Supplementary Materials

We supplement the proof of Lemma 4.1 in Section 4 of the main paper.

**Lemma 4.1.** If we can arrange $d$ outputs into a $Q$-mode tensor with an equal dimension in each mode, and $Qd^\frac{1}{Q} \leq N$, the time complexity of HOGPR is $O(N^2d)$ — linear to the number of outputs. Such $Q$ always exists when $d \leq e^{\frac{1}{e^2}}$.

**Proof.** First, since $d_1 = \ldots = d_Q$ and $d = \prod_{k=1}^{Q} d_k$, we have each $d_k = d^\frac{1}{Q}$. Then we have $\sum_{k=1}^{Q} d_k = Qd^\frac{1}{Q} \leq N$ and $Nd(\sum_{k=1}^{Q} d_k + N) \leq 2N^2d$. Therefore, the time complexity is $O(N^2d)$. Second, if we allow continuous $Q$, and take the gradient of the function, $f(Q) = Qd^\frac{1}{Q}$, we have

$$\nabla f(Q) \nabla Q = d^\frac{1}{Q}(1 - \log(d)Q).$$

When $Q \leq \log(d)$, the gradient will be non-positive, and the function is monotonically decreasing. The function reaches the minimum at $Q = \log(d)$, and the minimum value is $e \cdot \log(d)$. Hence as long as $N \geq e \cdot \log(d)$ — equivalently, $d \leq e^{\frac{1}{Q}}$, there must exists $Q$ such that $f(Q) \leq N$, namely, $Qd^\frac{1}{Q} \leq N$.  \qed