**Algorithm 1** Generalized tensor response regression with covariates on multiple modes

**Input:** Response tensor $\mathcal{Y} \in \mathbb{R}^{d_1 \times \cdots \times d_K}$, covariate matrices $\boldsymbol{X}_k \in \mathbb{R}^{d_k \times p_k}$ for $k = 1, \ldots, K$, target Tucker rank $\boldsymbol{r} = (r_1, \ldots, r_K)$, link function $f$, maximum norm bound $\alpha$

**Output:** Low-rank estimation for the coefficient tensor $\mathcal{B} \in \mathbb{R}^{p_1 \times \cdots \times p_K}$.

Calculate $\check{\mathcal{B}} = \mathcal{Y} \times_1 \left[(\boldsymbol{X}_1^T \boldsymbol{X}_1)^{-1} \boldsymbol{X}_1^T\right] \times_2 \cdots \times_K \left[(\boldsymbol{X}_K^T \boldsymbol{X}_K)^{-1} \boldsymbol{X}_K^T\right]$.

Initialize the iteration index $t = 0$. Initialize the core tensor $\mathcal{C}^{(0)}$ and factor matrices $\boldsymbol{M}_k^{(0)} \in \mathbb{R}^{p_k \times r_k}$ via rank-$\boldsymbol{r}$ Tucker approximation of $\check{\mathcal{B}}$, in the least-square sense.

**while** the relative increase in objective function $\mathcal{L}_{\mathcal{Y}}(\mathcal{B})$ is less than the tolerance **do**

  Update iteration index $t \leftarrow t + 1$.

  **for** $k = 1$ to $K$ **do**

    Obtain the factor matrix $\tilde{\boldsymbol{M}}_k^{(t+1)} \in \mathbb{R}^{p_k \times r_k}$ by a GLM with link function $f$.

    Perform QR factorization on $\tilde{\boldsymbol{M}}_k^{(t+1)} = \boldsymbol{Q}\boldsymbol{R}$, where $\boldsymbol{Q} \in \mathbb{R}^{p_k \times r_k}$ has orthogonal columns.

    Update $\boldsymbol{M}_k^{(t+1)} \leftarrow \boldsymbol{Q}$ and core tensor $\mathcal{C}^{(t+1)} \leftarrow \mathcal{C}^{(t+1)} \times_k \boldsymbol{R}$.

  **end for**

  Obtain the core tensor $\mathcal{C}^{(t+1)} \in \mathbb{R}^{r_1 \times \cdots \times r_K}$ by solving a GLM with $\text{vec}(\mathcal{Y})$ as response, $\otimes_{k=1}^{K}[\boldsymbol{X}_k \boldsymbol{M}_k^{(t)}]$ as covariates, and $f$ as link function. Here $\otimes$ denotes the kronecker product of matrices.

  Rescale the core tensor subject to the maximum norm constraint.

  Update $\mathcal{B}^{(t+1)} \leftarrow \mathcal{C}^{(t+1)} \times_1 \boldsymbol{M}_1^{(t+1)} \times_2 \cdots \times_K \boldsymbol{M}_K^{(t+1)}$.

**end while**