3)=k, \mathbf{h}_1(i_1)=t-k] = \frac{-2t^2+6(J-1)t+9J-3J^2-4}{2J^3}, & J \leq t \leq 2J-2 \\ \displaystyle\sum_{k=t-J+1}^{J-1} \mathrm{P}[\mathbf{h}_4(i_2,i_3)=k, \mathbf{h}_1(i_1)=t-k] = \frac{(3J-1-t)(3J-2-t)}{2J^3}, & 2J-1 \leq t \leq 3J-3 \end{cases}
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) = \mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)] \\
&= \frac{1}{(2J^3)^2}(\sum_{t=0}^{J-1}(t+1)^2(t+2)^2 + \sum_{t=J}^{2J-2}(-2t^2+6(J-1)t+9J-3J^2-4)^2 + \sum_{t=2J-1}^{3J-3}(3J-1-t)^2(3J-2-t)^2) \\
&= \frac{11J^4+5J^2+4}{20J^5}.
\end{aligned} \tag{16}$$

Hence,

$$\mathrm{S}_4 = 8 * \frac{11J^4+5J^2+4}{20J^5} \sum_{\substack{i_1,i_2,i_3\\ j_1,j_2,j_3}} \mathcal{M}_{i_1,i_2,i_3}^2\mathcal{N}_{j_1,j_2,j_3}^2 = \frac{22J^4+10J^2+8}{5J^5}\|\mathcal{M}\|_{\mathrm{F}}^2\|\mathcal{N}\|_{\mathrm{F}}^2. \tag{17}$$

As a result, we have

$$\begin{aligned}
\mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FCS}(\mathcal{M}), \mathrm{FCS}(\mathcal{N})\rangle^2] &= \mathrm{S}_1+\mathrm{S}_2+\mathrm{S}_3+\mathrm{S}_4 \\
&\leq \langle \mathcal{M},\mathcal{N}\rangle^2 + \frac{92J^4+30J^2+8}{5J^5}\|\mathcal{M}\|_{\mathrm{F}}^2\|\mathcal{N}\|_{\mathrm{F}}^2.
\end{aligned} \tag{18}$$

Therefore

$$\begin{aligned}
\mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FCS}(\mathcal{M}), \mathrm{FCS}(\mathcal{N})\rangle] &= \mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FCS}(\mathcal{M}), \mathrm{FCS}(\mathcal{N})\rangle^2] - \mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FCS}(\mathcal{M}), \mathrm{FCS}(\mathcal{N})\rangle]^2 \\
&\leq \frac{92J^4+30J^2+8}{5J^5}\|\mathcal{M}\|_{\mathrm{F}}^2\|\mathcal{N}\|_{\mathrm{F}}^2.
\end{aligned} \tag{19}$$

Combining (8, 19) the consistency is proved.

We further prove FCS computes a better estimator for tensor inner product than TS.

$$
\begin{aligned}
\langle \mathrm{TS}(\mathcal{M}), \mathrm{TS}(\mathcal{N}) \rangle^2 = \sum_{\substack{i_1,i_2,i_3 \\ j_1,j_2,j_3 \\ i'_1,i'_2,i'_3 \\ j'_1,j'_2,j'_3}} & \delta((\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)) \bmod J, (\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)) \bmod J) \\
& \delta((\mathbf{h}_1(i'_1)+\mathbf{h}_2(i'_2)+\mathbf{h}_3(i'_3)) \bmod J, (\mathbf{h}_1(j'_1)+\mathbf{h}_2(j'_2)+\mathbf{h}_3(j'_3)) \bmod J) \\
& \mathbf{s}_1(i_1)\mathbf{s}_2(i_2)\mathbf{s}_3(i_3)\mathbf{s}_1(j_1)\mathbf{s}_2(j_2)\mathbf{s}_3(j_3)\mathbf{s}_1(i'_1)\mathbf{s}_2(i'_2)\mathbf{s}_3(i'_3)\mathbf{s}_1(j'_1)\mathbf{s}_2(j'_2)\mathbf{s}_3(j'_3) \\
& \mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{j_1,j_2,j_3}\mathcal{M}_{i'_1,i'_2,i'_3}\mathcal{N}_{j'_1,j'_2,j'_3}.
\end{aligned} \tag{20}
$$

Define $\mathfrak{G}_1, \mathfrak{G}_2, \mathfrak{G}_3$ similarly as mentioned before. Then (20) can be grouped into following cases:

Case 1 $\mathfrak{G}_1, \mathfrak{G}_2, \mathfrak{G}_3$ all have best matches, then

$$
\mathrm{S}_1 = \sum_{\substack{i_1,i_2,i_3 \\ i'_1,i'_2,i'_3}} \mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{i_1,i_2,i_3}\mathcal{M}_{i'_1,i'_2,i'_3}\mathcal{N}_{i'_1,i'_2,i'_3} = \langle \mathcal{M}, \mathcal{N} \rangle^2 . \tag{21}
$$

Case 2 One out of $\mathfrak{G}_1, \mathfrak{G}_2, \mathfrak{G}_3$ does not have best match, take $\mathfrak{G}_1$ for example. Suppose $i_1 = i'_1 \neq j_1 = j'_1$, then

$$
\begin{aligned}
\mathrm{S}_2 = \binom{3}{1}\binom{2}{1} \sum_{\substack{i_1,i_2,i_3 \\ j_1,i'_2,i'_3}} & \mathrm{E}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J) \\
\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i'_2)+\mathbf{h}_3(i'_3) \bmod J, & \mathbf{h}_1(j_1)+\mathbf{h}_2(i'_2)+\mathbf{h}_3(i'_3) \bmod J)]\mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{j_1,i_2,i_3}\mathcal{M}_{i_1,i'_2,i'_3}\mathcal{N}_{j_1,i'_2,i'_3}.
\end{aligned} \tag{22}
$$

Obviously, $\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J) = 1$ only if $\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J = \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J$. Denote

$$
\begin{aligned}
\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) &:= k_1 J + m \\
\mathbf{h}_1(j_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) &:= k_2 J + m,
\end{aligned} \tag{23}
$$

we have

$$
\mathbf{h}_1(i_1) - \mathbf{h}_1(j_1) = (k_1 - k_2)J := kJ. \tag{24}
$$

Given $\mathbf{h}_1(i_1), \mathbf{h}_1(j_1) \in [J]$, (24) holds iff $k = 0$, i.e. $\mathbf{h}_1(i_1) = \mathbf{h}_1(j_1)$. On the other hand, when $\mathbf{h}_1(i_1) = \mathbf{h}_1(j_1)$, obviously $\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i'_2)+\mathbf{h}_3(i'_3) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i'_2)+\mathbf{h}_3(i'_3) \bmod J) = 1$. Hence

$$
\begin{aligned}
& \mathrm{E}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J) \\
& \delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i'_2)+\mathbf{h}_3(i'_3) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i'_2)+\mathbf{h}_3(i'_3) \bmod J)] \\
& = \mathrm{P}[\mathbf{h}_1(i_1) = \mathbf{h}_1(j_1)] = \frac{1}{J}.
\end{aligned} \tag{25}
$$

Therefore

$$
\mathrm{S}_2 = \binom{3}{1}\binom{2}{1}\frac{1}{J} \sum_{\substack{i_1,i_2,i_3 \\ j_1,i'_2,i'_3}} \mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{j_1,i_2,i_3}\mathcal{M}_{i_1,i'_2,i'_3}\mathcal{N}_{j_1,i'_2,i'_3} \leq \frac{6}{J}\|\mathcal{M}\|_{\mathrm{F}}^2\|\mathcal{N}\|_{\mathrm{F}}^2. \tag{26}
$$

Case 3 One out of $\mathfrak{G}_1, \mathfrak{G}_2, \mathfrak{G}_3$ has best match, take $\mathfrak{G}_1$ for example. Suppose $i_2 = i'_2 \neq j_2 = j'_2, i_3 = i'_3 \neq j_3 = j'_3$, then

$$
\begin{aligned}
\mathrm{S}_3 = \binom{3}{1}\binom{2}{1}\binom{2}{1} \sum_{\substack{i_1,i_2,i_3 \\ j_2,j_3,i'_1}} & \mathrm{E}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J, \mathbf{h}_1(i_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3) \bmod J) \\
\delta(\mathbf{h}_1(i'_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) \bmod J, & \mathbf{h}_1(i'_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3) \bmod J)]\mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{i_1,j_2,j_3}\mathcal{M}_{i'_1,i_2,i_3}\mathcal{N}_{i'_1,j_2,j_3}
\end{aligned}
$$

Denote

$$
\begin{aligned}
\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) &:= k_1 J + m \\
\mathbf{h}_1(i_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3) &:= k_2 J + m,
\end{aligned} \tag{27}
$$

we have

$$
\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)-\mathbf{h}_2(j_2)-\mathbf{h}_3(j_3) = (k_1-k_2)J := kJ. \tag{28}
$$

Similar to (24), we have $k = 0, \pm 1$:

(a) When $k = 0$, we have $\mathrm{P}[\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3) = \mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)] = \frac{2J^2+1}{3J^3}$ from (13).

(b) When $k=\pm 1$, we have $\mathrm{P}[\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)\pm J]=\sum_{t=J}^{2J-2}\mathrm{P}[\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=t]\,\mathrm{P}[\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)=t-J]=\sum_{t=1}^{J-1}\frac{t(J-t)}{J^4}=\frac{J^2-1}{6J^3}$.

Therefore, we have

$$\begin{aligned}\mathrm{S}_3&=\binom{3}{1}\binom{2}{1}\binom{2}{1}\left(\frac{2J^2+1}{3J^3}+2*\frac{J^2-1}{6J^3}\right)\sum_{\substack{i_1,i_2,i_3\\ j_2,j_3,i_1'}}\mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{i_1,j_2,j_3}\mathcal{M}_{i_1',i_2,i_3}\mathcal{N}_{i_1',j_2,j_3}\\&=\frac{12}{J}\sum_{\substack{i_1,i_2,i_3\\ j_2,j_3,i_1'}}\mathcal{M}_{i_1,i_2,i_3}\mathcal{N}_{i_1,j_2,j_3}\mathcal{M}_{i_1',i_2,i_3}\mathcal{N}_{i_1',j_2,j_3}\\&\le\frac{12}{J}\|\mathcal{M}\|_{\mathrm{F}}^2\|\mathcal{N}\|_{\mathrm{F}}^2.\end{aligned}\tag{29}$$

Case 4 None of $\mathfrak{G}_1,\mathfrak{G}_2,\mathfrak{G}_3$ has best matches. Suppose $i_1=i_1'\neq j_1=j_1', i_2=i_2'\neq j_2=j_2', i_3=i_3'\neq j_3=j_3'$, then

$$\mathrm{S}_4=\binom{2}{1}\binom{2}{1}\binom{2}{1}\sum_{\substack{i_1,i_2,i_3\\ j_1,j_2,j_3}}\mathrm{E}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)\bmod J,\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)\bmod J)]\mathcal{M}_{i_1,i_2,i_3}^2\mathcal{N}_{j_1,j_2,j_3}^2.\tag{30}$$

Denote

$$\begin{aligned}\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)&:=k_1J+m\\ \mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)&:=k_2J+m,\end{aligned}\tag{31}$$

we have

$$\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3):=\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)+kJ,\tag{32}$$

which breaks into following cases:

(a) When $k=0$, from (16): $\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)]=\frac{11J^4+5J^2+4}{20J^5}$.

(b) When $k=\pm 1$, we have

$$\begin{aligned}&\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)\pm J]\\&=\sum_{t=J}^{2J-2}\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=t,\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)=t-J]+\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=2J-1,\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)=J-1]\\&+\sum_{t=2J}^{3J-3}\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=t,\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)=t-J]\\&=\frac{13J^4-5J^2-8}{60J^5}.\end{aligned}\tag{33}$$

(c) When $k=\pm 2$, we have

$$\begin{aligned}&\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)\pm 2J]\\&=\sum_{t=2J}^{3J-3}\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)+\mathbf{h}_3(i_3)=t,\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)+\mathbf{h}_3(j_3)=t-2J]\\&=\frac{J^4-5J^2+4}{120J^5}.\end{aligned}\tag{34}$$

Therefore,

$$\begin{aligned}\mathrm{S}_4&=8*\left(\frac{11J^4+5J^2+4}{20J^5}+2*\frac{13J^4-5J^2-8}{60J^5}+2*\frac{J^4-5J^2+4}{120J^5}\right)\sum_{\substack{i_1,i_2,i_3\\ j_1,j_2,j_3}}\mathcal{M}_{i_1,i_2,i_3}^2\mathcal{N}_{j_1,j_2,j_3}^2\\&=\frac{8}{J}\sum_{\substack{i_1,i_2,i_3\\ j_1,j_2,j_3}}\mathcal{M}_{i_1,i_2,i_3}^2\mathcal{N}_{j_1,j_2,j_3}^2\le\frac{8}{J}\|\mathcal{M}\|_{\mathrm{F}}^2\|\mathcal{N}\|_{\mathrm{F}}^2.\end{aligned}\tag{35}$$

As a result, we have

$$\mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle\mathrm{TS}(\mathcal{M}),\mathrm{TS}(\mathcal{N})\rangle^2]=\mathrm{S}_1+\mathrm{S}_2+\mathrm{S}_3+\mathrm{S}_4\le\langle\mathcal{M},\mathcal{N}\rangle^2+\frac{26}{J}\|\mathcal{M}\|_{\mathrm{F}}^2\|\mathcal{N}\|_{\mathrm{F}}^2.\tag{36}$$

Therefore

$$\mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle\mathrm{TS}(\mathcal{M}),\mathrm{TS}(\mathcal{N})\rangle]=\mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle\mathrm{TS}(\mathcal{M}),\mathrm{TS}(\mathcal{N})\rangle^2]-\mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle\mathrm{TS}(\mathcal{M}),\mathrm{TS}(\mathcal{N})\rangle]^2\le\frac{26}{J}\|\mathcal{M}\|_{\mathrm{F}}^2\|\mathcal{N}\|_{\mathrm{F}}^2.\tag{37}$$

Comparing (19) and (37), clearly we have $\mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FCS}(\mathcal{M}), \mathrm{FCS}(\mathcal{N})\rangle] \leq \mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{TS}(\mathcal{M}), \mathrm{TS}(\mathcal{N})\rangle]$, which means FCS provides a more accurate estimator than TS for tensor inner product, especially when J is small.

B. *Proof for Corollary 1*

(19) can be represented as $\mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FCS}(\mathcal{M}), \mathrm{FCS}(\mathcal{N})\rangle] = \mathrm{O}(\frac{\|\mathcal{M}\|_{\mathrm{F}}^2\|\mathcal{N}\|_{\mathrm{F}}^2}{J})$. By substituting $\mathcal{M} = \mathcal{T}$, $\mathcal{N} = \mathbf{u} \circ \mathbf{u} \circ \mathbf{u}$, and $\mathcal{M} = \mathcal{T}$, $\mathcal{N} = \mathbf{e}_i \circ \mathbf{u} \circ \mathbf{u}$, respectively, from Chebychev's inequality we prove the corollary immediately, since $\mathbf{u}$ is unit vector and hence both $\mathbf{u} \circ \mathbf{u} \circ \mathbf{u}$ and $\mathbf{e}_i \circ \mathbf{u} \circ \mathbf{u}$ equal 1.