

Chapter 2

Submodular Information Flow Models for Multicast Communication

This chapter introduces and compares different models for multicast rate regions on an abstract level. It focuses entirely on the structure of multicast rate regions as sets and their representations. Connections of these multicast models to wireless communication models and to information theoretic source and channel models are established in Chap. 4 et seq. We distinguish between cut models and flow models. Cut models bound the multicast rates of all sources in any subset of nodes directly using real-valued set functions referred to as cut rate functions. Flow models bound the multicast rates of all sources indirectly through information flows, which represent the flow of the multicast rates of all sources to the terminals. Information flows need to be conserved at all nodes, and outgoing information flows at each node are locally bounded using either set functions referred to as hyperarc rate functions (graph and hypergraph model) or set functions referred to as broadcast rate functions (polymatroid broadcast model).

Section 2.1 formulates multicast rate regions in the cut model. The model is primarily inspired by the cut-set outer bound from information theory due to El Gamal [1], cf. [2, 3], and Chap. 4, and the min-cut bounds of the maximal flow in directed graphs with capacitated arcs [4, 5]. Subsequently, the cut model is refined to submodular cut rate functions. This refinement is motivated by the submodularity of the cut rate functions that model the cut-set outer bound for some classes of networks, see Chap. 4 et seq. and [6], and the submodularity of the cut rate functions that are associated with the flow models in the following sections.

Section 2.2 connects the submodular cut model from Sect. 2.1 the multicast max-flow min-cut theorem for capacitated graphs due to Ahlswede et al. [7]. In Sects. 2.3 and 2.4, the hypergraph and polymatroid broadcast models are introduced and the corresponding multicast max-flow min-cut results are established. The multicast max-flow min-cut for the hypergraph model has been indirectly established in [8] by

mapping a hypergraph onto an equivalent virtual graph, see also Sect. 2.3, and directly in [9, 10]. The multicast max-flow min-cut for the polymatroid broadcast model has been established in [11] based on the max-flow min-cut theorem for polymatroid flows in graphs due to [12–14]. Subsequently, Sect. 2.5 connects all four models with each other through linear transformations and establishes a chain of strict inclusion among these models. That is, the hypergraph model strictly generalizes the graph model, the polymatroid broadcast model strictly generalizes the hypergraph model, and the submodular cut model strictly generalizes the polymatroid broadcast model.

Sections 2.6 and 2.7 further generalize the (submodular) cut model and the polymatroid broadcast model to include multicast rate penalty terms, which prove extremely useful when modeling multicast capacity region inner bounds based on noisy network coding, cf. Chap. 4. In particular, the penalized polymatroid broadcast model immediately generalizes the multicast max-flow min-cut theorem for the polymatroid broadcast model. Section 2.8 discusses convexity and extreme points as well as downward comprehensiveness and Pareto efficiency of multicast rate regions within those generalized models. Section 2.9 analyzes the gap between two multicast rate regions based on two similar cut rate functions with their difference being a submodular set function. It establishes bounds on the gap, which are based entirely on the submodularity of the difference set function. Thus, Sect. 2.9 extends and generalizes results for Gaussian networks in [15] (cf. [16]). Finally, Sect. 2.10 briefly introduces a further extension to per-terminal cut models that allow for different cut rate functions for each terminal node. Models of this type are especially suited to represent achievable multicast rate regions for noisy network coding, where, loosely speaking, each terminal has access to its own channel output signal in addition to the quantized channel output signals of all nodes.

2.1 Cut Model

We begin with a general definition of a multicast network, a cut rate region on a multicast network, and its corresponding multicast rate region. These definitions are motivated by the cut-set outer bound [1] and the multicast capacity region of certain special classes of networks, e.g., graphical networks [7], networks of deterministic broadcast channels [17], and deterministic linear finite field networks [18].

Definition 2.1 A *multicast network* (N, T) consists of a nonempty finite set of *nodes* N and a nonempty subset of *terminals* $T \subset N$. A *cut rate region* $\mathcal{V} \subset \mathcal{C}_+^N$ on a multicast network (N, T) is a set of nonnegative set functions $v: 2^N \rightarrow \mathbb{R}$ satisfying $v(\emptyset) = v(N) = 0$, which are referred to as *cut rate functions*. A nonnegative *multicast rate vector* $r = (r_a : a \in N) \in \mathcal{R}_+^N$ is supported by a cut rate function v on a multicast network (N, T) if

$$\sum_{a \in A} r_a \leq v(A) \quad \forall t \in T, A \subset \{t\}^c. \quad (2.1)$$

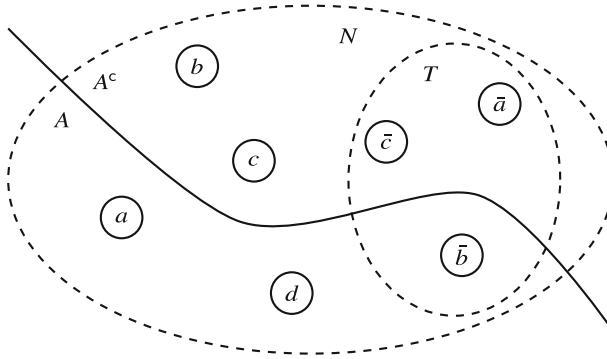


Fig. 2.1 Exemplary multicast network with seven nodes $N = \{a, b, c, d, \bar{c}, \bar{b}, \bar{a}\}$ and three terminals $T = \{\bar{c}, \bar{b}, \bar{a}\}$. All nodes can be sources and all sources are independent. The cut $A = \{a, d, \bar{b}\}$ (source side) with its complement $A^c = \{b, c, \bar{c}, \bar{a}\}$ (terminal side) is depicted, which bounds the multicast sum rate of all nodes in A by the cut rate function $v(A)$, i.e., $r_a + r_d + r_{\bar{b}} \leq v(\{a, d, \bar{b}\})$

The *multicast rate region* $\mathcal{R}(\mathcal{V}) \subset \mathcal{R}_+^N$ of the multicast network (N, T) with cut rate region \mathcal{V} is defined as the set of nonnegative multicast rate vectors supported by \mathcal{V} , i.e.,¹

$$\mathcal{R}(\mathcal{V}) = \bigcup_{v \in \mathcal{V}} \left(\bigcap_{t \in T} \left\{ r \in \mathcal{R}_+^N : \sum_{a \in A} r_a \leq v(A) \quad \forall A \subset \{t\}^c \right\} \right). \quad (2.2)$$

Definition 2.1 considers a multicast network with multiple independent data sources. Each node $a \in N$ may potentially act as information source. Each active source node injects data at rate $r_a > 0$ into the network, whereas $r_a = 0$ means that node a does not act as source and just assists by relaying data from other sources if necessary. The multicast nature of the communication stems from the condition that all terminals² $t \in T$ are interested in *all data* from *all sources* simultaneously. Figure 2.1 visualizes an exemplary multicast network with seven nodes and three terminals.

Each cut rate function $v \in \mathcal{V}$ quantifies for each cut $A \subset N$ the maximal total multicast rate that can traverse this cut from nodes $a \in A$, the source side of the cut, to any node $b \in A^c$, the terminal side of the cut, under the given policy or parameter choice corresponding to that particular cut rate function. The cut rate region \mathcal{V} represents all available policies and parameters, e.g., transmission and coding strategies, channel input distributions, etc., by their corresponding cut rate functions $v \in \mathcal{V}$. Since each terminal needs to recover all data from all sources, all cuts separating any terminal from any set of sources bound the total multicast rate

¹The multicast network (N, T) is not explicitly indicated in the notation of the cut and multicast rate regions \mathcal{V} and $\mathcal{R}(\mathcal{V})$ since, except for some proofs, N and T denote the node and terminal set, respectively, throughout this book.

²Terminals are occasionally also referred to as destinations or information sinks.

of the respective set of sources. These cut-set constraints define the structure of the multicast rate region $\mathcal{R}(\mathcal{V})$ in (2.2).

The multicast rate region $\mathcal{R}(\mathcal{V})$ in (2.2) is expressed in a *min-cut* formulation. That is, for each cut $A \subset N$, the multicast rates that can be conveyed from the source side A to the terminal side A^c is bounded by $v(A)$. This interpretation becomes clear if we suppose that $s \in N$ is a distinguished source node of the multicast and consider the face of the multicast rate region $\mathcal{R}(\mathcal{V})$ where $r_a = 0$ for all $a \in \{s\}^c$. The single-source multicast rate $r_s \geq 0$ of s is feasible with respect to (2.2) if³

$$r_s \leq \max_{v \in \mathcal{V}} \min_{t \in T} \min_{A: s \in A \subset \{t\}^c} v(A). \quad (2.3)$$

That is, the maximal single-source multicast rate of $s \in N$ with respect to each cut rate function v is equivalent to the minimum with respect to $t \in T$ of the *minimum cut* separating the source s from terminal t . Note that the maximum over all $v \in \mathcal{V}$ corresponds to choosing the best policy for this multicast communication represented by its corresponding cut rate functions.

The single-source multicast rate expression (2.3) leads to an equivalent reformulation of $\mathcal{R}(\mathcal{V})$ given by

$$\mathcal{R}(\mathcal{V}) = \bigcup_{v \in \mathcal{V}} \left(\bigcap_{t \in T} \left\{ r \in \mathcal{R}_+^N : \sum_{a \in A} r_a \leq \min_{A \subset C \subset \{t\}^c} v(C) \quad \forall A \subset \{t\}^c \right\} \right), \quad (2.4)$$

which follows from combining the cut constraints in the original definition of the multicast rate region (2.2) with the nonnegativity of the rate vectors $r \in \mathcal{R}_+^N$. This formulation is especially useful if either a certain source subset $S \subset N$ is considered and we are interested only in the face of the rate region for which $r_a = 0$ for all $a \in S^c$, or if the set function $\min_{A \subset C \subset \{t\}^c} v(C)$ is simpler to evaluate than the original cut rate function $v(A)$.

The cut model is well suited to model wireless communication scenarios since the cut rate function allows for arbitrary dependence among all nodes' transmissions and receptions. In particular, it assumes neither a decomposition across the transmitters, as for example the graph, hypergraph, and the proposed polymatroid broadcast models, nor a decomposition across the receivers, which is assumed by the graph model. However, this generality comes at the price that the model has no locally decomposable structure nor any other mathematical structure that can be exploited to simplify the characterization of the multicast rate region. This situation can be alleviated by considering submodular cut rate functions.

Definition 2.2 A cut rate region $\mathcal{V} \subset \mathcal{C}_+^N$ on a multicast network (N, T) is called *submodular* if all cut rate functions $v \in \mathcal{V}$ are submodular set functions on N , i.e.,

$$v(A) + v(B) \geq v(A \cup B) + v(A \cap B) \quad \forall A, B \subset N. \quad (2.5)$$

³We assume that $T \neq \{s\}$ and that the maximum over v in (2.3) exists.

The convex cone of all submodular cut rate functions is denoted by $\mathcal{K}_+^N \subset \mathcal{C}_+^N$.

Submodular cut rate regions are exceptionally important since they can be dealt with much easier than general cut rate regions. For example, the minimum cut problem in (2.3) with a distinguished source involves minimization problems on set functions. Grötschel et al. [19] proposed a polynomial algorithm for submodular function minimization.⁴ However, the general set function minimization problem is NP-hard.⁵ Furthermore, optimizing a linear function over the multicast rate region by a submodular cut rate region turns out to be a special case of the *submodular flow problem* [24–26], which can be solved using the greedy algorithm for linear optimization over polymatroid polyhedra and submodular polyhedra [26–29]. Chapter 3 introduces a general dual decomposition approach for multicast utility optimization with convex submodular cut rate regions, which exploits the greedy algorithm to efficiently deal with the exponential number of constraints in the definition of the multicast rate region (2.2).

Submodular cut rate regions appear in the context of outer and inner bounds on the multicast capacity region for discrete memoryless networks with independent noise across all nodes, see Chap. 4. Furthermore, the following sections introduce the graph, hypergraph, and polymatroid broadcast models. All three models lead to particular subclasses of submodular cut models through appropriate linear transformations, whose structures and properties are analyzed in Sect. 2.5.

2.2 Graph Model

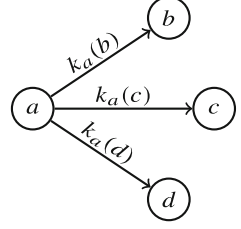
The simplest and most frequently used model in wired and wireless communication networks is the graph model, see for example [30–32] for an application of the graph model to wireless multicast. The network is represented by a node set N and directed arcs (a, b) consisting of a tail node $a \in N$ and a head node $b \in N$. The arc (a, b) models information transfer from a single transmitter a to a single receiver b . The information that is communicated over all arcs is considered independent across all arcs. The maximum rate at which information can be transmitted over an arc is referred to as the flow capacity of the arc. We represent the flow capacities of all arcs leaving a node a by the arc rate function k_a such that $k_a(b)$ denotes the flow capacity of the arc (a, b) . In order to keep the notation simple, we assume without loss of generality that the network includes all possible arcs $(a, b) \in N \times N$.⁶ Fig. 2.2

⁴McCormick [20] and Iwata [21] provide extensive surveys on this problem and solution algorithms.

⁵General set function minimization includes many standard NP-hard problems as special cases, e.g., maximum independent set in graphs [22], maximum cut in graphs and hypergraphs [22], and monotone submodular set function maximization with cardinality constraints [23].

⁶Whether an arc is not present in the network or is present but has zero flow capacity does not influence any of the results in this book. Similarly, multiple arcs with the same tail and head node can easily be combined to one arc with its flow capacity given by the sum of the flow capacities of the combined arcs.

Fig. 2.2 Graph model for wireless communication showing only those arcs originating at a with nonzero flow capacities



depicts a local view of the graph model showing only the arcs originating at a with nonzero flow capacity.

Definition 2.3 A function $k_a \in \mathcal{R}^N$ on N , i.e., $k_a: N \rightarrow \mathbb{R}$, originating at node $a \in N$ is called an *arc rate function* if it satisfies the following two conditions:

$$k_a(b) \geq 0 \quad \forall b \in N, \quad (\text{nonnegative})$$

$$k_a(a) = 0. \quad (\text{loop-free})$$

The *nonnegativity* condition ensures that $k_a(b)$ represents a reasonable flow capacity expression that bounds nonnegative information rates on the arc (a, b) . The *loop-free* condition just excludes loops, i.e., arcs that have the same head and tail node, since a transmitter cannot gain any new information from its own transmitted messages. The neighborhood of a in the graph model is defined as the set of all nodes that receive information from a , i.e.,

$$N_a(k_a) = \{b \in N : k_a(b) > 0\}. \quad (2.6)$$

Clearly, $a \notin N_a(k_a)$ due to the loop-free condition. Figure 2.2 shows the local arcs originating at a directed towards neighbors of a .

Definition 2.4 A vector $k = (k_a : a \in N) \in \mathcal{F}^N$ is called an *arc rate vector* if each element k_a , $a \in N$, is an arc rate function. The convex cone of all arc rate vectors is denoted by $\mathcal{A}_+^N \subset \mathcal{F}_+^N$. A subset $\mathcal{K} \subset \mathcal{A}_+^N$ is referred to as *arc rate region*.

Each arc rate vector $k \in \mathcal{K}$ quantifies the flow capacities of all arcs in the network and represents simultaneously achievable information rates on all arcs for the considered wireless network model. The arc rate region \mathcal{K} represents all available policies and parameters, e.g., transmission and coding strategies, channel input distributions, etc., by their corresponding arc rate vector $k \in \mathcal{K}$. Figure 2.3 shows a graph model for the example network in Fig. 2.1, showing only arcs with nonzero flow capacity for one particular arc rate vector k .

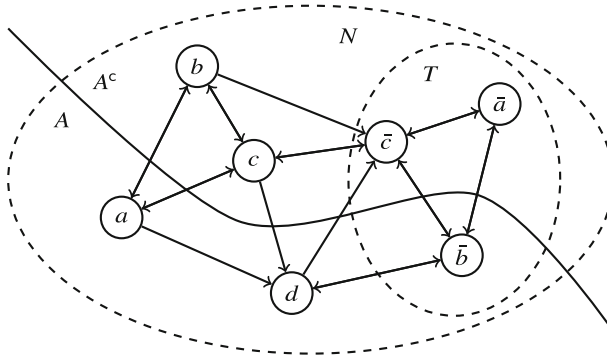


Fig. 2.3 Graph model for an exemplary multicast network with seven nodes $N = \{a, b, c, d, \bar{c}, \bar{b}, \bar{a}\}$ and three terminals $T = \{\bar{c}, \bar{b}, \bar{a}\}$. The *arrows* indicate arcs whose flow capacities are nonzero for the arc rate vector k . Two-sided *arrows* represent two arcs, one in each direction, potentially with different flow capacities. The depicted cut $A = \{a, d, \bar{b}\}$ with terminal side $A^c = \{b, c, \bar{c}, \bar{a}\}$ offers a cut rate $v(A) = k_a(b) + k_a(c) + k_d(\bar{c}) + k_{\bar{b}}(\bar{c}) + k_{\bar{b}}(\bar{a})$ according to (2.7) dropping all flow capacities that are zero from the sum and considering only the nonzero flow capacities

Definition 2.5 A cut rate region $\mathcal{V} \subset \mathcal{C}_+^N$ is generated by an arc rate region $\mathcal{K} \subset \mathcal{A}_+^N$ if

$$\mathcal{V} = \bigcup_{k \in \mathcal{K}} \left\{ v \in \mathcal{C}_+^N : v(A) = \sum_{a \in A} \sum_{b \in A^c} k_a(b) \quad \forall A \subset N \right\}. \quad (2.7)$$

The corresponding multicast rate region $\mathcal{R}(\mathcal{V})$ of a multicast network (N, T) is also denoted by $\mathcal{R}(\mathcal{K})$.

Proposition 2.1 A cut rate region $\mathcal{V} \subset \mathcal{C}_+^N$ generated by an arc rate region $\mathcal{K} \subset \mathcal{A}_+^N$ is submodular.

The cut rate function v associated with the arc rate vector k is defined for each cut A as the sum of the arc rates, representing the flow capacities, of all arcs crossing the respective cut from the source side A to the terminal side A^c , i.e., all arcs (a, b) such that $a \in A$ and $b \in A^c$, see Fig. 2.3 for an example. This is consistent with the classical definition of the cut value in capacitated directed graphs and the classical *max-flow min-cut* theorem [4, 5], which states that for graphs with capacitated arcs the maximum flow from a single source node to a single terminal node is equal to the minimum value of all cuts separating the source from the terminal. Applying this cut rate definition to all $k \in \mathcal{K}$ yields the cut rate region \mathcal{V} corresponding to the arc rate region \mathcal{K} according to (2.7). Finally, the special structure of the cut rate regions \mathcal{V} defined by (2.7) guarantees their submodularity, see Proposition 2.1, which is a corollary to Proposition 2.5 and as such provides a first example of a model that fits into the submodular cut model framework.

Using a random network coding argument, Ahlswede et al. [7] established that the same cut values determine the multicast capacity in a deterministic graphical net-

work where each arc represents an identity map from input to output and where its flow capacity corresponds to the logarithm of the input alphabet size. The multicast capacity is precisely formulated by (2.3) with the cut rate function given by (2.7). The result in [7] leads to an equivalent flow formulation of the multicast capacity problem and is thus a multicast *max-flow min-cut* theorem. The same result holds for general networks of independent point-to-point channels with the flow capacity of each arc given by the Shannon capacity of the underlying point-to-point communication channel [33]. The single-source max-flow min-cut result extends directly to multiple sources serving the same multicast terminal set [34]. The following theorem restates these results as a multicast max-flow min-cut theorem for capacitated directed graphs.

Theorem 2.1 (Graph max-flow min-cut) *Let (N, T) be a multicast network and $\mathcal{V} \subset \mathcal{C}_+^N$ be a cut rate region on (N, T) generated by an arc rate region $\mathcal{K} \subset \mathcal{A}_+^N$. The multicast rate region $\mathcal{R}(\mathcal{K})$ is given by*

$$\mathcal{R}(\mathcal{K}) = \bigcup_{k \in \mathcal{K}} \left(\bigcap_{t \in T} \left\{ r \in \mathcal{R}_+^N : x^t \in \mathcal{F}_+^N, \quad x_{ab}^t \leq k_a(b) \quad \forall a \in N, b \in N, \right. \right. \\ \left. \left. \sum_{b \in N} x_{ab}^t - \sum_{b \in N} x_{ba}^t = r_a \quad \forall a \in \{t\}^c \right\} \right). \quad (2.8)$$

At the core of the max-flow formulation (2.8) are the nonnegative flows x^t for all $t \in T$, representing the information flows through the graph from all sources directed to each terminal $t \in T$ individually. In the context of network coding, these are often referred to as *virtual information flows* or simply virtual flows [35, 36] since they describe how the source information propagates through the coding network to each terminal. The second line in (2.8) represents the *flow conservation law*, which states that the total outgoing flow at each node $a \in \{t\}^c$ must equal the total incoming flow at a and the multicast source rate of a . The last part in the first line in (2.8) is the *flow capacity constraint*, which bounds the information flow x_{ab}^t on each arc (a, b) by the corresponding arc rate $k_a(b)$. Note that the *loop-free* condition for arc rate functions (Definition 2.3) ensures that the flows x_{aa}^t are zero for all $a \in N$.

The flow conservation and the flow capacity constraints at each node are local in the sense that both constraints apply only to flows and multicast rates immediately connected with this particular node. This is in contrast to the min-cut formulation of the multicast rate region (2.2) with a cut rate region generated from a graph model (2.7), where the multicast rates are directly bounded for each cut by the sum of the arc rates over all arcs crossing the cut. Nevertheless, the max-flow min-cut theorem ensures that all flow conservation and flow capacity constraints jointly have the same bounding effect for the multicast rates as the cut constraints in (2.2) for cut rate function of the form (2.7).

We remark that the intersection over all terminals is due to the multicast nature of the communication. This result is due to Ahlswede et al. [7], who showed that each terminal independently bounds the multicast rate for a single source in graphical

networks and thus the multicast rate is in total bounded by the “worst” terminal. This property can be identically found in the min-cut formulation of the multicast rate region (2.2). Finally, the union over all arc rate vectors $k \in \mathcal{K}$, as the union over all cut rate functions in (2.2), represents all available policies, transmitter operations, etc. for the considered wireless network and channel model.

2.3 Hypergraph Model

The graph model is not suitable to model central aspects of wireless communication networks, in particular, the wireless broadcast advantage. This is due to the structure of the cut rate function (2.7), which sums over the flow capacities of all outgoing arcs from all transmitters that cross the considered cut, see Fig. 2.3. The structure inherently assumes that the information on all outgoing arcs of each node is independent. This is in contrast to the broadcast nature of the wireless medium, where each receiver gets a different view of the same transmit signal, which leads to a dependence among the received information at all receivers.

The hypergraph model partially alleviates this problem of the graph model by replacing capacitated arcs with capacitated *hyperarcs*, also referred to as hyperedges or hyperlinks. It was introduced as an extension to the graph model in [9, 10]. A *hyperarc* is a point-to-multipoint link represented by an ordered pair (a, B) consisting of a tail node $a \in N$ and a set of head nodes $B \subset N$, see Fig. 2.4 (left). It models the broadcast of common information from a transmitter a to all receivers $B \subset N$ but to none of the other nodes B^c . Wu et al. [8] introduced the hypergraph model indirectly by augmenting the graph model with virtual nodes, each of which corresponds to a hyperarc. The equivalent virtual graph model of the hyperarc in Fig. 2.4 (left) is depicted in Fig. 2.4 (right). Since the hypergraph model considers potential dependencies of the received information of all receivers, it is much better suited to model the wireless broadcast advantage in wireless networks than the graph model.

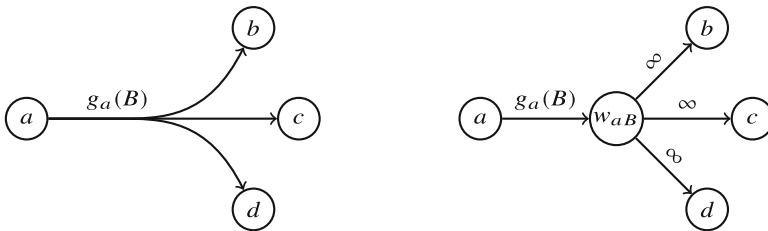


Fig. 2.4 Hyperarc model for wireless communication showing only the hyperarc (a, B) with $B = \{b, c, d\}$ (left) and the equivalent virtual graph model for the hyperarc (a, B) with virtual node w_{aB} (right). The hyperarc rate $g_a(B)$ (left) is represented by the depicted flow capacities on all virtual arcs originating at or directed to w_{aB} , i.e., the arc from a to w_{aB} has flow capacity $g_a(B)$, all arcs originating at w_{aB} have infinite or arbitrarily high flow capacities

Multiple hyperarcs may originate at each transmitter. Without loss of generality, one hyperarc for each potential receiver subset $B \subset N$. A hyperarc flow capacity $g_a(B)$ is associated with each hyperarc. It quantifies the maximum rate at which information can be communicated from a to the receiver set B such that all $b \in B$ obtain the transmitted information associated with this hyperarc. However, the transmitted information associated with each hyperarc is independent across all hyperarcs. This is particularly suited to models with separated network and channel coding, where multiple messages are encoded at the transmitter and simultaneously broadcast to different receiver groups such that each receiver of each group can perfectly recover those messages directed to it but none of the other messages (channel coding). Each node then combines its own information with those perfectly recovered messages from all other nodes for its own transmissions (network coding). The flow capacities of all hyperarcs originating at a node $a \in N$ are collected in the hyperarc rate function g_a defined as follows:

Definition 2.6 A set function $g_a \in \mathcal{S}^N$ on N , i.e., $g_a: 2^N \rightarrow \mathbb{R}$, originating at node $a \in N$ is called a *hyperarc rate function* if it satisfies the following three conditions:

$$g_a(\emptyset) = 0, \quad (\text{normalized})$$

$$g_a(B) \geq 0 \quad \forall B \subset N, \quad (\text{nonnegative})$$

$$g_a(B \cup \{a\}) = 0 \quad \forall B \subset N. \quad (\text{loop-free})$$

These conditions admit a natural interpretation in the context of independent messages on each hyperarc: The *normalization* simply forces that there is no message that is received by no nodes at all. The *nonnegativity* condition requires that hyperarc rates are nonnegative, i.e., they represent the rates of information messages. Finally, the *loop-free* condition just excludes all hyperarcs whose tail node a is also a head node to ensure that a does not gain any new information from its own transmitted broadcasts. The neighborhood in the hypergraph model is defined as all nodes that receive some information from a , i.e.,

$$N_a(g_a) = \{b \in N : b \in B \subset N, g_a(B) > 0\}. \quad (2.9)$$

Clearly, $a \notin N_a(g_a)$ due to the loop-free condition.

Definition 2.7 A vector $g = (g_a : a \in N) \in \mathcal{S}^N$ is called a *hyperarc rate vector* if each element $g_a, a \in N$, is a hyperarc rate function. The convex cone of all hyperarc rate vectors is denoted by $\mathcal{H}_+^N \subset \mathcal{S}_+^N$. A subset $\mathcal{G} \subset \mathcal{H}_+^N$ is referred to as *hyperarc rate region*.

Each hyperarc rate vector $g \in \mathcal{G}$ quantifies the flow capacities of all hyperarcs in the network and represents simultaneously achievable information rates on all hyperarcs for the considered wireless network model. The hyperarc rate region \mathcal{G}

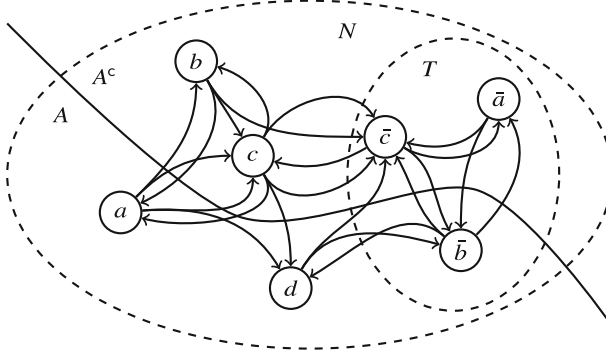


Fig. 2.5 Hypergraph model for an exemplary multicast network with seven nodes $N = \{a, b, c, d, \bar{c}, \bar{b}, \bar{a}\}$ and three terminals $T = \{\bar{c}, \bar{b}, \bar{a}\}$. Each single-tail multi-head arrow indicates a hyperarc originating at the node where the arrow is rooted with receiver set indicated by the arrow heads, e.g., two hyperarcs originate at a with receiver sets $\{b, c\}$ and $\{c, d\}$. Only hyperarcs with nonzero flow capacities according to the hyperarc rate vector g are drawn. The depicted cut $A = \{a, d, \bar{b}\}$ with terminal side $A^c = \{b, c, \bar{c}, \bar{a}\}$ offers a cut rate $v(A) = g_a(\{b, c\}) + g_a(\{c, d\}) + g_d(\{\bar{c}, \bar{b}\}) + g_{\bar{b}}(\{d, \bar{c}\}) + g_{\bar{b}}(\{\bar{a}\})$ according to (2.10) dropping all flow capacities that are zero from the sum and considering only the nonzero flow capacities

represents all available policies and parameters, e.g., transmission and coding strategies, channel input distributions, etc., by their corresponding hyperarc rate vector $g \in \mathcal{G}$. Figure 2.5 shows a hypergraph model for the example network in Fig. 2.1, showing only hyperarcs with nonzero flow capacities for one particular hyperarc rate vector g .

Definition 2.8 A cut rate region $\mathcal{V} \subset \mathcal{C}_+^N$ is generated by a hyperarc rate region $\mathcal{G} \subset \mathcal{X}_+^N$ if

$$\mathcal{V} = \bigcup_{g \in \mathcal{G}} \left\{ v \in \mathcal{C}_+^N : v(A) = \sum_{a \in A} \sum_{B \cap A^c \neq \emptyset} g_a(B) \quad \forall A \subset N \right\}. \quad (2.10)$$

The corresponding multicast rate region $\mathcal{R}(\mathcal{V})$ of a multicast network (N, T) is also denoted by $\mathcal{R}(\mathcal{G})$.

Proposition 2.2 A cut rate region $\mathcal{V} \subset \mathcal{C}_+^N$ generated by a hyperarc rate region $\mathcal{G} \subset \mathcal{X}_+^N$ is submodular.

The cut rate region generated from a hyperarc rate region can be interpreted as the direct generalization of the cut rate region for directed graphs (2.7) to directed hypergraphs, where the arc rates $k_a(b)$ are replaced by hyperarc rates $g_a(B)$. Each receiver $b \in B$ of a hyperarc (a, B) gets the same information over that hyperarc. Therefore, each hyperarc rate $g_a(B)$ is counted exactly once towards the cut rate $v(A)$ if the hyperarc crosses the cut, i.e., $a \in A$ and $B \cap A^c \neq \emptyset$. Similar to the graph model, independent information is communicated on all hyperarcs. As a

result, the total cut rate is computed by summing up all hyperarc rates of hyperarcs that cross the cut from left to right, see Fig. 2.5 for an example. The hypergraph model represents a decomposition of the wireless broadcast of each transmitter into independent pieces of information, each of which is conveyed to one particular subgroup of receivers. Furthermore, any cut rate region \mathcal{V} that is generated by hyperarc rate region \mathcal{G} according to (2.10) is submodular solely due to the structure of the transformation (2.10) and the nonnegativity of hyperarc rate vectors. This results is corollary to Proposition 2.5.

The hypergraph model clearly includes the classical graph model as a special case. Hyperarc rate functions of the form

$$g_a(B) = \begin{cases} k_a(b) & \text{if } B = \{b\}, \\ 0 & \text{otherwise} \end{cases} \quad (2.11)$$

in (2.10) yield exactly the cut rate function (2.7). That is, (2.11) defines the natural embedding of the cone of arc rate vectors \mathcal{A}_+^N into the cone of hyperarc rate vectors \mathcal{H}_+^N such that the cut rate function (2.7) generated by any arc rate vector $k \in \mathcal{A}_+^N$ matches the cut rate function (2.10) generated by the corresponding hyperarc rate vector $g \in \mathcal{H}_+^N$ according to (2.11).

Based on the definition of the hypergraph model, we state the max-flow min-cut theorem for multicast rate regions on hypergraphs, which extends the multicast min-cut max-flow theorem for the graph model (Theorem 2.1) towards the hypergraph model. This results has directly and indirectly been established in [8–10, 37].

Theorem 2.2 (Hypergraph max-flow min-cut) *Let (N, T) be a multicast network and $\mathcal{V} \subset \mathcal{C}_+^N$ be a cut rate region on (N, T) generated by a hyperarc rate region $\mathcal{G} \subset \mathcal{H}_+^N$. The multicast rate region $\mathcal{R}(\mathcal{G})$ is given by*

$$\begin{aligned} \mathcal{R}(\mathcal{G}) = \bigcup_{g \in \mathcal{G}} \bigg(\bigcap_{t \in T} \bigg\{ r \in \mathcal{R}_+^N : x^t \in (\mathcal{R}^N \otimes \mathcal{S}^N \otimes \mathcal{R}^N)_+, \\ \sum_{b \in C} x_{aCb}^t \leq g_a(C) \wedge \sum_{b \in C^c} x_{aCb}^t = 0 \quad \forall a \in N, C \subset N, \\ \sum_{b \in N} \sum_{C \subset N} x_{aCb}^t - \sum_{b \in N} \sum_{C \subset N} x_{bCa}^t = r_a \quad \forall a \in \{t\}^c \bigg\} \bigg). \end{aligned} \quad (2.12)$$

At the core of the max-flow formulation of the multicast rate region in the hypergraph model (2.12) are the nonnegative flows $x^t \in (\mathcal{R}^N \otimes \mathcal{S}^N \otimes \mathcal{R}^N)_+$ for all $t \in T$, representing the information flows through the hypergraph from all sources directed to each terminal $t \in T$ individually. $\mathcal{R}^N \otimes \mathcal{S}^N \otimes \mathcal{R}^N$ denotes the tensor product of \mathcal{R}^N , \mathcal{S}^N , and \mathcal{R}^N , i.e., real-valued functions on $N \times 2^N \times N$ or, equivalently, vectors indexed by a node $a \in N$, a subset of nodes $C \subset N$, and another node $b \in N$. $(\mathcal{R}^N \otimes \mathcal{S}^N \otimes \mathcal{R}^N)_+$ denotes the elementwise nonnegative cone of the tensor product $\mathcal{R}^N \otimes \mathcal{S}^N \otimes \mathcal{R}^N$. In particular, x_{aCb}^t represents the virtual information flow directed

to terminal t over the hyperarc (a, C) to the receiver b . As for the graph model, the *flow conservation law* (third line) states that the total outgoing flow at each node $a \in N$ must equal the total incoming flow at a and the multicast source rate of a . The *flow capacity constraint* (second line) ensures that, for each terminal $t \in T$, each hyperarc (a, C) can in total support an information flow to all its receivers $b \in C$ of no more than its flow capacity $g_a(C)$ and that it cannot support any positive flow to any other node $b \notin C$. Note that the loop-free condition together with the third line in (2.12) ensure that $x_{aCa}^t = 0$ for all $a \in N$ and $C \subset N$.

As in the cut rate formulation of the multicast rate region (2.12), the intersection with respect to all terminals $t \in T$ in (2.12) is due to the multicast nature of the communication. For single-source multicast, this means that only the worst terminal bounds the multicast rate of the source [9, 10], which is analogous to the same observation for graph models [7]. Finally, the union over all hyperarc rate vectors $g \in \mathcal{G}$, as the union over all cut rate functions in (2.2), represents all available policies, transmitter operations, etc. for the considered wireless network and channel model.

Theorem 2.2 can be proved by either applying the Ford-Fulkerson max-flow min-cut theorem extended to multicast on directed hypergraphs [9, 10, 37] or by replacing each hyperarc by a virtual node and multiple virtual arcs [8] and applying the Ford-Fulkerson max-flow min-cut theorem for multicast on directed graphs to the resulting augmented virtual graph. The proof is omitted in this work.

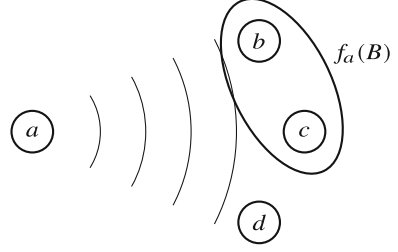
A simplified hypergraph flow formulation for the multicast source rate region was found in [38]. This formulation exploits the inherent polymatroid structure of the flow capacity constraint in (2.12) (second line) and the Minkowski sum characterization of polymatroid polyhedra to significantly reduce the number of flow variables. The resulting flow formulation is closely related to the flow formulation corresponding to the polymatroid broadcast model in Sect. 2.4. The relation of these two flow formulations is discussed in Sect. 2.5.

2.4 Polymatroid Broadcast Model

Splitting the broadcast into point-to-multipoint links (hypergraph model) instead of point-to-point links (graph model) is still not sufficient to fully capture the wireless broadcast advantage. The transmitter-centric approach of splitting the information into independent messages across different receiver groups and forcing each receiver of each group to perfectly decode the corresponding messages imposes a one-hop all-or-nothing structure with respect to these information messages into the wireless broadcast model. This approach is too restrictive to capture that all receivers see a similar but differently degraded version of the transmit signal. Therefore, the applicability of the hypergraph model is limited to a few types of wireless network models, cf. Chap. 4 et seq.

As a generalization of the hypergraph model, we propose the polymatroid broadcast model in this section, which removes this particular structure by focusing on the received information at all groups of receivers, i.e., a receiver-centric model. To

Fig. 2.6 Broadcast rate function model for wireless communication showing only the broadcast rate $f_a(B)$ from a to B with $B = \{b, c\}$ and its local cut interpretation



this end, we define a broadcast rate function f_a for each transmitter $a \in N$ that quantifies the rate $f_a(B)$ at which each subgroup of receivers $B \subset N$ can extract information from a transmission of node a . This means that for any cut A such that $a \in A$, $f_a(A^c)$ quantifies the contribution of a to the cut rate function $v(A)$. It can therefore be interpreted as a local version of the cut rate function $v(A)$, cf. Fig. 2.6. The special feature of the proposed broadcast function model is that it preserves the submodularity of the cut rate functions by enforcing a polymatroid structure on the broadcast functions.

Definition 2.9 A set function $f_a \in \mathcal{S}^N$ on N , i.e., $f_a: 2^N \rightarrow \mathbb{R}$, originating at node $a \in N$ is called a *broadcast rate function* if it satisfies the following four conditions:

$$f_a(\emptyset) = 0, \quad (\text{normalized})$$

$$f_a(B) \leq f_a(C) \quad \forall B \subset C \subset N, \quad (\text{monotone})$$

$$f_a(B) + f_a(C) \geq f_a(B \cup C) + f_a(B \cap C) \quad \forall B, C \subset N, \quad (\text{submodular})$$

$$f_a(B) = f_a(B - \{a\}) \quad \forall B \subset N. \quad (\text{loop-free})$$

The normalization, monotonicity, and submodularity conditions in Definition 2.9 define a *rank function* of a polymatroid (N, f_a) over the ground set N [27]. These conditions admit a natural interpretation in the context of wireless broadcast, where $f_a(B)$ represents the total information that the group of nodes B can extract from a broadcast of node a . The *normalization* and *monotonicity* conditions ensure that the total information available to any group of receivers is nonnegative and no smaller than the total information available to any of its subgroups. This means that adding further receivers to a group cannot decrease the total information available to the group. The *submodularity*⁷ condition in Definition 2.9 is equivalent to a *diminishing returns* condition [27]:

$$f_a(B \cup D) - f_a(B) \geq f_a(C \cup D) - f_a(C) \quad \forall D \subset N, B \subset C \subset D^c. \quad (2.13)$$

⁷Submodularity is considered the “discrete analogue of convexity” [39] due to the Lovász extension [26, 29, 39].

That is, the marginal broadcast rate that any group of receiver nodes D adds to a group B is at least as large as the gain which D contributes to any larger group C containing B . We remark that these conditions are tightly connected to mutual information expressions associated with the information transfer in networks of independent broadcast channels, in particular, deterministic, erasure, and Gaussian broadcast channels, cf. Chap. 4 et seq.

Finally, the *loop-free* condition ensures that a gains no information from its own broadcast. This is a technical condition and comes without loss of generality for the forthcoming results. It is equivalent to neglecting loops and hyper-loops, i.e., (hyper)arcs that start and end at the same node, in the graph and hypergraph models. It also gives rise to a general definition of the *neighbors* $N_a(f_a) \subset N$ of a node a with respect to a broadcast rate function f_a as

$$N_a(f_a) = \{b \in N : f_a(\{b\}) > 0\}. \quad (2.14)$$

The loop-free condition implies that $a \notin N_a(f_a)$. Note that it is sufficient to characterize a broadcast rate function f_a only for neighbors of a since any broadcast rate function satisfies

$$f_a(B) = f_a(B \cap N_a(f_a)) \quad (2.15)$$

due to $f_a(B - \{b\}) \leq f_a(B) \leq f_a(\{b\}) + f_a(B - \{b\})$ for all $b \in B \subset N$ and $f_a(\{b\}) = 0$ for all $b \notin N_a(f_a)$. Since for larger networks the set of neighbors is usually much smaller than the network, we refer to (2.15) as the *locality* property of the broadcast.

We remark that a polymatroid rank function f_a over N is tightly connected to the *polymatroid polyhedron* in \mathcal{R}_+^N defined as

$$P(f_a) = \left\{ x_a \in \mathcal{R}_+^N : \sum_{b \in B} x_{ab} \leq f_a(B) \quad \forall B \subset N \right\}, \quad (2.16)$$

see Fig. 2.7 for an exemplary visualization. Note that the loop-free condition ensures $x_{aa} = 0$ for all $x_a \in P(f_a)$. It can be interpreted as the set of information flow vectors x_a from a to all other nodes that can be supported by a particular broadcast rate function f_a . Polymatroid polyhedra are a fundamental part of the max-flow characterization of wireless broadcast networks. The next step towards a broadcast model is the definition of vectors of broadcast rate functions and their relation to cut rate functions.

Definition 2.10 A vector $f = (f_a : a \in N) \in \mathcal{X}^N$ is called a *broadcast rate vector* if each element f_a , $a \in N$, is a broadcast rate function. The convex cone of all broadcast rate vectors is denoted by $\mathcal{B}_+^N \subset \mathcal{X}_+^N$. A subset $\mathcal{F} \subset \mathcal{B}_+^N$ is referred to as *broadcast rate region*.

Analogous to the graph and hypergraph models, each broadcast rate vector $f \in \mathcal{F}$ quantifies the information transfer of all transmitters to their receivers in the network

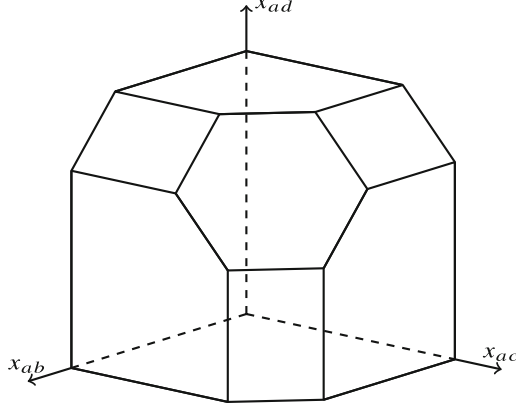


Fig. 2.7 Three dimensional visualization of the polymatroid polyhedron $P(f_a)$ corresponding to a broadcast function f_a with three neighbors $N_a(f_a) = \{b, c, d\}$

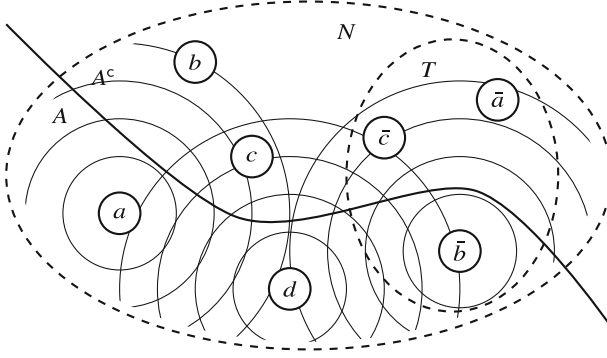


Fig. 2.8 Polymatroid broadcast model for an exemplary multicast network with seven nodes $N = \{a, b, c, d, \bar{c}, \bar{b}, \bar{a}\}$ and three terminals $T = \{\bar{c}, \bar{b}, \bar{a}\}$. The cut $A = \{a, d, \bar{b}\}$ with terminal side $A^c = \{b, c, \bar{c}, \bar{a}\}$ is depicted. The thin circles indicate the broadcasts of the nodes on the source side of the cut A , whose cut rate is given by $v(A) = f_a(A^c) + f_d(A^c) + f_{\bar{b}}(A^c)$ according to (2.17) for the considered broadcast rate vector f

simultaneously. The broadcast rate region \mathcal{F} represents all available policies and parameters, e.g., transmission and coding strategies, channel input distributions, etc., by their corresponding broadcast rate vector $f \in \mathcal{F}$. Figure 2.8 shows a broadcast model for the example network in Fig. 2.1. Only the broadcasts of nodes in $A = \{a, d, \bar{b}\}$ are indicated in the figure.

Definition 2.11 A cut rate region $\mathcal{V} \subset \mathcal{C}_+^N$ is generated by a broadcast rate region $\mathcal{F} \subset \mathcal{X}_+^N$ if

$$\mathcal{V} = \bigcup_{f \in \mathcal{F}} \left\{ v \in \mathcal{C}_+^N : v(A) = \sum_{a \in A} f_a(A^c) \quad \forall A \subset N \right\}. \quad (2.17)$$

The corresponding multicast rate region $\mathcal{R}(\mathcal{V})$ of a multicast network (N, T) is also denoted by $\mathcal{R}(\mathcal{F})$.

The cut rate region (2.17) is again a generalization of the cut rate region generated by the hypergraph model (2.10). For each node on the source side of the cut A , the sum of the crossing hyperarc rates $\sum_{B \cap A^c \neq \emptyset} g_a(B)$ is replaced by the corresponding broadcast rate function $f_a(A^c)$. This alludes to the following connection between hyperarc rate functions and broadcast rate functions:

$$f_a(B) = \sum_{\substack{C \subseteq N: \\ C \cap B \neq \emptyset}} g_a(C). \quad (2.18)$$

Section 2.5 shows that the linear map defined by (2.18) indeed maps any hyperarc rate function onto a unique broadcast rate function according to Definition 2.9. However, the reverse assertion does not hold, i.e., there are broadcast rate functions that cannot be represented by hyperarc rate functions according to this map. Consequently, the polymatroid broadcast model generalizes the hypergraph model as it does not enforce an explicit transmitter side decomposition of the broadcast into different messages each directed to a particular receiver set. Instead, it directly quantifies the total amount of received information of any particular receiver subset and thus the contribution of the broadcast to the cut. This means that the polymatroid broadcast model is a local cut model where the contribution of each transmitter to a cut is separately quantified by that transmitter's broadcast function. The decomposition across all transmitters is analogous to the graph and hypergraph models.

Local cut approaches have previously been proposed for specific physical layer models in [40, 41], but, so far, they have not been structurally analyzed and generalized to larger classes of wireless network models. The central advantage of the proposed polymatroid broadcast model compared to other local cut approaches is that submodularity of both the local cut contribution, namely the broadcast rate functions, and the global cut value, namely the cut rate functions, is ensured. The latter is formally stated in the following proposition, which is a corollary to Proposition 2.5. It shows that one central aspect of the graph model with capacitated arcs and the hypergraph model with capacitated hyperarcs, i.e., the submodularity of the associated cut rate functions, is inherited by the polymatroid broadcast model.

Proposition 2.3 *A cut rate region $\mathcal{V} \subset \mathcal{C}_+^N$ generated by a broadcast rate region $\mathcal{F} \subset \mathcal{B}_+^N$ is submodular.*

Further important aspects of the graph and hypergraph models are their max-flow formulations of the multicast rate region and the respective multicast max-flow min-cut theorems. An analogous multicast max-flow min-cut theorem can also be stated for the polymatroid broadcast model based on the definitions of broadcast rate functions, broadcast rate vectors, and broadcast generated cut rate regions.

Theorem 2.3 (Polymatroid max-flow min-cut) *Let (N, T) be a multicast network and $\mathcal{V} \subset \mathcal{C}_+^N$ be a cut rate region on (N, T) generated by a broadcast rate region*

$\mathcal{F} \subset \mathcal{B}_+^N$. The multicast rate region $\mathcal{R}(\mathcal{F})$ is given by

$$\mathcal{R}(\mathcal{F}) = \bigcup_{f \in \mathcal{F}} \left(\bigcap_{t \in T} \left\{ r \in \mathcal{R}_+^N : x^t \in \mathcal{F}_+^N, \sum_{b \in B} x_{ab}^t \leq f_a(B) \quad \forall a \in N, B \subset N, \right. \right. \\ \left. \left. \sum_{b \in N} x_{ab}^t - \sum_{b \in N} x_{ba}^t = r_a \quad \forall a \in \{t\}^c \right\} \right). \quad (2.19)$$

We remark that the theorem is solely a consequence of the polymatroid structure of the broadcast rate functions f_a . In Sect. 2.11.1, the theorem is proved based on the general polymatroid max-flow min-cut theorem for polymatroid flow networks [12–14]. The general theorem relies heavily on the submodularity of the resulting cut rate function. Alternatively, the result can directly be proved by exploiting the linear programming duality, the submodularity of the cut rate functions, the polymatroid structure of the broadcast functions, and the greedy algorithm for linear optimization over polymatroids [27, 28] and submodular polyhedra [29].

The vector $x^t = (x_{ab}^t : a, b \in N)$ represents the virtual information flow vector from all sources except t to the terminal t . Whenever the multicast rate region $\mathcal{R}(\mathcal{F})$ represents an achievable rate region, i.e., an inner bound to the multicast capacity region, the flow x^t is closely connected to the amount of coded information that nodes need to inject with network coding. A multicast rate vector r can be supported by a broadcast rate vector f if such a flow exists for each terminal $t \in T$. Similar to the max-flow formations for graphs and hypergraphs, the formulation (2.19) is based on the *flow conservation law* (second line) and *flow capacity constraint* (last part of the first line). The flow conservation law is identical to the one in the max-flow formulation for graphs. The flow capacity constraint defines a polymatroid polyhedron according to (2.16) for each transmitter $a \in N$. They are a local version of the cut constraints in the cut rate model (2.2) with (2.17) since the polymatroid polyhedron associated with each node a bounds the outgoing flow of node a to any group of receivers B by the maximum contribution that a transmission of a can contribute to the cut with terminal side $A^c = B$.

In analogy to the graph and hypergraph model, the intersection with respect to all terminals $t \in T$ in (2.19), as in the cut rate formulation (2.2), is due to the multicast nature of the communication. Finally, the union over all broadcast rate vectors $f \in \mathcal{F}$, as the union over all cut rate functions in (2.2), represents all available policies, transmitter operations, etc. for the considered wireless network and channel model.

2.5 Transformation of Models

We have introduced four different models for network coded multicast rate regions, namely a cut rate region model and three flow-based models, an arc rate region model, a hyperarc rate region model, and a broadcast rate region model. Simple linear

transformations connect all three flow models to the cut model with submodular cut rate regions. Furthermore, max-flow min-cut theorems show that the flow-based multicast rate region formulations coincide with the corresponding formulations based on the corresponding cut rate regions according to these linear transformations. In this section, we study these transformations in detail and derive existence and uniqueness theorems for these flow models.

Definition 2.12 (*Model transformations*)

- (a) The *arc-to-hyperarc rate transformation* is defined as the map $\Upsilon: \mathcal{F}^N \rightarrow \mathcal{Z}^N$ with $g = \Upsilon k$ for any $k \in \mathcal{F}^N$ such that for all $a \in N$ and $B \subset N$

$$g_a(B) = \begin{cases} k_a(b) & \text{if } B = \{b\}, b \in N, \\ 0 & \text{otherwise.} \end{cases} \quad (2.20)$$

- (b) The *hyperarc-to-broadcast rate transformation* is defined as the map $\Gamma: \mathcal{Z}^N \rightarrow \mathcal{Z}^N$ with $f = \Gamma g$ for any $g \in \mathcal{Z}^N$ such that for all $a \in N$ and $B \subset N$

$$f_a(B) = \sum_{A \subset N: A \cap B \neq \emptyset} g_a(A). \quad (2.21)$$

- (c) The *arc-to-cut rate transformation* is defined as the map $\Xi: \mathcal{F}^N \rightarrow \mathcal{C}^N$ with $v = \Xi k$ for any $k \in \mathcal{F}^N$ such that for all $A \subset N$

$$v(A) = \sum_{a \in A} \sum_{b \in A^c} k_a(b). \quad (2.22)$$

- (d) The *hyperarc-to-cut rate transformation* is defined as the map $\Omega: \mathcal{Z}^N \rightarrow \mathcal{C}^N$ with $v = \Omega g$ for any $g \in \mathcal{Z}^N$ such that for all $A \subset N$

$$v(A) = \sum_{a \in A} \sum_{B \subset N: B \cap A^c \neq \emptyset} g_a(B). \quad (2.23)$$

- (e) The *broadcast-to-cut rate transformation* is defined as the map $\Lambda: \mathcal{Z}^N \rightarrow \mathcal{C}^N$ with $v = \Lambda f$ for any $f \in \mathcal{Z}^N$ such that for all $A \subset N$

$$v(A) = \sum_{a \in A} f_a(A^c). \quad (2.24)$$

These transformations formalize the definitions of hyperarc rate regions generated from arc rate regions, i.e., $\mathcal{G} = \Upsilon(\mathcal{K})$, broadcast rate regions generated from hyperarc rate regions, i.e., $\mathcal{F} = \Gamma(\mathcal{G})$, cut rate regions generated from arc rate regions, hyperarc rate regions, and broadcast rate regions, i.e., $\mathcal{V} = \Xi(\mathcal{K})$, $\mathcal{V} = \Omega(\mathcal{G})$, and $\mathcal{V} = \Lambda(\mathcal{F})$, respectively. The following proposition establishes the fundamental properties

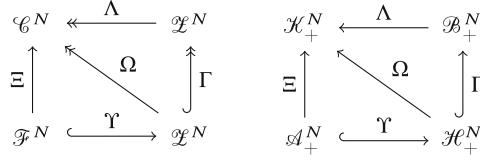


Fig. 2.9 Commutative diagrams of the transformations Υ , Γ , Ξ , Ω , and Λ . A linear map is represented by a simple arrow, a surjective linear map by a double-headed arrow, an injective linear map by a hooked-tail arrow, and a bijective linear map by a hooked-tail double-headed arrow. The *left* diagram relates the transformations defined on their vector space domains \mathcal{Z}^N and \mathcal{F}^N and their codomains \mathcal{Z}^N and \mathcal{C}^N . The *right* diagram relates these transformations when they are restricted to the domains \mathcal{A}_+^N (arc rate functions), \mathcal{H}_+^N (hyperarc rate functions), and \mathcal{B}_+^N (broadcast rate functions) and codomains \mathcal{H}_+^N , \mathcal{B}_+^N , and \mathcal{K}_+^N (nonnegative, two-sided normalized, submodular cut rate functions). The surjectivity and injectivity properties of all maps are drawn for $|N| \geq 4$. The diagrams are also valid for $|N| = 3$ if the arrow associated with Γ in the *right* diagram is replaced by a hooked-tail double-headed arrow (bijective linear map)

of these transformations, in particular, whether they are injective,⁸ surjective,⁹ or bijective,¹⁰ and their relations among each other. The left part of Fig. 2.9 summarizes these results in a commutative diagram. The proof is given in Sect. 2.11.2.

Proposition 2.4 *Let the node set N contain at least three nodes, i.e., $|N| \geq 3$. The transformations Υ , Γ , Ξ , Ω , and Λ are linear, consistent in the sense that the left diagram in Fig. 2.9 is commutative, and they have the following properties:*

- (a) $\Upsilon: \mathcal{F}^N \rightarrow \mathcal{Z}^N$ is injective but not surjective.
- (b) $\Gamma: \mathcal{Z}^N \rightarrow \mathcal{Z}^N$ is bijective and its inverse $\Gamma^{-1}: \mathcal{Z}^N \rightarrow \mathcal{Z}^N$ with $g = \Gamma^{-1}f$ is given by

$$g_a(B) = \sum_{A \subset B} (-1)^{|A|+1} f_a(A \cup B^c) \quad (2.25)$$

for all $a \in N$ and $B \subset N$ with $B \neq \emptyset$.

- (c) $\Lambda: \mathcal{Z}^N \rightarrow \mathcal{C}^N$ is surjective but not injective.
- (d) $\Omega: \mathcal{Z}^N \rightarrow \mathcal{C}^N$ is surjective but not injective.
- (e) $\Xi: \mathcal{F}^N \rightarrow \mathcal{C}^N$ is neither injective nor surjective.

This proposition shows that the transformations in Definition 2.12 are linear and consistent with each other, i.e., $\Omega = \Lambda\Gamma$ and $\Xi = \Omega\Upsilon = \Lambda\Gamma\Upsilon$, which can both be easily verified by plugging the respective definitions into each other. Furthermore, it establishes the invertibility properties of these transformations, in particular, the invertibility of Γ on \mathcal{Z}^N and the invertibility of Υ on its range $\Upsilon(\mathcal{F}^N)$, which are key features of the connections between the models.

⁸A map $\Phi: \mathcal{X} \rightarrow \mathcal{Y}$ is injective (one-to-one) if for all $\xi, \xi' \in \mathcal{X}$, $\xi \neq \xi'$ implies $\Phi(\xi) \neq \Phi(\xi')$.

⁹A map $\Phi: \mathcal{X} \rightarrow \mathcal{Y}$ is surjective (onto) if for any $\eta \in \mathcal{Y}$, there exists $\xi \in \mathcal{X}$ such that $\eta = \Phi(\xi)$.

¹⁰A map $\Phi: \mathcal{X} \rightarrow \mathcal{Y}$ is bijective (one-to-one and onto) if it is both injective and surjective.

The analysis of the relations between cut rate functions, arc rate functions, hyperarc rate functions, and broadcast rate functions requires to study of the restrictions of the maps in Definition 2.12 to the cones \mathcal{A}_+^N , \mathcal{H}_+^N and \mathcal{B}_+^N . The following proposition, which is also proved in Sect. 2.11.2 using the results from Proposition 2.4 and appropriate counter examples, establishes the ranges of the restrictions of these maps to the cones of interest and states their invertibility properties for nontrivial node sets N . The results are summarized by the commutative diagram in the right part of Fig. 2.9.

Proposition 2.5 *Let the node set N contain at least three nodes, i.e., $|N| \geq 3$. The restrictions of Υ , Γ , Ξ , Ω , and Λ to \mathcal{A}_+^N , \mathcal{H}_+^N , \mathcal{A}_+^N , \mathcal{H}_+^N , and \mathcal{B}_+^N , respectively, are well-defined and consistent such that the right diagram in Fig. 2.9 is commutative, i.e., $\Upsilon(\mathcal{A}_+^N) \subset \mathcal{H}_+^N$, $\Gamma(\mathcal{H}_+^N) \subset \mathcal{B}_+^N$, and $\Xi(\mathcal{A}_+^N) \subset \Omega(\mathcal{H}_+^N) \subset \Lambda(\mathcal{B}_+^N) \subset \mathcal{K}_+^N$. Furthermore, the restricted transformations have the following properties:*

- (a) $\Upsilon: \mathcal{A}_+^N \rightarrow \mathcal{H}_+^N$ is injective but not surjective.
- (b) $\Gamma: \mathcal{H}_+^N \rightarrow \mathcal{B}_+^N$ is bijective if $|N| = 3$ and injective but not surjective if $|N| \geq 4$.
- (c) $\Lambda: \mathcal{B}_+^N \rightarrow \mathcal{K}_+^N$ is neither injective nor surjective.
- (d) $\Omega: \mathcal{H}_+^N \rightarrow \mathcal{K}_+^N$ is neither injective nor surjective.
- (e) $\Xi: \mathcal{A}_+^N \rightarrow \mathcal{K}_+^N$ is neither injective nor surjective.

Proposition 2.5 implies a simple model hierarchy in terms of strict generalization as a consequence of the relations among the restricted maps and their properties. The following corollary to Proposition 2.5 explicitly states this hierarchy and establishes Propositions 2.1, 2.2, and 2.3.

Corollary 2.1 *If the node set N contains at least four nodes, then the following strict inclusions hold: $\Upsilon(\mathcal{A}_+^N) \subsetneq \mathcal{H}_+^N$, $\Gamma(\mathcal{H}_+^N) \subsetneq \mathcal{B}_+^N$, and $\Lambda(\mathcal{B}_+^N) \subsetneq \mathcal{K}_+^N$. That is, the hypergraph model strictly generalizes the graph model, the polymatroid broadcast model strictly generalizes the hypergraph models, and the submodular cut model strictly generalizes the polymatroid broadcast model. Furthermore, cut rate functions that are generated by arc rate functions $\Xi(\mathcal{A}_+^N)$, by hyperarc rate functions $\Omega(\mathcal{H}_+^N)$, or by broadcast rate functions $\Lambda(\mathcal{B}_+^N)$ are submodular, i.e., all three sets are strictly contained in \mathcal{K}_+^N .*

The model hierarchy also implies that the graph model and the hypergraph model have an inherent polymatroid broadcast structure by applying the maps $\Gamma\Upsilon$ and Γ to arc rate vectors $k \in \mathcal{A}_+^N$ and hyperarc rate vectors $g \in \mathcal{H}_+^N$, respectively. We further remark that Proposition 2.5 also implies that each hyperarc rate vector can uniquely be represented by a broadcast rate vector and that each arc rate vector can uniquely be represented by a hyperarc rate vector. However, there are multiple arc, hyperarc, and broadcast rate vectors that lead to the same submodular cut rate functions. Furthermore, the hierarchical inclusion of the models is strict for $|N| \geq 4$ in the sense that there are submodular cut rate functions that cannot be represented by broadcast rate vectors, broadcast rate vectors that cannot be represented by hyperarc rate vectors, and hyperarc rate vectors that cannot be represented by arc rate vectors.

Due to the closed convex conic nature of the sets \mathcal{A}_+^N , \mathcal{H}_+^N , \mathcal{B}_+^N , and \mathcal{K}_+^N , these strict inclusion results directly imply that none of the strictly included sets can approximate the surrounding set. That is, the surrounding sets contain vectors that are arbitrarily far away from the included sets. This is formally stated in the following proposition, which is proved in Sect. 2.11.2.

Proposition 2.6 *If the node set $|N|$ contains at least four nodes, then none of the strict inclusions can be approximated, i.e.,*

$$\sup_{g \in \mathcal{H}_+^N} \inf_{k \in \mathcal{A}_+^N} \|g - \Upsilon k\| = \infty, \quad (2.26)$$

$$\sup_{f \in \mathcal{B}_+^N} \inf_{g \in \mathcal{H}_+^N} \|f - \Gamma g\| = \infty, \quad (2.27)$$

$$\sup_{v \in \mathcal{K}_+^N} \inf_{f \in \mathcal{B}_+^N} \|v - \Lambda f\| = \infty. \quad (2.28)$$

Corollary 2.1 does not only connect the models on the cut, broadcast, hyperarc, and arc rate level, but it also connects the multicast rate regions that have been defined in (2.2), (2.8), (2.12), and (2.19). This connection between the models is shown in Fig. 2.10. The figure indicates the theorems that state the equivalence or inclusion of the different multicast rate region formulations based on the three models. It is again clear that the cut rate model with submodular cut functions is the most general of all four models.

We also observe that the broadcast flow model can be derived from the hyperarc flow model, provided it exists, in two different ways: The first approach essentially combines the main results from Sects. 2.3, 2.4, and 2.5, using first the hyperarc max-flow min-cut theorem, the model transformation theorem, and finally the polymatroid broadcast max-flow min-cut theorem. The alternative approach derives the polymatroid flow model directly from the hyperarc flow model based on the Minkowski sum of polymatroids, see Theorem 2.4 and [26, 27] for its proof. This theorem has been exploited in [38] to simplify the hyperarc flow formulation for wireless packet erasure networks.

Theorem 2.4 *Let f, f' be polymatroid rank functions on a common ground set N . The Minkowski sum of the polymatroid polyhedra $P(f)$ and $P(f')$ is a polymatroid polyhedron with rank function $f + f'$ on the ground set N , i.e., $P(f) + P(f') = P(f + f')$.*

The hyperarc capacity constraint in (2.12) and the nonnegativity of the flow define simple polymatroids with rank functions

$$f_{aC}(B) = \begin{cases} g_a(C) & \text{if } B \cap C \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \quad (2.29)$$

Note that although the polymatroid $P(f_{aC})$ is defined by $2^{|N|}$ inequalities, all but two are redundant due to the special structure of the rank function $f_{aC}(B)$. Furthermore,

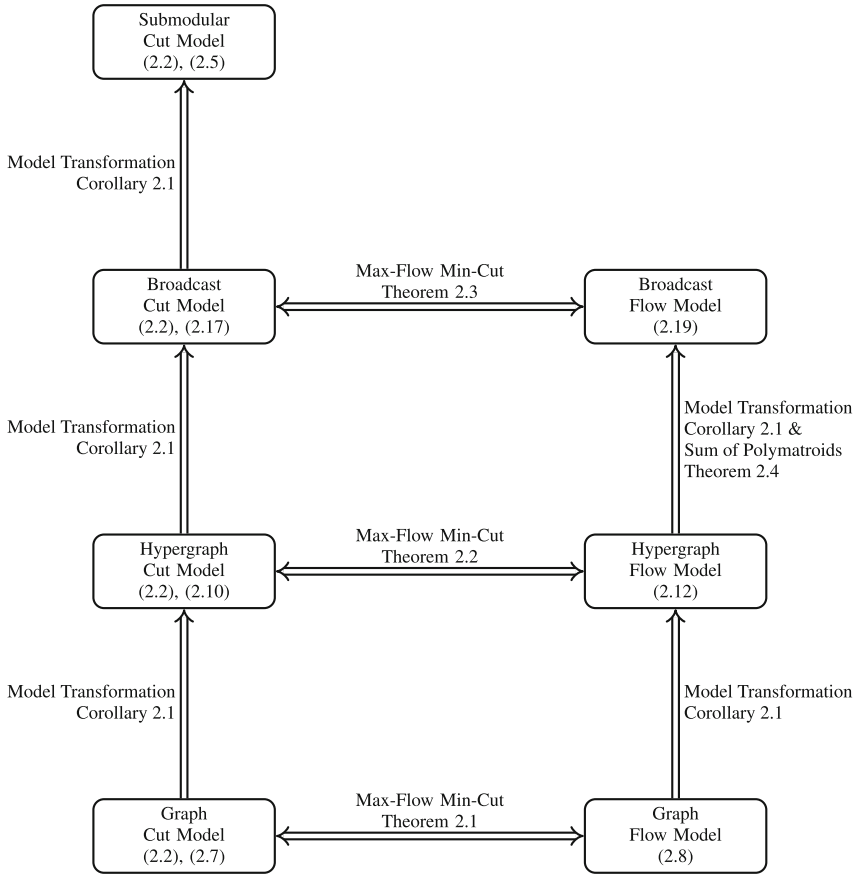


Fig. 2.10 Relation of multicast rate regions defined by flow and cut models generated by the graph model, the hypergraph model, the polymatroid broadcast model, and the submodular cut model. The arrows indicate which theorems are used to connect the multicast rate region formulations of the models. Two-sided arrows indicate equivalence, whereas one-sided arrows indicate that a transformation works in general only in one direction

we observe that

$$f_a(B) = \sum_{C \subset N} f_{aC}(B) = \sum_{C \cap B \neq \emptyset} g_a(C). \quad (2.30)$$

Applying Theorem 2.4, which states that the Minkowski sum of finitely many polymatroids is equivalent to the polymatroid defined by the sum of the rank functions, yields

$$P(f_a) = P\left(\sum_{C \subset N} f_{aC}\right) = \sum_{C \subset N} P(f_{aC}). \quad (2.31)$$

Therefore, $P(f_a)$ defines the same polyhedron as $\sum_{C \subset N} P(f_{aC})$, and consequently, the hyperarc and polymatroid broadcast flow formulations define the same multicast rate regions. Note that this conclusion implies the polymatroid max-flow min-cut theorem only for broadcast functions that are generated from hyperarc rate functions, but not in general.

Finally, we remark that Proposition 2.6 does not state anything on approximating multicast rate regions by either of the models. This means that the theorem states that there are $v \in \mathcal{H}_+^N$ and $f \in \mathcal{B}_+^N$ that cannot be approximated by any $f \in \mathcal{B}_+^N$ and $g \in \mathcal{H}_+^N$, respectively, but there may still exist sets \mathcal{F} and \mathcal{G} such that $\mathcal{R}(\{v\})$ is approximated by $\mathcal{R}(\mathcal{F})$ and $\mathcal{R}(\{f\})$ by $\mathcal{R}(\mathcal{G})$, respectively. Whether such sets exist and how they can be constructed remains open. Nevertheless, Theorem 2.6 implies that, if such sets exist, they must contain more than a single or a few elements.

2.6 Generalized Cut Model

Before concluding the discussion on rate region models for wireless multicast, the generalized cut model for multicast rate regions is introduced. It subsumes all previously presented models and leads to a more general wireless broadcast model, which combine broadcast rate functions with penalty rates. These models are suited to represent inner bounds on multicast capacity regions where each node spends some amount of its multicast rate to transmit side information for decoding to the terminals. This rate reduces the achievable rate that the broadcast rate vectors support and is referred to as penalty rate. Noisy network coding [15] defines a multicast rate region which exhibits such penalty rate, see Chap. 4.

Definition 2.13 A normalized set function $u \in \mathcal{N}^N$ is called a (*generalized*) *cut rate function*. The set of all (generalized) cut rate functions is the vector space \mathcal{N}^N . A subset $\mathcal{U} \subset \mathcal{N}^N$ is referred to as (*generalized*) *cut rate region*. A nonnegative multicast rate vector $r \in \mathcal{R}_+^N$ is supported by a cut rate function $u \in \mathcal{N}^N$ on a multicast network (N, T) if

$$\sum_{a \in A} r_a \leq u(A) \quad \forall t \in T, \quad A \subset \{t\}^c. \quad (2.32)$$

The multicast rate region $\mathcal{R}(\mathcal{U}) \subset \mathcal{R}_+^N$ of the multicast network (N, T) with cut rate region \mathcal{U} is defined as the set of nonnegative multicast rate vectors supported by \mathcal{U} , i.e.,

$$\mathcal{R}(\mathcal{U}) = \bigcup_{u \in \mathcal{U}} \left(\bigcap_{t \in T} \left\{ r \in \mathcal{R}_+^N : \sum_{a \in A} r_a \leq u(A) \quad \forall A \subset \{t\}^c \right\} \right). \quad (2.33)$$

We remark that the generalized definition is almost identical to the original definition of cut rate functions, which were required to be nonnegative and two-sided normalized. In particular, the cone \mathcal{C}_+^N is obviously a subset of the cone

of generalized cut rate functions \mathcal{N}^N . Furthermore, the value $u(N)$ of all $u \in \mathcal{U}$ is irrelevant for the definition of the multicast rate region $\mathcal{R}(\mathcal{U})$. Contrary to multicast rate regions that stem from nonnegative cut rate regions, where the rate vector $r = 0$ is always included, generalized cut rate regions may lead to empty multicast rate regions since negative cut rates cannot support nonnegative multicast rates. In order to analyze this effect, consider the alternative formulation of the multicast rate region $\mathcal{R}(\mathcal{U})$ given by

$$\mathcal{R}(\mathcal{U}) = \bigcup_{u \in \mathcal{U}} \left(\bigcap_{t \in T} \left\{ r \in \mathcal{R}_+^N : \sum_{a \in A} r_a \leq \min_{A \subset C \subset \{t\}^c} u(C) \quad \forall A \subset \{t\}^c \right\} \right). \quad (2.34)$$

It is the analogous formulation to (2.4) for nonnegative cut rate functions and follows by combining the cut constraints in the original definition of the multicast rate region (2.33) with the nonnegativity of the rate vectors, namely $r \in \mathcal{R}_+^N$. This formulation reveals that $\mathcal{R}(\mathcal{U})$ is nonempty if and only if there exists $u \in \mathcal{U}$ such that

$$u(A) \geq 0 \quad \forall A \subset N : A^c \cap T \neq \emptyset. \quad (2.35)$$

For example, a sufficient condition for $\mathcal{R}(\mathcal{U})$ being nonempty is that the zero cut rate function θ defined by $\theta(A) = 0$ for all $A \subset N$ is contained in \mathcal{U} . Note that although (2.35) requires that a generalized cut rate function is nonnegative for most cuts to support any multicast rates at all, generalized cut rate regions are still an important generalization of nonnegative cut rate regions since there are achievable multicast rate regions, most notably noisy network coding (Chap. 4), which cannot be modeled by nonnegative cut rate functions.

Definition 2.14 A generalized cut rate region $\mathcal{U} \subset \mathcal{N}^N$ is called *submodular* if all cut rate functions $u \in \mathcal{U}$ are submodular set functions on N , i.e.,

$$u(A) + u(B) \geq u(A \cup B) + u(A \cap B) \quad \forall A, B \subset N. \quad (2.36)$$

The convex cone of submodular cut rate functions is denoted by \mathcal{K}^N .

One particularly interesting construction of a cut rate function that is not necessarily nonnegative is the combination of a nonnegative cut rate function with a nonpositive modular function generated by a penalty rate vector.

Definition 2.15 A cut rate function $u \in \mathcal{N}^N$ is called a *penalized cut rate function* if it can be decomposed into a nonnegative cut rate function $v \in \mathcal{C}_+^N$ and a nonnegative rate vector $q \in \mathcal{R}_+^N$, called *penalty rate vector*, such that

$$u(A) = v(A) - \sum_{a \in A} q_a \quad (2.37)$$

for all $A \subset N$.

The rate vector q in the construction of the cut rate function u is referred to as penalty rate vector since it reduces the multicast rate vectors r that can be supported by a nonnegative cut rate function v . An equivalent construction for the multicast rate region $\mathcal{R}(\{u\})$ of a penalized cut rate function u based on the multicast rate region of the nonnegative cut rate function v , which clarifies the interpretation of q as a penalty rate vector, is

$$\mathcal{R}(\{u\}) = \left\{ r \in \mathcal{R}_+^N : r + q \in \mathcal{R}(\{v\}) \right\}. \quad (2.38)$$

Such penalty terms are especially useful to model the required information rate to transmit side information, e.g., quantized received signals, from the left side of the cut A to the nodes on the right side A^c . The penalized cut rate function u describes the available cut rate to the nodes after subtracting the penalty rate to deliver the side information to the terminals.

Note that $u \notin \mathcal{C}_+^N$ is no cut rate function in the strict sense of \mathcal{C}_+^N , except for the trivial case $q = 0$, since $v(N) = 0$ and $q \geq 0$ imply $u(N) = -\sum_{a \in N} q_a < 0$ if $q \neq 0$. Thus, the set of all penalized cut rate functions that are elementwise nonnegative is equivalent to the cone of nonnegative cut rate functions. Nevertheless, Definition 2.13 shows that the multicast rate region definition based on nonnegative cut rate functions can be naturally extended from nonnegative cut rate functions to penalized cut rate functions since the value of $u(N)$ is irrelevant for the multicast rate region even if $T = N$. This originates in the definition of $\mathcal{R}(\mathcal{U})$ where a cut A needs at least one terminal node opposite to any set of source nodes, i.e., A^c must be nonempty.

Finally, Definition 2.15 implies that a penalized cut rate function u is obviously submodular if and only if it can be decomposed into a submodular cut rate function v and a penalty rate vector since q according to (2.37) defines a modular (sub- and supermodular) function $\sum_{a \in A} q_a$ in the decomposition. In particular, if there are multiple different decompositions of the form (2.37), then the cut rate function must be submodular for all of them.

2.7 Penalized Polymatroid Broadcast Model

The penalized broadcast model presented in this section is one particular example for a model that cannot be represented by nonnegative cut rate functions. The model is obtained by augmenting the broadcast function model from Sect. 2.4 with a penalty rate vector in the same way as for the penalized cut rate function model in the previous section. The resulting model shares many properties with the original broadcast rate function model, especially the submodularity of the generated cut rate functions and the max-flow min-cut interpretation. As a special case of this model, we could also define a penalized hyperarc model, which shares the same properties with the

original hyperarc model in Sect. 2.3. However, since this model is included in the corresponding broadcast model, we refrain from presenting it in this book.

Definition 2.16 A vector $(f, h) \in \mathcal{Z}^N \times \mathcal{R}^N$ is called a *penalized broadcast rate vector* if $f \in \mathcal{B}_+^N$ is a broadcast rate vector and $h \in \mathcal{R}_+^N$ is a *penalty rate vector*. The convex cone of all penalized broadcast rate vectors is denoted by $\mathcal{P}_+^N = \mathcal{B}_+^N \times \mathcal{R}_+^N$. A subset $\mathcal{E} \subset \mathcal{P}_+^N$ is referred to as *penalized broadcast rate region*. A *cut rate region* $\mathcal{U} \subset \mathcal{N}^N$ is *generated by a penalized broadcast rate region* $\mathcal{E} \subset \mathcal{P}_+^N$ if

$$\mathcal{U} = \Lambda_p(\mathcal{E}) = \left\{ u \in \mathcal{N}^N : u(A) = \sum_{a \in A} f_a(A^c) - \sum_{a \in A} h_a \quad \forall A \subset N, (f, h) \in \mathcal{E} \right\}. \quad (2.39)$$

The corresponding multicast rate region $\mathcal{R}(\mathcal{U})$ of a multicast network (N, T) according to (2.33) is also denoted by $\mathcal{R}(\mathcal{E})$.

The transformation from penalized broadcast rate vectors to cut rate functions according to (2.39) is a linear map $\Lambda_p: \mathcal{Z}^N \times \mathcal{R}^N \rightarrow \mathcal{N}^N$ defined by $u = \Lambda_p(f, h)$ with

$$u(A) = \sum_{a \in A} f_a(A^c) - \sum_{a \in A} h_a \quad \forall A \subset N. \quad (2.40)$$

We observe that this cut rate function is a penalized cut rate function according to Definition 2.15 where the nonnegative cut rate region part $v = \Lambda(f)$ is generated by the broadcast to cut rate transformation, cf. Sect. 2.5, and the penalty rate vector of the cut rate model q is identical to the penalty rate vector of the broadcast model h . This structure leads directly to the following result on the submodularity of $u = \Lambda_p(f, h)$:

Corollary 2.2 A cut rate region $\mathcal{U} \subset \mathcal{N}^N$ generated by a penalized broadcast rate region $\mathcal{E} \subset \mathcal{P}_+^N$ is submodular, i.e., $\mathcal{U} = \Lambda_p(\mathcal{E}) \subset \mathcal{K}^N$.

This corollary characterizes the submodularity of cut rate regions generated by penalized broadcast rate regions as a direct consequence of Proposition 2.5. It can be equivalently stated as $\Lambda_p(\mathcal{P}_+^N) \subset \mathcal{K}^N$, which means that the cone of cut rate functions that are generated by penalized broadcast rate vectors is a subset of the cone of submodular cut rate functions \mathcal{K}^N . Since the model transformation carries over to the extended models with penalty terms, the max-flow min-cut theorem for the polymatroid broadcast model directly leads to a similar theorem for the penalized broadcast model.

Corollary 2.3 (Penalized polymatroid max-flow min-cut) *Let (N, T) be a multicast network and $\mathcal{U} \subset \mathcal{N}^N$ be a cut rate region on (N, T) generated by a penalized broadcast rate region $\mathcal{E} \subset \mathcal{P}_+^N$. The multicast rate region $\mathcal{R}(\mathcal{E})$ is given by*

$$\mathcal{R}(\mathcal{E}) = \bigcup_{(f,h) \in \mathcal{E}} \left(\bigcap_{t \in T} \left\{ r \in \mathcal{R}_+^N : x^t \in \mathcal{F}_+^N, \sum_{b \in B} x_{ab}^t \leq f_a(B) \quad \forall a \in N, B \subset N, \right. \right. \\ \left. \left. \sum_{b \in N} x_{ab}^t - \sum_{b \in N} x_{ba}^t - h_a = r_a \quad \forall a \in \{t\}^c \right\} \right). \quad (2.41)$$

The relation to the polymatroid max-flow min-cut theorem (Theorem 2.3) is obvious from the characterization of the multicast rate region for generalized cut rate functions (2.33). For a penalized cut rate function $u = \Lambda_p(f, h)$ that can be decomposed into $v = \Lambda f$ and $q = h$, the penalty rate vector q just reduces the supported multicast rate vectors r of a cut rate function v , i.e., $r + h \in \mathcal{R}(\{v\})$, to account for side information and similar overhead that has to be communicated to the terminals of the multicast. The penalized flow formulation obviously differs only slightly from the polymatroid flow formulation (2.19) since the outgoing flow at each node a must support the penalty rate h_a , which represents the total side information that a needs to forward, in addition to the incoming flow $\sum_{b \in N} x_{ba}$ and the multicast rate r_a . An analogous flow formulation can be obtained for the hyperarc model, but is omitted since the hyperarc model is subsumed by the polymatroid broadcast model in both the normal and penalized versions. In particular, the multicast rate region transformation results and max-flow min-cut relations summarized in Fig. 2.10 hold for penalized models as well as for the corresponding normal models from Sects. 2.1, 2.3, and 2.4.

2.8 Rate Region Properties and Equivalence

This section introduces a notion of equivalence among cut rate regions with respect to the multicast rate regions they generate. Based on this equivalence relation, the topological properties of cut and broadcast rate regions and their implications for the generated multicast rate regions are studied. All proofs for results in this section are given in Sect. 2.11.3.

Definition 2.17 Two cut rate regions $\mathcal{U}, \mathcal{U}' \subset \mathcal{N}^N$ are *equivalent* if $\mathcal{R}(\mathcal{U}) = \mathcal{R}(\mathcal{U}')$ for all $T \subset N$.

Proposition 2.7 Let (N, T) be a multicast network and $\mathcal{U}, \mathcal{U}' \subset \mathcal{N}^N$ be two equivalent cut rate regions. Then, $\mathcal{U} \cup \mathcal{U}'$ is equivalent to \mathcal{U} , there exists a unique largest cut rate region $\mathcal{U}'' \subset \mathcal{N}^N$ equivalent to \mathcal{U} , and \mathcal{U}'' contains all cut rate regions equivalent to \mathcal{U} .

Proposition 2.7 shows that the notion of equivalence introduced in Definition 2.17 is stable with respect to unions and inclusions of equivalent regions, i.e., the union of two equivalent cut rate regions is again equivalent. In the remainder of this section, we discuss various properties of cut rate regions and their relations to the multicast

rate region and cut rate region equivalence. In particular, we focus on compactness, comprehensiveness, and convexity.

Proposition 2.8 *Let (N, T) be a multicast network and $\mathcal{U} \subset \mathcal{N}^N$ be a compact cut rate region.*

- (a) $\mathcal{R}(\mathcal{U})$ is closed.
- (b) $\mathcal{R}(\mathcal{U}) \cap \{r \in \mathcal{R}_+^N : r_t = 0\}$ is compact for all $t \in T$.
- (c) $\mathcal{R}(\mathcal{U})$ is compact if $|T| \geq 2$.

Proposition 2.8 asserts that if a cut rate region is compact, then the generated multicast rate region is also essentially compact. This means that it is compact for $|T| \geq 2$ and can safely be replaced by the compact region $\mathcal{R}(\mathcal{U}) \cap \{r \in \mathcal{R}_+^N : r_t = 0\}$ if $T = \{t\}$ for some $t \in N$ since the terminals' multicast rate r_t is irrelevant and can be set to zero without loss of generality. Combining these results with the definition of cut rate region equivalence (Definition 2.17) leads to slightly relaxed prerequisites in Proposition 2.8, i.e., not \mathcal{U} itself needs to be compact, but it is sufficient that \mathcal{U} is equivalent to some compact cut rate region \mathcal{U}' .

The analysis of Pareto efficiency and comprehensiveness for cut rate functions requires definitions of what more efficient means for multicast rates and cut rates. In particular, the comprehensive hull, which includes all vectors that are less or equally efficient than any vector of the original region, should only include valid rate vectors, where the meaning of valid depends on the type of the region. To this end, we define the comprehensive hull and the Pareto efficient set of a multicast rate region $\mathcal{R} \subset \mathcal{R}_+^N$ as

$$\text{comp } \mathcal{R} = \{r' \in \mathcal{R}_+^N : r' \leq r \in \mathcal{R}\}, \quad (2.42)$$

$$\text{par } \mathcal{R} = \{r \in \mathcal{R} : \mathcal{R} \cap \{r' : r' \geq r\} = \{r\}\} \quad (2.43)$$

and of a cut rate region $\mathcal{U} \subset \mathcal{N}^N$ as

$$\text{comp } \mathcal{U} = \{u' \in \mathcal{N}^N : u' \leq u \in \mathcal{U}\}, \quad (2.44)$$

$$\text{par } \mathcal{U} = \{u \in \mathcal{U} : \mathcal{U} \cap \{u' : u' \geq u\} = \{u\}\}. \quad (2.45)$$

The inequalities in the above definitions are elementwise or pointwise inequalities, i.e., $r' \leq r$ if and only if $r'_a \leq r_a$ for all $a \in N$ and $u' \leq u$ if and only if $u'(A) \leq u(A)$ for all $A \subset N$. Note that we do not distinguish in notation between the different definitions of comprehensive hull and Pareto efficient set operators since it is clear from the context to which type of region they are applied and which definition is therefore appropriate.

Proposition 2.9 *Let (N, T) be a multicast network and $\mathcal{U} \subset \mathcal{N}^N$ be a cut rate region.*

- (a) The multicast rate region $\mathcal{R}(\mathcal{U})$ is comprehensive, i.e., $\mathcal{R}(\mathcal{U}) = \text{comp } \mathcal{R}(\mathcal{U})$.
- (b) \mathcal{U} and $\text{comp } \mathcal{U}$ are equivalent.

- (c) If \mathcal{U} is closed and upper bounded, i.e., there exists $\bar{u} \in \mathcal{N}^N$ such that $u \leq \bar{u}$ for all $u \in \mathcal{U}$, then \mathcal{U} and $\text{par } \mathcal{U}$ are equivalent.

Proposition 2.9 draws connections between the comprehensiveness of multicast rate regions and comprehensiveness and Pareto efficiency of cut rate regions. For example, we can simply state

$$\mathcal{R}(\mathcal{U}) = \text{comp } \mathcal{R}(\mathcal{U}) = \mathcal{R}(\text{comp } \mathcal{U}) \quad (2.46)$$

for any arbitrary cut rate region $\mathcal{U} \subset \mathcal{N}^N$ and terminal set $T \subset N$. Therefore, only the comprehensive hull of the cut rate region determines the shape of the multicast rate region. In particular, combining this result with Proposition 2.7 yields that the largest cut rate region equivalent to \mathcal{U} must be comprehensive, but may nevertheless be larger than the comprehensive hull of \mathcal{U} . For closed and upper bounded cut rate regions, the comprehensive hull is uniquely determined by their Pareto efficient set. Since $\text{par } \mathcal{U} \subset \mathcal{U} \subset \text{comp } \mathcal{U}$ holds, we can think of $\text{par } \mathcal{U}$ and $\text{comp } \mathcal{U} \cap \mathcal{N}^N$ as the *smallest* and *largest* equivalent cut rate regions, respectively, obtained from exploiting only elementwise monotonicity. However, this does not mean that there are no other smaller or larger equivalent cut rate regions in general, and the notion of smallest cut rate region may not even be well-defined. The results in Proposition 2.9 are important when studying the convexity of the multicast rate region and for network utility optimization problems and dual decomposition approaches as for instance discussed in Chap. 3.

The final result of this section regarding cut rate regions provides a relation between the convexity of multicast and cut rate regions and identifies some minimal set of cut rate functions whose convex combinations generate the multicast rate regions corresponding to a particular cut rate region. It does so by combining the extreme points of a set with Pareto efficiency. To this end, let $\text{conv } \mathcal{U}$ and $\text{ext } \mathcal{U}$ denote the convex hull and the extreme points, respectively, of a cut rate region \mathcal{U} , and define those operators analogously for multicast rate regions.

Proposition 2.10 *Let (N, T) be a multicast network and $\mathcal{U} \subset \mathcal{N}^N$ be a cut rate region.*

- (a) *If \mathcal{U} is convex, then $\mathcal{R}(\mathcal{U})$ is convex.*
 (b) *If \mathcal{U} is convex, then $\text{comp } \mathcal{U}$ is convex.*
 (c) *If \mathcal{U} is closed, convex, and upper bounded, i.e., there exists $\bar{u} \in \mathcal{N}^N$ such that $u \leq \bar{u}$ for all $u \in \mathcal{U}$, then $\text{conv}(\text{ext } \mathcal{U} \cap \text{par } \mathcal{U})$ is equivalent to \mathcal{U} .*

The first result shows that the convexity of the multicast rate region is a consequence of the convexity of the supporting cut rate region. Additionally, it suffices to check convexity of the comprehensive hull of a cut rate region since the comprehensive hull of a convex region is convex and the comprehensive hull of a region is equivalent to the original region. The last result essentially states that for closed convex upper bounded cut rate regions, only a very particular set of points, i.e., the Pareto efficient extreme points, are necessary to generate the original multicast rate

region from their convex combinations. This is a consequence of the representation theorem for closed convex sets containing no line [42, Theorem 18.5] and the combination of upper boundedness and the irrelevancy of the comprehensive hull for the shape of the multicast rate region, which implies that only Pareto efficient points determine the shape of the multicast rate region. This result is particularly useful for the analysis of the submodular dual decomposition approach in Chap. 3, where the optimization subproblem which describes the cut rate region related part of the optimization problem depends only on the Pareto efficient extreme points of these regions.

The properties of cut rate regions and the results connecting these properties to the structure of the multicast rate region can be transferred straightforwardly to the penalized broadcast model. The requirements for this transfer are suitable notions of equivalence, comprehensiveness, and Pareto efficiency for penalized broadcast rate regions with respect to the corresponding concepts for cut rate regions and the transformation from penalized broadcast rate vectors to cut rate functions. To this end, equivalence of penalized broadcast rate regions is defined by the equivalence of the corresponding generated cut rate regions. As a consequence, Proposition 2.7 can directly be applied to penalized broadcast rate regions. Furthermore, since the transformation Λ_p is continuous, a compact penalized broadcast rate region generates a compact cut rate region. Therefore, Proposition 2.8 applies identically to penalized broadcast rate regions and their corresponding multicast rate regions.

In order to transfer the results on comprehensiveness and Pareto efficiency, a suitable definition of the comprehensive hull and the Pareto efficient set of a penalized broadcast rate region $\mathcal{E} \subset \mathcal{P}_+^N$ is required, which takes into account that the broadcast rate functions and the penalty rates contribute to the corresponding cut rate function with opposite signs. Therefore, we define

$$\text{comp } \mathcal{E} = \{(f', h') \in \mathcal{P}_+^N : f' \leq f, h' \geq h, (f, h) \in \mathcal{E}\}, \quad (2.47)$$

$$\text{par } \mathcal{E} = \{(f, h) \in \mathcal{E} : \mathcal{E} \cap \{(f', h') : f' \geq f, h' \leq h\} = \{(f, h)\}\} \quad (2.48)$$

for a penalized broadcast rate region $\mathcal{E} \subset \mathcal{P}_+^N$. The inequalities are again element-wise inequalities, i.e., $h' \leq h$ if and only if $h'_a \leq h_a$ for all $a \in N$ and $f' \leq f$ if and only if $f'_a(B) \leq f_a(B)$ for all $a \in N$ and $B \subset N$. The next proposition restates Propositions 2.9 and 2.10 for penalized broadcast rate regions.

Proposition 2.11 *Let (N, T) be a multicast network, $\mathcal{E} \subset \mathcal{P}_+^N$ be a penalized broadcast rate region, and $\mathcal{U} \subset \mathcal{N}^N$ be the cut rate region generated by \mathcal{E} .*

- (a) *If \mathcal{E} is closed and has upper bounded broadcast rates, i.e., there exists $\bar{f} \in \mathcal{B}_+^N$ such that $f \leq \bar{f}$ for all $(f, h) \in \mathcal{E}$, then \mathcal{U} is closed and upper bounded.*
- (b) *If \mathcal{E} is convex, then \mathcal{U} is convex.*
- (c) *\mathcal{E} and $\text{comp } \mathcal{E}$ are equivalent.*
- (d) *If \mathcal{E} is closed and has upper bounded broadcast rates, then \mathcal{E} and $\text{par } \mathcal{E}$ are equivalent.*
- (e) *If \mathcal{E} is convex, then $\mathcal{R}(\mathcal{E})$ is convex.*

- (f) If \mathcal{E} is convex, then $\text{comp } \mathcal{E}$ is convex.
- (g) If \mathcal{E} is closed, convex, and has upper bounded broadcast rates, then $\text{conv}(\text{ext } \mathcal{E} \cap \text{par } \mathcal{E})$ is equivalent to \mathcal{E} .

The proof of Proposition 2.11 is omitted since (a) and (b) follow directly from the structure of the linear transformation Λ_p and since the proofs of the remaining statements are similar to the proofs for the corresponding statements involving cut rate regions instead of penalized broadcast rate regions.

2.9 Cut Rate Sandwiched Multicast Source Rate Regions

The multicast capacity region \mathcal{C} of general discrete memoryless networks (Chap. 4) or Gaussian networks (Chap. 7) has not been characterized. The typical outer and inner bounds, i.e., the cut-set outer bound (Sect. 4.1.1) and the noisy network coding inner bound (Sect. 4.1.2), are rather similar in their structure. This has been exploited in [15] to establish a bound on the gap between the inner and outer bounds on the multicast capacity regions of Gaussian networks that is independent of channel gains and transmit powers.

In this section, we extend this type of analysis to inner and outer bounds to general inner and outer bounds $\mathcal{R}(\mathcal{U})$ and $\mathcal{R}(\hat{\mathcal{U}})$ on the multicast capacity \mathcal{C} region that are formulated using the (generalized) cut rate regions \mathcal{U} and $\hat{\mathcal{U}}$ such that $\mathcal{R}(\hat{\mathcal{U}}) \subset \mathcal{C} \subset \mathcal{R}(\mathcal{U})$. The gap between inner and outer bound on the multicast capacity region is analyzed based on gap between the cut rate regions \mathcal{U} and $\hat{\mathcal{U}}$. We focus on the case where both inner and outer bound are singleton cut rate regions, i.e., $\mathcal{U} = \{u\}$ and $\hat{\mathcal{U}} = \{\hat{u}\}$. This analysis relies again on submodularity, but not of the cut rate regions themselves but on the submodularity of a set function $\Delta \geq u - \hat{u}$ which bounds the gap between these cut rate regions. This kind of analysis proves particularly fruitful if the gap between the two cut rate regions does not depend on certain channel or network parameters. Such gap functions can be found for certain important types of networks including (noisy) linear finite field networks, erasure broadcast networks, and Gaussian networks.

For the analysis of the gap between two multicast rate regions, we consider $u, \hat{u} \in \mathcal{N}^N$ such that $u(A) \geq \hat{u}(A)$ for all $A \subset N$, which implies $\mathcal{R}(\{\hat{u}\}) \subset \mathcal{R}(\{u\})$, and any nonnegative gap function $\Delta \in \mathcal{S}_+^N$, not necessarily normalized, satisfying

$$\Delta(A) \geq u(A) - \hat{u}(A) \quad (2.49)$$

for all $A \subset N$. The goal of this analysis is to establish an inner bound on the inner bound multicast rate region $\mathcal{R}(\{\hat{u}\})$ based on the outer bound $\mathcal{R}(\{u\})$ and the gap function Δ . For submodular gap functions Δ , we have the following result, which is proved in Sect. 2.11.4:

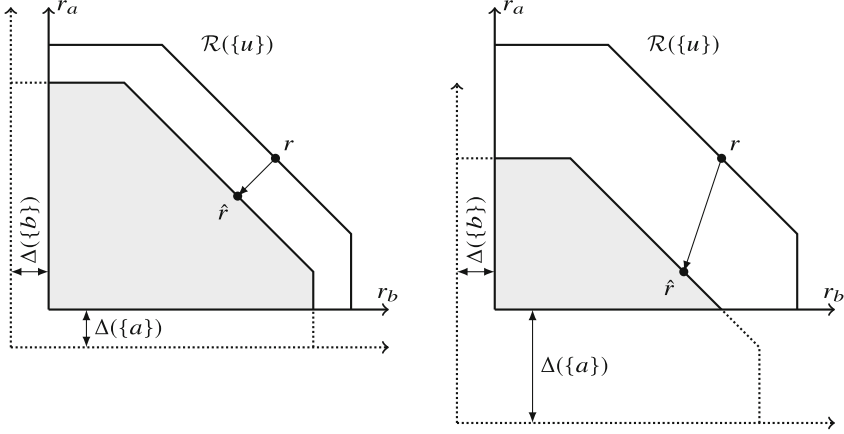


Fig. 2.11 Schematic two-dimensional representations of the inner bound on the multicast rate region $\mathcal{R}(\{\hat{u}\})$ based on two exemplary submodular gap functions: $\Delta(\{a\}) = \Delta(\{b\})$ (left) and $\Delta(\{a\}) \gg \Delta(\{b\})$ (right). The shaded areas represent the inner bounds obtained from $\mathcal{R}(\{u\})$ by defining \hat{r} as in Theorem 2.5 for all $r \in \mathcal{R}(\{u\})$, i.e., they are obtained by shifting $\mathcal{R}(\{u\})$ by $(-\Delta(\{a\}) : a \in N)$ and intersecting the result with \mathcal{R}_+^N . The thin one- and two-headed arrows indicate the shifts and their amounts in each dimension, respectively

Theorem 2.5 Let $u, \hat{u} \in \mathcal{N}^N$ and $\Delta \in \mathcal{S}^N$ such that $\Delta \geq u - \hat{u} \geq 0$, and Δ is submodular. For any $r \in \mathcal{R}(\{u\})$ such that $r_a \geq \Delta(\{a\})$ for all $a \in N$, $\hat{r} \in \mathcal{R}_+^N$ defined as $\hat{r}_a = r_a - \Delta(\{a\})$ for all $a \in N$ satisfies $\hat{r} \in \mathcal{R}(\{\hat{u}\})$.

Theorem 2.5 establishes that the set of rate vectors $\mathcal{R}(\{u\})$ shifted by the vector $(-\Delta(\{a\}) : a \in N)$ and intersected with \mathcal{R}_+^N is an inner bound on the rate region $\mathcal{R}(\{\hat{u}\})$. This shifted rate region is depicted in Fig. 2.11 for two exemplary gap functions. We remark that the shifted region depends only on values $\Delta(\{a\})$ for $a \in N$ but not on any value $\Delta(A)$ for $|A| \geq 2$. This is a consequence of the nonnegativity and submodularity of Δ , which implies the bound

$$u(A) - \hat{u}(A) \leq \Delta(A) \leq \sum_{a \in A} \Delta(\{a\}) \quad (2.50)$$

on the gap between the cut rate functions for all cuts $A \subset N$. Consequently, the gap between the two multicast rate regions $\mathcal{R}(\{u\})$ and $\mathcal{R}(\{\hat{u}\})$ can be characterized solely in terms of $\Delta(\{a\})$ for all $a \in N$. Additionally, note that although the submodularity of Δ plays a central role in Theorem 2.5, neither u nor \hat{u} need to be submodular themselves. Consequently, the result applies to any pair of cut rate functions whose difference can be bounded by a nonnegative submodular gap function.

Theorem 2.5 does not establish any bound with respect to the faces of $\mathcal{R}(\{u\})$ where some rates are zero, in particular, the axes where only one node acts as a source. To this end, we extend Theorem 2.5 to include all faces of $\mathcal{R}(\{u\})$ as follows:

Theorem 2.6 Let $u, \hat{u} \in \mathcal{N}^N$ and $\Delta \in \mathcal{S}^N$ such that $\Delta \geq u - \hat{u} \geq 0$, and Δ is submodular. For any $r \in \mathcal{R}(\{u\})$ for a multicast to $T \subset N$, define $\hat{r} \in \mathcal{R}_+^N$ as

$$\hat{r}_a = \left[r_a - \Delta(\{a\}) - \max_{B \subset S^c} \Delta(B) \right]_+ \quad \forall a \in N, \quad (2.51)$$

where $[\cdot]_+ = \max\{0, \cdot\}$ and $S \subset N$ is the set of supported sources defined as the unique largest set satisfying the implicit equation

$$S = \left\{ a \in N : r_a \geq \Delta(\{a\}) + \max_{B \subset S^c} \Delta(B) \right\}. \quad (2.52)$$

If $\hat{u}(A) \geq 0$ for all $A \subset N$ such that $A^c \cap T \neq \emptyset$, then $\hat{r} \in \mathcal{R}(\{\hat{u}\})$.

This theorem, which is proved in Sect. 2.11.4, establishes an inner bound on $\mathcal{R}(\{\hat{u}\})$ based on any outer bound rate vector $r \in \mathcal{R}(\{u\})$ which has at least one large enough entry, namely, rate vectors such that the supported source set S defined in (2.52) has at least one element. Therefore, Theorem 2.5 is an immediate corollary to Theorem 2.6 and applies to all rate vectors $r \in \mathcal{R}(\{u\})$ such that $S = N$. Theorem 2.6 is necessary if one asks for the inner bound generated by a single-source outer-bound rate vector r with $r_a = 0$ for all $a \neq s$ and some source node $s \in N$. For such rate vectors, the supported source set S is either $S = \emptyset$, which means that no inner bound is supported by r , or $S = \{s\}$, which means that the rate vector $\hat{r} \in \mathcal{R}_+^N$ with $\hat{r}_a = 0$ for all $a \neq s$ and \hat{r}_s given by (2.51) for $S = \{s\}$ is supported by r .

We remark that the nonnegativity and submodularity of Δ imply a loosened and simplified version of Theorem 2.6, where $r \in \mathcal{R}(\{u\})$ supports \hat{r} defined as

$$\hat{r}_a = \left[r_a - \Delta(\{a\}) - \sum_{b \in S^c} \Delta(\{b\}) \right]_+ \quad \forall a \in N, \quad (2.53)$$

where $S \subset N$ is the unique largest set satisfying

$$S = \left\{ a \in N : r_a \geq \Delta(\{a\}) + \sum_{b \in S^c} \Delta(\{b\}) \right\}. \quad (2.54)$$

This result can be obtained by observing that

$$\max_{B \subset S^c} \Delta(B) \leq \max_{B \subset S^c} \sum_{b \in B} \Delta(\{b\}) = \sum_{b \in S^c} \Delta(\{b\}) \quad (2.55)$$

since Δ is a normalized submodular set function. In particular, if the gap function is a modular function defined as $\Delta(A) = \sum_{a \in A} \Delta_a$ for some vector $(\Delta_a : a \in N) \in \mathcal{R}_+^N$, then this simplified version is equivalent to Theorem 2.6 since the inequality in (2.55) holds with equality.

2.10 Extension to Per-terminal Cut Models

This section briefly introduces a generalization of the cut model for multicast where an individual cut rate function is associated with each (potential) terminal node. Whereas the cut-set outer bound for discrete memoryless communication networks can be nicely modeled within the nonnegative cut model in Sect. 2.1, this extension turns out to be particularly useful for modeling the noisy network coding achievable multicast rate region, see Chap. 3. Its natural cut model representation uses mutual information based cut rate functions on collections of random variables that depend on both the cut A and the terminal t , instead of only the cut A as this is the case for the cut-set outer bound.

Definition 2.18 A vector of normalized set functions $u = (u^t : t \in N) \in \prod_{t \in N} \mathcal{N}^N$ is called a *per-terminal cut rate function*. A subset $\mathcal{U}^N \subset \prod_{t \in N} \mathcal{N}^N$ is referred to as *per-terminal cut rate region*. A nonnegative multicast rate vector $r \in \mathcal{R}_+^N$ is supported by a per-terminal cut rate function $u \in \prod_{t \in N} \mathcal{N}^N$ on a multicast network (N, T) if

$$\sum_{a \in A} r_a \leq u^t(A) \quad \forall t \in T, A \subset \{t\}^c. \quad (2.56)$$

The *multicast rate region* $\mathcal{R}(\mathcal{U}^N) \subset \mathcal{R}_+^N$ of the multicast network (N, T) with per-terminal cut rate region \mathcal{U}^N is defined as the set of nonnegative multicast rate vectors supported by \mathcal{U}^N , i.e.,

$$\mathcal{R}(\mathcal{U}^N) = \bigcup_{u \in \mathcal{U}^N} \left(\bigcap_{t \in T} \left\{ r \in \mathcal{R}_+^N : \sum_{a \in A} r_a \leq u^t(A) \quad \forall A \subset \{t\}^c \right\} \right). \quad (2.57)$$

We remark that the only difference between the definition of the multicast rate region based on the generalized cut model (2.33) and the per-terminal cut model (2.57) is that the former uses the same cut rate function for all terminals, whereas the latter uses an individual cut rate function for each terminal. Nevertheless, the cut type of the bounds on the multicast rates is the same for both regions. Consequently, if $u^t = u^s$ for all $t, s \in N$ and all $u \in \mathcal{U}^N$, then the per-terminal cut rate region \mathcal{U}^N can be represented by an ordinary generalized cut rate region.¹¹ In this case, the multicast rate regions defined by (2.33) and (2.57) match. Due to the similarities between the definitions of the multicast rate regions (2.33) and (2.57), there is a natural extension of cut rate region submodularity to per-terminal cut rate regions.

Definition 2.19 A per-terminal cut rate region $\mathcal{U}^N \subset \prod_{t \in N} \mathcal{N}^N$ is called *submodular* if all cut rate functions u^t for all $u \in \mathcal{U}^N$ are submodular set functions on $\{t\}^c$, i.e.,

¹¹For any multicast network (N, T) , both multicast rate region definitions are actually equivalent if $u^t = u^s$ for all $t, s \in T$ and all $u \in \mathcal{U}^N$ since $\mathcal{R}(\mathcal{U}^N)$ is independent of u^t for all $t \notin T$.

$$u^t(A) + u^t(B) \geq u^t(A \cup B) + u^t(A \cap B) \quad \forall A, B \subset \{t\}^c, t \in T. \quad (2.58)$$

The notable difference to submodular cut rate functions according to Definition 2.14 is that for each element u^t of $u \in \mathcal{U}^N$, submodularity is only required on the subsets of $\{t\}^c$. This is sufficient for all purposes in this work since the multicast rate region generated by a per-terminal cut rate region depends only on values $u^t(A)$ for $A \subset \{t\}^c$, see Definition 2.19.

The results in Sect. 2.8 can be easily transferred to the per-terminal cut model by naturally extending the underlying concepts. In particular, comprehensiveness and Pareto efficiency are defined by applying (2.44) and (2.45) in an elementwise manner to per-terminal cut rate vectors $u = (u^t : t \in N) \subset \prod_{t \in N} \mathcal{N}^N$. Furthermore, per-terminal versions of the graph, hypergraph, and (penalized) polymatroid broadcast models can be straight forwardly defined by applying the respective transformations to the cut model on a per-terminal basis. As a result, one also obtains per-terminal versions of the respective flow models and max-flow min-cut theorems. These extensions are not covered in this book.

2.11 Proofs

2.11.1 Polymatroid Max-Flow Min-Cut Theorem

Proof (Proof of Theorem 2.3) Let $\hat{\mathcal{R}}(\mathcal{F})$ and $\check{\mathcal{R}}(\mathcal{F})$ denote the min-cut formulation (2.2) and the max-flow formulation (2.19), respectively, of the multicast rate region generated by the broadcast rate region \mathcal{F} . We observe that $\hat{\mathcal{R}}(\mathcal{F})$ and $\check{\mathcal{R}}(\mathcal{F})$ can be equivalently written as

$$\hat{\mathcal{R}}(\mathcal{F}) = \bigcup_{f \in \mathcal{F}} \left(\bigcap_{t \in T} \hat{\mathcal{R}}^t(f) \right), \quad (2.59)$$

$$\check{\mathcal{R}}(\mathcal{F}) = \bigcup_{f \in \mathcal{F}} \left(\bigcap_{t \in T} \check{\mathcal{R}}^t(f) \right) \quad (2.60)$$

with appropriate multicast rate regions per terminal and broadcast rate vectors given by

$$\hat{\mathcal{R}}^t(f) = \left\{ r \in \mathcal{R}_+^N : \sum_{a \in A} r_a \leq \sum_{a \in A} f_a(A^c) \quad \forall A \subset \{t\}^c \right\}, \quad (2.61)$$

$$\check{\mathcal{R}}^t(f) = \left\{ r \in \mathcal{R}_+^N : x^t \in \mathcal{F}_+^N, \right.$$

$$\left. \begin{aligned} \sum_{b \in N} x_{ab}^t - \sum_{b \in N} x_{ba}^t &= r_a \quad \forall a \in \{t\}^c, \\ \sum_{b \in B} x_{ab}^t &\leq f_a(B) \quad \forall a \in N, B \subset N \end{aligned} \right\}. \quad (2.62)$$

Therefore, it suffices to show that $\hat{\mathcal{R}}^t(f) = \check{\mathcal{R}}^t(f)$ for arbitrary polymatroid broadcast rate vectors $f \in \mathcal{B}$. Note that r_t can be arbitrary nonnegative in both regions and that it does not influence the constraints on any other rates r_a . We restrict our analysis to rate vectors r with $r_t = 0$ without loss of generality.

First, assume $r \in \check{\mathcal{R}}^t(f)$ with $r_t = 0$ and x^t is a flow corresponding to the rate vector r in $\check{\mathcal{R}}^t(f)$. For any $A \subset \{t\}^c$, we have

$$\begin{aligned} \sum_{a \in A} r_a &= \sum_{a \in A} \sum_{b \in N} x_{ab}^t - \sum_{a \in A} \sum_{b \in N} x_{ba}^t \\ &= \sum_{a \in A} \sum_{b \in A^c} x_{ab}^t - \sum_{a \in A} \sum_{b \in A^c} x_{ba}^t. \end{aligned} \quad (2.63)$$

Nonnegativity of the flow x^t and the polymatroid broadcast constraint at each node $a \in N$ imply

$$\sum_{a \in A} r_a \leq \sum_{a \in A} \sum_{b \in A^c} x_{ab}^t \leq \sum_{a \in A} f_a(A^c). \quad (2.64)$$

Consequently, $r \in \hat{\mathcal{R}}^t(f)$, which proves $\check{\mathcal{R}}^t(f) \subset \hat{\mathcal{R}}^t(f)$.

To show the reverse inclusion, we consider $r \in \hat{\mathcal{R}}^t(f)$ with $r_t = 0$. In order to reduce the multi-source problem to a single-source setup, to which we can apply the polymatroid max-flow min-cut theorem due to [12–14], we extend the network by introducing a virtual source node s and define the extended node set $\bar{N} = N \cup \{s\}$. Furthermore, we define the extended broadcast rate vector \bar{f} on the extended node set \bar{N} such that

$$\bar{f}_a(B) = f_a(B \cap N) \quad \forall a \in N, B \subset \bar{N}, \quad (2.65)$$

i.e., no node can transmit information to the virtual source node s , and the broadcast function $\bar{f}_s: 2^{\bar{N}} \rightarrow \mathbb{R}$ at the virtual source node is given by

$$\bar{f}_s(B) = \sum_{b \in B} r_b, \quad (2.66)$$

which corresponds to independent arcs (s, a) with maximal data rates r_a for all $a \in N$.

Denote by R the value of the minimum cut separating s from t in the extended network. That is,

$$R \leq \sum_{a \in A \cup \{s\}} \bar{f}_a(\bar{N} - (A \cup \{s\})) = \sum_{a \in A} f_a(N - A) + \sum_{b \in N - A} r_b \quad (2.67)$$

for all $A \subset N - \{t\}$ with equality for some $A^* \subset N - \{t\}$ such that $A^* \cup \{s\}$ denotes the minimum cut. Since $r \in \check{\mathcal{R}}^t(f)$ and $r_t = 0$, we have $\sum_{a \in A^*} r_a \leq \sum_{a \in A^*} f_a(N - A)$, and therefore

$$R \geq \sum_{a \in A^*} r_a + \sum_{b \in N - A^*} r_b = \sum_{a \in N} r_a. \quad (2.68)$$

Combining (2.67) for $A = \emptyset$ with (2.68) yields $R = \sum_{a \in N} r_a$.

We apply the polymatroid max-flow min-cut theorem [12–14] to the extended network. It states that there exists a flow $\bar{x}^t \in \mathcal{F}_+^{\bar{N}}$ such that

$$R = \sum_{a \in N} \bar{x}_{sa}^t = \sum_{a \in \bar{N} - \{t\}} \bar{x}_{at}^t, \quad (2.69)$$

$$\sum_{b \in \bar{N}} \bar{x}_{ab}^t = \sum_{b \in \bar{N}} \bar{x}_{ba}^t \quad \forall a \in \bar{N} - \{s, t\}, \quad (2.70)$$

$$\sum_{b \in B} \bar{x}_{ab}^t \leq \bar{f}_a(B) \quad \forall a \in \bar{N}, B \subset \bar{N}. \quad (2.71)$$

Combining (2.69) with the min-cut result $R = \sum_{a \in N} r_a$ yields $\sum_{a \in N} \bar{x}_{sa}^t = \sum_{a \in N} r_a$. Applying the definition of \bar{f}_s (2.66) and (2.71) to this result gives $\bar{x}_{sa}^t = r_a$ for all nodes $a \in N$.

Furthermore, the definition of the extended broadcast functions \bar{f}_a at all other nodes $a \in N$ implies

$$\sum_{b \in B} \bar{x}_{ab}^t \leq f_a(B), \quad \forall a \in N, \forall B \subset N, \quad (2.72)$$

and $\bar{x}_{as}^t = 0$ for all $a \in N$. Using these results in (2.69) and (2.70) yields

$$r_a = \sum_{b \in N} \bar{x}_{ab}^t - \sum_{b \in N} \bar{x}_{ba}^t \quad (2.73)$$

for all nodes $a \in N - \{t\}$. Consequently, the flow x^t on the original nodes set N defined as the restriction of \bar{x}^t to the node set N , i.e., $x_{ab}^t = \bar{x}_{ab}^t$ for all $a, b \in N$, satisfies the constraints in $\check{\mathcal{R}}^t(f)$ for the rate vector r , and therefore $r \in \check{\mathcal{R}}^t(f)$, which proves $\hat{\mathcal{R}}^t(f) \subset \check{\mathcal{R}}^t(f)$.

2.11.2 Transformation of Models

Proof (Proof of Proposition 2.4) The linearity of the maps follows immediately from their definitions, which involve only sums of elements of the respective arguments. The consistency of the maps with each other, i.e., the commutativity of the left diagram in Fig. 2.9, follows directly from the definitions. In particular, plugging

(2.20) into (2.23) yields $\Xi = \Omega\Upsilon$, plugging (2.23) into (2.24) yields $\Omega = \Lambda\Gamma$, and combining both results yields $\Xi = \Lambda\Gamma\Upsilon$. It remains to show the properties of these maps.

(a) The map Υ is the natural embedding of vectors of functions on N into the space of vectors of set functions on N , and as such injective. Furthermore, since

$$\dim \mathcal{X}^N = |N|(2^{|N|} - 1) > |N|^2 = \dim \mathcal{F}^N \quad (2.74)$$

for all $|N| > 1$, Υ cannot be surjective for $|N| > 1$.

(b) Let $B \neq \emptyset$ and denote by $i: 2^N \rightarrow \{0, 1\}$ the nonempty set indicator function defined as $i(\emptyset) = 0$ and $i(A) = 1$ if $A \neq \emptyset$. To show the inversion formula, we plug (2.21) into (2.25), which yields

$$\begin{aligned} g_a(B) &= \sum_{A \subset B} (-1)^{|A|+1} \sum_{C \cap (A \cup B^c) \neq \emptyset} g_a(C) \\ &= \sum_{C \subset N} g_a(C) \sum_{A \subset B} (-1)^{|A|+1} i((A \cup B^c) \cap C) \\ &= \sum_{C \subset N} g_a(C) \kappa(B, C) \end{aligned} \quad (2.75)$$

with

$$\kappa(B, C) = \sum_{A \subset B} (-1)^{|A|+1} i((A \cap C) \cup (B^c \cap C)). \quad (2.76)$$

It remains to show that, for nonempty N , $\kappa(B, C) = 1$ if $B = C \neq \emptyset$, and $\kappa(B, C) = 0$ otherwise. Suppose $C \cap B^c \neq \emptyset$, which implies $C \neq B$. Then, $i((A \cap C) \cup (B^c \cap C)) = 1$ for all $A \subset B$ and, consequently,

$$\kappa(B, C) = \sum_{A \subset B} (-1)^{|A|+1} = \sum_{n=0}^{|B|} (-1)^{n+1} \binom{|B|}{n} = 0. \quad (2.77)$$

Suppose $C \cap B^c = \emptyset$, which implies $C \subset B$. Then, $i((A \cap C) \cup (B^c \cap C)) = i(A \cap C)$ for all $A \subset B$ and $i(A \cap C) = 0$ for all $A \subset B - C$. Therefore,

$$\kappa(B, C) = \sum_{A \subset B} (-1)^{|A|+1} - \sum_{A \subset B-C} (-1)^{|A|+1} = \begin{cases} 1 & \text{if } B = C \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \quad (2.78)$$

This proves that Γ^{-1} is well-defined and the inverse of Γ on \mathcal{X}^N . Therefore, Γ is bijective on \mathcal{X}^N .

(c) $\Lambda: \mathcal{X}^N \rightarrow \mathcal{C}^N$ is clearly not injective since

$$\dim \mathcal{X}^N = |N|(2^{|N|} - 1) > 2^{|N|} - 2 = \dim \mathcal{C}^N \quad (2.79)$$

for any nonempty node set N . To show that Λ is surjective, consider $v \in \mathcal{C}^N$ and define the set function vector f by

$$f_a(B) = \frac{v(B^c \cup \{a\})}{|B^c \cup \{a\}|}. \quad (2.80)$$

Since $f_a(\emptyset) = |N|^{-1}v(N) = 0$, we have $f \in \mathcal{Z}^N$. Furthermore,

$$(\Lambda f)(A) = \sum_{a \in A} f_a(A^c) = \sum_{a \in A} \frac{v(A \cup \{a\})}{|A \cup \{a\}|} = \sum_{a \in A} \frac{v(A)}{|A|} = v(A). \quad (2.81)$$

This shows that for any $v \in \mathcal{C}^N$, there exists some $f \in \mathcal{Z}^N$ such that $v = \Lambda f$. Consequently, Λ is surjective onto \mathcal{C}^N .

(d) The composition $\Omega = \Lambda \Gamma$ of a surjective but not injective linear map Λ with a bijective linear map Γ is itself surjective but not injective. This follows from the rank-nullity theorem (see for example [43]).

(e) Consider the transposed of Ξ with respect to the standard elementwise Euclidean inner product. $\Xi^T: \mathcal{C}^N \rightarrow \mathcal{F}^N$ with $k = \Xi^T v$ is given by

$$k_a(b) = \sum_{\substack{A \subset N: \\ a \in A, b \in A^c}} v(A). \quad (2.82)$$

Applying Ξ^T to the cut rate function

$$v(A) = \begin{cases} 1 & \text{if } A = \{a\}, a \in N, \\ -1 & \text{if } A = \{a\}^c, a \in N, \\ 0 & \text{otherwise,} \end{cases} \quad (2.83)$$

which is well-defined if $|N| \geq 3$, yields $k_a(b) = 0$ for all $a, b \in N$. This means that Ξ^T has a nontrivial null space and, therefore, Ξ cannot be surjective. Furthermore, any nonzero arc rate vector that satisfies $k_a(b) = 0$ for all $a, b \in N$ with $a \neq b$ is mapped onto the zero cut rate function θ defined by $\theta(A) = 0$ for all $A \subset N$. Consequently, Ξ is not injective.

Proof (Proof of Proposition 2.5)

(a) $\Upsilon(\mathcal{A}_+^N) \subset \mathcal{H}_+^N$ follows directly from the definition (2.20) since $k_a(b) \geq 0$ for all $a, b \in N$ implies $g_a(B) \geq 0$ for all $a \in N$ and $B \subset N$ and $k_a(a) = 0$ for all $a \in N$ implies $g_a(B) = 0$ for all $a \in N$ and $B \subset N$ such that $a \in B$. Therefore, the restriction $\Upsilon: \mathcal{A}_+^N \rightarrow \mathcal{H}_+^N$ is well-defined. Furthermore, if $|N| \geq 3$, the hyperarc rate vector $g \in \mathcal{H}_+^N$ defined by $g_a(B) = 1$ for all $a \in N$ and $B \subset \{a\}^c$ satisfies $g \notin \Upsilon(\mathcal{A}_+^N)$. Consequently, $\Upsilon: \mathcal{A}_+^N \rightarrow \mathcal{H}_+^N$ is not surjective for $|N| \geq 3$. Since the unrestricted map is injective, the restricted map is also injective.

(b) Showing $\Gamma(\mathcal{H}_+^N) \subset \mathcal{B}_+^N$ requires the verification that for all $a \in N$, the set function f_a generated from a hyperarc rate function g_a according to (2.21) satisfies

the conditions in Definition 2.9. Normalization, nonnegativity, and monotonicity of f_a follow directly from the transformation (2.21) and the normalization and nonnegativity of g_a . Submodularity follows from the nonnegativity of g_a and the following chain of (in)equalities¹²:

$$\begin{aligned}
 f_a(B) + f_a(C) &= \sum_{A \cap B \neq \emptyset} g_a(A) + \sum_{A \cap C \neq \emptyset} g_a(A) \\
 &= \sum_{A \cap (B \cap C) \neq \emptyset} g_a(A) + \sum_{A \cap (B \cup C) \neq \emptyset} g_a(A) + \sum_{\substack{A \cap B \neq \emptyset, A \cap C \neq \emptyset, \\ A \cap (B \cap C) = \emptyset}} g_a(A) \\
 &\geq \sum_{A \cap (B \cap C) \neq \emptyset} g_a(A) + \sum_{A \cap (B \cup C) \neq \emptyset} g_a(A) \\
 &= f_a(B \cap C) + f_a(B \cup C).
 \end{aligned} \tag{2.84}$$

for all $B, C \subset N$. Furthermore, the loop-free condition for f_a follows from $g_a(A) = 0$ for all $A \ni a$, since for any $B \ni a$ we have

$$f_a(B) - f_a(B - \{a\}) = \sum_{A \cap B \neq \emptyset} g_a(A) - \sum_{A \cap (B - \{a\}) \neq \emptyset} g_a(A) = \sum_{A \subset (B - \{a\})^c} g_a(A) = 0. \tag{2.85}$$

Therefore, $\Gamma(g) \in \mathcal{B}_+^N$ so that the restriction $\Gamma: \mathcal{H}_+^N \rightarrow \mathcal{B}_+^N$ is well-defined. Furthermore, $\Gamma: \mathcal{H}_+^N \rightarrow \mathcal{B}_+^N$ is injective since the unrestricted map $\Gamma: \mathcal{X}^N \rightarrow \mathcal{X}^N$ is injective.

For $|N| = 3$, i.e., $N = \{a, b, c\}$, $f = \Gamma g$ is given by

$$f_a(\emptyset) = 0, \tag{2.86}$$

$$f_a(\{b\}) = f_a(\{a, b\}) = g_a(\{b\}) + g_a(\{b, c\}), \tag{2.87}$$

$$f_a(\{c\}) = f_a(\{a, c\}) = g_a(\{c\}) + g_a(\{b, c\}), \tag{2.88}$$

$$f_a(\{b, c\}) = f_a(\{a, b, c\}) = g_a(\{b\}) + g_a(\{c\}) + g_a(\{b, c\}), \tag{2.89}$$

and the analogous equations with a exchanged with b and c . Note that the loop-free condition $g_a(B \cup \{a\}) = 0$ for all $B \subset N$ has been included in these equations. Combining these equations with the nonnegativity condition on $g_a(B)$ for all $B \subset N$ and eliminating g_a from the resulting system of (in)equalities yields precisely the condition for f_a being a broadcast rate function. The same results follows also for f_b and f_c . Consequently, $\Gamma: \mathcal{H}_+^N \rightarrow \mathcal{B}_+^N$ is surjective if $|N| = 3$, i.e., $\Gamma(\mathcal{H}_+^N) = \mathcal{B}_+^N$ if $|N| = 3$. Since $\Gamma: \mathcal{H}_+^N \rightarrow \mathcal{B}_+^N$ is also injective, it is bijective if $|N| = 3$.

Finally, to show that the restricted map $\Gamma: \mathcal{H}_+^N \rightarrow \mathcal{B}_+^N$ is not surjective for $|N| \geq 4$, we consider a node set N containing the nodes $\{a, b, c, d\}$. Consider a broadcast rate vector $f \in \mathcal{B}_+^N$ such that

¹²[38] used a similar chain of inequalities as part of the simplification of the hypergraph flow model for packet erasure networks—see Sect. 2.5 for a discussion of this simplification.

$$f_a(B) = \begin{cases} 0 & \text{if } |B \cap \{b, c, d\}| = 0, \\ 1 & \text{if } |B \cap \{b, c, d\}| = 1, \\ 2 & \text{if } |B \cap \{b, c, d\}| \geq 2 \end{cases} \quad (2.90)$$

for all $B \subset N$. Since the unrestricted map $\Gamma: \mathcal{Z}^N \rightarrow \mathcal{Z}^N$ is invertible, the unique $g \in \mathcal{Z}^N$ such that $f = \Gamma g$ is given by $g = \Gamma^{-1} f$ and satisfies

$$g_a(C) = \begin{cases} 1 & \text{if } C \subset \{b, c, d\}, |C| = 2, \\ -1 & \text{if } C \subset \{b, c, d\}, |C| = 3, \\ 0 & \text{otherwise} \end{cases} \quad (2.91)$$

for all $C \subset N$. The set function g_a clearly violates the nonnegativity condition for hyperarc rate functions in Definition 2.6, which implies $g = \Gamma^{-1} f \notin \mathcal{H}_+^N$. Therefore, the restricted map $\Gamma: \mathcal{H}_+^N \rightarrow \mathcal{B}_+^N$ is not surjective. This means that for any $|N| \geq 4$, there are broadcast rate vectors that cannot be represented by hyperarc rate vectors, i.e., $\Gamma(\mathcal{H}_+^N) \subsetneq \mathcal{B}_+^N$ for $|N| \geq 4$.

(c) First, we need to show $\Lambda(\mathcal{B}_+^N) \subset \mathcal{H}_+^N$ so that the restricted map $\Lambda: \mathcal{B}_+^N \rightarrow \mathcal{H}_+^N$ is well-defined.¹³ Consider $f \in \mathcal{B}_+^N$ and $v = \Lambda f$. Then, v is nonnegative since f is nonnegative. Furthermore, v is submodular since the following chain of equalities and inequalities holds:

$$\begin{aligned} v(A) + v(B) &= \sum_{a \in A} f_a(A^c) + \sum_{a \in B} f_a(B^c) \\ &= \sum_{a \in A \cap B} f_a(A^c) + f_a(B^c) + \sum_{a \in A - B} f_a(A^c) + \sum_{a \in B - A} f_a(B^c) \\ &\geq \sum_{a \in A \cap B} f_a(A^c \cup B^c) + f_a(A^c \cap B^c) \\ &\quad + \sum_{a \in A - B} f_a(A^c \cap B^c) + \sum_{a \in B - A} f_a(A^c \cap B^c) \\ &= \sum_{a \in A \cap B} f_a((A \cap B)^c) + f_a((A \cup B)^c) \\ &\quad + \sum_{a \in A - B} f_a((A \cup B)^c) + \sum_{a \in B - A} f_a((A \cup B)^c) \\ &= \sum_{a \in A \cap B} f_a((A \cap B)^c) + \sum_{a \in A \cup B} f_a((A \cup B)^c) \\ &= v(A \cap B) + v(A \cup B), \end{aligned} \quad (2.92)$$

where the inequality is due to submodularity and monotonicity of f_a for all $a \in N$. Consequently, $v = \Lambda f \in \mathcal{H}_+^N$ and the restricted map $\Lambda: \mathcal{B}_+^N \rightarrow \mathcal{H}_+^N$ is well-defined.

¹³Note that this part proves Proposition 2.3.

Next, we show that the restricted map Λ is not surjective for any $|N| \geq 3$. Let N such that $\{a, b, c\} \subset N$ and let $v \in \mathcal{X}_+^N$ be defined as

$$v(A) = \begin{cases} 2 & \text{if } |A \cap \{a, b, c\}| = 1, \\ 1 & \text{if } |A \cap \{a, b, c\}| = 2, \\ 0 & \text{otherwise.} \end{cases} \quad (2.93)$$

Submodularity of v can easily be verified. Any $f \in \mathcal{L}^N$ such that $v = \Lambda f$ must satisfy

$$f_a(\{a\}^c) = 2, \quad (2.94)$$

$$f_b(\{b\}^c) = 2, \quad (2.95)$$

$$f_c(\{c\}^c) = 2 \quad (2.96)$$

and

$$f_b(\{b, c\}^c) + f_c(\{b, c\}^c) = 1, \quad (2.97)$$

$$f_a(\{a, c\}^c) + f_c(\{a, c\}^c) = 1, \quad (2.98)$$

$$f_a(\{a, b\}^c) + f_b(\{a, b\}^c) = 1. \quad (2.99)$$

On the other hand, submodularity of f_a , f_b , and f_c requires

$$f_a(\{a, b\}^c) + f_a(\{a, c\}^c) \geq f_a(\{a\}^c) + f_a(\{a, b, c\}^c), \quad (2.100)$$

$$f_b(\{a, b\}^c) + f_b(\{b, c\}^c) \geq f_b(\{b\}^c) + f_b(\{a, b, c\}^c), \quad (2.101)$$

$$f_c(\{a, c\}^c) + f_c(\{b, c\}^c) \geq f_c(\{c\}^c) + f_c(\{a, b, c\}^c), \quad (2.102)$$

and nonnegativity of f_a , f_b , and f_c requires

$$f_a(\{a, b, c\}^c) \geq 0, \quad (2.103)$$

$$f_b(\{a, b, c\}^c) \geq 0, \quad (2.104)$$

$$f_c(\{a, b, c\}^c) \geq 0. \quad (2.105)$$

The two sets of conditions (2.94)–(2.99) and (2.100)–(2.105) are mutually exclusive, as the following argument shows: Consider the sum of the conditions (2.100)–(2.102), i.e.,

$$\begin{aligned} & f_a(\{a, b\}^c) + f_a(\{a, c\}^c) \\ & + f_b(\{a, b\}^c) + f_b(\{b, c\}^c) \\ & + f_c(\{a, c\}^c) + f_c(\{b, c\}^c) \\ & \geq f_a(\{a\}^c) + f_a(\{a, b, c\}^c) \\ & + f_b(\{b\}^c) + f_b(\{a, b, c\}^c) \\ & + f_c(\{c\}^c) + f_c(\{a, b, c\}^c). \end{aligned} \quad (2.106)$$

Plugging the conditions (2.94)–(2.99) into this inequality yields

$$f_a(\{a, b, c\}^c) + f_b(\{a, b, c\}^c) + f_c(\{a, b, c\}^c) \leq -3, \quad (2.107)$$

which clearly contradicts (2.103)–(2.105). Therefore, $f \notin \mathcal{B}_+^N$ for all f such that $v = \Lambda f$. This means that the restricted map $\Lambda: \mathcal{B}_+^N \rightarrow \mathcal{K}_+^N$ is not surjective, and consequently, $\Lambda(\mathcal{B}_+^N) \subsetneq \mathcal{K}_+^N$.

Finally, we show that the restricted map Λ is also not injective for any $|N| \geq 3$. Let $\{a, b, c\} \subset N$ and define the two broadcast rate vectors $f \in \mathcal{B}_+^N$ and $f' \in \mathcal{B}_+^N$ such that

$$f_a(\{b\}) = f_b(\{c\}) = f_c(\{a\}) = 1, \quad (2.108)$$

$$f_a(\{c\}) = f_b(\{a\}) = f_c(\{b\}) = 1, \quad (2.109)$$

$$f_a(\{b, c\}) = f_b(\{a, c\}) = f_c(\{a, b\}) = 2 \quad (2.110)$$

and

$$f'_a(\{b\}) = f'_b(\{c\}) = f'_c(\{a\}) = 2, \quad (2.111)$$

$$f'_a(\{c\}) = f'_b(\{a\}) = f'_c(\{b\}) = 0, \quad (2.112)$$

$$f'_a(\{b, c\}) = f'_b(\{a, c\}) = f'_c(\{a, b\}) = 2, \quad (2.113)$$

respectively, and

$$f_a(B \cap \{a, b, c\}) = f_a(B), \quad (2.114)$$

$$f_b(B \cap \{a, b, c\}) = f_b(B), \quad (2.115)$$

$$f_c(B \cap \{a, b, c\}) = f_c(B) \quad (2.116)$$

and

$$f'_a(B \cap \{a, b, c\}) = f'_a(B), \quad (2.117)$$

$$f'_b(B \cap \{a, b, c\}) = f'_b(B), \quad (2.118)$$

$$f'_c(B \cap \{a, b, c\}) = f'_c(B), \quad (2.119)$$

respectively, for all $B \subset N$. Furthermore, f and f' are chosen such that $f_d(B) = 0$ for all $d \in \{a, b, c\}^c$ and $B \subset N$. It is easily verified that $v = \Lambda f = \Lambda f'$ with $v \in \mathcal{K}_+^N$ given by

$$v(A) = \begin{cases} 2 & \text{if } 1 \leq |A \cap \{a, b, c\}| \leq 2, \\ 0 & \text{otherwise.} \end{cases} \quad (2.120)$$

Consequently, the restricted map Λ is not injective if $|N| \geq 3$.

(d) Since $\Omega = \Lambda\Gamma$ and the restrictions of Λ and Γ are well-defined and compatible, the restriction $\Omega: \mathcal{H}_+^N \rightarrow \mathcal{K}_+^N$ is well-defined, i.e., $\Omega(\mathcal{H}_+^N) \subset \mathcal{K}_+^N$,¹⁴ and given by the corresponding composition of the restricted maps, i.e., $\Omega = \Lambda\Gamma$. Furthermore, since the restricted map Λ is not surjective, the restricted map Ω is also not surjective.

To show that the restricted map $\Omega: \mathcal{H}_+^N \rightarrow \mathcal{K}_+^N$ is also not injective, consider $|N| \geq 3$ with $\{a, b, c\} \subset N$ and define the two hyperarc rate vectors $g \in \mathcal{H}_+^N$ and $g' \in \mathcal{H}_+^N$ such that

$$g_a(\{b\}) = g_b(\{c\}) = g_c(\{a\}) = 1, \quad (2.121)$$

$$g_a(\{c\}) = g_b(\{a\}) = g_c(\{b\}) = 1, \quad (2.122)$$

$$g_a(\{b, c\}) = g_b(\{a, c\}) = g_c(\{a, b\}) = 0 \quad (2.123)$$

and

$$g'_a(\{b\}) = g'_b(\{c\}) = g'_c(\{a\}) = 2, \quad (2.124)$$

$$g'_a(\{c\}) = g'_b(\{a\}) = g'_c(\{b\}) = 0, \quad (2.125)$$

$$g'_a(\{b, c\}) = g'_b(\{a, c\}) = g'_c(\{a, b\}) = 0, \quad (2.126)$$

respectively, and $g_a(C) = g'_a(C) = 0$, $g_b(C) = g'_b(C) = 0$, and $g_c(C) = g'_c(C) = 0$ for all $C \cap \{b, c\}^c \neq \emptyset$, $C \cap \{a, c\}^c \neq \emptyset$, and $C \cap \{a, b\}^c \neq \emptyset$, respectively. Furthermore, g and g' are chosen such that $g_d(C) = g'_d(C) = 0$ for all $d \in \{a, b, c\}^c$ and $C \subset N$. We observe that $f = \Gamma g$ and $f' = \Gamma g'$ for f and f' defined in (2.108)–(2.113). As a consequence of (c), $v = \Omega g = \Omega g'$ with $v \in \mathcal{K}_+^N$ given by (2.120). Therefore, the restricted map $\Omega: \mathcal{H}_+^N \rightarrow \mathcal{K}_+^N$ is not injective if $|N| \geq 3$.

(e) Since $\Xi = \Omega\Upsilon$ and the restrictions of Ω and Υ are well-defined and compatible, the restriction $\Xi: \mathcal{A}_+^N \rightarrow \mathcal{K}_+^N$ is well-defined, i.e., $\Xi(\mathcal{A}_+^N) \subset \mathcal{K}_+^N$,¹⁵ and given by the corresponding composition of the restricted maps, i.e., $\Xi = \Omega\Upsilon$, and additionally $\Xi = \Lambda\Gamma\Upsilon$ by (d). Furthermore, since the restricted map Ω is not surjective, the restricted map Ξ is also not surjective.

To show that the restricted map $\Xi: \mathcal{A}_+^N \rightarrow \mathcal{K}_+^N$ is also not injective, we consider again $|N| \geq 3$ with $\{a, b, c\} \subset N$ and define two arc rate vectors $k \in \mathcal{A}_+^N$ and $k' \in \mathcal{A}_+^N$ such that

$$k_a(b) = k_b(c) = k_c(a) = 1, \quad (2.127)$$

$$k_a(c) = k_b(a) = k_c(b) = 1, \quad (2.128)$$

$$k_a(a) = k_b(b) = k_c(c) = 0 \quad (2.129)$$

and

$$k'_a(b) = k'_b(c) = k'_c(a) = 2, \quad (2.130)$$

¹⁴Note that this part proves Proposition 2.2.

¹⁵Note that this part proves Proposition 2.1.

$$k'_a(c) = k'_b(a) = k'_c(b) = 0, \quad (2.131)$$

$$k'_a(a) = k'_b(b) = k'_c(c) = 0, \quad (2.132)$$

respectively. Furthermore, k and k' are chosen such that $k_d(e) = k_e(d) = 0$ and $k'_d(e) = k'_e(d) = 0$ for all $d \in \{a, b, c\}^c$ and $e \in N$. We observe that $g = \Upsilon k$ and $g' = \Upsilon g'$ for g and g' defined in (2.121)–(2.126). As a consequence of (d), $v = \Xi g = \Xi g'$ with $v \in \mathcal{K}_+^N$ given by (2.120). Therefore, the restricted map $\Xi: \mathcal{A}_+^N \rightarrow \mathcal{K}_+^N$ is not injective if $|N| \geq 3$.

Proof (Proof of Proposition 2.6) Since \mathcal{A}_+^N , \mathcal{H}_+^N , \mathcal{B}_+^N , and \mathcal{K}_+^N are closed polyhedral convex cones and the maps Υ , Γ , and Λ are linear with nonnegative coefficients, the sets $\Upsilon(\mathcal{A}_+^N)$, $\Gamma(\mathcal{H}_+^N)$, and $\Lambda(\mathcal{B}_+^N)$ are also closed polyhedral convex cones. Due to the strict inclusions, it suffices to show that for two closed convex cones \mathcal{C} , \mathcal{D} such that $\mathcal{C} \subsetneq \mathcal{D}$, it holds that

$$\sup_{y \in \mathcal{D}} \inf_{x \in \mathcal{C}} \|y - x\| = \infty. \quad (2.133)$$

To this end, let $\bar{y} \in \mathcal{D}$. Since \mathcal{C} is closed and convex, there exists a closest point $\bar{x} \in \mathcal{C}$ to \bar{y} such that $\|\bar{x} - \bar{y}\| = \inf_{x \in \mathcal{C}} \|x - \bar{y}\|$. Furthermore, since \mathcal{D} is a cone, $\alpha \bar{y} \in \mathcal{D}$ for all $\alpha > 0$. Since \mathcal{C} is also a cone, we have

$$\inf_{x \in \mathcal{C}} \|x - \alpha \bar{y}\| = \alpha \inf_{x \in \mathcal{C}} \|\frac{1}{\alpha} x - \bar{y}\| = \alpha \inf_{x \in \mathcal{C}} \|x - \bar{y}\|. \quad (2.134)$$

Letting α go to infinity shows (2.133), and therefore, (2.26)–(2.28).

2.11.3 Rate Region Properties and Equivalence

Proof (Proof of Proposition 2.7) Let $\mathcal{U}, \mathcal{U}' \subset \mathcal{N}^N$. Since $\mathcal{R}(\mathcal{U}) = \bigcup_{u \in \mathcal{U}} \mathcal{R}(\{u\})$, it follows that $\mathcal{R}(\mathcal{U}) \cup \mathcal{R}(\mathcal{U}') = \mathcal{R}(\mathcal{U} \cup \mathcal{U}')$. Therefore, if \mathcal{U} and \mathcal{U}' are equivalent, then $\mathcal{U} \cup \mathcal{U}'$ is also equivalent to them. Let \mathfrak{C} be the collection of all cut rate regions equivalent to \mathcal{U} and define $\mathcal{U}'' = \bigcup_{\mathcal{U}' \in \mathfrak{C}} \mathcal{U}'$. Then, $\mathcal{U}' \subset \mathcal{U}''$ for all $\mathcal{U}' \in \mathfrak{C}$ and $\mathcal{U}'' \in \mathfrak{C}$, i.e.,

$$\mathcal{R}(\mathcal{U}) = \bigcup_{\mathcal{U}' \in \mathfrak{C}} \mathcal{R}(\mathcal{U}') = \mathcal{R}\left(\bigcup_{\mathcal{U}' \in \mathfrak{C}} \mathcal{U}'\right) = \mathcal{R}(\mathcal{U}''). \quad (2.135)$$

Therefore, \mathcal{U}'' is the unique largest cut rate region equivalent to \mathcal{U} and contains all cut rate regions $\mathcal{U}' \in \mathfrak{C}$ that are equivalent to \mathcal{U} .

Proof (Proof of Proposition 2.8)

(a) Let $r^n \in \mathcal{R}(\mathcal{U})$ be a sequence of multicast rate vectors that converges to $r \in \mathcal{R}_+^N$. There exists a sequence of cut rate functions $u^n \in \mathcal{U}$ such that u^n supports r^n for all n . Since \mathcal{U} is compact, u^n converges—if necessary by passing to a subsequence—

to some $u \in \mathcal{U}$. By continuity with respect to u and r of the inequalities defining the multicast rate region (2.33), it follows that u supports r . Consequently, $r \in \mathcal{R}(\mathcal{U})$.

(b) Since (a) implies that $\mathcal{R}(\mathcal{U})$ is closed, we need to prove only that $\mathcal{R}(\mathcal{U})$ is appropriately bounded, i.e., bounded except for at most one dimension. Let $\bar{u} \in \mathcal{N}^N$ denote a set function such that $\bar{u} \leq u$ for all $u \in \mathcal{U}$. Clearly, $\mathcal{R}(\mathcal{U}) \subset \mathcal{R}(\{\bar{u}\})$. For any $t \in T$, we have $r_a \leq \bar{u}(\{a\})$ for all $a \in \{t\}^c$. Therefore, $\mathcal{R}(\mathcal{U}) \cap \{r \in \mathcal{R}_+^N : r_t = 0\}$ is bounded.

(c) This result follows directly from the proof of (b).

Proof (Proof of Proposition 2.9) (a) Let $r \in \mathcal{R}(\mathcal{U})$ and denote by $u \in \mathcal{U}$ a cut rate function supporting r . Then, for each $r' \in \mathcal{R}_+^N$ such that $r' \leq r$, we have

$$\sum_{a \in A} r'_a \leq \sum_{a \in A} r_a \leq u(A) \quad \forall t \in T, A \subset \{t\}^c, \quad (2.136)$$

which implies $r' \in \mathcal{R}(\mathcal{U})$.

(b) Let $T \subset N$ be an arbitrary terminal set. Since $\mathcal{U} \subset \text{comp } \mathcal{U}$ implies $\mathcal{R}(\mathcal{U}) \subset \mathcal{R}(\text{comp } \mathcal{U})$, it suffices to show that $\mathcal{R}(\text{comp } \mathcal{U}) \subset \mathcal{R}(\mathcal{U})$. Let $r \in \mathcal{R}(\text{comp } \mathcal{U})$ and denote by $u \in \text{comp } \mathcal{U}$ a cut rate function supporting r . Then, there exists a cut rate function $u' \in \mathcal{U}$ such that $u \leq u'$. Therefore,

$$\sum_{a \in A} r_a \leq u(A) \leq u'(A) \quad \forall t \in T, A \subset \{t\}^c, \quad (2.137)$$

which implies that r is also supported by $u' \in \mathcal{U}$ and, consequently, $r \in \mathcal{R}(\mathcal{U})$.

(c) It suffices to show that $\mathcal{U} \subset \text{comp } \text{par } \mathcal{U}$ for closed and upper bounded \mathcal{U} due to (b). Let $u' \in \mathcal{U}$. Since \mathcal{U} is closed and upper bounded, $\mathcal{U}' = \{u' \in \mathcal{U} : u' \geq u\} \subset \mathcal{U}$ is compact. Therefore, there exists $u'' \in \text{par } \mathcal{U}'$ (see for example [44]) and $u'' \geq u'$ by construction, i.e., $u' \in \text{comp}\{u''\}$. Suppose $u'' \notin \text{par } \mathcal{U}$. Then, there exists $u''' \in \text{par}\{u \in \mathcal{U} : u \geq u''\}$ such that $u''' \neq u''$. However, this contradicts $u'' \in \text{par } \mathcal{U}'$. Therefore, $u'' \in \text{par } \mathcal{U}$ and $u' \in \text{comp } \text{par } \mathcal{U}$, which proves $\mathcal{U} \subset \text{comp } \text{par } \mathcal{U}$.

Proof (Proof of Proposition 2.10) (a) Let $\hat{r}, \check{r} \in \mathcal{R}(\mathcal{U})$ be multicast rate vectors and $\hat{u}, \check{u} \in \mathcal{U}$ be cut rate functions supporting \hat{r} and \check{r} , respectively. For any $\alpha \in [0, 1]$,

$$\sum_{a \in A} \alpha \hat{r}_a + (1 - \alpha) \check{r}_a \leq \alpha \hat{u}(A) + (1 - \alpha) \check{u}(A) \quad \forall t \in T, A \subset \{t\}^c, \quad (2.138)$$

i.e., $r = \alpha \hat{r} + (1 - \alpha) \check{r}$ is supported by the cut rate region $u = \alpha \hat{u} + (1 - \alpha) \check{u}$. Since \mathcal{U} is convex, $u \in \mathcal{U}$. Therefore, $r \in \mathcal{R}(\mathcal{U})$.

(b) Let $\hat{u}, \check{u} \in \text{comp } \mathcal{U}$ be cut rate functions. There exist $\hat{u}', \check{u}' \in \mathcal{U}$ such that $\hat{u} \leq \hat{u}'$ and $\check{u} \leq \check{u}'$. For any $\alpha \in [0, 1]$, defining $u = \alpha \hat{u} + (1 - \alpha) \check{u}$ and $u' = \alpha \hat{u}' + (1 - \alpha) \check{u}'$ yields $u \leq u' \in \mathcal{U}$ since \mathcal{U} is convex. Consequently, $u \in \text{comp } \mathcal{U}$.

(c) Since \mathcal{U} is closed and upper bounded, \mathcal{U} is equivalent to $\text{par } \mathcal{U}$ due to Proposition 2.9. Furthermore, since $\text{conv}(\text{ext } \mathcal{U} \cap \text{par } \mathcal{U}) \subset \mathcal{U}$ it suffices to show $\text{par } \mathcal{U} \subset \text{conv}(\text{ext } \mathcal{U} \cap \text{par } \mathcal{U})$.

Since \mathcal{U} is closed, convex, and upper bounded it contains no lines. Therefore, any $u \in \mathcal{U}$ can be represented as convex combination of extreme points and conic combination of extreme directions [42, Theorem 18.5]. The upper bound on \mathcal{U} implies that all extreme directions of \mathcal{U} are nonpositive. Therefore, any Pareto efficient point in \mathcal{U} is solely a convex combination of extreme points of \mathcal{U} . That is, for any $u \in \text{par } \mathcal{U}$, there exist some positive integer L , extreme points $\{u_l : l = 1, \dots, L\} \subset \text{ext } \mathcal{U}$, and coefficients $\alpha_l > 0$, $l = 1, \dots, L$, such that $\sum_{l=1}^L \alpha_l = 1$ and $u = \sum_{l=1}^L \alpha_l u_l$.

Suppose $u \notin \text{conv}(\text{ext} \cap \text{par } \mathcal{U})$. This implies that there is at least one extreme point in $\{u_l : l = 1, \dots, L\}$ that is not Pareto efficient. Without loss of generality, let $u_1 \notin \text{par } \mathcal{U}$. Using the same argument as in the proof of Proposition 2.9(c), there exists $\hat{u}_1 \in \text{par } \mathcal{U}$ such that $\hat{u}_1 \geq u_1$ and $\hat{u}_1 \neq u_1$. Defining \hat{u} as

$$\hat{u} = \alpha_1 \hat{u}_1 + \sum_{l=2}^L \alpha_l u_l \quad (2.139)$$

leads to $u \leq \hat{u}$ and $u \neq \hat{u}$. This contradicts the Pareto efficiency of u . Therefore, $u \in \text{conv}(\text{ext} \cap \text{par } \mathcal{U})$, which proves $\text{par } \mathcal{U} \subset \text{conv}(\text{ext } \mathcal{U} \cap \text{par } \mathcal{U})$.

2.11.4 Cut Rate Sandwiched Multicast Source Rate Regions

Proof (Proof of Theorem 2.5) The following chain of inequalities holds for all $A \subset N$ such that $A^c \cap T \neq \emptyset$:

$$\sum_{a \in A} \hat{r}_a = \sum_{a \in A} (r_a - \Delta(\{a\})) \stackrel{\text{a}}{\leq} u(A) - \sum_{a \in A} \Delta(\{a\}) \stackrel{\text{b}}{\leq} u(A) - \Delta(A) \stackrel{\text{c}}{\leq} \hat{u}(A). \quad (2.140)$$

Inequality (a) is due to $r \in \mathcal{R}(\{u\})$ and inequality (b) due to the submodularity and nonnegativity of the gap function Δ . That is,

$$\Delta(B) + \Delta(C) \geq \Delta(\emptyset) + \Delta(B \cup C) \geq \Delta(B \cup C) \quad \forall B, C \subset N : B \cap C \neq \emptyset \quad (2.141)$$

is recursively applied to $\sum_{a \in A} \Delta(\{a\})$. Inequality (c) holds as a consequence of the definition of the gap function Δ . Note that (2.140) directly implies $\hat{u}(A) \geq 0$ for all $A \subset N$ with $A^c \cap T \neq \emptyset$ and, therefore, also $\mathcal{R}(\{\hat{u}\}) \neq \emptyset$.

Proof (Proof of Theorem 2.6) The first step is to prove that the set of sources S is well-defined and unique. Let $S', S'' \subset N$ be sets satisfying (2.52), and suppose without loss of generality that

$$\max_{B \subset S'^c} \Delta(B) \leq \max_{B \subset S''^c} \Delta(B). \quad (2.142)$$

Together with (2.52), this implies

$$0 \leq r_a - \Delta(\{a\}) - \max_{B \subset S''^c} \Delta(B) \leq r_a - \Delta(\{a\}) - \max_{B \subset S'^c} \Delta(B) \quad (2.143)$$

for all $a \in S''$. Consequently, we have $S'' \subset S'$. Suppose $S \subset N$ is the set that minimizes $\max_{B \subset S^c} \Delta(B)$ among all sets satisfying (2.52). Therefore, S contains all sets satisfying (2.52), which establishes that S exists and is the unique largest set satisfying (2.52).

The following chain of inequalities for all $A \subset S$ such that $A \neq \emptyset$ and $A^c \cap T \neq \emptyset$ shows that $\hat{r} \in \mathcal{R}(\{\hat{u}\})$ if $S \neq \emptyset$:

$$\begin{aligned} \sum_{a \in A} \hat{r}_a &= \sum_{a \in A} \left(r_a - \Delta(\{a\}) - \max_{B \subset S^c} \Delta(B) \right) \\ &\stackrel{\text{a}}{\leq} \min_{\substack{B \subset S^c: \\ (A \cup B)^c \cap T \neq \emptyset}} u(A \cup B) - \sum_{a \in A} \Delta(\{a\}) - |A| \max_{B \subset S^c} \Delta(B) \\ &\stackrel{\text{b}}{\leq} \min_{\substack{B \subset S^c: \\ (A \cup B)^c \cap T \neq \emptyset}} \left\{ \hat{u}(A \cup B) + \Delta(A \cup B) \right\} - \sum_{a \in A} \Delta(\{a\}) - |A| \max_{B \subset S^c} \Delta(B) \\ &\stackrel{\text{c}}{\leq} \min_{\substack{B \subset S^c: \\ (A \cup B)^c \cap T \neq \emptyset}} \hat{u}(A \cup B) + \max_{B \subset S^c} \Delta(A \cup B) - \sum_{a \in A} \Delta(\{a\}) - |A| \max_{B \subset S^c} \Delta(B) \\ &\stackrel{\text{d}}{\leq} \min_{\substack{B \subset S^c: \\ (A \cup B)^c \cap T \neq \emptyset}} \hat{u}(A \cup B) + \sum_{a \in A} \Delta(\{a\}) + \max_{B \subset S^c} \Delta(B) - \sum_{a \in A} \Delta(\{a\}) - |A| \max_{B \subset S^c} \Delta(B) \\ &\stackrel{\text{e}}{\leq} \min_{\substack{B \subset S^c: \\ (A \cup B)^c \cap T \neq \emptyset}} \hat{u}(A \cup B) \end{aligned} \quad (2.144)$$

Inequality (a) follows from $r \in \mathcal{R}(\{u\})$ since

$$\sum_{a \in \bar{A}} r_a \leq u(\bar{A}) \quad \forall \bar{A} \subset N : \bar{A}^c \cap T \neq \emptyset \quad (2.145)$$

combined with $r \geq 0$ implies

$$\sum_{a \in A} r_a \leq u(A \cup B) \quad \forall A \subset S, B \subset S^c : A^c \cap T \neq \emptyset, (A \cup B)^c \cap T \neq \emptyset. \quad (2.146)$$

Inequality (b) is due to the definition of the gap function Δ , whereas inequality (c) follows from upper bounding the minimum of the sum of two function with respect to B by summing the minimum of the first function and the maximum of the second function. Then, the submodularity and the nonnegativity of Δ are exploited in inequality (d) by repeatedly applying (2.141), which yields

$$\Delta(A \cup B) \leq \sum_{a \in A} \Delta(\{a\}) + \Delta(B) \quad \forall A \subset S, B \subset S^c. \quad (2.147)$$

Finally, inequality (e) follows from the assumption $A \neq \emptyset$. Applying the same argument as for inequality (a) to \hat{r} and the right hand side of inequality (e) in combination with the requirement $\hat{u}(A) \geq 0$ for all $A \subset N : A^c \cap T \neq \emptyset$, which is necessary since $A = \emptyset$ is excluded in (2.144), yields $\hat{r} \in \mathcal{R}(\{\hat{u}\})$.

References

1. El Gamal A (1981) On Information flow in relay networks. In: IEEE national telecommunications conference, vol 2, New Orleans, LA, USA, pp D4.1.1–D4.1.4
2. El Gamal A, Kim YH (2011) Network information theory. Cambridge University Press, Cambridge
3. Cover TM, Thomas JA (2006) Elements of information theory, 2nd edn. Wiley, New York
4. Ford LR Jr, Fulkerson DR (1962) Flows in networks. Princeton University Press, Princeton
5. Elias P, Feinstein A, Shannon C (1956) A note on the maximum flow through a network. IRE Trans Inf Theory 2(4):117–119
6. Parvaresh F, Etkin R (2014) Efficient capacity computation and power optimization for relay networks. IEEE Trans Inf Theory 60(3):1782–1792
7. Ahlswede R, Cai N, Li SYR, Yeung RW (2000) Network information flow. IEEE Trans Inf Theory 46(4):1204–1216
8. Wu Y, Chou PA, Zhang Q, Jain K, Zhu W, Kung SY (2005b) Network planning in wireless Ad Hoc networks: a cross-layer approach. IEEE J Sel Areas Commun 23(1):136–150
9. Lun D, Ratnakar N, Médard M, Kötter R, Karger D, Ho T, Ahmed E, Zhao F (2006) Minimum-cost multicast over coded packet networks. IEEE Trans Inf Theory 52(6):2608–2623
10. Lun D, Médard M, Kötter R, Effros M (2008) On coding for reliable communication over packet networks. Phys Commun 1(1):3–20
11. Riemensberger M, Utschick W (2014) A polymatroid flow model for network coded multicast in wireless networks. IEEE Trans Inf Theory 60(1):443–460
12. Hassin R (1978) On Network Flows. Phd thesis, Yale University
13. Hassin R (1982) Minimum cost flow with set-constraints. Networks 12(1):1–21
14. Lawler EL, Martel CU (1982) Computing maximal "polymatroidal" network flows. Math Oper Res 7(3):334–347
15. Lim S, Kim YH, El Gamal A, Chung SY (2011) Noisy network coding. IEEE Trans Inf Theory 57(5):3132–3152
16. Riemensberger M, Gerdes L, Utschick W (2014) Submodular structure and optimal quantization in gaussian multiple access relay networks. In: IEEE workshop on signal processing advances in wireless communications (SPAWC), pp 319–323
17. Ratnakar N, Kramer G (2006) The multicast capacity of deterministic relay networks with no interference. IEEE Trans Inf Theory 52(6):2425–2432
18. Avestimehr A, Diggavi S, Tse D (2011) Wireless network information flow: a deterministic approach. IEEE Trans Inf Theory 57(4):1872–1905
19. Grötschel M, Lovász L, Schrijver A (1981) The ellipsoid method and its consequences in combinatorial optimization. Combinatorica 1(2):169–197
20. McCormick ST (2006) Handbook on discrete optimization, Chap Submodular Function Minimization, Elsevier, pp 321–391
21. Iwata S (2008) Submodular function minimization. Math Program 112:45–64
22. Papadimitriou C, Yannakakis M (1991) Optimization, approximation, and complexity classes. J Comput Syst Sci 43(3):425–440
23. Nemhauser G, Wolsey L, Fisher M (1978) An Analysis of Approximations for Maximizing Submodular Set Functions—I. Mathematical Programming 14(1):265–294

24. Edmonds J, Giles R (1977) A min-max relation for submodular functions on graphs. In: Hammer P, Johnson E, Korte B, Nemhauser G (eds) *Studies in integer programming*, annals of discrete mathematics, vol 1, Elsevier, pp 185–204
25. Zimmermann U (1982) Minimization on submodular flows. *Discrete Appl Math* 4(4):303–323
26. Fujishige S (2005) *Submodular functions and optimization*, annals of discrete mathematics, vol 58, 2nd edn. Elsevier
27. Edmonds J (1970) Submodular functions, matroids, and certain polyhedra. In: Guy R, Hanani H, Sauer N, Schönheim J (eds) *Combinatorial structures and their applications*. Gordon and Breach, New York, pp 69–87
28. Shapley L (1971) Cores of convex games. *Int J Game Theory* 1(1):11–26
29. Lovász L (1983) Submodular functions and convexity. In: Korte B, Grötschel M, Bachem A (eds) *Mathematical programming the state of the art*. Springer, Heidelberg, pp 235–257
30. Lin X, Shroff N, Srikant R (2006) A tutorial on cross-layer optimization in wireless networks. *IEEE J Sel Areas Commun* 24(8):1452–1463
31. Georgiadis L, Neely MJ, Tassiulas L (2006) Resource allocation and cross-layer control in wireless networks. *Found Trends® Netw* 1(1):1–144
32. Yuan J, Li Z, Yu W, Li B (2006) A cross-layer optimization framework for multihop multicast in wireless mesh networks. *IEEE J Sel Areas Commun* 24(11):2092–2103
33. Song L, Yeung RW, Cai N (2006) A separation theorem for single-source network coding. *IEEE Trans Inf Theory* 52(5):1861–1871
34. Kötter R, Médard M (2003) An algebraic approach to network coding. *IEEE/ACM Trans Netw* 11(5):782–795
35. Ho T, Viswanathan H (2009) Dynamic algorithms for multicast with intra-session network coding. *IEEE Trans Inf Theory* 55(2):797–815
36. Riemensberger M, Utschick W (2015) On carrier sense multiple access in coded wireless packet networks. In: *International ITG conference on systems, communications and coding (SCC)*
37. Ho T, Lun D (2008) *Network coding: an introduction*. Cambridge University Press, Cambridge
38. Traskov D, Heindlmaier M, Médard M, Kötter R (2012) Scheduling for network-coded multicast. *IEEE/ACM Trans Netw* 20(5):1479–1488
39. Murota K (2003) *Discrete convex analysis: monographs on discrete mathematics and applications*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, SIAM
40. Sagduyu YE, Ephremides A (2007) On joint MAC and network coding in wireless Ad Hoc networks. *IEEE Trans Inf Theory* 53(10):3697–3713
41. Wu Y, Chiang M, Kung SY (2006) Distributed utility maximization for network coding based multicasting: a critical cut approach. In: *International symposium on modeling and optimization in mobile, Ad Hoc, and wireless networks (WiOpt)*, pp 1–6
42. Rockafellar R (1970) *Convex analysis*. Princeton University Press, Princeton
43. Roman S (2008) *Advanced linear algebra, graduate texts in mathematics*, vol 135, 3rd edn. Springer, New York
44. Hartley R (1978) On cone-efficiency, cone-convexity and cone-compactness. *SIAM J Appl Math* 34(2):211–222

Submodular Rate Region Models for Multicast
Communication in Wireless Networks

Riemensberger, M.

2018, XXIII, 281 p. 52 illus., Hardcover

ISBN: 978-3-319-65231-3