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 $\mathcal{T} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \cdots \circ \mathbf{u}^{(N)} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, assign $l = \sum_{n=1}^{N}(i_n - 1)\prod_{j=1}^{n-1} I_j + 1$, we have $\mathrm{vec}(\mathcal{T})_l = \mathbf{u}^{(1)}_{i_1} \cdots \mathbf{u}^{(N)}_{i_N}$. By definition of HCS, we have:

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

where $\mathcal{P}_{\mathrm{CS}_n(\mathbf{u}^{(n)})}(\mathrm{w})$ is the polynomial form of $\mathrm{CS}_n(\mathbf{u}^{(n)})$ 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

$$\begin{aligned}
&\sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} \mathrm{vec}(\mathcal{T})_l \mathbf{s}_1(i_1) \cdots \mathbf{s}_N(i_N) \mathrm{w}^{\mathbf{h}_1(i_1)+\cdots+\mathbf{h}_N(i_N)-N} \\
&= \sum_{l} \mathrm{vec}(\mathcal{T})_l \mathbf{s}_{N+1}(l) \mathrm{w}^{\mathbf{h}_{N+1}(l)-1} \\
&= \mathcal{P}_{\mathrm{HCS}(\mathrm{vec}(\mathcal{T}))}(\mathrm{w}) \\
&= \mathcal{P}_{\mathrm{FHCS}(\mathcal{T})}(\mathrm{w}).
\end{aligned} \tag{2}$$

Therefore, we have

$$\mathcal{P}_{\mathrm{FHCS}(\mathcal{T})}(\mathrm{w}) = \mathcal{P}_{\mathrm{CS}_1(\mathbf{u}^{(1)})}(\mathrm{w}) \cdots \mathcal{P}_{\mathrm{CS}_N(\mathbf{u}^{(N)})}(\mathrm{w}). \tag{3}$$

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

$$\begin{aligned}
\mathrm{FHCS}(\mathcal{T}) &= \mathrm{CS}_1(\mathbf{u}^{(1)}) \circledast \cdots \circledast \mathrm{CS}_N(\mathbf{u}^{(N)}) \\
&= \mathbb{F}^{-1}(\mathbb{F}(\mathrm{CS}_1(\mathbf{u}^{(1)})) * \cdots * \mathbb{F}(\mathrm{CS}_N(\mathbf{u}^{(N)}))),
\end{aligned} \tag{4}$$

which completes the proof.

II. Proofs for Proposition 1 and Corollary 1

A. Proof for Proposition 1

First we prove the consistency of FHCS. We proceed by proving that for any 2-order[1] tensors $\mathbf{M}$, $\mathbf{N}$ with the same size, $\langle \mathrm{FHCS}(\mathbf{M}), \mathrm{FHCS}(\mathbf{N}) \rangle$ is a consistent estimator of $\langle \mathbf{M}, \mathbf{N} \rangle$. To this end we prove it is unbiased with bounded variance.

1) Unbiasedness:

$$\begin{aligned}
\langle \mathrm{FHCS}(\mathbf{M}), \mathrm{FHCS}(\mathbf{N}) \rangle &= \sum_t \mathrm{FHCS}(\mathbf{M})_t \mathrm{FHCS}(\mathbf{N})_t \\
&= \sum_t \Big(\sum_{\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2)=t} \mathbf{s}_1(i_1)\mathbf{s}_2(i_2)\mathbf{M}_{i_1,i_2} \Big) \Big(\sum_{\mathbf{h}_1(j_1)+\mathbf{h}_2(j_2)=t} \mathbf{s}_1(j_1)\mathbf{s}_2(j_2)\mathbf{N}_{j_1,j_2} \Big),
\end{aligned} \tag{5}$$

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

[1]For simplicity we focus on 2-order tensors, i.e. matrices. The conclusion can be easily generalized to tensors with any order.

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

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

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

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

2) Variance boundedness:

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

Denote $\mathfrak{G}_1 := \{i_1, j_1, i'_1, j'_1\}, \mathfrak{G}_2 := \{i_2, j_2, i'_2, j'_2\}$, summation variables $\mathrm{S}_1, \mathrm{S}_2, \mathrm{S}_3$. We say $\mathfrak{G}_k$ has *best match* if $i_k = j_k, i'_k = j'_k$ for $k \in [2]$. then (9) can be grouped into the following cases:

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

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

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

$$\begin{aligned} \mathrm{S}_2 &= \binom{2}{1}\binom{2}{1} \sum_{\substack{i_1,i_2 \\ j_1,i'_2}} \mathrm{E}[\delta(\mathbf{h}_1(i_1), \mathbf{h}_1(j_1))^2] \mathbf{M}_{i_1,i_2}\mathbf{N}_{j_1,i_2}\mathbf{M}_{i_1,i'_2}\mathbf{N}_{j_1,i'_2} \\ &= \frac{4}{J} \sum_{\substack{i_1,i_2 \\ j_1,i'_2}} \mathbf{M}_{i_1,i_2}\mathbf{N}_{j_1,i_2}\mathbf{M}_{i_1,i'_2}\mathbf{N}_{j_1,i'_2} \\ &= \frac{4}{J} \sum_{i_1,j_1} \langle \mathbf{M}(\mathbf{e}_{i_1}, \mathbf{I}), \mathbf{N}(\mathbf{e}_{j_1}, \mathbf{I}) \rangle^2 \\ &\leq \frac{4}{J} \sum_{i_1,j_1} \|\mathbf{M}(\mathbf{e}_{i_1}, \mathbf{I})\|_{\mathrm{F}}^2 \|\mathbf{N}(\mathbf{e}_{j_1}, \mathbf{I})\|_{\mathrm{F}}^2 \\ &= \frac{4}{J} \|\mathbf{M}\|_{\mathrm{F}}^2 \|\mathbf{N}\|_{\mathrm{F}}^2 \end{aligned} \tag{11}$$

holds due to the Cauchy-Schwartz inequality.

Case 3 Neither of $\mathfrak{G}_1, \mathfrak{G}_2$ has best matches. Suppose $i_1 = i'_1 \neq j_1 = j'_1, i_2 = i'_2 \neq j_2 = j'_2$, then

$$\mathrm{S}_3 = \binom{2}{1}\binom{2}{1} \sum_{\substack{i_1,i_2 \\ j_1,j_2}} \mathrm{E}[\delta(\mathbf{h}_1(i_1) + \mathbf{h}_2(i_2), \mathbf{h}_1(j_1) + \mathbf{h}_2(j_2))] \mathbf{M}_{i_1,i_2}^2 \mathbf{N}_{j_1,j_2}^2 \tag{12}$$

From 2-wise independence of $\mathbf{h}_1$ and $\mathbf{h}_2$, we have

$$\mathrm{P}[\mathbf{h}_1(i_1) + \mathbf{h}_2(i_2) = t] = \begin{cases} \displaystyle\sum_{k=0}^{t} \mathrm{P}[\mathbf{h}_1(i_1) = k, \mathbf{h}_2(i_2) = t - k] = \frac{t+1}{J^2}, & t \leq J - 1 \\ \displaystyle\sum_{k=t-J+1}^{J-1} \mathrm{P}[\mathbf{h}_1(i_1) = k, \mathbf{h}_2(i_2) = t - k] = \frac{2J - 1 - t}{J^2}, & t \geq J \end{cases}$$

Therefore

$$\mathrm{P}[\delta(\mathbf{h}_1(i_1) + \mathbf{h}_2(i_2), \mathbf{h}_1(j_1) + \mathbf{h}_2(j_2) = 1] = (\frac{1}{J^2})^2 + (\frac{2}{J^2})^2 + \cdots + (\frac{J}{J^2})^2 + (\frac{J-1}{J^2})^2 + \cdots + (\frac{1}{J^2})^2 = \frac{2J^2+1}{3J^3}.$$

Then we have

$$\mathrm{S}_3 = \binom{2}{1}\binom{2}{1} \frac{2J^2+1}{3J^3} \sum_{\substack{i_1,i_2 \\ j_1,j_2}} \mathbf{M}_{i_1,i_2}^2 \mathbf{N}_{j_1,j_2}^2 = \frac{4(2J^2+1)}{3J^3} \|\mathbf{M}\|_{\mathrm{F}}^2 \|\mathbf{N}\|_{\mathrm{F}}^2. \tag{13}$$

As a result, we have

$$\mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FHCS}(\mathbf{M}), \mathrm{FHCS}(\mathbf{N})\rangle^2] = \mathrm{S}_1 + \mathrm{S}_2 + \mathrm{S}_3 \leq \langle \mathbf{M}, \mathbf{N}\rangle^2 + \frac{20J^2+4}{3J^3}\|\mathbf{M}\|_{\mathrm{F}}^2\|\mathbf{N}\|_{\mathrm{F}}^2. \tag{14}$$

Therefore

$$\begin{aligned}\mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FHCS}(\mathbf{M}), \mathrm{FHCS}(\mathbf{N})\rangle] &= \mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FHCS}(\mathbf{M}), \mathrm{FHCS}(\mathbf{N})\rangle^2] - \mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FHCS}(\mathbf{M}), \mathrm{FHCS}(\mathbf{N})\rangle]^2 \\ &\leq \frac{20J^2+4}{3J^3}\|\mathbf{M}\|_{\mathrm{F}}^2\|\mathbf{N}\|_{\mathrm{F}}^2.\end{aligned} \tag{15}$$

Combining (8, 15) the consistency is proved.

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

$$\begin{aligned}\langle \mathrm{TS}(\mathbf{M}), \mathrm{TS}(\mathbf{N})\rangle^2 = &\sum_{\substack{i_1,i_2\\ j_1,j_2\\ i_1',i_2'\\ j_1',j_2'}} \delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(j_2) \bmod J)\delta(\mathbf{h}_1(i_1')+\mathbf{h}_2(i_2') \bmod J, \mathbf{h}_1(j_1')+\mathbf{h}_2(j_2') \bmod J) \\ &\mathbf{s}_1(i_1)\mathbf{s}_2(i_2)\mathbf{s}_1(j_1)\mathbf{s}_2(j_2)\mathbf{s}_1(i_1')\mathbf{s}_2(i_2')\mathbf{s}_1(j_1')\mathbf{s}_2(j_2')\mathbf{M}_{i_1,i_2}\mathbf{N}_{j_1,j_2}\mathbf{M}_{i_1',i_2'}\mathbf{N}_{j_1',j_2'}.\end{aligned} \tag{16}$$

Define $\mathfrak{G}_1, \mathfrak{G}_2$ similarly as mentioned before. Then (16) can be grouped into following cases:

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

$$\mathrm{S}_1 = \sum_{\substack{i_1,i_2\\ i_1',i_2'}} \mathbf{M}_{i_1,i_2}\mathbf{N}_{i_1,i_2}\mathbf{M}_{i_1',i_2'}\mathbf{N}_{i_1',i_2'} = \langle \mathbf{M}, \mathbf{N}\rangle^2. \tag{17}$$

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

$$\begin{aligned}\mathrm{S}_2 = &\binom{2}{1}\binom{2}{1}\sum_{\substack{i_1,i_2\\ j_1,i_2'}} \mathrm{E}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2) \bmod J)\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2') \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2') \bmod J)] \\ &\mathbf{M}_{i_1,i_2}\mathbf{N}_{j_1,i_2}\mathbf{M}_{i_1,i_2'}\mathbf{N}_{j_1,i_2'}.\end{aligned} \tag{18}$$

Obviously $\mathrm{E}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2) \bmod J)] = 1$ iff $\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J = \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2) \bmod J$. Denote

$$\begin{aligned}\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J &:= k_1J + m \\ \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2) \bmod J &:= k_2J + m,\end{aligned} \tag{19}$$

we have

$$\mathbf{h}_1(i_1) - \mathbf{h}_1(j_1) = (k_1-k_2)J := kJ. \tag{20}$$

Given $\mathbf{h}_1(i_1), \mathbf{h}_1(j_1) \in [J]$, (20) holds iff $k = 0$, i.e. $\mathbf{h}_1(i_1) = \mathbf{h}_1(j_1)$. Hence

$$\begin{aligned}&\mathrm{E}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2) \bmod J)\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2') \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(i_2') \bmod J)] = 1 \\ &= \mathrm{P}[\mathbf{h}_1(i_1) = \mathbf{h}_1(j_1)] = \frac{1}{J}.\end{aligned} \tag{21}$$

Therefore

$$\mathrm{S}_2 = \binom{2}{1}\binom{2}{1}\frac{1}{J}\sum_{\substack{i_1,i_2\\ j_1,i_2'}} \mathbf{M}_{i_1,i_2}\mathbf{N}_{j_1,i_2}\mathbf{M}_{i_1,i_2'}\mathbf{N}_{j_1,i_2'} \leq \frac{4}{J}\|\mathbf{M}\|_{\mathrm{F}}^2\|\mathbf{N}\|_{\mathrm{F}}^2. \tag{22}$$

Case 3 Neither of $\mathfrak{G}_1, \mathfrak{G}_2$ has best matches. Suppose $i_1 = i_1' \neq j_1 = j_1', i_2 = i_2' \neq j_2 = j_2'$, then

$$\mathrm{S}_3 = \binom{2}{1}\binom{2}{1}\sum_{\substack{i_1,i_2\\ j_1,j_2}} \mathrm{E}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(j_2) \bmod J)]\mathbf{M}_{i_1,i_2}^2\mathbf{N}_{j_1,j_2}^2. \tag{23}$$

From 2-wise independence of $\mathbf{h}_1$ and $\mathbf{h}_2$, we have

$$\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J = t] = \begin{cases}\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) = t] + \mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) = t+J] = \frac{1}{J}, & t < J-1 \\ \mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) = t] = \frac{1}{J}, & t = J-1\end{cases}$$

i.e. $\mathrm{P}[\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J = t] = \frac{1}{J}$ for $t = 0, \cdots, J-1$. Hence

$$\mathrm{P}[\delta(\mathbf{h}_1(i_1)+\mathbf{h}_2(i_2) \bmod J, \mathbf{h}_1(j_1)+\mathbf{h}_2(j_2) \bmod J) = 1] = (\frac{1}{J})^2 J = \frac{1}{J}.$$

Therefore

$$\mathrm{S}_3 = \binom{2}{1}\binom{2}{1}\frac{1}{J}\sum_{\substack{i_1,i_2\\j_1,j_2}} \mathbf{M}^2_{i_1,i_2}\mathbf{N}^2_{j_1,j_2} = \frac{4}{J}\|\mathbf{M}\|_\mathrm{F}^2\|\mathbf{N}\|_\mathrm{F}^2. \tag{24}$$

As a result, we have

$$\mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{TS}(\mathbf{M}), \mathrm{TS}(\mathbf{N})\rangle^2] = \mathrm{S}_1 + \mathrm{S}_2 + \mathrm{S}_3 \leq \langle \mathbf{M}, \mathbf{N}\rangle^2 + \frac{8}{J}\|\mathbf{M}\|_\mathrm{F}^2\|\mathbf{N}\|_\mathrm{F}^2. \tag{25}$$

Therefore

$$\mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{TS}(\mathbf{M}), \mathrm{TS}(\mathbf{N})\rangle] = \mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{TS}(\mathbf{M}), \mathrm{TS}(\mathbf{N})\rangle^2] - \mathrm{E}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{TS}(\mathbf{M}), \mathrm{TS}(\mathbf{N})\rangle]^2 \leq \frac{8}{J}\|\mathbf{M}\|_\mathrm{F}^2\|\mathbf{N}\|_\mathrm{F}^2. \tag{26}$$

Since the Hash length $J \geq 1$, clearly we have $\mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FHCS}(\mathbf{M}), \mathrm{FHCS}(\mathbf{N})\rangle] \leq \mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{TS}(\mathbf{M}), \mathrm{TS}(\mathbf{N})\rangle]$, which means FHCS provides a more accurate estimator than TS for tensor inner product, especially when J is small. It can be proved in a similar way that as the tensor order increases, this inequality becomes more pronounced.

B. Proof for Corollary 1

(15) can be represented as $\mathrm{V}_{\mathbf{h},\mathbf{s}}[\langle \mathrm{FHCS}(\mathbf{M}), \mathrm{FHCS}(\mathbf{N})\rangle] = \mathrm{O}(\frac{\|\mathbf{M}\|_\mathrm{F}^2\|\mathbf{N}\|_\mathrm{F}^2}{J})$. By substituting $\mathbf{M} = \mathbf{T}$, $\mathbf{N} = \mathbf{u} \circ \mathbf{u}$, and $\mathbf{M} = \mathbf{T}$, $\mathbf{N} = \mathbf{e}_i \circ \mathbf{u}$, respectively, from Chebychev's inequality we prove the corollary immediately since $\mathbf{u}$ is a unit vector and hence both $\mathbf{u} \circ \mathbf{u}$ and $\mathbf{e}_i \circ \mathbf{u}$ equals 1. The conclusion also generalizes to tensors of any order.