APPENDIX A
ERROR BOUND FOR REDUCTION

Given a vector $\mathbf{x} \in \mathbb{R}^{1\times n}$, let $\mathbf{x}' \in \mathbb{R}^{1\times n}$ be its perturbed version. Define $\mathbf{y} = \mathbf{x}\mathbf{A}$ and $\mathbf{y}' = \mathbf{x}'\mathbf{A}$. The condition number $\kappa(\mathbf{A})$ provides an upper bound on the amplification induced by linear transformation:

$$\frac{\|\mathbf{y}' - \mathbf{y}\|}{\|\mathbf{y}\|} \leq \kappa(\mathbf{A}) \cdot \frac{\|\mathbf{x}' - \mathbf{x}\|}{\|\mathbf{x}\|}. \tag{14}$$

In the reduction, the destination tends to decode the summation of the messages of the transmitters, i.e., $\sum_{t\in\mathcal{T}} \mathbf{x}_t$. Analyzing the bound on the distortion rate of this summation is challenging, as $\|\sum_{t\in\mathcal{T}} \mathbf{x}_t\|$ may be close to zero. To simplify the analysis, we consider the distortion rate of decoding the message $\mathbf{x}_t$ by assuming that the symbols transmitted by other sink nodes are zero. This distortion rate provides an upper bound on the distortion rate of $\|\sum_{t\in\mathcal{T}} \mathbf{x}_t\|$.

Recall that the transmitted message of node $t \in \mathcal{T}$ on its outgoing links is given by

$$\mathbf{s}_t = \mathbf{x}_t((\overline{\mathbf{F}}_t^{-1})^T + \mathbf{E}_t^{(1)}),$$

where

$$\mathbf{E}_t^{(1)} = (\mathbf{E}_t^{(rep)})^T + \mathbf{E}_t^{(mul)}.$$

After transmitted by t, $\mathbf{s}_t$ is recoded by the intermediate nodes and finally decoded at ρ.

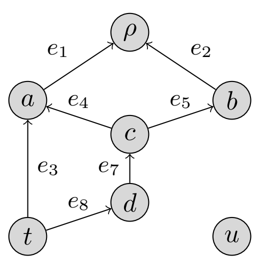

Fig. 5: Three paths from t to ρ

Fig. 5 illustrates an example of transmitting the message of t to the sink node ρ in a butterfly network. There are three paths: $t \to e_3 \to e_1 \to \rho$, $t \to e_8 \to e_7 \to e_4 \to e_1 \to \rho$, and $t \to e_8 \to e_7 \to e_5 \to e_2 \to \rho$, which are not mutually disjoint. Ideally, we can use vector-matrix multiplication to represent the process of recoding and decoding:

$$\begin{bmatrix} s_{t,1} & s_{t,2} \end{bmatrix} \begin{bmatrix} \beta^*_{e_3,e_1} & 0 \\ \beta^*_{e_8,e_7}\beta^*_{e_7,e_4}\beta^*_{e_4,e_1} & \beta^*_{e_8,e_7}\beta^*_{e_7,e_5}\beta^*_{e_5,e_2} \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$
$$= \begin{bmatrix} s_{t,1} & s_{t,2} \end{bmatrix} \overline{\mathbf{F}}_t^T = \mathbf{s}_t \overline{\mathbf{F}}_t^T.$$

Note that $\beta^*_{e_i,e_j} = \beta_{e_j,e_i}$ and $\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$ is the transpose of the generator matrix $\mathbf{A}$.

Ideally, the decoded message is denoted as

$$\mathbf{x}_t^{(0)} = \mathbf{s}_t \overline{\mathbf{F}}_t^T.$$

The inequality (14) and (7) suggest that the error between $\mathbf{x}_t^{(0)}$ and the original message $\mathbf{x}_t = [\mathbf{x}_t(\overline{\mathbf{F}}_t^{-1})^T](\overline{\mathbf{F}}_t^T)$ is lower bounded by:

$$\begin{aligned} \frac{\|\mathbf{x}_t^{(0)} - \mathbf{x}_t\|}{\|\mathbf{x}_t\|} &\leq \kappa(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{s}_t - \mathbf{x}_t(\overline{\mathbf{F}}_t^{-1})^T\|}{\|\mathbf{x}_t(\overline{\mathbf{F}}_t^{-1})^T\|} \\ &\leq \kappa^2(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{E}_t^{(1)}\|_{op}}{\|\overline{\mathbf{F}}_t^{-1}\|_{op}}. \end{aligned} \tag{15}$$

Additionally, the following inequality holds:

$$\begin{aligned} \frac{\|\mathbf{x}_t^{(0)}\|}{\|\mathbf{x}_t\|} &\leq \frac{\|\mathbf{x}_t^{(0)} - \mathbf{x}_t\| + \|\mathbf{x}_t\|}{\|\mathbf{x}_t\|} \\ &\leq 1 + \kappa^2(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{E}_t^{(1)}\|_{op}}{\|\overline{\mathbf{F}}_t^{-1}\|_{op}}. \end{aligned} \tag{16}$$

In practice, recoding and decoding involve floating-point arithmetic, and round-off errors are unavoidable. We model these errors by introducing an error term $\mathbf{E}_t^{(2)}$. The decoded message, denoted by $\mathbf{x}_{\mathbf{t}}^{(\mathbf{1})}$, is then modeled as:

$$\mathbf{x}_t^{(1)} = \mathbf{s}_t(\overline{\mathbf{F}}_t^T + \mathbf{E}_t^{(2)}).$$

It can be derived that:

$$\frac{\|\mathbf{x}_t^{(1)} - \mathbf{x}_t^{(0)}\|}{\|\mathbf{x}_t^{(0)}\|} \leq \kappa(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{E}_t^{(2)}\|_{op}}{\|\overline{\mathbf{F}}_t\|_{op}}. \tag{17}$$

By combining inequalities (17) and (16), we derive the upper bound for the distortion rate $\frac{\|\mathbf{x}_t^{(1)} - \mathbf{x}_t\|}{\|\mathbf{x}_t\|}$ as follows:

$$\begin{aligned} &\frac{\|\mathbf{x}_t^{(1)} - \mathbf{x}_t\|}{\|\mathbf{x}_t\|} \\ \leq &\frac{\|\mathbf{x}_t^{(1)} - \mathbf{x}_t^{(0)}\| + \|\mathbf{x}_t^{(0)} - \mathbf{x}_t\|}{\|\mathbf{x}_t\|} \\ = &\frac{\|\mathbf{x}_t^{(1)} - \mathbf{x}_t^{(0)}\|}{\|\mathbf{x}_t^{(0)}\|} \cdot \frac{\|\mathbf{x}_t^{(0)}\|}{\|\mathbf{x}_t\|} + \frac{\|\mathbf{x}_t^{(0)} - \mathbf{x}_t\|}{\|\mathbf{x}_t\|} \\ \leq &\kappa(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{E}_t^{(2)}\|_{op}}{\|\overline{\mathbf{F}}_t\|_{op}} \cdot (1 + \kappa^2(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{E}_t^{(1)}\|_{op}}{\|\overline{\mathbf{F}}_t^{-1}\|_{op}}) \\ &\quad + \kappa^2(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{E}_t^{(1)}\|_{op}}{\|\overline{\mathbf{F}}_t^{-1}\|_{op}} \\ \leq &\kappa^3(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{E}_t^{(1)}\|_{op} \cdot \|\mathbf{E}_t^{(2)}\|_{op}}{\|\overline{\mathbf{F}}_t^{-1}\|_{op} \cdot \|\overline{\mathbf{F}}_t\|_{op}} \\ &\quad + \kappa^2(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{E}_t^{(1)}\|_{op}}{\|\overline{\mathbf{F}}_t^{-1}\|_{op}} + \kappa(\overline{\mathbf{F}}_t) \cdot \frac{\|\mathbf{E}_t^{(2)}\|_{op}}{\|\overline{\mathbf{F}}_t\|_{op}}. \end{aligned} \tag{18}$$

APPENDIX B
EXAMPLE OF NETWORK CODES CONSTRUCTION

Consider multicast in a combination network, where the source node is ρ and the sink nodes are $\{d, e, f\}$, as illustrated in Fig. 6. We use this example to illustrate how to construct a coding scheme using float32 by applying **Algorithm I** and **Algorithm II**, along with the optimization technique.

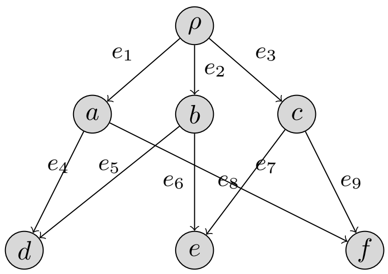

Fig. 6: Combination network

For this multicast, we know that $\min_{t\in\mathcal{T}} \text{MaxFlow}(\rho, t) = r = 2$. We first construct a graph G^0, which consists of the source node ρ and its adjacent nodes a, b, c, and its outgoing links, as illustrated by Fig. 7.

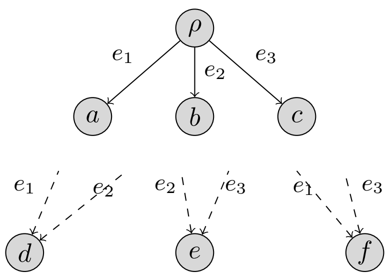

Fig. 7: Graph G^0

We construct the generator matrix $\mathbf{A}$ using the approach outlined in Appendix F, as given below

$$\mathbf{A} = \begin{pmatrix} 1.2987896 & 1.1710904 & -0.12769923 \\ -0.60240215 & 0.8235837 & 1.4259859 \end{pmatrix}$$

The transfer matrices are then given by

$$\widetilde{\mathbf{F}}_d^0 = \begin{pmatrix} 1.2987896 & 1.1710904 \\ -0.60240215 & 0.8235837 \end{pmatrix},$$
$$\widetilde{\mathbf{F}}_e^0 = \begin{pmatrix} 1.1710904 & -0.12769923 \\ 0.8235837 & 1.4259859 \end{pmatrix},$$
$$\widetilde{\mathbf{F}}_f^0 = \begin{pmatrix} 1.2987896 & -0.12769923 \\ -0.60240215 & 1.4259859 \end{pmatrix}.$$

It can be verified that all the transfer matrix are full rank.

In the next step, edge e_4 is added to the graph. The graph G^1 is illustrated in Fig. 8.

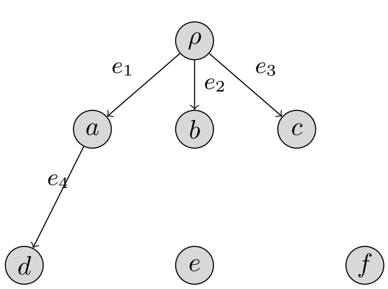

Fig. 8: Graph G^1

Since $\text{In}(a) = \{e_1\}$, we have $\mathbf{c}_{e_4} = \beta_{e_1,e_4}\mathbf{c}_{e_1}$. We pick

$$\widetilde{\mathbf{F}}_d^1 = \begin{pmatrix} 1.2987896 & 1.1710904 \\ -0.60240215 & 0.8235837 \end{pmatrix}, \beta_{e_1,e_4} = 1.$$

Next, edge e_5 is added to the graph. The graph G^2 is illustrated in Fig. 9.

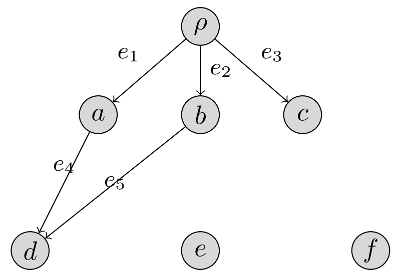

Fig. 9: Graph G^2

We pick

$$\widetilde{\mathbf{F}}_d^2 = \begin{pmatrix} 1.2987896 & 1.1710904 \\ -0.60240215 & 0.8235837 \end{pmatrix}, \beta_{e_2,e_5} = 1.$$

The remaining steps are omitted, as they are similar. We finally obtain the following coding scheme.

- Generator Matrix:

$$\mathbf{A} = \begin{pmatrix} 1.2987896 & 1.1710904 & -0.12769923 \\ -0.60240215 & 0.8235837 & 1.4259859 \end{pmatrix}$$

- Transfer Matrix and Condition Number:

$$\overline{\mathbf{F}}_d = \begin{pmatrix} 1.2987896 & 1.1710904 \\ -0.60240215 & 0.8235837 \end{pmatrix},\ \kappa(\overline{\mathbf{F}}_d) = 1.73.$$
$$\overline{\mathbf{F}}_e = \begin{pmatrix} 1.1710904 & -0.12769923 \\ 0.8235837 & 1.4259859 \end{pmatrix},\ \kappa(\overline{\mathbf{F}}_e) = 1.73.$$
$$\overline{\mathbf{F}}_f = \begin{pmatrix} 1.2987896 & -0.12769923 \\ -0.60240215 & 1.4259859 \end{pmatrix},\ \kappa(\overline{\mathbf{F}}_f) = 1.73.$$

- Coefficient:

$$\beta_{e_1,e_4} = 1,\ \beta_{e_1,e_8} = 1,\ \beta_{e_2,e_5} = 1,$$
$$\beta_{e_2,e_6} = 1,\ \beta_{e_3,e_7} = 1,\ \beta_{e_3,e_9} = 1.$$

Below is the dual coding scheme for implementing the reduction from nodes $t \in \mathcal{T}$ to node ρ.

- Decoding Matrix:

$$\mathbf{A}^T = \begin{pmatrix} 1.2987896 & -0.60240215 \\ 1.1710904 & 0.8235837 \\ -0.12769923 & 1.4259859 \end{pmatrix}$$

- Generator Matrix of each Source:

$$\mathbf{A}_d = \begin{pmatrix} 0.463957 & 0.33935675 \\ -0.65972114 & 0.7316592 \end{pmatrix}.$$
$$\mathbf{A}_e = \begin{pmatrix} 0.8033138 & -0.463957 \\ 0.07193799 & 0.65972114 \end{pmatrix}.$$
$$\mathbf{A}_f = \begin{pmatrix} 0.8033138 & 0.33935675 \\ 0.07193798 & 0.7316591 \end{pmatrix}.$$

- Coefficient:

$$\beta_{e_4,e_1}=1,\ \beta_{e_8,e_1}=1,\ \beta_{e_5,e_2}=1,$$
$$\beta_{e_6,e_2}=1,\ \beta_{e_7,e_3}=1,\ \beta_{e_9,e_3}=1.$$

Appendix C
Proof of Lemma 3

Note that $\phi(\boldsymbol{\alpha})=\prod_{i=1}^{m}(\boldsymbol{\alpha}^T\boldsymbol{\alpha}+\gamma_i^2)^{r/2}$ and $f(\boldsymbol{\alpha})=\prod_{i=1}^{m}|\boldsymbol{b}_i^T\boldsymbol{\alpha}|$. Assume that $r\geq 2$. If $r=1$, then a direct transmission strategy is sufficient, and the network coding strategy provides no additional benefit.

We denote the domain of $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}$ as $\mathcal{D}=\{\boldsymbol{\alpha}:f(\boldsymbol{\alpha})\neq 0\}$. Fix an arbitrary point $\boldsymbol{\alpha}'\in\mathcal{D}$. Considering $0<f(\boldsymbol{\alpha})\leq\prod_{i=1}^{m}(\|\mathbf{b}_i\|\cdot\|\boldsymbol{\alpha}\|)$ and $\phi(\boldsymbol{\alpha})\geq\prod_{i=1}^{m}\gamma_i^r>0$, it follows that $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}\geq\prod_{i=1}^{m}\frac{\gamma_i^r}{\|\mathbf{b}_i\|\cdot\|\boldsymbol{\alpha}\|}$. Then there exists $l>0$ such that for any $\boldsymbol{\alpha}$ satisfying $\|\boldsymbol{\alpha}\|\leq l$, we have $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}\geq\prod_{i=1}^{m}\frac{\gamma_i^r}{\|\mathbf{b}_i\|\cdot l}>\frac{\phi(\boldsymbol{\alpha}')}{f(\boldsymbol{\alpha}')}$. In addition, there exists $\epsilon>0$ such that for any $\boldsymbol{\alpha}$ satisfying $f(\boldsymbol{\alpha})\leq\epsilon$, we have $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}\geq\frac{\prod_{i=1}^{m}\gamma_i^r}{\epsilon}>\frac{\phi(\boldsymbol{\alpha}')}{f(\boldsymbol{\alpha}')}$. On the other hand, as $\boldsymbol{\alpha}\to\infty$, $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}$ becomes unbounded. Since $\phi(\boldsymbol{\alpha})>\prod_{i=1}^{m}(\boldsymbol{\alpha}^T\boldsymbol{\alpha})^{r/2}=\prod_{i=1}^{m}\|\boldsymbol{\alpha}\|^2$, it follows that $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}>\prod_{i=1}^{m}\frac{\|\boldsymbol{\alpha}\|}{\|\mathbf{b}_i\|}$. As a result, there exists $u>0$ such that for any $\boldsymbol{\alpha}$ satisfying $\|\boldsymbol{\alpha}\|\geq u$, we have $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}>\prod_{i=1}^{m}\frac{u}{\|\mathbf{b}_i\|}>\frac{\phi(\boldsymbol{\alpha}')}{f(\boldsymbol{\alpha}')}$.

Define $\mathcal{D}_{\boldsymbol{\alpha}'}=\{\boldsymbol{\alpha}:l\leq\|\boldsymbol{\alpha}\|\leq u, f(\boldsymbol{\alpha})\geq\epsilon\}$. It follows that the global minimum of $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}$ over $\mathcal{D}$ is either $\boldsymbol{\alpha}'$ or lies within $\mathcal{D}_{\boldsymbol{\alpha}'}$. If the global minimum is $\boldsymbol{\alpha}'$, then the proof is done. Otherwise, by the Extreme Value Theorem, since $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}$ is continuous and $\mathcal{D}_{\boldsymbol{\alpha}'}$ is closed and bounded, there exists $\boldsymbol{\alpha}^*$ such that $\boldsymbol{\alpha}^*=\min_{\boldsymbol{\alpha}\in\mathcal{D}_{\boldsymbol{\alpha}'}}\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}=\min_{\boldsymbol{\alpha}\in\mathcal{D}}\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}$.

Appendix D
Proof of Theorem 1

In this section, we show that the coding scheme created by **Algorithm I and II** ensures that $\overline{\mathbf{F}}_t$ for $t\in\mathcal{T}$ has rank r.

We first show that $\widetilde{\mathbf{F}}_t^i$ always maintains rank r during the updating, provided that the desired coding coefficients specified in the algorithm can be found.

Initially, the generator matrix $\mathbf{A}\in\mathbb{F}^{r\times|\operatorname{Out}(\rho)|}$ can be constructed such that $\widetilde{\mathbf{F}}_t^0$ for $t\in\mathcal{T}$ has rank r. The condition $r\leq|\operatorname{Out}(\rho)|$ is sufficient for the existence of such $\mathbf{A}$.

Assume that at the $(i-1)$-th step, $\widetilde{\mathbf{F}}_t^{i-1}$ has rank r. In the i-th step, edge e with tail v is added to the graph. We show that the update in the i-th step of the algorithm ensures that $\widetilde{\mathbf{F}}_t^i$ also has rank r. Let $\mathcal{T}_e$ denote the set of sink nodes with e on one of its paths from the source node. Let $U=\langle\mathbf{c}_{e'}:e'\in\operatorname{In}(v)\rangle$ denote the space span by vectors $\{\mathbf{c}_{e'}:e'\in\operatorname{In}(v)\}$. Suppose e is part of the j-th path of node t. Let $\widetilde{\mathbf{f}}_{t,j}^{i-1}$ be the column of $\widetilde{\mathbf{F}}_t^{i-1}$ corresponding to the j-th path, and $\widetilde{\mathbf{F}}_{t,-j}^{i-1}$ be the matrix obtained from $\widetilde{\mathbf{F}}_t^{i-1}$ by removing $\widetilde{\mathbf{f}}_{t,j}^{i-1}$. Let $V_t=\langle$columns of $\widetilde{\mathbf{F}}_{t,-j}^{i-1}\rangle$. We just need to show that there exists $\mathbf{c}_e\in U\setminus\cup_{t\in\mathcal{T}_e}(U\cap V_t)$.

Let's view this problem from the perspective of linear subspaces. For any $t\in\mathcal{T}_e$, since $\widetilde{\mathbf{f}}_{t,j}^{i-1}\notin V_t$ and $\widetilde{\mathbf{f}}_{t,j}^{i-1}\in U$, we have $\dim(U\cap V_t)\leq\dim(U)-1<\dim(U)$. It follows that $\mathbf{c}_e\in U\setminus\cup_{t\in\mathcal{T}_e}(U\cap V_t)$ exists. In fact, since $|\mathcal{T}_e|$ is finite, nearly all of the points in linear space U lie in the set $U\setminus\cup_{t\in\mathcal{T}_e}(U\cap V_t)$. This implies that, when using floating-point numbers, randomly choosing $\beta_{e',e}$ for each $e'\in\operatorname{In}(v)$ is sufficient to obtain a $\mathbf{c}_e\in U\setminus\cup_{t\in\mathcal{T}_e}(U\cap V_t)$. We construct $\widetilde{\mathbf{F}}_t^i$ from $\widetilde{\mathbf{F}}_t^{i-1}$ by replacing $\widetilde{\mathbf{f}}_{t,j}^{i-1}$ with $\mathbf{c}_e$. This ensures that $\widetilde{\mathbf{F}}_t^i$ is full rank. Similar randomized approach can be utilized to construct the matrix $\mathbf{A}$.

In each step i of the algorithm, $\widetilde{\mathbf{f}}_{t,j}^i$ always takes the value of $\mathbf{c}_e$ where e is the most downstream edge in the truncation in G^i along the j-th edge-disjoint path from the source node ρ to the sink node t. Therefore, when the algorithm stops, we have $\bar{\mathbf{F}}_t=\widetilde{\mathbf{F}}_t$, which has full rank.

Appendix E
Proof of Theorem 2

Suppose that

$$\boldsymbol{\alpha}_0=\arg\max_{\boldsymbol{\alpha}:\boldsymbol{\alpha}^T\boldsymbol{\alpha}=1}f(\boldsymbol{\alpha})$$

and

$$\boldsymbol{\alpha}_1=\arg\min_{\boldsymbol{\alpha}:f(\boldsymbol{\alpha})\neq 0}\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}.$$

We have

$$\frac{\phi(\|\boldsymbol{\alpha}_1\|\boldsymbol{\alpha}_0)}{f(\|\boldsymbol{\alpha}_1\|\boldsymbol{\alpha}_0)}=\frac{\phi(\boldsymbol{\alpha}_1)}{f(\|\boldsymbol{\alpha}_1\|\boldsymbol{\alpha}_0)}\leq\frac{\phi(\boldsymbol{\alpha}_1)}{f(\boldsymbol{\alpha}_1)}$$

where the first equality holds as $\boldsymbol{\alpha}_0^T\boldsymbol{\alpha}_0=1$, and the second inequality holds as

$$f(\|\boldsymbol{\alpha}_1\|\boldsymbol{\alpha}_0)=\|\boldsymbol{\alpha}_1\|^m\cdot f(\boldsymbol{\alpha}_0)\geq\|\boldsymbol{\alpha}_1\|^m\cdot f(\frac{\boldsymbol{\alpha}_1}{\|\boldsymbol{\alpha}_1\|})=f(\boldsymbol{\alpha}_1).$$

On the other hand, since $\frac{\phi(\|\boldsymbol{\alpha}_1\|\boldsymbol{\alpha}_0)}{f(\|\boldsymbol{\alpha}_1\|\boldsymbol{\alpha}_0)}\geq\frac{\phi(\boldsymbol{\alpha}_1)}{f(\boldsymbol{\alpha}_1)}$, it follows that $\frac{\phi(\|\boldsymbol{\alpha}_1\|\boldsymbol{\alpha}_0)}{f(\|\boldsymbol{\alpha}_1\|\boldsymbol{\alpha}_0)}=\frac{\phi(\boldsymbol{\alpha}_1)}{f(\boldsymbol{\alpha}_1)}$. Therefore, we also have $f(\boldsymbol{\alpha}_0)=f(\frac{\boldsymbol{\alpha}_1}{\|\boldsymbol{\alpha}_1\|})$. As a result $\|\boldsymbol{\alpha}_1\|\boldsymbol{\alpha}_0$ minimizes $\frac{\phi(\boldsymbol{\alpha})}{f(\boldsymbol{\alpha})}$ and $\frac{\boldsymbol{\alpha}_1}{\|\boldsymbol{\alpha}_1\|}$ maximizes $f(\boldsymbol{\alpha})$.

Appendix F
Construction of Generator Matrix

To optimize the generator matrix $\mathbf{A}$, we propose a heuristic algorithm to minimize $\prod_{t\in\mathcal{T}}\kappa(\widetilde{\mathbf{F}}_t^0)$.

First, initialize $\mathbf{A}\in\mathbb{F}^{r\times|\operatorname{Out}(\rho)|}$ by sampling each entry uniformly from the interval $[0,2]$. With high probability $rank(\mathbf{A})=r$. Note that each column in $\mathbf{A}$ corresponds to a coefficient vector $\mathbf{c}_e$ for each $e\in\operatorname{Out}(\rho)$. The transfer matrix $\widetilde{\mathbf{F}}_t^0$ for $t\in\mathcal{T}$ is a submatrix of $\mathbf{A}$, consisting of the vectors $\mathbf{c}_e\in\{\mathbf{c}_e:e\in\operatorname{Out}(\rho)\}$ if edge e is part of its r paths. With high probability, $rank(\widetilde{\mathbf{F}}_t^0)=r$. If this condition is not met, we will resample $\mathbf{A}$ until it is satisfied. We then replace the columns of $\mathbf{A}$ that appear in only one transfer matrix $\widetilde{\mathbf{F}}_t^0$ with standard basis vectors from $\{\hat{\mathbf{e}}_i:i=1,\dots r\}$, ensuring that the full-rank property of $\widetilde{\mathbf{F}}_t^0$ is preserved.

Next, we optimize the columns of $\mathbf{A}$ that are shared by multiple $\widetilde{\mathbf{F}}_t^0$. Suppose column $\mathbf{c}_e$ is of these. Let $\mathcal{T}_e\subset\mathcal{T}$ denote the set of sink nodes where $\mathbf{c}_e$ appears in their transfer matrices. We optimize $\mathbf{c}_e$ by minimizing the upper bound of $\prod_{t \in \mathcal{T}_e} \kappa(\widetilde{\mathbf{F}}_t^0)$, as given in (10). Using our optimization framework, this involves determining $\boldsymbol{\alpha} \in \mathbb{F}^r$ for (11). Subsequently, $\mathbf{c}_e$ is updated to $\sum_{i=1}^{r} \alpha_i \hat{\mathbf{e}}_i$. This process is iterated for all shared columns, allowing for multiple cycles to improve results, until the decrease in $\prod_{t \in \mathcal{T}_e} \kappa(\widetilde{\mathbf{F}}_t^i)$ falls below a specified threshold.