# Weighted Graph Coloring for Quantized Computing

Derya Malak
Communication Systems Department, EURECOM, Biot Sophia Antipolis, 06904 France
derya.malak@eurecom.fr

***Abstract*—We consider the problem of distributed lossless computation of a function of two sources by one common user. To do so, we first build a bipartite graph, where two disjoint parts denote the individual source outcomes. We then project the bipartite graph onto each source to obtain an edge-weighted characteristic graph (EWCG), where edge weights capture the function's structure, by how much the source outcomes are to be distinguished, generalizing the classical notion of characteristic graphs. Via exploiting the notions of characteristic graphs, the fractional coloring of such graphs, and edge weights, the sources separately build multi-fold graphs that capture vector-valued source sequences, determine vertex colorings for such graphs, encode these colorings, and send them to the user that performs minimum-entropy decoding on its received information to recover the desired function in an asymptotically lossless manner. For the proposed EWCG compression setup, we characterize the fundamental limits of distributed compression, verify the communication complexity through an example, contrast it with traditional coloring schemes, and demonstrate that we can attain compression gains higher than $\%30$ over traditional coloring.**

## I. Introduction

Over the past years, we have been experiencing an ever-increasing demand for computationally-intensive tasks, motivating us to devise new parallel processing techniques to speed up and efficiently distribute computations across groups of servers. In modern distributed computing, a primary concern is communication cost. While parallel processing to distribute communication can reduce the need for coordination and alleviate this cost, reduction of the same communication cost is challenged due to issues of scalability [1], accuracy [2], low capacity edges [3], and stragglers [4] in distributed computing.

### A. Related Work

**Distributed coded computation.** There have been various efforts to mitigate the communication cost in distributed computing following Yao's seminal work in [5] on communication complexity. Some recent breakthroughs in this direction include coded computing [6]–[8], and distributed computation of, e.g., matrix products [9]–[12], matrix multiplication with stragglers [13], and linearly separable functions [14]–[16].

**Distributed source and functional compression.** Other attempts have been inspired from the seminal work of Slepian-Wolf [17] on distributed source compression, the rate-distortion coding models of Wyner-Ziv with side information [18], and for lossy source coding [19], toward function computation. These works include [20]–[23] that consider function computation over networks, as well as [23] and [24], considering the generalization to functional rate-distortion, and [25] and [26], focusing on hypergraph-based source coding and function approximation under maximal distortion. Recent works also include hyperbinning for distributed function quantization [27], generalizing the orthogonal binning ideas in Slepian-Wolf coding [28], and fractional coloring-based distributed computation [29] that reduces complexity of [21].

**Coding for specific functions and channels.** The communication cost is also affected by the nature of the computed function. Examples include Körner-Marton' encoding problem for computing *modulo-two sum of binary sources* [30], the generalization of Körner-Marton's problem to a two-terminal source coding scheme with *common sum reconstruction* [31], which has applications in distributed stochastic gradient descent, and Max-Lloyd's algorithms [32] to compute large-scale averages. Han and Kobayashi have established necessary and sufficient conditions on functions such that the Slepian-Wolf region is optimal for distributed lossless computing [33]. The authors in [34]–[36] have explored the *combinatorial aspects of source coding* to compress correlated sources separately or with decoder side information. The *joint source-channel scheme* of Cover, El Gamal, and Salehi uses the source correlations to achieve a collaborative gain and create channel input distributions adapted to the channel [37]. To that end, Nazer and Gastpar have devised designs for distributed computing over multiple access channels [38], and *structured coding* [39].

### B. Overview and Contributions

We focus on distributed computing of a function of two jointly distributed finite alphabet sources at a user. We pose this problem as an edge-weighted characteristic graph (EWCG) compression problem. To do so, we build a bipartite graph[1] where two disjoint parts denote the individual source outcomes, and the edges capture the joint source distribution.

Our main contributions can be summarized as follows:

- ***Edge-weighted $b$-fold compression.*** We propose an EWCG encoding scheme to provide low-complexity compression for computing, where we describe the weights by the joint source distribution and the function. An EWCG is a fractionally colored characteristic graph built by each source as an edge-weighted projection of the bipartite graph (Sect. II). To capture the unequal edge weights, the source devises $b$ characteristic graphs (one graph per source coordinate,

Funded by the European Union (ERC, SENSIBILITÉ, 101077361). Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

[1]Important classes of bipartite networks are the collaboration network and the opinion network. They are significant in information and economic systems, social networks, opinion networks and recommendation systems [40].

see [41, App. A]), where the edge weights in EWCG are quantized across these graphs (that we detail via Example 1). In an EWCG, a vertex captures a *b-fold, i.e., vector-valued, source value*, and is given $b$ colors out of $a$ available colors, where $b$ captures the quantization depth of each source. The edge weights are used to determine $a$, $b$, and the overlap of colorings for any vertex pair (Sect. III), upon which each source establishes and encodes the colorings of its EWCG.

- ***Edge-weighted fractional chromatic entropy.*** The fractional chromatic number $\chi_f$ – given by the limit in (13) in [41, App. A] – determines the communication complexity when the edges have unit weights. Using OR power graphs, we can exploit the gains in complexity through fractional coloring as the blocklength $n$ tends to infinity [42, Ch. 3]. To that end, we generalize the definition of $\chi_f$ via EWCGs to provide a lower communication complexity (Sect. III).
- ***Joint quantization and distributed functional compression via EWCGs.*** In the edge-weighted fractional coloring of *vector-valued sources*, $b$ is the quantization depth. The encoding rates for EWCGs are lower versus traditional or fractional coloring of graphs because the higher the value of $b$ is, the more refined the weights in an EWCG are, enabling a lower rate of compression per source coordinate. We characterize in (3) the number of disjoint colors between two vertices of an EWCG. We provide in Theorem 1 (Sect. III) the encoding rate for a $b$-fold fractional coloring of EWCGs.
- ***Numerical experiments.*** Contrasting it with the existing techniques via an example, EWCG exhibits significant savings in communication complexity by taking into account the structures of the sources (via the Slepian-Wolf theorem [17]) and the function (via the edge weights).

### C. Notation

For a random variable $X$ with a finite alphabet $\mathcal{X}$, $P_X$ denotes its probability mass function (PMF). Similarly, for variables $X_1$ and $X_2$, $P_{X_1,X_2}$ denotes the joint PMF. We denote the probability of an event $A$ by $\mathbb{P}(A)$. Let the entropy function of a PMF $\mathbf{p}$ be $h(\mathbf{p}) = -\sum_i p_i \log p_i$ where the logarithm is in base 2, $h(p)$ be the binary entropy function with parameter $p$, and $H(X) = \mathbb{E}[-\log P_X(X)]$ be the Shannon entropy of $X$. We denote by $\mathbf{X}_1^n = X_{11}, X_{12}, \dots, X_{1n} \in \mathcal{X}_1^n$ the length $n$ sequence of $X_1$ sampled from an $n$-fold finite alphabet $\mathcal{X}_1^n$. We let $[N] = \{1, 2, \dots, N\}$, $N \in \mathbb{Z}^+$.

## II. Model and Problem Statement

We pose the problem of distributed computation of a function $f(X_1, X_2)$ of the two sources $X_1$ and $X_2$ as a *compression problem for the edge-weighted projections of a bipartite graph that captures $P_{X_1,X_2}$*. For this partially distributed setting, we will devise an encoding scheme for EWCGs and quantify the sum rate for computing $f(X_1, X_2)$, by exploiting the notions of characteristic graphs and their entropy [20]–[23] and the concept of bipartite graph projection.

### A. Bipartite Graph Representation

We construct a bipartite graph representation $G_f = (\mathcal{X}_1, \mathcal{X}_2, E)$ to compute the function $f(X_1, X_2)$, whose partition has the parts $\mathcal{X}_1$ and $\mathcal{X}_2$, which correspond to the set of realizations of the sources $X_1$ and $X_2$, respectively, and $E$ denotes the set of edges of $G_f$. The bipartite graph $G_f$ is derived from the joint distribution $P_{X_1,\,X_2}$, and $E$ captures the correlation between $X_1$ and $X_2$. More specifically,

1) The set of vertices $\mathcal{X}_1$ and $\mathcal{X}_2$ that partition $G_f$ are disjoint.
2) $G_f$ is a balanced bipartite graph with $|\mathcal{X}_1| = |\mathcal{X}_2|$, i.e., the two subsets of vertices have the same cardinality.
3) There is an edge between vertices $u_k \in \mathcal{X}_1$ and $v_l \in \mathcal{X}_2$, i.e., $(u_k, v_l) \in E$, if and only if $\mathbb{P}(X_1 = u_k, X_2 = v_l) > 0$.
4) If $u_k \in \mathcal{X}_1$ and $v_l \in \mathcal{X}_2$ are connected, i.e., $(u_k, v_l) \in E$, and $(v_l, u_k) \in E$, then the symmetry of the edges does not imply that both edges yield the same function outcome.

If $G_f$ is complete, it has $|\mathcal{X}_1| \cdot |\mathcal{X}_2|$ edges and the number of distinct function outcomes is determined by the structure of $f(X_1, X_2)$. On the other hand, if $G_f$ is not connected, it may have more than one bipartition [43]. In that case, encoding of $f(X_1, X_2)$ is facilitated upon the extraction of the bipartition information. We note that the sources do not have the full knowledge of $E$, as determined by $P_{X_1,X_2}$, but only the weights jointly determined by $P_{X_1,X_2}$ and $f(X_1, X_2)$. We assume that the edge weights are available, and can be learned via feedback, the study of which is left as future work.

### B. Weighted Bipartite Graphs through Projections of $G_f$

Source one $X_1$ observes a weighted projection of $G_f$ onto a graph – the $X_1$ projection of $G_f$ – denoted by $G^w_{X_1}$, and similarly for source two. For the EWCG of $X_1$, given by $G^w_{X_1}$, the edge weight between $u_{k_1}, u_{k_2} \in \mathcal{X}_1$ of $G^w_{X_1}$, denoted by $w(u_{k_1}, u_{k_2})$, is set to be the weighted number of common neighbors in $X_2$. Hence, the notion $G^w_{X_1}$ generalizes the concept of the characteristic graph $G_{X_1}$ detailed in [41, App. A]. In this paper, we determine $\{w(u_{k_1}, u_{k_2}),\, u_{k_1}, u_{k_2} \in \mathcal{X}_1\}$ as

$$w(u_{k_1}, u_{k_2}) = \sum_{\substack{v_l \in \mathcal{X}_2 :\, f(u_{k_1}, v_l) \neq f(u_{k_2}, v_l) \\ \prod_{k \in \{k_1, k_2\}} P_{X_1,X_2}(u_k, v_l) > 0}} P_{X_1,X_2}([u_{k_1}, u_{k_2}], v_l)\ , \quad (1)$$

where $P_{X_1,X_2}([u_{k_1}, u_{k_2}], v_l) = \sum_{k \in \{k_1, k_2\}} P_{X_1,X_2}(u_k, v_l)$. The idea is similar for determining $w(v_{l_1}, v_{l_2})$ of $G^w_{X_2}$.

Similarly, towards realizing the limits of compression, for the $n$-th power graph of $G^w_{X_1}$, namely $G^{n,w}_{\mathbf{X}_1}$, the edge weight between the vertices $\mathbf{u}_i^n, \mathbf{u}_j^n \in \mathcal{X}_1^n$ of $G^{n,w}_{\mathbf{X}_1}$ is given as

$$w(\mathbf{u}_i^n, \mathbf{u}_j^n) = \sum_{\substack{\mathbf{v}_l^n \in \mathcal{X}_2^n :\ f(\mathbf{u}_i^n, \mathbf{v}_l^n) \neq f(\mathbf{u}_j^n, \mathbf{v}_l^n) \\ \prod_{k \in \{i,j\}} P_{\mathbf{X}_1^n, \mathbf{X}_2^n}(\mathbf{u}_k^n, \mathbf{v}_l^n) > 0}} P_{\mathbf{X}_1^n, \mathbf{X}_2^n}([\mathbf{u}_i^n, \mathbf{u}_j^n], \mathbf{v}_l^n)\ .$$

We can note that for the standard construction $G_{X_1}$ of $X_1$ [44], [21], as detailed in [41, App. A], the edge weights satisfy

$$1_{\left|\left\{v_l \in \mathcal{X}_2 :\, \prod_{k \in \{k_1, k_2\}} P_{X_1,X_2}(u_k, v_l) > 0,\, f(u_{k_1}, v_l) \neq f(u_{k_2}, v_l)\right\}\right| > 0}\ .$$

In distributed compression, exploiting the notion of jointly typical sequences, it is possible for the user to estimate the number of $\mathbf{X}_2^n$ sequences jointly typical with $\mathbf{X}_1^n$ given $\mathbf{X}_1^n$. Hence, as a simplification of this paper's model in (1), while still generalizing $G_{X_1}$, the weight $w(u_{k_1}, u_{k_2})$ for $u_{k_1}, u_{k_2} \in \mathcal{X}_1$ can be set as the number of common neighbors in $\mathcal{X}_2$:

$$\sum_{v_l \in \mathcal{X}_2} 1_{\prod_{k\in\{k_1,k_2\}} P_{X_1,X_2}(u_k,v_l)>0,\, f(u_{k_1},v_l)\neq f(u_{k_2},v_l)} \; . \quad (2)$$

The edge weights in (1) affect the quantization of the source outcomes through a $b$-tuple of graphs, which we detail next.

## III. Main Results

In this section, we provide an achievable encoding and decoding approach for asymptotically lossless distributed computation of $f(X_1, X_2)$, which is based on projecting the bipartite graph $G_f$ onto EWCGs and compressing the EWCGs.

### A. Valid Colorings of Edge-Weighted Graphs

In traditional coloring of an unweighted graph $G_{X_1}$, we note that given a pair of vertices $u_{k_1}, u_{k_2} \in \mathcal{X}_1$ such that $w(u_{k_1}, u_{k_2}) = 0$, it implies that the two vertices can have identical colors $c_{G_{X_1}}(u_{k_1}) = c_{G_{X_1}}(u_{k_2})$. On the other hand, $w(u_{k_1}, u_{k_2}) > 0$ implies $c_{G_{X_1}}(u_{k_1}) \neq c_{G_{X_1}}(u_{k_2})$.

In fractional coloring of EWCGs, prior to a valid coloring of vertices of $G^w_{X_1}$ and $G^w_{X_2}$, we normalize each weight in (1) by $\max\{w(u_{k_1}, u_{k_2}),\, u_{k_1}, u_{k_2} \in \mathcal{X}_1\}$, and similarly for $\{w(v_{l_1}, v_{l_2}),\, v_{l_1}, v_{l_2} \in \mathcal{X}_2\}$ of $G^w_{X_2}$.

We next let $c^f_{G^w_{X_1}}(u_{k_1})$, $u_{k_1} \in \mathcal{X}_1$ be a valid fractional coloring with a $b$-fold coloring, where $u_{k_1}$ is assigned $b$ colors out of $a$ available colors. Note that the distance between colors $c^f_{G^w_{X_1}}(u_{k_1})$ and $c^f_{G^w_{X_1}}(u_{k_2})$, i.e., $\mathrm{dist}(c^f_{G^w_{X_1}}(u_{k_1}), c^f_{G^w_{X_1}}(u_{k_2}))$, is an increasing function of $w(u_{k_1}, u_{k_2})$ [45]. To that end, we stretch Defns. (1) and (2) in [41, App. A] of the standard $a : b$ coloring, and adopt the following model. As in traditional coloring, for a given $u_{k_1}, u_{k_2} \in \mathcal{X}_1$, when $w(u_{k_1}, u_{k_2}) = 0$, then the $b$-fold colors $c^f_{G^w_{X_1}}(u_{k_1})$ and $c^f_{G^w_{X_1}}(u_{k_2})$ could be identical, i.e., $\mathrm{dist}(c^f_{G^w_{X_1}}(u_{k_1}), c^f_{G^w_{X_1}}(u_{k_2})) = 0$. On the other hand, when $w(u_{k_1}, u_{k_2}) = 1$, then the $b$-fold colors $c^f_{G^w_{X_1}}(u_{k_1})$ and $c^f_{G^w_{X_1}}(u_{k_2})$ can have no overlaps, i.e., $\mathrm{dist}(c^f_{G^w_{X_1}}(u_{k_1}), c^f_{G^w_{X_1}}(u_{k_2})) = b$. More generally, a valid $a : b$ coloring of $G^w_{X_1}$ is such that given $w(u_{k_1}, u_{k_2})$, the minimum number of disjoint colors between $u_{k_1}$ and $u_{k_2}$ of $G^w_{X_1}$ is

$$\mathrm{dist}(c^f_{G^w_{X_1}}(u_{k_1}), c^f_{G^w_{X_1}}(u_{k_2})) = \lceil w(u_{k_1}, u_{k_2}) \cdot b \rceil \; , \quad (3)$$

meaning that if $w(u_{k_1}, u_{k_2}) \in \left(\frac{b-(k+1)}{b}, \frac{b-k}{b}\right]$ for $k \in \{0\} \cup [b-1]$, then vertices $u_{k_1}$ and $u_{k_2}$ are assigned $b-k$ distinct colors, and only if $w(u_{k_1}, u_{k_2}) = 0$ they are assigned exactly the same $b$ colors. We note that the number of different colors between two vertices of $G^w_{X_1}$ changes as a function of the edge weight, as given in (3). The neighboring vertices in $G^w_{X_{1i}}$ have at least one different color, and the endpoints of edges with large weights have a higher number of disjoint colors. Clearly, this coloring scheme generalizes the notion of fractional chromatic number (Defn. 2 in [41, App. A]).

We next expand $G^w_{X_1}$ into a $b$-tuple of graphs represented by $G^w_{X_1(S)} = \{G^w_{X_{1i}} : i \in S,\, |S| = b\}$, where $G^w_{X_{1i}}$, $i \in S$ is an $i$-th replica of $G^w_{X_1}$. We jointly color the set of graphs $G^w_{X_1(S)}$ such that $c_{G_{X_1(S)}}(X_1(S)) = \{c_{G^w_{X_{1i}}}(X_{1i}) : i \in S,\, |S| = b\}$ and $w(u_{k_1}, u_{k_2}) = \frac{1}{b}\sum_{i\in S} w_i(u_{k_1}, u_{k_2})$ is split such that

$$w_i(u_{k_1}, u_{k_2}) = \min\Big\{1,\, \max\Big\{b \cdot w(u_{k_1}, u_{k_2}) - b\sum_{i'=1}^{i-1} w_{i'}(u_{k_1}, u_{k_2}),\, 0\Big\}\Big\} \quad (4)$$

denotes the weight between $u_{k_1}$ and $u_{k_2}$ of $G^w_{X_{1i}}$, $i \in S$, i.e., the $i$-th replica of $G^w_{X_1}$. Note that (4) yields a sequence of monotone decreasing edge weights $w_i(u_{k_1}, u_{k_2})$ for $i \in S$ that jointly determine the traditional colorings for the set of graphs $G^w_{X_1(S)}$. In Fig. 1, we show a joint coloring for an example $|S| = 2$-tuple EWCG. We will detail this example in Sect. III-C to indicate the achievable gains in compression.

We next explore the fundamental rate limits for distributed computing of $f(X_1, X_2)$, by exploiting the notions of characteristic graph entropy, and EWCGs, where we determine the weights according to (4), following the bipartite projection scheme. To that end, we next detail encoding and decoding of $G^w_{X_1}$ for asymptotically lossless compression of $f(\mathbf{X}_1^n, \mathbf{X}_2^n)$.

### B. An Achievable Coloring Scheme for Edge-Weighted Graphs

In this part, we detail the encoding and decoding principle of EWCGs for distributed computing of $f(X_1, X_2)$. We next describe the encoding of $b$-fold colors. Note that the computation of $f$ is lossless independent of the value of $b \in \mathbb{Z}^+$.

*a) Encoding:* Given $G_f$, the encoding phase includes the projections of $G_f$ onto $G^w_{X_1}$ and $G^w_{X_2}$ by determining the corresponding edge weights using (1) followed by their normalization. Each source then builds a $b$-tuple of characteristic graphs, $G^w_{X_1(S)}$ and $G^w_{X_2(S)}$, respectively, for $|S| = b$. The sources can then compress their weighted graphs asymptotically at rates $H^f_{G^w_{X_1}}(X_1)$ and $H^f_{G^w_{X_1}}(X_2)$, where we next give the conditional fractional graph entropy of the EWCG $G^w_{X_1}$.

**Theorem 1.** *The fractional graph entropy of $G^w_{X_1}$ is equal to*

$$H^f_{G^w_{X_1}}(X_1 \,|\, X_2) = \lim_{n\to\infty} \frac{1}{n} \inf_b \frac{1}{b} \min_{c^f_{G^{n,w}_{\mathbf{X}_1}}} \{H(c^f_{G^{n,w}_{\mathbf{X}_1}}(\mathbf{X}_1)) : c^f_{G^{n,w}_{\mathbf{X}_1}}(\mathbf{X}_1) \text{ is a valid } a : b \text{ coloring of } G^{n,w}_{\mathbf{X}_1} \,|\, \mathbf{X}_2\} \; , \quad (5)$$

*where $c^f_{G^{n,w}_{\mathbf{X}_1}}(\mathbf{X}_1)$ is a fractional coloring variable for $G^{n,w}_{\mathbf{X}_1}$ with an $a : b$ coloring of each vertex of $G^{n,w}_{\mathbf{X}_1}$.*

*Proof.* A proof sketch is given in [41, App. B]. □

*b) Decoding:* For lossless decoding, the user needs to be instructed on $P_{X_1,X_2}$, $f$, $b$, and the look-up table for recovering $f(\mathbf{X}_1^n, \mathbf{X}_2^n)$ using the received fractional colorings of the $b$-tuple of graphs from each source. The user first performs minimum-entropy decoding on its received information [46]. Via Slepian-Wolf decoding, it achieves the sequences $c^f_{G^{n,w}_{\mathbf{X}_1}}(\mathbf{X}_1)$ and $c^f_{G^{n,w}_{\mathbf{X}_2}}(\mathbf{X}_2)$ that model the $b$-fold color tuples. The user then uses a look-up table to compute $f(\mathbf{X}_1^n, \mathbf{X}_2^n)$.

To demonstrate the procedure for encoding and decoding of an EWCG $G^w_{X_1}$, determining the edge weights in (4), and

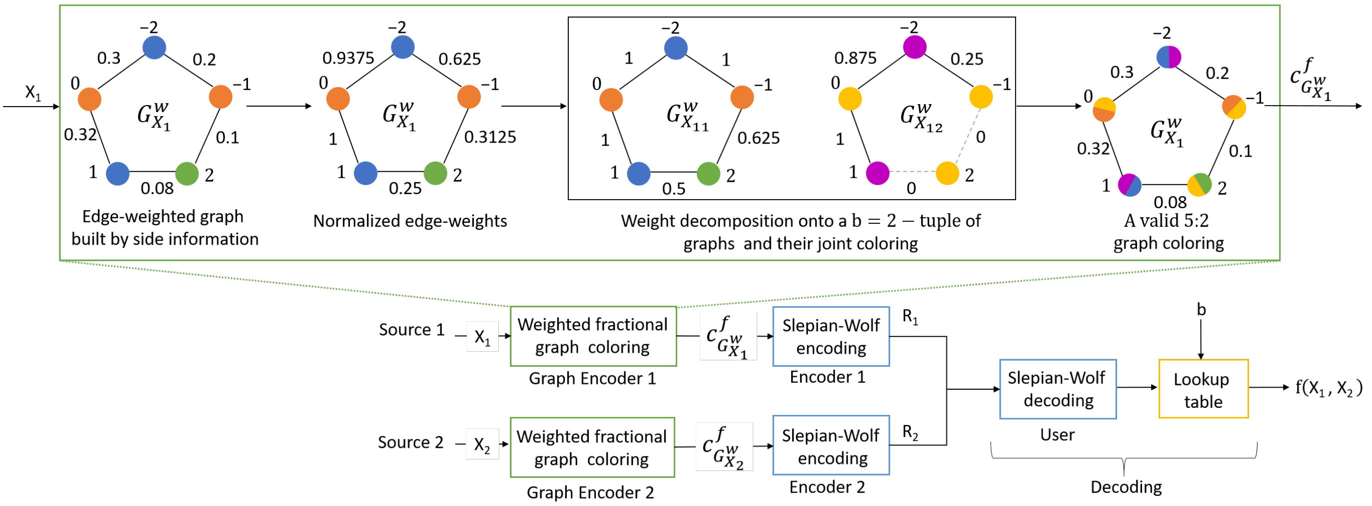


Fig. 1. Distributed computation of $f(X_1, X_2)$: an end-to-end multi-fold encoding and decoding scheme for EWCGs. The encoding phase consists of determining the EWCG tuples and their colorings, followed by Slepian-Wolf encoding on the $b$-fold colors. Decoding relies on recovering the $b$-fold colors using Slepian-Wolf decoding followed by recovering the outcomes using a look-up table. We note that $P_{X_1} = (0.2, 0.15, 0.32, 0.24, 0.09)$, and the edge weights are given in Example 1. In the bottom figure, the graph encoder for each source is independent – output is a $b$-tuple color sequence – with Slepian-Wolf encoding. In this example, the user uses 4 fractional colors received ($b = 2$ from each source per transmission) to reconstruct the function outcome.

sending a pair of $b$-tuples of coloring sequences for recovery of $f(\mathbf{X}_1^n, \mathbf{X}_2^n)$ by the user in an asymptotically lossless manner, we next detail an end-to-end distributed computing example with a $b = 2$-fold coloring of $G^w_{X_1}$, which is shown in Fig. 1.

### C. An Example toward Edge-Weighted Encoding-Decoding

We present an example to illustrate how to build an EWCG and how to encode and decode the coloring, to obtain the desired function outcomes. Through this example, we also contrast the performance of our scheme with that of traditional graph coloring that does not exploit the weight information.

**Example 1. An EWCG and its chromatic entropy.** *The source variables $X_1$ and $X_2$ share a common alphabet such that $\mathcal{X} = \{-2, -1, 0, 1, 2\}$. The ordered marginal PMFs are $X_1 \sim \mathbf{p}_1 = (0.2, 0.15, 0.32, 0.24, 0.09)$ and $X_2 \sim \mathbf{p}_2 = (0.2, 0.3, 0.32, 0.08, 0.1)$, and $P_{X_1,X_2}$ is given as follows:*

$$P_{X_1,X_2} = \begin{bmatrix} 0.1 & 0.1 & 0 & 0 & 0 \\ 0.1 & 0 & 0 & 0 & 0.05 \\ 0 & 0.2 & 0.12 & 0 & 0 \\ 0 & 0 & 0.2 & 0.04 & 0 \\ 0 & 0 & 0 & 0.04 & 0.05 \end{bmatrix} . \quad (6)$$

*We note that the entropy of $X_1$ satisfies $H(X_1) = h(0.2, 0.15, 0.32, 0.24, 0.09) = 2.2078 < H(X_{1,u}) = 2.32$, with $X_{1,u} \sim P_{X_{1,u}}$, where $P_{X_{1,u}}$ is uniform over $\mathcal{X}$.*

**Unweighted scenario.** *Without taking into account the edge weights, the minimum entropy coloring of $G_{X_1}$ is given as $H(c_{G_{X_1}}) = h(0.44, 0.47, 0.09) = 1.35$. The entropy with a $5 : 2$ fractional coloring with $\chi_f(G_{X_1}) = 2.5$ satisfies $\frac{1}{2}H(c^f_{G_{X_1}}) = \frac{1}{2}h(0.22, 0.235, 0.205, 0.145, 0.195) = 1.15$. Similarly, for $G^2_{\mathbf{X}_1}$, with an $8 : 1$ coloring, and a PMF [23]*

$$c_{G^2_{\mathbf{X}_1}} \sim (0.176, 0.188, 0.018, 0.176, 0.188, 0.036, 0.036, 0.182) ,$$

*we get $\frac{1}{2}H(c_{G^2_{\mathbf{X}_1}}) = 1.34$. For a $13 : 2$ coloring, $\frac{1}{4}H(c^f_{G^2_{\mathbf{X}_1}}) = 0.91$. For $X_{1,u}$ uniform, it holds that $\frac{1}{2}H(c^f_{G_{X_{1,u}}}) = \frac{1}{2}\log 5 = 1.16$, and for $\mathbf{X}_{1,u}$ uniform, $\frac{1}{4}H(c^f_{G^2_{\mathbf{X}_{1,u}}}) = \frac{1}{4}\log 13 = 0.92$.*

**Weighted scenario.** *We next take into account the edge weights. Using (1), the edge weights are $w(-2, -1) = 0.2$, $w(-2, 0) = 0.3$, $w(0, 1) = 0.32$, $w(1, 2) = 0.08$, and $w(-1, 2) = 0.1$. Note that for this specific example, $W \sim \mathbf{p}_2$.*

*We next decompose $G^w_{X_1}$ into $b = 2$ graphs, as shown in Fig. 1 (top row). Normalizing the edge weights, and then using (4), the weights are $w_1(-2, -1) = w_1(-2, 0) = w_1(0, 1) = 1$, $w_1(1, 2) = 0.5$, and $w_1(-1, 2) = 0.625$ for $G^w_{X_{11}}$, and $w_2(-2, -1) = 0.25$, $w_2(-2, 0) = 0.875$, $w_2(0, 1) = 1$, and $w_2(1, 2) = w_2(-1, 2) = 0$ for $G^w_{X_{12}}$. This yields a valid $5 : 2$ coloring of $G^w_{X_1(S)}$ for $|S| = 2$, as also shown in the top row.*

*Using the joint coloring information of $G^w_{X_1(S)}$, i.e., for $G^w_{X_{11}}$ and $G^w_{X_{12}}$, the color PMF for the $5 : 2$ fractional coloring of $G^w_{X_1}$ for the set of ordered colors $\{c_1 = Blue, c_2 = Orange, c_3 = Green, c_4 = Purple, c_5 = Yellow\}$ satisfies*

$$P_{c^f_{G^w_{X_1}}}(c_1) = \frac{1}{2}(\mathbf{p}_1(-2) + \mathbf{p}_1(1)) = \frac{0.44}{2} = 0.22 ,$$

$$P_{c^f_{G^w_{X_1}}}(c_2) = \frac{1}{2}(\mathbf{p}_1(-1) + \mathbf{p}_1(0)) = \frac{0.47}{2} = 0.235 ,$$

$$P_{c^f_{G^w_{X_1}}}(c_3) = \frac{1}{2}\mathbf{p}_1(2) = 0.045 ,$$

$$P_{c^f_{G^w_{X_1}}}(c_4) = \frac{1}{2}(\mathbf{p}_1(-2) + \mathbf{p}_1(1)) = \frac{0.44}{2} = 0.22 ,$$

$$P_{c^f_{G^w_{X_1}}}(c_5) = \frac{1}{2}(\mathbf{p}_1(-1) + \mathbf{p}_1(0) + \mathbf{p}_1(2)) = \frac{0.56}{2} = 0.28 , \quad (7)$$

*which yields from (5) that $\frac{1}{2}H(c^f_{G^w_{X_1}}) = 1.08 < \frac{1}{2}H(c^f_{G_{X_1}}) = 1.15 < H(c_{G_{X_1}}) = 1.35$. Hence, for $b = 2$, capturing the edge weights yields a saving of %16 over traditional coloring and does not offer enhancement over standard fractional coloring.*

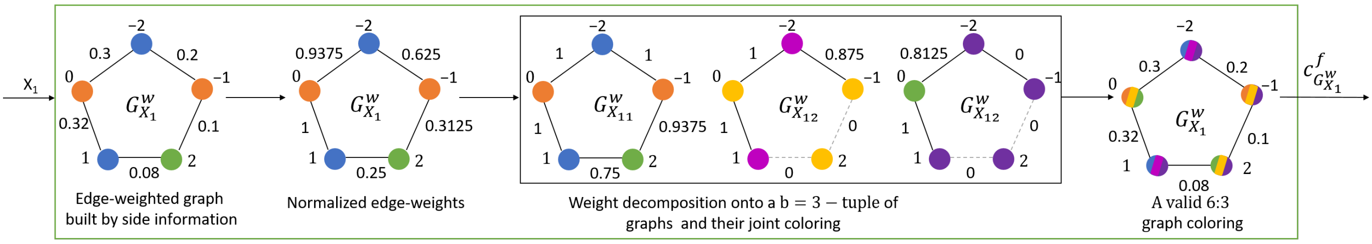


Fig. 2. A fractional coloring scheme for distributed computation of $f(X_1, X_2)$ with a $6:3$ coloring.

*For the same example, with $b = 3$ and with the inclusion of a sixth color, where $c_6 = Violet$, we can achieve a $6:3$ coloring as shown in Fig. 2, and the coloring PMF is*

$$P_{c^f_{G^w_{X_1}}}(c_1) = P_{c^f_{G^w_{X_1}}}(c_4) = \frac{1}{3}(\mathbf{p}_1(-2) + \mathbf{p}_1(1)) = \frac{0.44}{3} ,$$
$$P_{c^f_{G^w_{X_1}}}(c_2) = \frac{1}{3}(\mathbf{p}_1(-1) + \mathbf{p}_1(0)) = \frac{0.47}{3} ,$$
$$P_{c^f_{G^w_{X_1}}}(c_3) = \frac{1}{3}(\mathbf{p}_1(2) + \mathbf{p}_1(0)) = \frac{0.41}{3} ,$$
$$P_{c^f_{G^w_{X_1}}}(c_5) = \frac{1}{3}(\mathbf{p}_1(-1) + \mathbf{p}_1(0) + \mathbf{p}_1(2)) = \frac{0.56}{3} ,$$
$$P_{c^f_{G^w_{X_1}}}(c_6) = \frac{1}{3}(1 - \mathbf{p}_1(0)) = \frac{0.68}{3} . \quad (8)$$

*Then, a valid $6:3$ coloring of $G^w_{X_1}$ yields $\frac{1}{3}H(c^f_{G^w_{X_1}}) = 0.85$, providing a saving of $\%37$ over traditional coloring. Hence, a larger $b$ can capture the edge weights more accurately.*

*We next consider the second power graph $G^{2,w}_{\mathbf{X}_1}$. We note that $\chi_f(G^w_{X_1}) = 2.5$, and $\chi_f(G^{2,w}_{\mathbf{X}_1}) = \chi_f^2(G^w_{X_1}) = 6.25$. Hence, a $12:2$ coloring is not possible for $n = 2$. We show a valid $13:2$ coloring of $G^{2,w}_{\mathbf{X}_1}$ in Fig. 3, given the ordered set $\{c_1 = Blue, c_2 = Yellow, c_3 = Green, c_4 = Orange, c_5 = Purple, c_6 = LightBlue, c_7 = Brown, c_8 = Violet, c_9 = BrickRed, c_{10} = DarkGreen, c_{11} = Black, c_{12} = Gray, c_{13} = Navy\}$. Its coloring PMF can be derived from that for $G^w_{X_1}$ and can be shown to satisfy*

$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_m) = \frac{2}{5}P_{c^f_{G^w_{X_1}}}(c_1) = 0.088 , \quad m \in \{1, 11, 12\} ,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_2) = \frac{2}{5}P_{c^f_{G^w_{X_1}}}(c_5) = 0.112 ,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_3) = \frac{1}{5}(P_{c^f_{G^w_{X_1}}}(c_3) + P_{c^f_{G^w_{X_1}}}(c_2)) = 0.056 ,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_4) = \frac{2}{5}P_{c^f_{G^w_{X_1}}}(c_2) = 0.094 ,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_5) = \frac{1}{5}(P_{c^f_{G^w_{X_1}}}(c_4) + P_{c^f_{G^w_{X_1}}}(c_3)) = 0.053 ,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_6) = \frac{1}{5}(P_{c^f_{G^w_{X_1}}}(c_4) + P_{c^f_{G^w_{X_1}}}(c_5)) = 0.1 ,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_7) = \frac{1}{5}(P_{c^f_{G^w_{X_1}}}(c_1) + P_{c^f_{G^w_{X_1}}}(c_2)) = 0.091 ,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_8) = \frac{2}{5}P_{c^f_{G^w_{X_1}}}(c_3) = 0.018,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_9) = \frac{1}{5}(P_{c^f_{G^w_{X_1}}}(c_3) + P_{c^f_{G^w_{X_1}}}(c_5)) = 0.065 ,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_{10}) = \frac{1}{5}(P_{c^f_{G^w_{X_1}}}(c_2) + P_{c^f_{G^w_{X_1}}}(c_5)) = 0.103 ,$$
$$P_{c^f_{G^{2,w}_{\mathbf{X}_1}}}(c_{13}) = \frac{1}{5}P_{c^f_{G^w_{X_1}}}(c_4) = 0.044 ,$$

*which yields from (5) that $\frac{1}{4}H(c^f_{G^{2,w}_{\mathbf{X}_1}}(\mathbf{X}_1)) = 0.9 = \frac{1}{4}H(c^f_{G^2_{\mathbf{X}_1}}) = 0.91 < \frac{1}{2}H(c_{G^2_{\mathbf{X}_1}}) = 1.34$. Hence, capturing the edge weights yields a saving of $\%32$ over traditional coloring, and does not have much gain over the fractional coloring approach that does not capture the weights. Increasing $b$ allows us to capture the edge weights more accurately.*

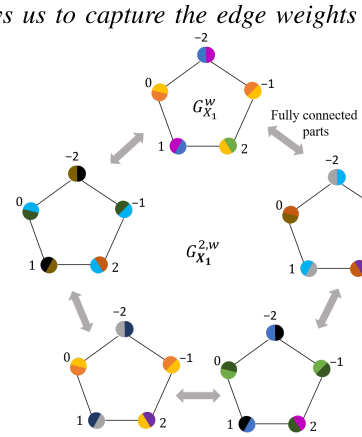


Fig. 3. A valid $13:2$ fractional coloring of $G^{2,w}_{\mathbf{X}_1}$ for Example 1, where $\chi_f(G^{2,w}_{\mathbf{X}_1}) = \chi_f^2(G^w_{X_1}) = (2.5)^2 = 6.25$.

*Similarly, we can determine the compression rate for general $n$. Exploiting [42, Cor. 3.4.3], $\chi(G^n_{\mathbf{X}_1}) \approx \chi_f^n(G_{X_1})$ as $n$ goes to infinity. Hence, we can derive the $n$-th power graph, $G^{n,w}_{\mathbf{X}_1}$, along with its $a:b$ fractional coloring, $c^f_{G^{n,w}_{\mathbf{X}_1}}(\mathbf{X}_1)$.*

From Example 1, as $b$ increases, we have a finer-grained quantization of the graph edge weights. As the skew of the edge weights increases, the efficiency in compressing the $b$-tuples of $G^w_{X_1}$ increases (e.g., in Fig. 1 some edges have relatively low weights, e.g., $w(1, 2) = 0.08$, and $w(-1, 2) = 0.1$, yielding a fewer number of distinct colors between these two end vertices). As the value of $b$ increases, the edge weights will be captured with greater precision, leading to a more refined fractional coloring (more skewed) and a reduced total number of colors and smaller graph entropy $H^f_{G^w_{X_1}}(X_1)$ given by (5).

When the *total bit budget for quantization and compression* is limited, there is a tradeoff between $b$ that determines the fold of coloring, and the complexity of encoding the characteristic graph. That is, the number of bits spent on quantizing the edge weights determines the attainable gains in compression.

## REFERENCES


[1] M. Soleymani, H. Mahdavifar, and A. S. Avestimehr, "Analog lagrange coded computing," *IEEE J. Sel. Areas Inf. Theory*, vol. 2, no. 1, pp. 283–295, Feb. 2021.

[2] J. Wang, Z. Jia, and S. A. Jafar, "Price of precision in coded distributed matrix multiplication: A dimensional analysis," in *Proc., IEEE Inf. Theory Wksh.*, Virtual Conference, Oct. 2021, pp. 1–6.

[3] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Sep. 2006.

[4] A. Behrouzi-Far and E. Soljanin, "Efficient replication for straggler mitigation in distributed computing," *arXiv preprint arXiv:2006.02318*, Jun. 2020.

[5] A. C.-C. Yao, "Some complexity questions related to distributive computing (preliminary report)," in *Proc. ACM Symp. Theory of Computing*, Atlanta, GA, Apr. 1979, pp. 209–213.

[6] Q. Yu, S. Li, N. Raviv, S. M. M. Kalan, M. Soltanolkotabi, and S. A. Avestimehr, "Lagrange coded computing: Optimal design for resiliency, security, and privacy," in *Proc., Int. Conf. Artificial Intelligence and Statistics*. Naha, Okinawa, Japan: PMLR, Apr. 2019, pp. 1215–1225.

[7] A. Reisizadeh, S. Prakash, R. Pedarsani, and A. S. Avestimehr, "Coded computation over heterogeneous clusters," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4227–4242, Mar. 2019.

[8] S. Prakash, A. Reisizadeh, R. Pedarsani, and A. S. Avestimehr, "Coded computing for distributed graph analytics," *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6534–6554, Jun. 2020.

[9] Z. Jia and S. A. Jafar, "On the capacity of secure distributed batch matrix multiplication," *IEEE Trans. Inf. Theory*, vol. 67, no. 11, pp. 7420–7437, Sep. 2021.

[10] K. Wan, H. Sun, M. Ji, D. Tuninetti, and G. Caire, "Cache-aided matrix multiplication retrieval," *IEEE Trans. Inf. Theory*, Mar. 2022.

[11] W.-T. Chang and R. Tandon, "On the capacity of secure distributed matrix multiplication," in *Proc., IEEE Global Commun. Conf.*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[12] Z. Chen, Z. Jia, Z. Wang, and S. A. Jafar, "GCSA codes with noise alignment for secure coded multi-party batch matrix multiplication," *IEEE J. Sel. Areas Inf. Theory*, vol. 2, no. 1, pp. 306–316, Jan. 2021.

[13] W. Li, Z. Chen, Z. Wang, S. A. Jafar, and H. Jafarkhani, "Flexible constructions for distributed matrix multiplication," in *Proc., IEEE Int. Symp. Inf. Theory*, Virtual Conference, Jul. 2021, pp. 1576–1581.

[14] A. Khalesi and P. Elia, "Multi-user linearly-separable distributed computing," *arXiv preprint arXiv:2206.11119*, Jun. 2022.

[15] K. Wan, H. Sun, M. Ji, and G. Caire, "Distributed linearly separable computation," *IEEE Trans. Inf. Theory*, vol. 68, no. 2, pp. 1259–1278, Nov. 2021.

[16] ——, "On secure distributed linearly separable computation," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 3, pp. 912–926, Jan. 2022.

[17] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

[18] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theoy*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[19] H. Yamamoto, "Wyner-Ziv theory for a general function of the correlated sources," *IEEE Trans. Inf. Theory*, vol. 28, no. 5, pp. 803–7, Sep. 1982.

[20] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, p. 903–917, Mar. 2001.

[21] J. Körner, "Coding of an information source having ambiguous alphabet and the entropy of graphs," in *Proc., 6th Prague Conf. Inf. Theory*, Prague, Czech Republic, Sep. 1973, pp. 411–425.

[22] N. Alon and A. Orlitsky, "Source coding and graph entropies," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1329–1339, Sep. 1996.

[23] S. Feizi and M. Médard, "On network functional compression," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5387–5401, Sep. 2014.

[24] V. Doshi, D. Shah, and M. Médard, "Source coding with distortion through graph coloring," in *Proc., IEEE Int. Symp. Inf. Theory*, Nice, France, Jun. 2007, pp. 1501–1505.

[25] S. Basu, D. Seo, and L. R. Varshney, "Hypergraph-based source codes for function computation under maximal distortion," *arXiv preprint arXiv:2204.02586*, Apr. 2022.

[26] ——, "Functional epsilon entropy," in *Proc., Data Compression Conf.*, Virtual Conference, Mar. 2020, pp. 332–341.

[27] D. Malak and M. Médard, "A distributed computationally aware quantizer design via hyper binning," *IEEE Trans. Signal Proces.,*, Jan. 2023.

[28] T. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources (corresp.)," *IEEE Trans. Inf. Theory*, vol. 21, no. 2, pp. 226–228, Mar. 1975.

[29] D. Malak, "Fractional graph coloring for functional compression with side information," in *Proc., IEEE Inf. Theory Wksh.*, Mumbai, India, Nov. 2022.

[30] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources (corresp.)," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 219–221, Mar. 1979.

[31] T. Adikari and S. Draper, "Two-terminal source coding with common sum reconstruction," in *Proc., IEEE Int. Symp. Inf. Theory*, Espoo, Finland, Jun. 2022, pp. 1420–1424.

[32] A. T. Suresh, X. Y. Felix, S. Kumar, and H. B. McMahan, "Distributed mean estimation with limited communication," in *Proc., Int. Conf. Machine Learning*. Sydney, Australia: PMLR, Jul. 2017, pp. 3329–3337.

[33] T. Han and K. Kobayashi, "A dichotomy of functions F (X, Y) of correlated sources (X, Y)," *IEEE Trans. Inf. Theory*, vol. 33, no. 1, pp. 69–76, Jan. 1987.

[34] H. Witsenhausen, "The zero-error side information problem and chromatic numbers (corresp.)," *IEEE Trans. Inf. Theory*, vol. 22, no. 5, pp. 592–593, Sep. 1976.

[35] R. Ahlswede, "Coloring hypergraphs: A new approach to multi-user source coding," *J. Comb.*, vol. 4, no. 1, pp. 76–115, 1979.

[36] J. Körner and A. Orlitsky, "Zero-error information theory," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2207–2229, Oct. 1998.

[37] T. Cover, A. E. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. 26, no. 6, pp. 648–657, Nov. 1980.

[38] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Sep. 2007.

[39] ——, "Lattice coding increases multicast rates for gaussian multiple-access networks," in *Proc., Allerton Conf.*, Monticello, IL, Sep. 2007.

[40] T. Zhou, J. Ren, M. Medo, and Y.-C. Zhang, "Bipartite network projection and personal recommendation," *Physical Review E*, vol. 76, no. 4, p. 046115, Oct. 2007.

[41] D. Malak, "Weighted graph coloring for quantized computing," May 2023. [Online]. Available: https://www.dropbox.com/s/mmeyl7fm0hvhb1u/ISIT_2023_Weighted_Graph_Coloring_for_Quantized_Computing_with_Appendix.pdf?dl=0

[42] E. R. Scheinerman and D. H. Ullman, *Fractional graph theory: a rational approach to the theory of graphs*. Courier Corporation, 2011.

[43] G. Chartrand and P. Zhang, *Chromatic Graph Theory*. CRC press, 2019.

[44] C. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, Sep. 1956.

[45] H. Toivonen, F. Zhou, A. Hartikainen, and A. Hinkka, "Compression of weighted graphs," in *Proc., ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Diego, CA, Aug. 2011, pp. 965–973.

[46] I. Csiszár and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, Jun. 2011.