The authors assumed that the equations given only in Section 4.1 need to be discussed in detail, since they demonstrate and justify the main contribution and are crucial for understanding the proposed method. Due to the compact representation of the Kruskal convolution introduced in eqs. (5) and (6), we were able to considerably simplify the basic governing equations in Section 4.1. and present them in a very elegant way. This form of presentation is based on our original concept, and in our opinion, it is a better way of presentation than using direct element-based operations. Hence, we still prefer our original way of presenting these equations, but to dissipate any doubts we also provide the standard (entry-based) presentation of the same mathematical operations. For the sake of completeness, we demonstrate this alternative way of presentation below.

In this approach, we present the mathematical formula from Section 4.1 using direct operations onto matrix and tensor entries. We use the following notation: $x_{i,j}$ denotes the entry of matrix $\boldsymbol{X} \in \mathbb{R}^{I \times J}$ in its i-th row and the j-th column. Equivalently, matrix $\boldsymbol{X}$ composed of all entries $x_{i,j}$ is expressed by the form: $\boldsymbol{X} = [x_{i,j}]$. The same style of notation applies to tensors.

Let $\mathcal{W} = [w_{t,c,d_1,d_2}] \in \mathbb{R}^{T \times C \times D_1 \times D_2}$ be the kernel weight tensor, and $\mathcal{X} = [x_{i_1,i_2,c}] \in \mathbb{R}^{I_1 \times I_2 \times C}$ be the input activation tensor. After applying the mode-permutation operation to $\mathcal{W}$ (as shown in (9) in the manuscript), we get tensor $\tilde{\mathcal{W}} = [\tilde{w}_{c,d_1,t,d_2}] \in \mathbb{R}^{C \times D_1 \times T \times D_2}$. Thus, the second mode in $\mathcal{W}$ became the first mode in $\tilde{\mathcal{W}}$, and the first mode from $\mathcal{W}$ shifted to the third mode.

The Kruskal convolution given in (11) can be equivalently expressed in the following element-based form:

$$y_{\tilde{i}_1,\tilde{i}_2,t} = \sum_{c=1}^{C} \sum_{d_1=1}^{D_1} \sum_{d_2=1}^{D_2} \tilde{w}_{c,d_1,t,d_2} x_{i_1(d_1),i_2(d_2),c}. \tag{1}$$

The HT-2 model given in (10) can also be expressed in the similar form:

$$\tilde{w}_{c,d_1,t,d_2} = \sum_{r_1=1}^{R_1} \sum_{r_3=1}^{R_3} \sum_{r_{13}=1}^{R_{13}} \tilde{g}^{(1)}_{r_1,d_1,r_{13}} \tilde{g}^{(2)}_{r_3,d_2,r_{13}} u^{(1)}_{c,r_1} u^{(3)}_{t,r_3}, \tag{2}$$

where

$$\begin{aligned} \tilde{\mathcal{G}}^{(1)} &= [\tilde{g}^{(1)}_{r_1,d_1,r_{13}}] \in \mathbb{R}^{R_1 \times D_1 \times R_{13}}, \\ \tilde{\mathcal{G}}^{(2)} &= [\tilde{g}^{(2)}_{r_3,d_2,r_{13}}] \in \mathbb{R}^{R_3 \times D_2 \times R_{13}}, \\ \boldsymbol{U}^{(1)} &= [u^{(1)}_{c,r_1}] \in \mathbb{R}^{C \times R_1}, \end{aligned}$$

and

$$\boldsymbol{U}^{(3)} = [u^{(3)}_{t,r_3}] \in \mathbb{R}^{T \times R_3}.$$

Using the scalar multiplication and addition property, and inserting model (2) to model (1), we have:

$$\begin{aligned} y_{\tilde{i}_1,\tilde{i}_2,t} &= \sum_{c=1}^{C} \sum_{d_1=1}^{D_1} \sum_{d_2=1}^{D_2} \left(\sum_{r_1=1}^{R_1} \sum_{r_3=1}^{R_3} \sum_{r_{13}=1}^{R_{13}} \tilde{g}^{(1)}_{r_1,d_1,r_{13}} \tilde{g}^{(2)}_{r_3,d_2,r_{13}} u^{(1)}_{c,r_1} u^{(3)}_{t,r_3} \right) \\ &\times x_{i_1(d_1),i_2(d_2),c}. \end{aligned} \tag{3}$$

The summations in (4) can be reordered to the following form:

$$\begin{aligned} y_{\tilde{i}_1,\tilde{i}_2,t} &= \sum_{r_3=1}^{R_3} u_{t,r_3}^{(3)} \left[\sum_{d_1=1}^{D_1} \sum_{d_2=1}^{D_2} \sum_{r_1=1}^{R_1} \sum_{r_{13}=1}^{R_{13}} \tilde{g}_{r_1,d_1,r_{13}}^{(1)} \tilde{g}_{r_3,d_2,r_{13}}^{(2)} \right. \\ &\times \left. \left(\sum_{c=1}^{C} x_{i_1(d_1),i_2(d_2),c} u_{c,r_1}^{(1)} \right) \right]. \end{aligned} \tag{4}$$

Note that eq. (4) is the same as the second equation in (12). Computing first the last factor in (4), we have:

$$z_{i_1(d_1),i_2(d_2),r_1} = \sum_{c=1}^{C} x_{i_1(d_1),i_2(d_2),c} u_{c,r_1}^{(1)}, \tag{5}$$

and eq. (5) is equivalent to (13), assuming $\mathcal{Z} = [z_{i_1(d_1),i_2(d_2),r_1}]$. Considering (5), eq. (4) simplifies to:

$$y_{\tilde{i}_1,\tilde{i}_2,t} = \sum_{r_3=1}^{R_3} u_{t,r_3}^{(3)} \left[\sum_{d_1=1}^{D_1} \sum_{d_2=1}^{D_2} \sum_{r_1=1}^{R_1} \sum_{r_{13}=1}^{R_{13}} \tilde{g}_{r_1,d_1,r_{13}}^{(1)} \tilde{g}_{r_3,d_2,r_{13}}^{(2)} z_{i_1(d_1),i_2(d_2),r_1} \right]. \tag{6}$$

Then, we can also perform the next reordering of the summations in (6) to the following form:

$$y_{\tilde{i}_1,\tilde{i}_2,t} = \sum_{r_3=1}^{R_3} u_{t,r_3}^{(3)} \left[\sum_{r_{13}=1}^{R_{13}} \sum_{d_2=1}^{D_2} \tilde{g}_{r_3,d_2,r_{13}}^{(2)} \left(\sum_{r_1=1}^{R_1} \sum_{d_1=1}^{D_1} \tilde{g}_{r_1,d_1,r_{13}}^{(1)} z_{i_1(d_1),i_2(d_2),r_1} \right) \right]. \tag{7}$$

Note that the expression given in the parenthesis in (7) expresses the 1D Kruskal convolution, and executing it first, we have:

$$z_{\tilde{i}_1,i_2(d_2),r_{13}}^{(V)} = \sum_{r_1=1}^{R_1} \sum_{d_1=1}^{D_1} \tilde{g}_{r_1,d_1,r_{13}}^{(1)} z_{i_1(d_1),i_2(d_2),r_1}. \tag{8}$$

Eq. (8) is equivalent to (14) in the manuscript, where $\mathcal{Z}^{(V)} = [z_{\tilde{i}_1,i_2(d_2),r_{13}}^{(V)}]$. Thus, eq. (7) can be rewritten to the form:

$$y_{\tilde{i}_1,\tilde{i}_2,t} = \sum_{r_3=1}^{R_3} u_{t,r_3}^{(3)} \left[\sum_{r_{13}=1}^{R_{13}} \sum_{d_2=1}^{D_2} \tilde{g}_{r_3,d_2,r_{13}}^{(2)} z_{\tilde{i}_1,i_2(d_2),r_{13}}^{(V)} \right]. \tag{9}$$

Next, applying the similar approach as above, the expression in the bracket in (9) can be presented by:

$$z_{\tilde{i}_1,\tilde{i}_2,r_3}^{(V,H)} = \sum_{r_{13}=1}^{R_{13}} \sum_{d_2=1}^{D_2} \tilde{g}_{r_3,d_2,r_{13}}^{(2)} z_{\tilde{i}_1,i_2(d_2),r_{13}}^{(V)}. \tag{10}$$

Note that eq. (10) is equivalent to (15) in the manuscript, where $\mathcal{Z}^{(V,H)} = [z^{(V,H)}_{\tilde{i}_1,\tilde{i}_2,r_3}]$. Thus, eq. (9) takes the form:

$$y_{\tilde{i}_1,\tilde{i}_2,t} = \sum_{r_3=1}^{R_3} z^{(V,H)}_{\tilde{i}_1,\tilde{i}_2,r_3} u^{(3)}_{t,r_3}, \tag{11}$$

and this is an equivalent form to the last equation in (12) in the manuscript.