

Edge removal in undirected networks

Michael Langberg

Michelle Effros

Abstract—The edge-removal problem asks whether the removal of a λ -capacity edge from a given network can decrease the communication rate between source-terminal pairs by more than λ . We prove that for undirected networks, removing a λ capacity edge decreases the rate by $O(\lambda)$. Through previously known reductive arguments, here newly applied to undirected networks, our result implies that the zero-error capacity region of an undirected network equals its vanishing-error capacity region. Whether it is possible to prove similar results for directed networks remains an open question.

I. INTRODUCTION

The *edge removal problem*, defined and studied in [2], [3], aims to quantify the loss in capacity that results from the removal of a single edge (i.e., a point-to-point channel) from a given network coding instance. For some network coding instances, it is known that the removal of an edge of capacity λ can decrease the rate of communication for each source-receiver pair by at most λ [2], [3]. These instances include networks with colocated sources, networks in which we are restricted to perform linear encoding, networks in which the edges removed are connected to terminals with no outgoing edges, as well as other families of network coding instances. However, whether the removal of an edge of capacity λ decreases the rate of communication for each source-receiver pair by at most λ for *any* network coding instance remains an intriguing open problem connected to a spectrum of (at times seemingly unrelated) questions in the context of network communication (see, e.g., [4], [5], [6], [7], [8], [9], [10], [11]).

In this work, we study the edge removal problem on *undirected* networks. In an undirected network, the information on any edge $e = (u, u')$, can travel from u to u' and/or from u' to u , as long as the sum of the rates in both directions do not exceed the edge capacity. Undirected networks have seen several studies in the context of network coding (e.g., [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24]). To date, the arguably most well known open question regarding network coding in undirected networks concerns the maximal potential benefit in rate that one can obtain in multiple-unicast instances when comparing communication via network coding to communication without coding, i.e., the *coding advantage* in undirected multiple-unicast networks. It is conjectured in [12], [13], [14] that no such advantage exists. This conjecture has been confirmed on several special cases

(e.g., [13], [16], [18]) but remains an open question in full generality. We note that, if true, the conjecture will imply that the removal of a λ capacity edge from a multiple-unicast undirected network will decrease the communication rate of each unicast connection by at most λ .

Our work is structured as follows. In Section II, we present our model and define a number of statements regarding the edge-removal problem. Our main results and analysis are given in Section III. In Theorem 1, we show that for any undirected network coding instance \mathcal{I} there exists a constant c such that the removal of an edge of capacity λ from \mathcal{I} reduces the rate between source-terminal pairs by at most $c\lambda$. We then derive two immediate corollaries to Theorem 1. In Corollary 1, we prove the so-called asymptotic edge removal statement on undirected instances. This statement asserts that removing an edge of negligible capacity has a negligible effect on the rate between source-terminal pairs. Building on prior work [6], in Corollary 2 we prove that the zero-error capacity region and the vanishing-error capacity region of undirected network coding instances are equal.

II. MODEL

Throughout the paper, the size of a finite set S is denoted by $|S|$. For any positive real k , $[k]$ denotes the set $\{1, \dots, [k]\}$. We use bold letters to denote vectors; for example, $\mathbf{R} = (R_1, \dots, R_k)$ is a vector of dimension k and R_i is the i^{th} element of vector \mathbf{R} . We define $\mathbf{R} - \gamma$ as $((R_1 - \gamma)^+, \dots, (R_k - \gamma)^+)$ where $(R - \gamma)^+ = \max\{0, R - \gamma\}$. For $\alpha > 0$ and a set \mathcal{R} of real vectors, the set $\alpha\mathcal{R}$ refers to the set obtained by multiplying each vector in \mathcal{R} by α .

A. Network Coding Instances and Network Codes

An undirected instance $\mathcal{I} = (G, S, D, M)$ of the network coding problem includes an undirected network $G = (V, E)$, a vector of k source nodes $S = (s_1, \dots, s_k) \in V^k$, a vector of terminal nodes $D = (d_1, \dots, d_r) \in V^r$, and a binary requirement matrix $M = [m_{ij}]$ in which $m_{ij} = 1$ if and only if the message of source s_i is requested by terminal d_j . Source node $s_i \in S$ holds message random variable W_i demanded by terminals $\{d_j \mid m_{ij} = 1\}$. Each edge $e \in E$ has an associated capacity λ_e .

We here assume that communication occurs in N rounds, and in each round, every edge $e \in E$ carries a message over an alphabet \mathcal{X}_e^n of size $2^{\lambda_e n}$. We call N the *outer blocklength* and n the *inner blocklength*. Namely, we think of communication over edge e in terms of symbols over the alphabet \mathcal{X}_e^n corresponding to n channel uses. We thus use the term *time-step* to refer to each round of communication.

M. Langberg is with the Department of Electrical Engineering at the University at Buffalo (State University of New York). Email: mikel@buffalo.edu

M. Effros is with the Department of Electrical Engineering at the California Institute of Technology. Email: effros@caltech.edu

This work is supported in part by NSF grants CCF-1817241 and CCF-1909451. The full version of this work appears in [1].

More formally, for an outer blocklength N , and an inner blocklength n , network code

$$(\mathcal{F}, \mathcal{G}) = (\{\overrightarrow{f_{e,t}}\}, \{\overleftarrow{f_{e,t}}\}, \{g_j\})$$

is an assignment of encoding functions $\{\overrightarrow{f_{e,t}}\}$ and $\{\overleftarrow{f_{e,t}}\}$ for every time step $t \in [N]$ and each edge $e \in E$ and a decoding function g_j to each terminal $d_j \in D$. At each time step t and for each edge $e = (u, u')$ the alphabet \mathcal{X}_e^n is represented by two sets

$$\overrightarrow{\mathcal{X}_{e,t}^n} \text{ and } \overleftarrow{\mathcal{X}_{e,t}^n} \text{ such that } |\overrightarrow{\mathcal{X}_{e,t}^n}| \cdot |\overleftarrow{\mathcal{X}_{e,t}^n}| \leq |\mathcal{X}_e^n|.$$

At each time step t and for each edge $e = (u, u')$ the edge message $\overrightarrow{X_{e,t}^n} \in \overrightarrow{\mathcal{X}_{e,t}^n}$ from u to u' and the message $\overleftarrow{X_{e,t}^n} \in \overleftarrow{\mathcal{X}_{e,t}^n}$ from u' to u are equal to the evaluation of encoding functions $\{\overrightarrow{f_{e,t}}\}$ and $\{\overleftarrow{f_{e,t}}\}$ on inputs $X_{\text{In}(u),[t-1]}^n$ and $X_{\text{In}(u'),[t-1]}^n$, respectively. Here, for a generic node u_0 , and time t ,

$$X_{\text{In}(u_0),[t]}^n = (\overrightarrow{X_{e',t}^n} : e' = (v, u_0) \in E, t' \leq t), (W_i : u_0 = s_i)$$

captures all information available to node u at time t . The evaluation of decoding function g_j on the vector of random variables $X_{\text{In}(d_j),[N]}^n$ equals the reproduction of message random variables $(W_i : m_{i,j} = 1)$ requested at terminal node $d_j \in D$.

Suppose that we are given rate vector $\mathbf{R} = (R_1, \dots, R_k)$, constant $\varepsilon \in [0, 1]$, and positive integers n, N . Instance \mathcal{I} of the network coding problem is said to be $(\varepsilon, \mathbf{R}, n, N)$ -feasible if for W_i uniformly distributed over $[2^{R_i N n}]$ (for $i \in [k]$) there exists a network code $(\mathcal{F}, \mathcal{G})$ with inner-blocklength n and outer-blocklength N such that, with probability at least $1 - \varepsilon$, for each $d_j \in D$ the output of decoding function g_j equals $(W_i : m_{i,j} = 1)$.

Definition 1 (Capacity region): The capacity region of \mathcal{I} , denoted by $\mathcal{R}(\mathcal{I})$, is the set of all rate vectors \mathbf{R} such that for all $\varepsilon > 0$ and all $\Delta > 0$ there exist infinitely many blocklengths n and infinitely many blocklengths N such that \mathcal{I} is $(\varepsilon, \mathbf{R} - \Delta, n, N)$ -feasible.

Definition 2 (Zero-error capacity region): The zero-error capacity region of \mathcal{I} , denoted by $\mathcal{R}_0(\mathcal{I})$, is the set of all rate vectors \mathbf{R} such that for all $\Delta > 0$ there exist infinitely many blocklengths n and infinitely many blocklengths N such that \mathcal{I} is $(0, \mathbf{R} - \Delta, n, N)$ -feasible.

Some remarks are in place. For directed acyclic networks, our Definitions 1 and 2, which use both inner and outer blocklengths, are equivalent to the standard definitions of capacity, e.g., [25], in which for a single blocklength parameter \tilde{n} , each edge e of capacity λ_e can communicate a message in $[2^{\lambda_e \tilde{n}}]$. In this equivalence, the blocklength \tilde{n} equals the product Nn .

Our notion of inner and outer blocklengths stems from two aspects of cyclic networks. Primarily, given the cyclic dependence of information flowing through the network, communication is often defined in rounds, in which each round of communication depends on the information obtained through



Fig. 1: A schematic description of Theorem 1. The edge $e = (u, u')$ is marked as a dotted line. The first case (a) in which the graph G of instance \mathcal{I} is disconnected and adding the edge $e = (u, u')$ connects between two components U and U' . The second case (b) in which there exists in G a path connecting u and u' . Here, the path is $u, u_2, u_3, \dots, u_7, u'$ and γ is the value of the minimum capacity edge (u_4, u_5) in the path.

previous rounds. Hence we employ the outer blocklength N . Secondly, to accommodate networks with edge capacities λ_e for which $2^{\lambda_e} < 1$ (equivalently, $\lfloor 2^{\lambda_e} \rfloor = 0$), e.g., the bounding model for a binary symmetric channel from [26], we consider communication over *sub-rounds* in which outgoing edge messages are aggregated over an inner blocklength of size n . The rate R is normalized by the product Nn .

Operationally speaking, our notion of inner and outer blocklengths governs the cyclic dependence of coding operations over time, where for the inner-blocklength n the cyclic dependence is temporarily broken. Therefore, our definitions imply tradeoffs between the outer-blocklength N and inner-blocklength n . For example, if all edge capacities are integral, then any network code that is $(\varepsilon, \mathbf{R} - \Delta, n, N)$ -feasible is also $(\varepsilon, \mathbf{R} - \Delta, 1, nN)$ -feasible, but the other direction does not necessarily hold.

We now address two technical lemmas that are useful in our analysis (both proven in the full version of this work [1]).

Lemma 1 ([6]): Let $\mathcal{I} = (G, S, D, M)$. Let $\mathbf{R} \in \mathcal{R}(\mathcal{I})$. Then for any $\Delta > 0$ there exist infinitely many blocklengths n and infinitely many blocklengths N such that \mathcal{I} is $(\varepsilon, \mathbf{R} - \Delta, n, N)$ -feasible with $\varepsilon \leq 1/\max(n^2, N^2)$.

Lemma 2: Let $\mathcal{I} = (G, S, D, M)$. Let $\alpha > 0$, and define $\alpha\mathcal{I} = (\alpha G, S, D, M)$ to be the instance obtained by multiplying each edge capacity in G by α (to obtain a graph here described as αG). Then $\alpha\mathcal{R}(\mathcal{I}) = \mathcal{R}(\alpha\mathcal{I})$ and $\alpha\mathcal{R}_0(\mathcal{I}) = \mathcal{R}_0(\alpha\mathcal{I})$.

B. Edge Removal

Throughout the discussions in this work, we use the term “edge-removal statement,” often shortened to “edge removal,” to refer to the mathematical statement defined here. Let $\mathcal{I} = (G, S, D, M)$. Let $G^{\lambda,e}$ be the graph obtained by adding an (undirected) edge e of capacity $\lambda > 0$ to G .¹ Let $\mathcal{I}^{\lambda,e} = (G^{\lambda,e}, S, D, M)$ describe the resulting network instance. The edge-removal statements given below compare the rate vectors achievable over \mathcal{I} and $\mathcal{I}^{\lambda,e}$. We use notation stemming from [11] to define the following variants of edge removal.

¹Instead of starting with a network and then *removing* an edge as in [2], [3], it is more convenient for our presentation to start with a network and then *add* an edge as in [11].

Statement 1 (The edge-removal statements on instance \mathcal{I}):

- The *edge-removal statement* holds with function $f_{\mathcal{I}}(\lambda)$ on instance \mathcal{I} and edge capacity λ if for any edge $e \in V \times V$: $\mathbf{R} \in \mathcal{R}(\mathcal{I}^{\lambda,e}) \implies \mathbf{R} - f_{\mathcal{I}}(\lambda) \in \mathcal{R}(\mathcal{I})$.
- The *zero-error edge-removal statement* holds with function $f_{\mathcal{I}}(\lambda)$ on instance \mathcal{I} and edge capacity λ if for any edge $e \in V \times V$: $\mathbf{R} \in \mathcal{R}_0(\mathcal{I}^{\lambda,e}) \implies \mathbf{R} - f_{\mathcal{I}}(\lambda) \in \mathcal{R}_0(\mathcal{I})$.
- The *vanishing-edge-removal statement* holds on instance \mathcal{I} if for any edge $e \in V \times V$: $\mathcal{R}(\mathcal{I}) = \lim_{\lambda \rightarrow 0} \mathcal{R}(\mathcal{I}^{\lambda,e})$.
- The *zero-error vanishing-edge-removal statement* holds on instance \mathcal{I} if for any edge $e \in V \times V$: $\mathcal{R}_0(\mathcal{I}) = \lim_{\lambda \rightarrow 0} \mathcal{R}_0(\mathcal{I}^{\lambda,e})$.

III. MAIN RESULTS

We now present the main result of this work.

Theorem 1: Let $\mathcal{I} = (G, S, D, M)$ be an undirected network instance. Let $\lambda > 0$. The edge-removal statement and the zero-error edge-removal statement hold with function $f_{\mathcal{I}}(\lambda) = c\lambda$ for some constant c that depends only on the edge capacities in \mathcal{I} .

The proof of Theorem 1 applies the follow lemma.

Lemma 3: Let $\mathcal{I} = (G, S, D, M)$. Let $e = (u, u')$ be an edge in G of capacity λ . Let $\mathcal{I}_{path} = (G_{path}, S, T, M)$ be the instance obtained by modifying graph G to yield a new graph G_{path} as follows. Starting with G , we remove the edge $e = (u, u')$ and replace it with a path $u = u_1, u_2, u_3, \dots, u_{\ell-1}, u_{\ell} = u'$ of length $\ell - 1$ and capacity λ , where u_i for $i = 2, 3, \dots, \ell - 1$ are nodes in G . Here, edges (u_i, u_{i+1}) that appeared in the original graph G will increase their capacity in G_{path} by λ and edges (u_i, u_{i+1}) that are not originally in G will appear in G_{path} with capacity λ . Then $\mathcal{R}(\mathcal{I}) \subseteq \mathcal{R}(\mathcal{I}_{path})$ and $\mathcal{R}_0(\mathcal{I}) \subseteq \mathcal{R}_0(\mathcal{I}_{path})$.

We start by proving Theorem 1 using Lemma 3. We then prove Lemma 3. The proof of Theorem 1 uses the conceptually simple idea of *rerouting* the information of the removed edge e on an alternative path as expressed in Lemma 3. Such rerouting gives rise to a number of technical challenges, resulting from the notion of delay and the time varying nature of $\mathcal{X}_{e,t}^n$ and $\overleftarrow{\mathcal{X}}_{e,t}^n$, which are addressed in detail below.

Proof: (of Theorem 1) Let $\mathcal{I} = (G, S, D, M)$. For any nodes u and u' in V , let $e = (u, u')$ be an edge of capacity λ to be added to G . Let $\mathcal{I}^{\lambda,e} = (G^{\lambda,e}, S, D, M)$. Let $\mathbf{R}^{\lambda,e} = (R_1^{\lambda,e}, \dots, R_k^{\lambda,e}) \in \mathcal{R}(\mathcal{I}^{\lambda,e})$. We show that $\mathbf{R}^{\lambda,e} - c\lambda \in \mathcal{R}(\mathcal{I})$ for some constant c that depends only on the edge capacities in G . We consider two cases depicted in Figure 1.

In the first case, we assume that the graph G is disconnected and that adding the edge $e = (u, u')$ connects two unconnected components of G . Let U and U' be a partition of the vertex set V of G such that $u \in U$, $u' \in U'$, and subsets U and U' are disconnected in G . In this case, the added edge e acts as a bridge of capacity λ between (subsets of) U and U' .

We first consider sources $s_i \in U$ for which there exist terminals d_j in U' such that $m_{ij} = 1$. As e is a bridge between U and U' , then using the cut set bound (e.g., Cor. 25 of [18]) it follows that $R_i^{\lambda,e} \leq \lambda$. The same holds if $s_i \in U'$ and $d_j \in U$.

We now consider all sources $s_i \in U$ such that all d_j for which $m_{ij} = 1$ satisfy $d_j \in U$. Denote this set of sources by S_U . Let $D_U = D \cap U$ be the set of terminals in U . Let G_U be the subgraph of G induced by the vertices in U . Finally, let M_U be the minor of M induced on columns and rows of M corresponding to U . Consider the instance $\mathcal{I}_U = (G_U, S_U, D_U, M_U)$. We now claim that $\mathbf{R}_U = (R_i^{\lambda,e} : i \in S_U) \in \mathcal{R}(\mathcal{I}_U)$. We prove this claim using the following averaging argument. Let $(\mathcal{F}, \mathcal{G})$ be an $(\varepsilon, \mathbf{R}^{\lambda,e} - \Delta, n, N)$ -feasible network code for $\mathcal{I}^{\lambda,e}$. By an averaging argument on the source messages $\mathbf{W} = (W_1, \dots, W_k)$, there exist fixed values $(w_i : i \notin S_U)$ for $(W_i : i \notin S_U)$ for which the probability of successful communication using $(\mathcal{F}, \mathcal{G})$ conditioned on $(W_i = w_i : i \notin S_U)$ is at least $1 - \varepsilon$. Moreover, under the condition $(W_i = w_i : i \notin S_U)$ the network code $(\mathcal{F}, \mathcal{G})$ can be *simulated* on \mathcal{I}_U . That is, there exists a network code $(\mathcal{F}_U, \mathcal{G}_U)$ for \mathcal{I}_U in which, for any time step $t \in [N]$ and for any edge $e' \in G_U$, the values transmitted over e' in $\mathcal{I}^{\lambda,e}$ using $(\mathcal{F}, \mathcal{G})$ are also transmitted over edge e' in \mathcal{I}_U using $(\mathcal{F}_U, \mathcal{G}_U)$. This follows from the fact that node u can simulate all incoming information from node u' in $(\mathcal{F}, \mathcal{G})$ given the knowledge that $(W_i : i \notin S_U) = (w_i : i \notin S_U)$. Thus, $(\mathcal{F}_U, \mathcal{G}_U)$ is an $(\varepsilon, \mathbf{R}_U - \Delta, n, N)$ -feasible network code for \mathcal{I}_U . As the argument applies for any $\varepsilon > 0$ and $\Delta > 0$, this implies that $\mathbf{R}_U \in \mathcal{R}(\mathcal{I}_U)$. Similarly, one can define $\mathcal{I}_{U'}$ and show that $\mathbf{R}_{U'} = (R_i^{\lambda,e} : i \in S_{U'}) \in \mathcal{R}(\mathcal{I}_{U'})$. Thus, we conclude that the rate vector $\mathbf{R} = (R_1, \dots, R_k)$ for which $R_i = R_i^{\lambda,e}$ if $i \in S_U \cup S_{U'}$ and $R_i = 0$ otherwise satisfies $\mathbf{R} \in \mathcal{R}(\mathcal{I})$. This follows by running codes over \mathcal{I}_U and $\mathcal{I}_{U'}$ in parallel. As, for each $i = 1, \dots, k$, the above analysis implies that $R_i \geq R_i^{\lambda,e} - \lambda$ we conclude the assertion of the theorem for the case under study with $f_{\mathcal{I}}(\lambda) = \lambda$.

In the second case, we assume that there exists in G a path connecting u and u' . Let $u = u_1, u_2, u_3, u_{\ell-1}, u_{\ell} = u'$ be one such path and let γ be the capacity of the minimum capacity edge in the path. Let $\delta > 0$ satisfy $\lambda = \delta\gamma$. Consider the graph G_{path} obtained from $G^{\lambda,e}$ by removing the edge $e = (u, u')$ and increasing the capacity of all edges in the path $u_1, u_2, u_3, u_{\ell-1}, u_{\ell}$ by λ . Let \mathcal{I}_{path} be the instance (G_{path}, S, D, M) . By Lemma 3, if $\mathbf{R}^{\lambda,e} = (R_1^{\lambda,e}, \dots, R_k^{\lambda,e}) \in \mathcal{R}(\mathcal{I}^{\lambda,e})$ then $\mathbf{R}^{\lambda,e} \in \mathcal{R}(\mathcal{I}_{path})$. Let $\alpha = \frac{\gamma}{\gamma + \lambda} = \frac{1}{1 + \delta}$. Consider the instance $\alpha\mathcal{I}_{path} = (\alpha G_{path}, S, D, M)$. By Lemma 2, $\mathbf{R}^{\lambda,e} \in \mathcal{R}(\mathcal{I}_{path})$ implies $\alpha\mathbf{R}^{\lambda,e} \in \mathcal{R}(\alpha\mathcal{I}_{path})$. Notice that the capacity of every edge in αG_{path} is at most the capacity of the corresponding edge in G . This is clearly true for edges in G_{path} that are not on the path $u_1, u_2, u_3, u_{\ell-1}, u_{\ell}$, and holds for path-edge (u_i, u_{i+1}) of capacity $\gamma' \geq \gamma$ as $\alpha(\gamma' + \lambda) \leq \gamma'$. Thus, $\alpha\mathbf{R}^{\lambda,e} \in \mathcal{R}(\alpha\mathcal{I}_{path}) \subseteq \mathcal{R}(\mathcal{I})$.

Notice that $\mathbf{R}^{\lambda,e} - f_{\mathcal{I}}(\lambda) \leq \alpha\mathbf{R}^{\lambda,e}$, and thus $\mathbf{R}^{\lambda,e} - f_{\mathcal{I}}(\lambda) \in \mathcal{R}(\mathcal{I})$ for

$$f_{\mathcal{I}}(\lambda) = \frac{\delta}{1 + \delta} \max_i R_i^{\lambda,e} \leq \delta(W(\mathcal{I}) + \lambda) \leq 2\delta W(\mathcal{I})$$

where $W(\mathcal{I})$ is the sum of all edge capacities in G and $(W(\mathcal{I}) + \lambda) \leq 2W(\mathcal{I})$ is the sum of all edge capacities in $G^{\lambda,e}$. Here, we assume without loss of generality that $\lambda \leq W(\mathcal{I})$ (as

otherwise, $\forall i, R_i^{\lambda,e} \leq 2\lambda$, a setting in which the proof of the theorem is immediate). Thus, $f_{\mathcal{I}}(\lambda) \leq 2\delta W(\mathcal{I}) = \frac{2W(\mathcal{I})}{\gamma} \lambda$. Let $w(\mathcal{I})$ be the minimum edge capacity over all edges in G , then $f_{\mathcal{I}}(\lambda) \leq \frac{2W(\mathcal{I})}{w(\mathcal{I})} \lambda = c\lambda$ for a constant $c = \frac{2W(\mathcal{I})}{w(\mathcal{I})}$ that only depends on the capacities of edges in \mathcal{I} .

An identical proof holds for the zero-error case. ■

We now prove Lemma 3.

Proof: The proof follows the line of proof given in [27], [26], in which it is shown that adding constant time delays in network communication has no impact on capacity. Throughout, to simplify our presentation, we consider the instance $\mathcal{I}_{path}^* = (G_{path}^*, S, D, M)$ in which we define G_{path}^* (similar to G_{path}) by starting with G , removing the edge e , and replacing it with a path of length $\ell - 1$ and capacity λ consisting of nodes $u = u_1, u_2, u_3, \dots, u_{\ell-1}, u_{\ell} = u'$. However, in G_{path}^* the nodes u_i for $i = 2, 3, \dots, \ell - 1$ are new nodes that do not originally appear in G . As any network code, for sources S , terminals D , and demands M , on \mathcal{I}_{path}^* can be implemented on \mathcal{I}_{path} , it holds that $\mathcal{R}(\mathcal{I}_{path}^*) \subseteq \mathcal{R}(\mathcal{I}_{path})$. To conclude our proof, we show that $\mathcal{R}(\mathcal{I}) \subseteq \mathcal{R}(\mathcal{I}_{path}^*)$.

The proof that $\mathcal{R}(\mathcal{I}) \subseteq \mathcal{R}(\mathcal{I}_{path}^*)$ proceeds as follows. For $\mathbf{R} \in \mathcal{R}(\mathcal{I})$ and $\Delta > 0$, by Lemma 1, consider an $(\varepsilon, \mathbf{R} - \Delta, n, N)$ feasible code $(\mathcal{F}, \mathcal{G})$ on \mathcal{I} with n and N sufficiently large and with $\varepsilon \leq 1/N^2$. *Interleaving* N such codes on N independent sub-messages from each source in S , as in [27], we obtain a new code $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ that is $(N\varepsilon, \mathbf{R} - \Delta, n, N^2)$ feasible as follows. $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ executes N independent sessions of the original $(\varepsilon, \mathbf{R} - \Delta, n, N)$ feasible-code $(\mathcal{F}, \mathcal{G})$ on N independent sub-messages. The sessions operate in a time-interleaved manner. In time steps $t = 1, \dots, N$ of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$, the first time step of all independent sessions of $(\mathcal{F}, \mathcal{G})$ is executed. (Time step 1 of independent session j of $(\mathcal{F}, \mathcal{G})$ operates in time step j of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$.) In general, in time steps $t = (i - 1)N + 1, \dots, iN$ of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$, the i 'th time step of each independent session of $(\mathcal{F}, \mathcal{G})$ is executed. (Time step i of independent session j of $(\mathcal{F}, \mathcal{G})$ operates in time step $(i - 1)N + j$ of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$.) After N^2 time steps, the N independent sessions of $(\mathcal{F}, \mathcal{G})$ are completed, implying an $(N\varepsilon, \mathbf{R} - \Delta, n, N^2)$ feasible code $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ for \mathcal{I} . Here, we bound the error by a union bound over the individual independent sessions of $(\mathcal{F}, \mathcal{G})$. The given $(N\varepsilon, \mathbf{R} - \Delta, n, N^2)$ feasible code $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ for \mathcal{I} satisfies $\tilde{\mathcal{X}}_{e,t}^n = \tilde{\mathcal{X}}_{e,t'}^n$ for any $t = (i - 1)N + j$ and $t' = (i - 1)N + j'$ with $j, j' \in [N]$. That is, for any time steps t and t' of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ in the same *sub-block* of length N , we have $\tilde{\mathcal{X}}_{e,t}^n = \tilde{\mathcal{X}}_{e,t'}^n$, as in that sub-block we are executing N independent session of the same time step i in the original code $(\mathcal{F}, \mathcal{G})$. Similarly, $\tilde{\mathcal{X}}_{e,t}^n = \tilde{\mathcal{X}}_{e,t'}^n$ for all $t, t' \in [(i - 1)N + 1, iN]$. Here and in what follows, we refer to time steps $t = (i - 1)N + j$ for $j \in [N]$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ as the i 'th sub-block of time steps of code $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$.

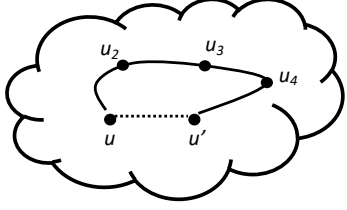
We now use the $(N\varepsilon, \mathbf{R} - \Delta, n, N^2)$ code $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ to construct an $(N\varepsilon, \mathbf{R} - \Delta, n, N(N + \ell))$ code for \mathcal{I}_{path}^* . That is, we use a code for the network that includes edge $e = (u, u')$ to build a code for the network in which edge $e = (u, u')$ is removed and replaced with a path of length $\ell - 1$ of the same

capacity. As N is chosen above to be sufficiently large, we assume here that ℓ is significantly smaller than N , say $\ell = \delta N$ for any constant $\delta > 0$ of our choice. The code $(\mathcal{F}_{path}, \mathcal{G}_{path})$ for \mathcal{I}_{path}^* is a slight modification of the code $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$. In $(\mathcal{F}_{path}, \mathcal{G}_{path})$ we still communicate with inner-blocklength n . However the outer-blocklength is set to $N(N + \ell)$ instead of N^2 . Like code $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$, code $(\mathcal{F}_{path}, \mathcal{G}_{path})$ operates in sub-blocks, where for each $i \in [N]$, sub-block i here takes the form $t = (i - 1)(N + \ell) + j$ for $j \in [N + \ell]$. Roughly speaking, the i 'th sub-block of code $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ determines the i 'th sub-block of code $(\mathcal{F}_{path}, \mathcal{G}_{path})$. The first N time steps in each sub-block i of $(\mathcal{F}_{path}, \mathcal{G}_{path})$ perform the same operations as are performed in sub-block i of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$. The last ℓ time steps in each sub-block in $(\mathcal{F}_{path}, \mathcal{G}_{path})$ are needed to transmit the information sent across edge $e = (u, u')$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ along the path of length $\ell - 1$ that replaced edge e in $(\mathcal{F}_{path}, \mathcal{G}_{path})$.

We first describe the coding operations of $(\mathcal{F}_{path}, \mathcal{G}_{path})$ on edges e' that are not on the path $u = u_1, u_2, u_3, \dots, u_{\ell-1}, u_{\ell} = u'$. For any such edge $e' = (v, v')$, the transmitted message from v to v' at time step $t = (i - 1)(N + \ell) + j$ in $(\mathcal{F}_{path}, \mathcal{G}_{path})$ equals the transmitted message from v to v' in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ in time step $t = (i - 1)N + j$. That is, the message over e' in the j 'th time step of the i 'th sub-block in $(\mathcal{F}_{path}, \mathcal{G}_{path})$ equals the message over e' in the j 'th time step of the i 'th sub-block in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$. In the remaining time steps in each sub-block of $(\mathcal{F}_{path}, \mathcal{G}_{path})$, no information is transmitted in either direction over edge e' . Namely, for each sub-block i , in time steps $t = (i - 1)(N + \ell) + (N + 1)$ through $t = i(N + \ell)$ a predetermined fixed message is transmitted over e' .

We now describe the coding operations of $(\mathcal{F}_{path}, \mathcal{G}_{path})$ on edges $e' = (u_r, u_{r+1})$ on the path $u = u_1, u_2, u_3, \dots, u_{\ell-1}, u_{\ell} = u'$. Roughly speaking, these edges “pipe-line” the message transmitted over the removed edge $e = (u, u')$ from u to u' and from u' to u , in G . (See Figure 2.) For edge $e' = (u_r, u_{r+1})$, the transmitted message from u_r to u_{r+1} in time step $t = (i - 1)(N + \ell) + j + r - 1$ in $(\mathcal{F}_{path}, \mathcal{G}_{path})$ equals the transmitted message from u to u' over $e = (u, u')$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ in time step $t = (i - 1)N + j$. In addition, the transmitted message from u_{r+1} to u_r in time step $t = (i - 1)(N + \ell) + j + \ell - r - 1$ of $(\mathcal{F}_{path}, \mathcal{G}_{path})$ equals the transmitted message from u' to u over $e = (u', u)$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ in time step $t = (i - 1)N + j$. That is, the message over $e = (u, u')$ from u to u' in the j 'th time step of the i 'th block of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ traverses the path $u = u_1, u_2, u_3, \dots, u_{\ell-1}, u_{\ell} = u'$ in time steps $t = (i - 1)(N + \ell) + j, \dots, (i - 1)(N + \ell) + j + \ell - 2$ in $(\mathcal{F}_{path}, \mathcal{G}_{path})$. In the other direction, the message over $e = (u', u)$ from u' to u in the j 'th time step of the i 'th block of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ traverses the path $u' = u_{\ell}, u_{\ell-1}, \dots, u_2, u_1 = u$ in time steps $t = (i - 1)(N + \ell) + j, \dots, (i - 1)(N + \ell) + j + \ell - 2$.

We now show that the above communication scheme of $(\mathcal{F}_{path}, \mathcal{G}_{path})$ is feasible on \mathcal{I}_{path}^* . We first show that any message on edge $e' = (v, v')$ from v to v' in time step t in $(\mathcal{F}_{path}, \mathcal{G}_{path})$ can be computed in G_{path}^* from the information available to node v prior to time step t . If the edge $e = (u, u')$ removed from \mathcal{I} is not an incoming edge to



Time\Edge	(u, u_2)	(u_2, u_3)	(u_3, u_4)	(u_4, u')
$(i-1)(N+\ell)+j$	\longrightarrow			\longleftarrow
$(i-1)(N+\ell)+j+1$	$\cdots\cdots\cdots\longrightarrow$	\longrightarrow	\longleftarrow	$\cdots\cdots\cdots\longleftarrow$
$(i-1)(N+\ell)+j+2$		\longleftarrow	\longrightarrow	
$(i-1)(N+\ell)+j+3$	\longleftarrow	$\cdots\cdots\cdots\longrightarrow$	$\cdots\cdots\cdots\longrightarrow$	\longrightarrow
$(i-1)(N+\ell)+j+4$	$\cdots\cdots\cdots\longleftarrow$			$\cdots\cdots\cdots\longrightarrow$

Fig. 2: An illustration of how information sent across edge $e = (u, u')$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ traverses the path $u = u_1, u_2, u_3, u_4, u_5 = u'$ in $(\mathcal{F}_{path}, \mathcal{G}_{path})$. In this example, $\ell = 5$. The solid right arrows (in black) represent information sent from u to u' at time $t = (i-1)(N+\ell) + j$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$. The dotted right arrows (in green) represent information sent from u to u' at time $t = (i-1)(N+\ell) + j + 1$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$. The solid left arrows (in red) represent information sent from u' to u at time $t = (i-1)(N+\ell) + j$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$. The dotted left arrows (in purple) represent information sent from u' to u at time $t = (i-1)(N+\ell) + j + 1$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$.

v in \mathcal{I} (i.e., if $v \neq u$ and $v \neq u'$), this follows directly by the feasibility of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ over \mathcal{I} . Otherwise, we take into account the delay incurred by replacing e with the path of length $\ell - 1$. Due to the interleaved structure of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$, this delay does not impact the feasibility of $(\mathcal{F}_{path}, \mathcal{G}_{path})$. Specifically, the structure of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ ensures that any message transmitted over an edge incoming to v in sub-block i of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ will be used by node v as input to subsequent encoding only in the next sub-block $i+1$ of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$. The same holds for $(\mathcal{F}_{path}, \mathcal{G}_{path})$. Moreover, by our definitions, any message transmitted over an edge incoming to v in sub-block i of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ must also be transmitted to v in sub-block i of $(\mathcal{F}_{path}, \mathcal{G}_{path})$. We conclude that any message on edge $e' = (v, v')$ from v to v' in time step t in $(\mathcal{F}_{path}, \mathcal{G}_{path})$ can be computed in G_{path}^* from the information available to v prior to time step t .

Secondly, we show that the alphabet $\mathcal{X}_{e'}^n$ of edges $e' = (v, v')$ in G_{path}^* support the code $(\mathcal{F}_{path}, \mathcal{G}_{path})$ in the sense that the messages transmitted from v to v' and from v' to v in time step t have support $\overrightarrow{\mathcal{X}_{e',t}^n}$ and $\overleftarrow{\mathcal{X}_{e',t}^n}$ respectively, that satisfy $|\overrightarrow{\mathcal{X}_{e',t}^n}| \cdot |\overleftarrow{\mathcal{X}_{e',t}^n}| \leq |\mathcal{X}_{e'}^n|$. For edges e' that are not on the path u_1, \dots, u_ℓ , this follows by the feasibility of the code $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$. For edges $e' = (u_r, u_{r+1})$ along the path, the alphabet $\mathcal{X}_{e'}^n$ corresponding to e' equals \mathcal{X}_e^n of (the removed edge) $e = (u, u')$ in G . Moreover, as discussed above, we have that $\overrightarrow{\mathcal{X}_{e,t}^n} = \overrightarrow{\mathcal{X}_{e,t'}^n}$ and $\overleftarrow{\mathcal{X}_{e,t}^n} = \overleftarrow{\mathcal{X}_{e,t'}^n}$ for any t and t' in the i 'th sub-block of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$. We thus denote the alphabets $\overrightarrow{\mathcal{X}_{e,t}^n}$ and $\overleftarrow{\mathcal{X}_{e,t}^n}$ in $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ for any t in the i 'th sub-block of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ (i.e., $t = (i-1)N + j$ for $j \in [N]$) by $\overrightarrow{\mathcal{X}_{e, \text{block } i}^n}$ and $\overleftarrow{\mathcal{X}_{e, \text{block } i}^n}$ respectively. Thus, for any t in the i 'th sub-block of $(\mathcal{F}_{path}, \mathcal{G}_{path})$, and any edge $e' = (u_r, u_{r+1})$ on the path $u = u_1, u_2, u_3, \dots, u_{\ell-1}, u_\ell = u'$, we define alphabets $\overrightarrow{\mathcal{X}_{e',t}^n}$ and $\overleftarrow{\mathcal{X}_{e',t}^n}$ in $(\mathcal{F}_{path}, \mathcal{G}_{path})$ to be equal to $\overrightarrow{\mathcal{X}_{e, \text{block } i}^n}$ and $\overleftarrow{\mathcal{X}_{e, \text{block } i}^n}$ of $(\tilde{\mathcal{F}}, \tilde{\mathcal{G}})$ respectively. Such a definition for $\overrightarrow{\mathcal{X}_{e',t}^n}$ and $\overleftarrow{\mathcal{X}_{e',t}^n}$ allows $e' = (u_r, u_{r+1})$ to support the messages defined previously by $(\mathcal{F}_{path}, \mathcal{G}_{path})$ during time steps $t =$

$i(N+\ell) + j$ for $j \in [N+\ell]$. See, Figure 2.

We therefore conclude that $(\mathcal{F}_{path}, \mathcal{G}_{path})$ is an $(N\varepsilon, \frac{N^2}{N(N+\ell)}(\mathbf{R} - \Delta), n, N(N+\ell))$ -feasible code for \mathcal{I}_{path}^* . As $\varepsilon \leq 1/N^2$, $\Delta > 0$ can be chosen to be arbitrarily small, and for any $\delta > 0$ we can choose N sufficiently large such that $\ell = \delta N$, we conclude that $\mathbf{R} \in \mathcal{R}(\mathcal{I}_{path}^*)$.

An almost identical proof (with very slight modifications) holds for the zero-error case as well. ■

Corollary 1 follows immediately from Theorem 1.

Corollary 1: The vanishing edge-removal statement and the zero-error vanishing-edge-removal statement hold for undirected network instances.

Using the connections outlined in [11], Corollary 2 also follows from Theorem 1. Proof appears in the full version [1].

Corollary 2: Let \mathcal{I} be an undirected network instance, then $\mathcal{R}(\mathcal{I}) = \mathcal{R}_0(\mathcal{I})$.

IV. CONCLUSIONS

In this work, we study the edge removal problem on undirected networks. Using the conceptually simple idea of rerouting information on the removed edge (if possible in the given topology) we show that the asymptotic version of the edge-removal statement holds. That is, we show that removing an edge of negligible capacity in undirected networks has only a negligible impact on the capacity region. This, in turn, implies that the zero-error capacity region of an undirected network equals its vanishing-error capacity region. Whether similar results are true for directed networks is an intriguing open problem. In addition, in light of the multiple-unicast coding advantage conjecture on undirected networks, it would be interesting to prove Theorem 1 with a constant $c = 1$. Finally, as common in the network coding literature, our model assumes that each message W_i is held by a single network node s_i (before communication starts). Extending our results to the case in which messages W_i can be held by multiple network nodes involve challenges that are subject to future studies.

REFERENCES

- [1] M. Langberg and M. Effros, "Edge removal in undirected networks," *Manuscript, appears in <https://arxiv.org/abs/2005.10315>*, 2020.
- [2] T. Ho, M. Effros, and S. Jalali, "On equivalences between network topologies," in *Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing*, 2010.
- [3] S. Jalali, M. Effros, and T. Ho, "On the impact of a single edge on the network coding capacity," in *Information Theory and Applications Workshop (ITA)*, 2011.
- [4] T. Chan and A. Grant, "On capacity regions of non-multicast networks," in *International Symposium on Information Theory*, 2010, pp. 2378 – 2382.
- [5] —, "Network coding capacity regions via entropy functions," *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5347–5374, 2014.
- [6] M. Langberg and M. Effros, "Network coding: Is zero error always possible?" in *Proceedings of Forty-Ninth Annual Allerton Conference on Communication, Control, and Computing*, 2011, pp. 1478–1485.
- [7] —, "Source coding for dependent sources." In *proceedings of IEEE Information Theory Workshop (ITW)*, 2012.
- [8] M. F. Wong, M. Langberg, and M. Effros, "On a Capacity Equivalence between Network and Index Coding and the Edge Removal Problem," in *proceedings of International Symposium on Information Theory*, 2013, pp. 972 – 976.
- [9] —, "On an equivalence of the reduction of k-unicast to 2-unicast capacity and the edge removal property," in *IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 371–375.
- [10] —, "On the tightness of an entropic region outer bound for network coding and the edge removal property," in *IEEE International Symposium on Information Theory*, 2016, pp. 1769–1773.
- [11] M. Langberg and M. Effros, "The edge-removal problem's connections to the zero-error and δ -dependence problems in network coding," *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 900 – 913, 2020.
- [12] Z. Li and B. Li, "Network coding in undirected networks," in *Proceedings of the 38th Annual Conference on Information Sciences and Systems*, 2004.
- [13] —, "Network coding: The case of multiple unicast sessions," in *Proceedings of the 42nd Allerton Conference on Communications, Control, and Computing*, 2004.
- [14] N. J. Harvey, R. D. Kleinberg, and A. R. Lehman, "Comparing network coding with multicommodity flow for the k-pairs communication problem," *Tech. report.*, 2004.
- [15] Z. Li, B. Li, and L. C. Lau, "On achieving maximum multicast throughput in undirected networks," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2467–2485, 2006.
- [16] K. Jain, V. V. Vazirani, and G. Yuval, "On the capacity of multiple unicast sessions in undirected graphs," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2805–2809, 2006.
- [17] C. Chekuri, C. Fragouli, and E. Soljanin, "On average throughput and alphabet size in network coding," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2410–2424, 2006.
- [18] N. J. A. Harvey, R. Kleinberg, and A. R. Lehman, "On the capacity of information networks," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2345–2364, 2006.
- [19] Z. Li, B. Li, and L. C. Lau, "A constant bound on throughput improvement of multicast network coding in undirected networks," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 1016–1026, 2009.
- [20] S. Maheshwar, Z. Li, and B. Li, "Bounding the coding advantage of combination network coding in undirected networks," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 570–584, 2012.
- [21] M. Langberg and M. Médard, "On the multiple unicast network coding conjecture," in *Proceedings of the 47th Annual Allerton Conference on Communication, Control, and Computing*. IEEE, 2009, pp. 222–227.
- [22] K. Cai and G. Han, "On network coding advantage for multiple unicast networks," in *Proceedings of the IEEE International Symposium on Information Theory*, 2015, pp. 366–370.
- [23] B. Haeupler, D. Wajc, and G. Zuzic, "Network coding gaps for completion times of multiple unicasts," *arXiv preprint arXiv:1905.02805*, 2019.
- [24] M. Braverman, S. Garg, and A. Schwartzman, "Network coding in undirected graphs is either very helpful or not helpful at all," *arXiv preprint arXiv:1608.06545*, 2016.
- [25] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network Information Flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.
- [26] R. Koetter, M. Effros, and M. Médard, "A theory of network equivalence—Part I: Point-to-point channels," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 972–995, 2011.
- [27] M. Effros, "On dependence and delay: Capacity bounds for wireless networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, 2012, pp. 550–554.