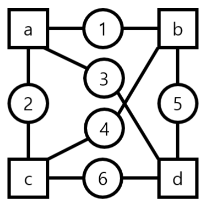
# Supplement:

## Gauged Mini-Bucket Elimination for Approximate Inference

## A Example of gauge transformations




$$f_a(x_1, x_2, x_3) = \left[\begin{bmatrix} 2432 & 832 \\ 4672 & 640 \end{bmatrix} \begin{bmatrix} 4864 & 384 \\ 5120 & 4160 \end{bmatrix}\right]$$

$$f_a(x_1, x_2, x_3; \mathbf{G}_a) = \left[\begin{bmatrix} 2837 & 1559 \\ 3591 & 2077 \end{bmatrix} \begin{bmatrix} 3631 & 2005 \\ 4261 & 2077 \end{bmatrix}\right]$$

$$f_b(x_1, x_4, x_5) = \left[\begin{bmatrix} 1088 & 128 \\ 4928 & 4608 \end{bmatrix} \begin{bmatrix} 448 & 1664 \\ 3264 & 1344 \end{bmatrix}\right]$$

$$f_b(x_1, x_4, x_5; \mathbf{G}_b) = \left[\begin{bmatrix} 2142 & 1434 \\ 4634 & 4558 \end{bmatrix} \begin{bmatrix} 966 & 1490 \\ 1490 & 758 \end{bmatrix}\right]$$

$$f_c(x_2, x_4, x_6) = \left[\begin{bmatrix} 1216 & 5440 \\ 768 & 1856 \end{bmatrix} \begin{bmatrix} 5568 & 896 \\ 640 & 512 \end{bmatrix}\right]$$

$$f_c(x_1, x_4, x_5; \mathbf{G}_c) = \left[\begin{bmatrix} 3960 & 8808 \\ -1608 & -2520 \end{bmatrix} \begin{bmatrix} -328 & -3288 \\ 6520 & 8296 \end{bmatrix}\right]$$

$$f_d(x_3, x_5, x_6) = \left[\begin{bmatrix} 5632 & 5632 \\ 6080 & 6208 \end{bmatrix} \begin{bmatrix} 5568 & 896 \\ 640 & 512 \end{bmatrix}\right]$$

$$f_d(x_3, x_5, x_6; \mathbf{G}_d) = \left[\begin{bmatrix} 2408 & 9160 \\ 10760 & 9192 \end{bmatrix} \begin{bmatrix} 14536 & -6232 \\ -7448 & -1208 \end{bmatrix}\right]$$

$$G_{1a}, G_{2a}, G_{3a}, G_{4b}, G_{5b}, G_{6c} = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}$$

$$G_{1b}, G_{2c}, G_{3d}, G_{4c}, G_{5d}, G_{6d} = \begin{bmatrix} 1.5 & -0.5 \\ -0.5 & 1.5 \end{bmatrix}$$

Figure 5: Example of gauge transformations on the complete graph (with respect to factors) of size 4. Arrays follow row-column major indexing, e.g., $f_a(1, 1, 2) = 4864$ and $f_a(1, 2, 1) = 832$.

## B Proof of Theorem 1

We prove reparameterization with respect to $\boldsymbol{\theta} = \{\theta_{v\alpha}(x_v) = 0, \ \forall (v, \alpha) \in E, x_v\}$ is optimal at GM with symmetric factors in the following optimization:

$$\begin{aligned} \underset{\boldsymbol{\theta}}{\text{minimize}} \quad & \sum_{\bar{x}_{\bar{1}}}^{\bar{w}_{\bar{n}}} \cdots \sum_{\bar{x}_1}^{\bar{w}_1} \prod_{\alpha \in F} f_\alpha(\bar{\mathbf{x}}_\alpha; \boldsymbol{\theta}_\alpha), \\ \text{subject to} \quad & \prod_{\alpha \in N(v)} \exp(\theta_{v\alpha}(x_v)) = 1 \quad \forall \ v \in X, x_v. \end{aligned}$$

The optimization is convex, and assuming $\theta_{v\beta} + \theta_{v\alpha} = 0$ from the constraint, $\partial \log Z_{\text{WMBE}} / \partial \theta_{v\alpha} = 0$ implies optimality of the solution. To this end, the derivative is expressed as:

$$\frac{\partial \log Z_{\text{WMBE}}}{\partial \bar{\theta}_\alpha(\mathbf{x}_{v\alpha})} = \sum_{\mathbf{x}_{\alpha \setminus v}} q(\bar{\mathbf{x}}_\alpha) - \sum_{\bar{\mathbf{x}}_{\beta \setminus v}} q(\bar{\mathbf{x}}_\beta).$$

When factors are symmetric, it immediately follows that

$$\sum_{\mathbf{x}_{\alpha \setminus v}} q(\bar{\mathbf{x}}_\alpha) = \sum_{\bar{\mathbf{x}}_{\beta \setminus v}} q(\bar{\mathbf{x}}_\beta) = 0.5,$$

since $q$ is expressed via weighted absolute sum and normalization operation of factors, which both preserve symmetry. Hence marginals are also symmetric, implying uniform distribution. Hence the optimality condition is satisfied.