Fractional Graph Coloring for Functional Compression with Side Information

Derya Malak
Communication Systems Department, EURECOM
derya.malak@eurecom.fr

***Abstract*—We describe a rational approach to reduce the computational and communication complexities of lossless point-to-point compression for computation with side information. The traditional method relies on building a characteristic graph with vertices representing the source symbols and with edges that assign a source symbol to a collection of independent sets to be distinguished for the exact recovery of the function. Our approach uses fractional coloring for a b-fold coloring of characteristic graphs to provide a linear programming relaxation to the traditional coloring method and achieves coding at a fine-grained granularity. We derive the fundamental lower bound for compression, given by the fractional characteristic graph entropy, through generalizing the notion of Körner's graph entropy. We demonstrate the coding gains of fractional coloring over traditional coloring via a computation example. We conjecture that the integrality gap between fractional coloring and traditional coloring approaches the smallest b that attains the fractional chromatic number to losslessly represent the independent sets for a given characteristic graph, up to a linear scaling which is a function of the fractional chromatic number.**

I. Introduction

We consider the problem of point-to-point compression for computing a function with decoder side information. Traditionally, this problem is referred to as Körner's graph coloring problem [1]. This problem stems from source coding (compression), which has been the subject of extensive study in information theory dating back to the seminal work of Shannon [2], and its many extensions.

A. Coding for Compression

In the traditional compression approach, for a point to point compression of a source variable X_1 drawn from distribution P_{X_1}, the source coding theorem, as demonstrated by Shannon [2], states that in the limit, as the length of a stream of independent and identically-distributed (i.i.d.) random variable data tends to infinity, the best rate of compression of X_1 (quantified in the average number of bits per symbol) is the Shannon entropy of the source, $H(X_1) = \mathbb{E}[-\log P_{X_1}(X_1)]$.

A natural extension of the source coding theorem for networked settings is the problem of *distributed compression*. The problem of distributed lossless compression dates back to the seminal work of Slepian and Wolf [3], instantiated by random binning of the typical source sequences [4]. For concreteness, consider two random variables X_1 and X_2, jointly distributed according to P_{X_1,X_2}. Given two sequences $\mathbf{X}_1^n = (X_{11}, X_{12}, \dots, X_{1n})$ and $\mathbf{X}_2^n = (X_{21}, X_{22}, \dots, X_{2n})$ drawn i.i.d. from P_{X_1,X_2}, Slepian-Wolf Theorem gives a theoretical bound for the lossless coding rate of distributed coding [3]: To recover a joint source $(\mathbf{X}_1^n, \mathbf{X}_2^n)$ drawn from P_{X_1,X_2} at a receiver that has access to side information $\mathbf{X}_2^n$, it is both necessary and sufficient to encode the source $\mathbf{X}_1^n$ up to the rate $H(X_1|X_2)$ [3]. Slepian-Wolf problem is a special case of the general distributed function compression problem.

Practical schemes for Slepian-Wolf compression have been proposed by several authors, including [5]–[8]. The generalization of the distributed compression scheme of Slepian-Wolf to trees and to networks beyond depth one uses random linear coding, as shown by Ho *et al.* in [8]. Distributed communication has also been considered in [8], and by Ahlswede *et al.* [9] and Yeung in [10] for multicasting under general network settings via random linear network coding.

B. Coding for Functional Compression

Distributed compression of source variables for the purpose of computing a deterministic function across a network, is referred to as *distributed functional compression*. To that end, since the pioneering work of Slepian-Wolf [3], different techniques have been explored, e.g., computation with *decoder side information* and *functional distortion criterion* in Wyner-Ziv settings [11], compression for multiple descriptions of functions [12], special functions such as addition [13], and multiplication with side information [14]. Function-oriented recovery, or compression of $f(X_1, X_2)$ is better understood through the lens of characteristic graph-entropy $H_{G_{X_1}}(X_1)$ [1], which quantifies the minimum number of bits required to represent a function of random variables. This notion was initially devised by Körner [1] for point-to-point compression.

The zero-error side information problem and the rate regions for the functional compression problem have been investigated by Witsenhausen [15], along with the formal introduction of the characteristic graph of X_1, X_2 and f, by Orlitsky-Roche [16] when one source is fully available at the receiver via conditional graph entropy $H_{G_{X_1}}(X_1|X_2)$, and for restricted and unrestricted inputs by Alon-Orlitsky [17]. The problem of distributed lossless compression has been studied by Doshi *et al.* [18], and also extended to tree networks (Feizi-Médard [19] and Doshi *et al.* [20]) via generalizing distributed compression (Slepian-Wolf [3]) to *distributed functional compression*.

The requirement for structured coding for computing. For distributed compression of a general function, trimming of (independently encoded) Slepian–Wolf partitions may not be feasible. In other instances, good computation codes may only achieve marginal gains in computing capacity over separation-based codes (see e.g., [21] and [19]) at the expense of a significant computation *burden on the encoders and decoders*. There also exist approaches to compression of graphical data in sparse scenarios, e.g., [22]–[24], which may not capture *computations of general functions*. Hence, designing efficient function-oriented codebooks requires a different vision to alleviate the redundancy of data-oriented encoding.

The need for constructive techniques to compression. For existing graph entropy-based approaches, we refer the reader to [1], [16]–[18], [25]–[27] and the references therein. The proofs of functional compression are by means of coloring characteristics graphs of functions but provide no natural constructive approach to instantiate functional compression.

Our contributions in this paper are summarized as follows:

- In Sect. II, we provide a relaxation to the traditional graph coloring approach (NP-hard) for functional compression. This is possible through fractional coloring, by generalizing traditional graph coloring that assigns one color per vertex. An $a : b$ fractional graph coloring assigns b colors out of a total of a available colors to each vertex of a graph such that adjacent vertices have disjoint colors.
- We introduce the concept of the *fractional chromatic entropy* and characterize it in Prop 1 exploiting Han's theorem [28], which states that the average entropy decreases monotonically in the size of the subset. In Prop. 2, we derive the analytical expression for the fractional graph entropy, $H^f_{G_{X_1}}(X_1|X_2)$, that gives the achievable rate for functional compression with side information.
- Sect. III presents our main results providing a rate bound on functional compression using fractional coloring. Exploiting several properties of fractional chromatic number, we state that $H^f_{G_{X_1}}(X_1|X_2)$ lower bounds $H_{G_{X_1}}(X_1|X_2)$ (Lemma 1). This new notion provides a refinement in coloring such that on average less colors are spent and the communication complexity is reduced.
- Sect. IV provides several bounds on the integrality gap between fractional and traditional colorings. To contrast the potential rate savings of our approach over traditional coloring, we show that the integrality gap is a monotonically increasing function of the source sequence length n (Lemma 2), and approximate the integrality gap, and observe that (Prop. 3) it scales linearly with b, given a valid b-fold coloring. Hence, our approach yields *a reduced communication complexity* by a factor of b (up to a linear scaling), in the number of communication or exchanged bits [29], [30], versus the traditional approach.

II. Problem Setup

We consider the problem of lossless distributed functional compression with side information introduced in [16] (via generalizing [11] using a characteristic graph approach). The encoder has source X_1. The decoder has source X_2, which is not accessible at the encoder side. Given two statistically dependent i.i.d. finite-alphabet random sequences X_1^n and X_2^n, our goal is to give a theoretical bound for the lossless coding rate to encode X_1 for computing a function $f(X_1, X_2)$ to achieve arbitrarily small error probability for long sequences. Both the encoder and the decoder knows the function. In [16] and its extensions, this problem, for the zero-error setting, has been tackled using a traditional coloring approach and it has been proven that a zero-error compression up to a rate $H_{G_{X_1}}(X_1|X_2)$ is possible when X_2 is available at the receiver.

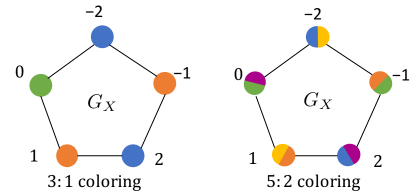

Fig. 1. (Left) G_{X_1} for Example 1 with $|\mathcal{X}_1| = 5$, where different colors on connected vertices indicate that those should be distinguished. (Right) A fractional $5:2$ coloring, which achieves χ_f [31].

A. Traditional Coloring of Characteristic Graphs

Let G_{X_1} be the characteristic graph the encoder builds for computing the function $f(X_1, X_2)$, determined as function of X_1, X_2, and f. The characteristic graph is denoted by $G_{X_1} = (V_{G_{X_1}}, E_{G_{X_1}})$, where $V_{G_{X_1}} = \mathcal{X}_1$ and an edge $(x_1^1, x_1^2) \in E_{G_{X_1}}$ if and only if there exists a $x_2^1 \in \mathcal{X}_2$ such that $p(x_1^1, x_2^1) \cdot p(x_1^2, x_2^1) > 0$ and $f(x_1^1, x_2^1) \neq f(x_1^2, x_2^1)$. We assign different codes (colors) to connected vertices, which corresponds to a graph coloring. Vertices that are not connected to each other can be assigned to the same or different colors. In this paper, we only consider vertex colorings. A valid coloring of a graph G_{X_1} is such that each vertex of G_{X_1} is assigned a color such that adjacent vertices receive disjoint colors.

To better motivate our approach and demonstrate the achievable description lengths, we illustrate the relevance of characteristic graph in compression via the following example.

Example 1. A characteristic graph and its entropy. *Random variables X_1 and X_2 are over the alphabet $\mathcal{X} = \{-2, -1, 0, 1, 2\}$. The joint distribution has ordered entries:*

$$P_{X_1, X_2} = \begin{bmatrix} 0.1 & 0.1 & 0 & 0 & 0 \\ 0.1 & 0 & 0 & 0 & 0.1 \\ 0 & 0.1 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 & 0.1 \end{bmatrix}, \tag{1}$$

where X_1 is uniformly distributed, and $f(X_1, X_2) = X_1 + X_2$ such that G_{X_1} denotes the characteristic graph the encoder builds, where $V_{G_{X_1}} = \mathcal{X}$, and $E_{G_{X_1}} = \{(-2,-1), (-2,0), (0,1), (1,2), (2,-1)\}$. A valid coloring of G_{X_1}, denoted by $c_{G_{X_1}}(X_1)$ and shown in Fig. 1 (Left), has a distribution $P(c_1) = P(c_2) = 0.4$ and $P(c_3) = 0.2$ over $\{c_1, c_2, c_3\}$. This yields an entropy $H(c_{G_{X_1}}(X_1)) \approx 1.52$.

Next, we encode a source sequence with length two, $\mathbf{X}_1^2 = (X_{11}, X_{12})$, which can take 25 values $\{(-2,-2), (-2,-1), (-2,0), \dots, (2,2)\}$. To construct the characteristic graph for $\mathbf{X}_1^2$, i.e., the second power graph $G^2_{\mathbf{X}_1}$, we connect two vertices if at least one of coordinates are connected in G_{X_1}. It is possible to color $G^2_{\mathbf{X}_1}$ using 8 colors. The entropy of this coloring satisfies $\frac{1}{2}H(c_{G^2_{\mathbf{X}_1}}) \approx 1.48 < H(c_{G_{X_1}}) \approx 1.52 < H(X_1) \approx 2.32$.

The chromatic number $\chi(G_{X_1})$ of a graph G_{X_1} is the minimum number of colors needed to color the vertices in such a way that no two adjacent vertices have the same color.

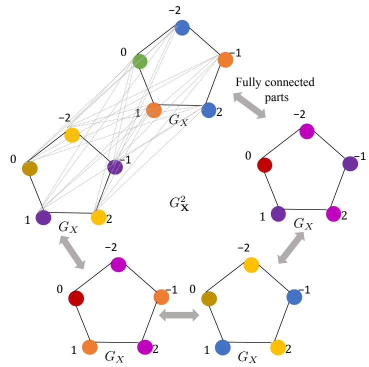

Fig. 2. The second power graph $G^2_{\mathbf{X_1}}$ for Example 1, where $a = 8$ is the minimum number of colors for which an $a : 1$ coloring is possible.

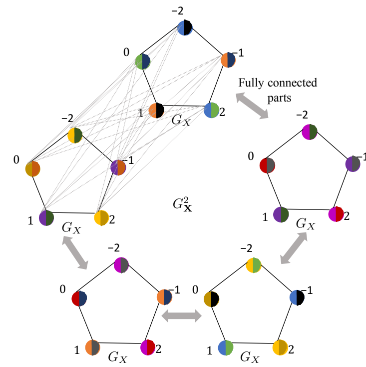

Fig. 3. A fractional coloring of $G^2_{\mathbf{X_1}}$ for Example 1, where $a = 13$ is the minimum number of colors for which an $a : 2$ coloring exists.

Definition 1. *(Chromatic entropy [17].) The chromatic entropy of a graph G_{X_1} is defined as*

$$H^{\chi}_{G_{X_1}}(X_1|X_2) = \min_{c_{G_{X_1}}} \mathcal{H}^{\chi}(c_{G_{X_1}}(X_1)) \ , \qquad (2)$$

where $\mathcal{H}^{\chi}(c_{G_{X_1}}(X_1)) = \{H(c_{G_{X_1}}(X_1)) : c_{G_{X_1}}(X_1)$ is a valid coloring of $G_{X_1}|X_2\}$ is the set of chromatic entropies over the set of valid colorings of G_{X_1}.

Let $G^n_{\mathbf{X_1}} = (V^n_{X_1}, E^n_{X_1})$ be the n-th power of a graph G_{X_1} such that $V^n_{X_1} = \mathcal{X}^n_1$ and $(\mathbf{x}^1_1, \mathbf{x}^2_1) \in E^n_{X_1}$, where $\mathbf{x}^1_1 = (x^1_{11}, x^1_{12}, \dots, x^1_{1n})$ and similarly for $\mathbf{x}^2_1$, when there exists at least one coordinate $i \in \{1, 2, \dots, n\}$ such that $(x^1_{1i}, x^2_{1i}) \in E_{X_1}$. We denote a coloring of $G^n_{\mathbf{X_1}}$ by $c_{G^n_{\mathbf{X_1}}}(\mathbf{X_1})$. Körner showed in [1] that, in the limit of large n, the chromatic entropy and the graph entropy are related as

$$H_{G_{X_1}}(X_1|X_2) = \lim_{n\to\infty} \frac{1}{n} \min_{c_{G^n_{\mathbf{X_1}}}} H(c_{G^n_{\mathbf{X_1}}}(\mathbf{X_1})|\mathbf{X_2}) \ . \qquad (3)$$

The entropy of $c_{G^n_{\mathbf{X_1}}}(\mathbf{X_1})$ characterizes the minimal description length needed to reconstruct with fidelity $f(X_1, X_2)$ [1]. The case of the identity function yields a complete graph.

B. Fractional Coloring of Characteristic Graphs

Fractional graph coloring is a natural extension of traditional coloring, where each vertex is assigned a set of colors and the adjacent vertices have disjoint sets. Traditional graph coloring problems may not be amenable to a linear programming approach. Solving (3) is equivalent to determining a coloring random variable which minimizes the entropy. However, finding the minimum entropy coloring of G_{X_1} is an NP-hard problem [32]. To solve the coloring problem losslessly in polynomial time, we exploit the fractional coloring relaxation.

Definition 2 (Scheinerman and Ullman [31])**.** *A valid b-fold coloring is an assignment of sets of size b to vertices such that adjacent vertices receive disjoint sets of colors. A valid $a : b$ coloring is a valid b-fold coloring out of a available colors. The notation $\chi_b(G)$ represents the b-fold chromatic number of graph G that is the least a such that an $a : b$ coloring exists.*

The chromatic number $\chi(G)$ is subadditive, i.e., $\chi_{a+b}(G) \leq \chi_a(G) + \chi_b(G)$. If $g : \mathbb{Z}^+ \to \mathbb{R}$ is subadditive and $g(b) \geq 0$ for all b, then from the sub-additivity lemma [31], the limit $\lim\limits_{b\to\infty} \frac{g(b)}{b}$ exists and is equal to the infimum of $\frac{g(b)}{b}$ $(b \in \mathbb{Z}^+)$.

Definition 3. *The fractional chromatic number is defined as*

$$\chi_f(G) := \liminf_{b\to\infty} \left\{ \frac{\chi_b(G)}{b} \right\} = \inf_b \frac{\chi_b(G)}{b} \ , \qquad (4)$$

where the existence of this limit follows from the sub-additivity of b-fold colorings, and the sub-additivity lemma [31].

From a probabilistic perspective, $\chi_f(G)$ represents the smallest k for which there is a distribution over the independent sets (an independent set is a set of vertices in a graph, no two of which are adjacent) of G such that for each vertex v, given an independent set I drawn from the distribution, $\mathbb{P}(v \in I) \geq \frac{1}{k}$. Let $\mathcal{I}(G)$ be the set of all independent sets of G, and $\mathcal{I}(G, x)$ be the set of all those independent sets which include vertex x, and $x_I \in \mathbb{R}^+$ for each independent set I. Then, the fractional chromatic number $\chi_f(G)$ can be obtained as a solution of the following linear program [31]:

$$\chi_f(G) = \min_{\forall x} \Big\{ \sum_{I\in\mathcal{I}(G)} x_I : \sum_{I\in\mathcal{I}(G,x)} x_I \geq 1, \ x_I \geq 0 \Big\} \ . \qquad (5)$$

This relaxation transforms traditional coloring, which is an integer programming problem (NP-complete), into a fractional coloring problem. We illustrate fractional coloring for Example 1 in Fig. 1 (right). A 5:2 coloring achieves the optimal solution of (5), where $|\mathcal{I}(G)| = 10$, with 5 of those sets have cardinality 2, and 5 sets have size 1, and $|\mathcal{I}(G, x)| = 3$ such that $x_I = 0.5$ for sets of size 2 and $x_I = 0$ for the sets with size 1. Hence, $\chi_b(G_{X_1}) = 5$ and from (4) $\chi_f(G_{X_1}) = 2.5$.

Example 1 indicates that assigning colors to sufficiently large power graphs, we can compress X_1 more. A $5 : 2$ coloring yields $H^{\chi_f}_{G_{X_1}}(X_1|X_2) = 1.16$, providing a saving of 0.36 bits over $H^{\chi}_{G_{X_1}}(X_1|X_2) = 1.52$. We sketch the traditional coloring for $G^2_{\mathbf{X_1}}$ in Fig. 2. The distribution of colors satisfies $P(c_k) = 0.16$, $k \in \{1, \dots, 5\}$ and $P(c_k) = 0.08$, $k \in \{6, 7\}$ and $P(c_8) = 0.04$. Hence, $H^{\chi}_{G^2_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2}) = 1.44$. For a $13 : 2$ coloring (Fig. 3), the colors satisfy $P(c_k) = 0.08$, $k \in \{1, \dots, 12\}$ and $P(c_{13}) = 0.04$. Hence, $H^{\chi_f}_{G^2_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2}) = 0.92$, i.e., 0.52 bits of savings from $H^{\chi}_{G^2_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2})$.

Due to space constraints, we next focus on the theoretical notions and the achievable gains pertaining to fractional coloring. We explain an example code construction in [33] to demonstrate the practical savings of the proposed model.

III. Fractional Chromatic Entropy

We next formalize the notion of fractional chromatic entropy of a set of valid fractional colorings via extending Defn. 1.

Definition 4. *(Fractional chromatic entropy.)* $c^f_{G_{X_1}}(X_1)$ *is a valid* $a:b$ *fractional coloring of* G_{X_1} *if it assigns* b *colors out of a total of* a *available colors to each* V_{X_1} *such that adjacent vertices have disjoint colors. We define* $\mathcal{H}^{\chi_f}(c^f_{G_{X_1}}(X_1)) = \{H(c^f_{G_{X_1}}(X_1)) : c^f_{G_{X_1}}(X_1)$ *is a valid a:b coloring of* $G_{X_1}\}$ *to be the collection of fractional chromatic entropies over the set of all valid* $a:b$ *colorings of* G_{X_1} *given* X_2.

Proposition 1. *The fractional chromatic entropy of a characteristic graph* G_{X_1}*, denoted by* $H^{\chi_f}_{G_{X_1}}(X_1|X_2)$*, is given as*

$$H^{\chi_f}_{G_{X_1}}(X_1|X_2) = \inf_b \frac{1}{b} \min_{c^f_{G_{X_1}}} \mathcal{H}^{\chi_f}(c^f_{G_{X_1}}(X_1)) \ . \tag{6}$$

Proof. Let $(Z_1, Z_2, \ldots, Z_n)$ be a collection of random variables. For every $S \subseteq \{1, 2, \ldots, n\}$, denote by $Z(S)$ the subset $\{Z_i : i \in S\}$. From [34, Ch. 16.5] the average entropy in bits per symbol of a randomly drawn b-element subset $Z(S)$ is

$$\frac{1}{\binom{n}{b}} \sum_{S:|S|=b} \frac{H(Z(S))}{b} \ . \tag{7}$$

Let $G_{X_1(S)} = \{G_{X_{1i}} : i \in S\}$ be an $b = |S|$-tuple of graphs where each $G_{X_{1i}}$ is a replica of G_{X_1}. We jointly color $G_{X_1(S)}$ such that $c_{G_{X_1(S)}}(X_1(S)) = \{c_{G_{X_{1i}}}(X_{1i}) : i \in S\}$. Using a colors in total and i.i.d. valid colorings across disjoint S, the entropy of a randomly drawn b-element subset of colorings is

$$H(c_{G_{X_1(S)}}(X_1(S))|\mathbf{X}_2) = H(c^f_{G_{X_1}}(X_1)|\mathbf{X}_2) \ , \tag{8}$$

where $c^b_{G_{X_1}}(X_1)$ is a valid coloring of $G_{X_1(S)}$, and equivalently $c^f_{G_{X_1}}(X_1)$ is a valid $a:b$ coloring of G_{X_1}.

The average entropy decreases monotonically in the size of the subset (Han [28]). As b increases the rate of functional compression via fractional coloring decreases. The minimum entropy of a fractional coloring can be found by minimizing across all valid $a:b$ colorings of G_{X_1}. Observing from (4) that $\chi_f(G_{X_1}) = \lim_{b\to\infty} \chi_b(G_{X_1})/b$ and (8), we obtain (6). ☐

To visualize the rate given in Prop. 1, consider Example 1. The $5:2$ coloring distribution satisfies $P(c_1) = P(c_2) = 2/5$ and $P(c_3) = 1/5$ and $P(c_4) = P(c_5) = 2/5$ (c_3 is repeated in $G_{X_{11}}$ and $G_{X_{12}}$). The distribution of $c_{G_{X_1(S)}}(X_1(S))$ across 2 graphs yields $H(c^f_{G_{X_1}})/2 = 1.16 < H(c_{G_{X_1}}) \approx 1.52$.

The following is an intuitive result due to a finer-grained granularity that the fractional graph coloring provides. Its proof follows from combining the definition in (6) and (2).

Corollary 1. *The fractional chromatic entropy of a graph* G_{X_1} *and chromatic entropy satisfy the following relation:*

$$H^{\chi_f}_{G_{X_1}}(X_1|X_2) \leq H^{\chi}_{G_{X_1}}(X_1|X_2) \ . \tag{9}$$

Exploiting [1] the fractional graph entropy satisfies

$$H^f_{G_{X_1}}(X_1|X_2) = \lim_{n\to\infty} \frac{1}{n} H^{\chi_f}_{G^n_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2}) \ , \tag{10}$$

where χ_f is the fractional chromatic number of $G^n_{\mathbf{X_1}}$.

Using (6) and (10) we can derive the fractional graph entropy, a natural generalization of (3). Prop. 2 is derived from Prop. 1 and (10), and we skip its proof.

Proposition 2. *The fractional graph entropy is given as*

$$H^f_{G_{X_1}}(X_1|X_2) = \lim_{n\to\infty} \frac{1}{n} \inf_b \frac{1}{b} \min_{c^f_{G^n_{\mathbf{X_1}}}} \{H(c^f_{G^n_{\mathbf{X_1}}}(\mathbf{X_1})) : c^f_{G^n_{\mathbf{X_1}}}(\mathbf{X_1}) \text{ is a valid a:b coloring of } G^n_{\mathbf{X_1}}|\mathbf{X}^n_2\} \ , \tag{11}$$

where $c^f_{G^n_{\mathbf{X_1}}}(\mathbf{X_1})$ *is a fractional coloring variable that assigns* b *colors to each vertex of* $G^n_{\mathbf{X_1}}$ *out of* $a \geq b$ *available colors.*

Lemma 1. *The following relation holds for the fractional graph entropy and graph entropy:*

$$H^f_{G_{X_1}}(X_1|X_2) \leq H_{G_{X_1}}(X_1|X_2) \ . \tag{12}$$

Proof. We use an important result that ties the fractional chromatic number to the n-th power of a graph.

The following relation between the n-th power of G and the fractional chromatic number holds [31, Corollary 3.4.3]:

$$\chi_f(G) = \inf_n \sqrt[n]{\chi(G^n)} = \lim_{n\to\infty} \sqrt[n]{\chi(G^n)} \ . \tag{13}$$

The relation (13) implies that $\chi(G^n) \approx \chi_f(G)^n$ as $n \to \infty$.

It also holds from [31, Corollary 3.4.2] that

$$\chi_f(G^n) = \chi_f(G)^n \ . \tag{14}$$

As a result of (13) and (14), we infer for the m-th power of G that $\chi_f(G^m) \overset{(14)}{=} \chi_f(G)^m \overset{(13)}{=} (\lim_{n\to\infty} \sqrt[n]{\chi(G^n)})^m$, and $\chi_f(G^m) \overset{(13)}{=} \lim_{n\to\infty} \sqrt[n]{\chi(G^{n\cdot m})} \leq \lim_{n\to\infty} \chi(G^n)^{\frac{m}{n}}$. The fractional coloring requires $\log \chi_f(G^n)$ bits which is less than $\log \chi(G^n)$ bits as required by the traditional coloring. ☐

The current paper aims to improve the compression rate by introducing fractional chromatic entropy. On the other hand, this approach does not outperform the independent set-based fundamental limit for graph entropy that establishes the optimal rate for lossless function computation $f(X_1, X_2)$ given side information X_2 [35, Theorem 21.2].

IV. Coding Gains of Fractional Coloring

We denote the integrality gap (IG), i.e., ratio of the solutions of the traditional coloring versus the fractional coloring problems for encoding $G^n_{\mathbf{X_1}}$, by IG_n. It is given as

$$IG_n = \frac{H^{\chi}_{G^n_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2})}{H^{\chi_f}_{G^n_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2})} \ , \quad n \geq 1 \ .$$

From Lemma 1, it is immediate that $IG_n \geq 1$ for $n \geq 1$.

Lemma 2. IG_n *is an increasing function of* n.

Proof. We rewrite IG_n as $IG_n = \frac{1}{n}\sum_{k=1}^n a_k \,/\, \frac{1}{n}\sum_{k=1}^n b_k$, where $a_k = H^{\chi}_{G^k_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2}) - H^{\chi}_{G^{k-1}_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2})$ and $b_k = H^{\chi_f}_{G^k_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2}) - H^{\chi_f}_{G^{k-1}_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2})$ which both decrease in k because the sequences $H^{\chi}_{G^k_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2})$ and $H^{\chi_f}_{G^k_{\mathbf{X_1}}}(\mathbf{X_1}|\mathbf{X_2})$ are increasing and concave in the sense of decreasing slope, which can be shown using Han's theorem [28]. Furthermore, $a_k \geq b_k$ for $k \in \{1, 2, \ldots, n\}$. Due to fractional coloring

b_k decreases with a higher rate versus a_k. We infer that $1 \leq a_1/b_1 \leq a_2/b_2 \leq \dots a_n/b_n$. Hence, we can show that

$$\frac{a_1}{b_1} \leq \frac{a_1+a_2}{b_1+b_2} \leq \dots \leq \frac{a_1+a_2+\dots+a_n}{b_1+b_2+\dots+b_n} = IG_n \ .$$

Hence, IG_n is an increasing function of n. □

In Example 1, $IG_1 = \frac{1.52}{1.16} = 1.31$, $IG_2 = \frac{1.44}{0.92} = 1.57$, and IG_n is higher for $n > 2$ (Lemma 2). The gain is due to cross-coding across graphs (Prop. 1). For the identity function, the G_{X_1} is complete, $b^* = 1$, and fractional coloring does not have savings. Significant gains are possible for sparse graphs.

For a valid a:b coloring of G_{X_1}, let $b^*_{G_{X_1}}$ be the smallest $b = |S|$ that achieves (6) for G_{X_1}, and $c_{G_{X_1}}(X_1)$ and $c^f_{G_{X_1}}(X_1)$ be the valid colorings with distributions $\mathbf{q} = (q_1, \dots, q_{\chi(G_{X_1})})$ and $\mathbf{r} = (r_1, \dots, r_{\chi_{b^*_{G_{X_1}}}(G_{X_1(S)})})$ that minimize the respective entropies of the colorings. Similarly, for $n > 1$, let $b^*_{G^n_{\mathbf{X_1}}}$ be the smallest $b = |S|$ for $G^n_{\mathbf{X_1}}$ satisfying (4), $c_{G^n_{\mathbf{X_1}}}(\mathbf{X_1})$ and $c^f_{G^n_{\mathbf{X_1}}}(\mathbf{X_1})$ be the valid colorings with $\mathbf{q}^n = (q_1, \dots, q_{\chi(G^n_{\mathbf{X_1}})})$, $\mathbf{r}^n = (r_1, \dots, r_{\chi_{b^*_{G^n_{\mathbf{X_1}}}}(G^n_{\mathbf{X}_1(S)})})$.

Proposition 3. *The fractional coloring scheme attains*

$$IG_n \geq \frac{b^*_{G^n_{\mathbf{X_1}}} H(\mathbf{q}^n)}{H(\mathbf{q}^n) + \Delta_{G^n}} \ ,$$

where

$$\Delta_{G^n} = \sum_{j \in \mathcal{J}_{G^n}} q_j \Big[h\Big(\frac{1}{b^*_{G^n_{\mathbf{X_1}}}}\Big) + \frac{b^*_{G^n_{\mathbf{X_1}}} - 1}{b^*_{G^n_{\mathbf{X_1}}}} \log(m_{G^n}(j)(b^*_{G^n_{\mathbf{X_1}}} - 1)) \Big],$$

$j \in \mathcal{J}_{G^n} = \{1, \dots, \chi_b(G^n_{\mathbf{X}_1(S)})\}$ *represents the coloring class, and* $m_{G^n}(j)$ *is the count of class* j *vertices.*

Proof. We first focus on $n = 1$. Then,

$$IG_1 = b^*_{G_{X_1}} \frac{H(c_{G_{X_1}}(X_1)|X_2)}{H(c^f_{G_{X_1}}(X_1)|X_2)} = b^*_{G_{X_1}} \frac{H(\mathbf{q})}{H(\mathbf{r})} \ .$$

Using the grouping property of entropy [34, Ch.2],

$$H(\mathbf{r}) = H(\mathbf{q}) + \sum_j q_j H\Big((\frac{r_l}{q_j} : \sum\nolimits_{l \in j} r_l = q_j) \Big) \ . \tag{15}$$

We next analyze the RHS of (15) for a given j, assuming that there exist $m_G(j)$ vertices in G_{X_1} in color class j. It then holds that in the b-fold coloring scheme q_j accumulates the probabilities of $m_G(j) \times b$ vertices in $G_{X_1(S)}$. The marginal distribution of colors in each $G_{X_{1i}} : i \in S$ is identical. Hence,

$$H\Big((\frac{r_l}{q_j} : \sum_{l \in j} r_l = q_j) \Big) \leq H\Big(\frac{1}{m_G(j)b}, \dots, \frac{1}{m_G(j)b}, \frac{1}{b} \Big) = h\Big(\frac{1}{b}\Big) + \frac{b-1}{b} \log(m_G(j)(b-1)) \ ,$$

where the colors of $G_{X_{11}}$ and G_{X_1} are the same, putting $1/b$ of the mass of q_i in $G_{X_1(S)}$, and leaving $(b-1)/b$ of the mass to the remaining $b-1$ graph replicas $\{G_{X_{1i}} : i \in S, i \neq 1\}$ with $m_G(j)(b-1)$ vertices. The RHS holds when the colors are uniformly split among $m_G(j)(b-1)$ vertices. Hence,

$$IG_1 \geq \frac{b^*_{G_{X_1}} H(\mathbf{q})}{H(\mathbf{q}) + \Delta_G} \ ,$$

where $\Delta_G = \sum_{j \in \mathcal{J}_G} q_j \Big[h\Big(\frac{1}{b^*_{G_{X_1}}}\Big) + \frac{b^*_{G_{X_1}} - 1}{b^*_{G_{X_1}}} \log(m_G(j)(b^*_{G_{X_1}} - 1)) \Big]$.

For $n > 1$, employing the grouping property we can obtain the final result. We note that the count of class j vertices $m_{G^n}(j)$ is less than $m_G(j)^n$ for $j \in \mathcal{J}_G = \{1, \dots, \chi(G_{X_1})\}$. □

From Prop. 3, for any given G_{X_1}, if not a complete graph, fractional coloring for $n > 1$ offers compression savings.

The following corollary implies theoretically that the achievable IG_n is lower bounded by $b^*_{G^n_{\mathbf{X_1}}}$ (up to a scaling), for a valid $b^*_{G^n_{\mathbf{X_1}}}$-fold coloring that is determined by the function to be computed and its n-th power graph $G^n_{\mathbf{X_1}}$.

Corollary 2. *Under the assumption that the colorings of* $G^n_{\mathbf{X_1}}$ *are uniform, the fractional coloring scheme attains*

$$IG_n \geq b^*_{G^n_{\mathbf{X_1}}} \cdot \log \chi_f(G_{X_1}) \Big/ \log \chi_{b^*_{G_{X_1}}}(G_{X_1}) \ .$$

Proof. Let $b^*_{G_{X_1}}$ and $b^*_{G^n_{\mathbf{X_1}}}$ be the smallest b values such that (4) holds for G_{X_1} and $G^n_{\mathbf{X_1}}$, respectively. Then, provided that the colorings of $G^n_{\mathbf{X_1}}$ are uniform, we can simplify the expressions for $H(\mathbf{q})$ and $H(\mathbf{r})$ and obtain

$$IG_n \geq \frac{\frac{1}{n} \log \chi_f(G_{X_1})^n}{\frac{1}{n} \cdot \frac{1}{b^*_{G^n_{\mathbf{X_1}}}} \cdot \log(b^*_{G^n_{\mathbf{X_1}}} \cdot \chi_f(G^n_{\mathbf{X_1}}))} \tag{16}$$

$$= b^*_{G^n_{\mathbf{X_1}}} \cdot \frac{\log \chi_f(G_{X_1})}{\log (b^*_{G^n_{\mathbf{X_1}}})^{1/n} + \log \chi_f(G_{X_1})} \tag{17}$$

$$\geq b^*_{G^n_{\mathbf{X_1}}} \cdot \log \chi_f(G_{X_1}) \Big/ \log \chi_{b^*_{G_{X_1}}}(G_{X_1}) \ , \tag{18}$$

where the lower bound (16) follows from using (13), and (17) is due to (14). Employing the relations $\chi_f(G_{X_1}) = \frac{\chi_b(G_{X_1})}{b}$, (14), which yields $\chi_f(G^n_{\mathbf{X_1}}) = \frac{\chi^n_b(G_{X_1})}{b^n}$, and $\chi^n_b(G_{X_1}) \geq \chi_b(G^n_{\mathbf{X_1}})$, we obtain $b^*_{G^n_{\mathbf{X_1}}} \leq (b^*_{G_{X_1}})^n$, which yields (18). □

Discussion and future directions. Fractional coloring exploits the possibility of cross-coding between graphs and provides coding gains (at infinite and finite source sequence lengths). This approach provides a reduced communication complexity versus traditional coloring via decreasing the number of bits to send roughly from $\log \chi$ to $\log \chi_f$. Lower bounding IG_n for different classes of functions and source distributions is of primary importance. Quantifying and upper bounding the IG in the limit of large n, i.e., the ratio of the bits required by fractional compression, $H^f_{G_{X_1}}(X_1|X_2)$, to $H_{G_{X_1}}(X_1|X_2)$, is left as future work.

While fractional coloring is a less combinatorial problem than traditional coloring and accepts a linear programming solution (solvable in polynomial time), finding an independent set is strongly NP-hard [36]. Hence, the relaxation is in the class of NP-hard problems. This issue can be alleviated via using *fractional edge coloring* (versus *fractional vertex coloring*) for which a polynomial-time solution exists [31]. The characteristic graph approach is concerned with functional compression of source sequences when the adjacency matrix is a (0,1)-matrix. A possible generalization includes edge-weighted graphs to capture the distortion in reconstruction.

ACKNOWLEDGMENT

The author is indebted to Dr. A. Beirami for the invaluable conversations and the constructive feedback on the paper.

References

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

[2] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

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

[4] 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.

[5] M. Pawlak, E. Rafajłowicz, and A. Krzyzak, "Postfiltering versus prefiltering for signal recovery from noisy samples," *IEEE Trans. Inf. Theory*, vol. 49, pp. 3195–3212, Dec. 2003.

[6] M. Sartipi and F. Fekri, "Distributed source coding using short to moderate length rate-compatible LDPC codes: the entire Slepian-Wolf rate region," *IEEE Trans. Commun.*, vol. 56, no. 3, pp. 400–411, Mar. 2008.

[7] C. Zhang, H. Fu, Q. Hu, X. Cao, Y. Xie, D. Tao, and D. Xu, "Generalized latent multi-view subspace clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 1, pp. 86–99, Oct. 2018.

[8] 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, Oct. 2006.

[9] R. Ahlswede, N. Cai, S. Y. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–16, Jul. 2000.

[10] R. W. Yeung, *A First Course in Information Theory*. Springer Science & Business Media, 2002.

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

[12] A. Gamal and T. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 851–857, Nov. 1982.

[13] 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.

[14] S. Watanabe, "The rate-distortion function for product of two sources with side-information at decoders," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5678–5691, Jun. 2013.

[15] 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.

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

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

[18] V. Doshi, D. Shah, M. Médard, and S. Jaggi, "Graph coloring and conditional graph entropy," in *Proc., IEEE Asilomar Conf. Signals, Systems and Computers*, Oct. 2006, pp. 2137–2141.

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

[20] V. Doshi, D. Shah, M. Médard, and M. Effros, "Functional compression through graph coloring," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3901–3917, Jul. 2010.

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

[22] P. Delgosha and V. Anantharam, "Universal lossless compression of graphical data," *IEEE Trans. Inf. Theory*, vol. 66, no. 11, pp. 6962–6976, May 2020.

[23] C. T. Li and V. Anantharam, "A unified framework for one-shot achievability via the Poisson matching lemma," in *Proc., IEEE Int. Symp. Inf. Theory*, Paris, France, Jul. 2019, pp. 942–946.

[24] A. Magner, K. Turowski, and W. Szpankowski, "Lossless compression of binary trees with correlated vertex names," *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6070–6080, Jun. 2018.

[25] M. Dehmer and A. Mowshowitz, "A history of graph entropy measures," *Information Sciences*, vol. 181, no. 1, pp. 57–78, Jan. 2011.

[26] H. Feng, M. Effros, and S. Savari, "Functional source coding for networks with receiver side information," in *Proc., Allerton Conference on Communication, Control, and Computing*, Sep. 2004, pp. 1419–1427.

[27] C. Shalizi, "Graph entropy," Oct. 2010, lecture 4: CSE533: Information Theory in Computer Science. [Online]. Available: https://homes.cs.washington.edu/~anuprao/pubs/CSE533Autumn2010/lecture4.pdf

[28] T. S. Han, "Nonnegative entropy measures of multivariate symmetric correlations," *Information and Control*, vol. 36, pp. 133–156, Feb. 1978.

[29] A. Rao and A. Yehudayoff, *Communication Complexity: and Applications*. Cambridge University Press, 2020.

[30] C.-C. Y. Andrew, "Some complexity questions related to distributed computing," in *Proc., ACM Symp. Theory of Computing (STOC)*, Atlanta, GA, Apr. – May 1979, pp. 209–213.

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

[32] J. Cardinal, S. Fiorini, and G. Joret, "Tight results on minimum entropy set cover," *Algorithmica*, vol. 51, no. 1, pp. 49–60, May 2008.

[33] D. Malak, "Fractional graph coloring for functional compression with side information," *arXiv preprint arXiv:2204.11927*, Apr. 2022.

[34] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.

[35] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[36] M. R. Garey and D. S. Johnson, ""Strong" NP-completeness results: Motivation, examples, and implications," *Journal of the ACM*, vol. 25, no. 3, pp. 499–508, Jul. 1978.