Abstract

In the minimum Multicut problem, the input is an edge-weighted supply graph $G = (V, E)$ and a demand graph $H = (V, F)$. Either $G$ and $H$ are directed (Dir-MulC) or both are undirected (Undir-MulC). The goal is to remove a minimum weight set of supply edges $E' \subseteq E$ such that in $G - E'$ there is no path from $s$ to $t$ for any demand edge $(s, t) \in F$. Undir-MulC admits $O(\log k)$-approximation where $k$ is the number of edges in $H$ while the best known approximation for Dir-MulC is $\min\{k, O(|V|^{11/23})\}$. These approximations are obtained by proving corresponding results on the multicommodity flow-cut gap. In this paper we consider the role that the structure of the demand graph plays in determining the approximability of Multicut. We obtain several new positive and negative results.

In undirected graphs our main result is a 2-approximation in $n^{O(t)}$ time when the demand graph excludes an induced matching of size $t$. This gives a constant factor approximation for a specific demand graph that motivated this work, and is based on a reduction to uniform metric labeling and not via the flow-cut gap.

In contrast to the positive result for undirected graphs, we prove that in directed graphs such approximation algorithms can not exist. We prove that, assuming the Unique Games Conjecture (UGC), that for a large class of fixed demand graphs Dir-MulC cannot be approximated to a factor better than the worst-case flow-cut gap. As a consequence we prove that for any fixed $k$, assuming UGC, Dir-MulC with $k$ demand pairs is hard to approximate to within a factor better than $k$. On the positive side, we obtain a $k$ approximation when the demand graph excludes certain graphs as an induced subgraph. This generalizes the known 2 approximation for directed Multiway Cut to a larger class of demand graphs.

1 Introduction

The minimum Multicut problem is a generalization of the classical s-t cut problem to multiple pairs. The input to the Multicut problem is an edge-weighted graph $G = (V, E)$ and $k$ source-sink pairs $(s_1, t_1), (s_2, t_2), \ldots, (s_k, t_k)$. The goal is to find a minimum weight subset of edges $E' \subseteq E$ such that all the given pairs are disconnected in $G - E'$; that is, for $1 \leq i \leq k$, there is no path from $s_i$ to $t_i$ in $G - E'$. In this paper we consider an equivalent formulation that exposes, more directly, the structure that the source-sink pairs may have.

The input now consists of an edge-weighted supply graph $G = (V, E)$ and a demand graph $H = (V, F)$. The goal is to find a minimum weight set of edges $E' \subseteq E$ such that for each edge $f = (s, t) \in F$, there is no path from $s$ to $t$ in $G - E'$. In other words the source-sink pairs are encoded in the form of the demand graph $H$. Either both $G$ and $H$ are directed in which case we refer to the problem as Dir-MulC (directed Multicut) or both are undirected in which case we refer to the problem as Undir-MulC (undirected Multicut).

Multicut in both directed and undirected graphs has been extensively studied. It is a natural cut problem, has several applications, and important connections to several other well-known problems such as sparsest cut and multicommodity flows. Undir-MulC and Dir-MulC are NP-Hard even in very restricted settings. For instance Undir-MulC is NP-Hard even when $H$ has 3 edges and it generalizes Vertex Cover even when $G$ is a tree. Dir-MulC is NP-Hard and APX-Hard even in the special case when $H$ is a cycle of length 2 which is better understood as removing a minimum weight set of edges to disconnect $s$ from $t$ and $t$ from $s$ in a directed graph. Consequently there has been substantial effort towards developing approximation algorithm for these problems as well as understanding special cases. We briefly summarize some of the known results. We use $k$ to denote the number of edges in the demand graph $H$. For Undir-MulC there is an $O(\log k)$-approximation [13] which improves to an $O(r)$-approximation if the supply graph $G$ excludes $K_r$. 

\*

Department of Computer Science, University of Illinois, Urbana, IL 61801. \{chekuri, vmadan2\}@illinois.edu. Work on this paper partly supported by NSF grant CCF-1319376.
Our preceding discussion has focused on the approximability of MultiCut when $H$ is restricted, with some improved results when $G$ is arbitrary. Here we are interested in the setting where $H$ is restricted and $G$ is arbitrary. Before we describe a concrete application that motivated us, we mention the well-known Multiway Cut problem in undirected graphs (Undir-Multiway-Cut) and directed graphs (Dir-Multiway-Cut). In Multiway Cut $H$ is the complete (directed) graph on a set of $k$ terminals. This problem has been extensively studied over the years. Undir-Multiway-Cut admits a 1.29 approximation [21] and Dir-Multiway-Cut admits a 2-approximation [19, 7]. In a recent work [5], motivated by connections to the problem of understanding the information capacity of networks with delay constraints, the following special case of MultiCut was considered. The authors of [5] referred to it as the Triangle-Cast problem since they were unaware of that the same demand graph had been studied previously in the fixed parameter tractability (FPT) literature as the Skew-MultiCut problem. For the sake of continuity we use the Skew-MultiCut nomenclature.

The demand graph: Our preceding discussion has focused on the approximability of MultiCut when $H$ is arbitrary, with some improved results when $G$ is restricted. Here we are interested in the set-

**Figure 2:** Demand graph for Skew-MultiCut

Two results. First, we obtain a 2-approximation for Skew-MultiCut and even Skew-MultiCut is hard to approximate to within an $\Omega(2^k)$ factor for some fixed $\delta > 0$. The best known approximation is $\min\{k, O(n^{11/23})\}$; here $n = |V|$. Note that a $k$-approximation is trivial.

We note that all the preceding positive results for Multicut are based on bounding the integrality gap of a natural LP relaxation shown in Figure 1. This is the standard cut formulation with a variable $x_e$ for each edge $e$ and an exponential set of constraints; the constraints specify that for each demand edge $st \in F$ the length of any path $p$ between $s$ and $t$ is at least 1. One can solve this LP in polynomial-time using the Ellipsoid method. One can also write a compact polynomial-sized formulation using distance variables and it is shown in the same figure. The dual is a maximum multicommodity flow LP. We henceforth refer to the integrality gap of this LP as the flow-cut gap and the LP as Distance-LP. Most multicut approximation algorithms are based on bounding the flow-cut gap.

**Question 1.** What is the approximability of Undir-Skew-MultiCut and Dir-Skew-MultiCut? Is the flow-cut gap $O(1)$ for Undir-Skew-MultiCut and even Dir-Skew-MultiCut?

Answering the preceding questions has not been easy. In fact we do not yet know whether the flow-cut gap is $O(1)$ for Undir-Skew-MultiCut. One can show APX-hardness and a constant factor lower bound on the flow-cut gap for Undir-Skew-MultiCut (and hence also for Dir-Skew-MultiCut) via a reduction from Undir-Multiway-Cut problem. In this paper we provide two results. First, we obtain a 2-approximation

$$\min \sum_{e \in E} w_e x_e$$

$$\sum_{e \in p} x_e \geq 1 \quad p \in P_{st}, st \in F$$

$$x_e \geq 0 \quad e \in E$$

Under the Unique Game Conjecture (UGC) it is known as a minor [11, 12, 14] (in particular this yields a constant factor approximation in planar graphs). In terms of inapproximability, Undir-MulC is at least as hard as Vertex Cover even in trees and hence APX-Hard. Under the Unique Game Conjecture (UGC) it is known to be super-constant hard [4]. Dir-MulC is known to be a harder problem. Assuming $NP \neq ZPP$ it is hard to approximate Dir-MulC to within a factor of $\Omega(2^{\log^{1/7} n})$ [9]; evidence is also presented in [9] that it could be hard to approximate to within an $\Omega(n^\delta)$ factor for some fixed $\delta > 0$. The best known approximation is $\min\{k, O(n^{11/23})\}$; here $n = |V|$. Note that a $k$-approximation is trivial.

**Figure 1:** Distance-LP for MultiCut written using exponential number of constraints (top) and in a compact fashion using additional variables (bottom)
for \textsc{Undir-Skew-MultiCut}; this result is not established via the \textsc{Distance-LP} relaxation and requires a new LP relaxation. Second, we show that under UGC, for any fixed constant \( k \), the hardness of approximation for \textsc{Dir-Skew-MultiCut} co-incides with its worst-case flow-cut gap. At this moment we only know that the flow-cut gap is \( O(\log k) \) and at least some fixed constant \( c > 1 \). We mention that \textsc{Dir-Skew-MultiCut} is approximation equivalent to another problem called \textsc{Lin-Cut} \cite{gh10}; here the demand graph \( H \) consists of \( k \) terminals \( s_1, \ldots, s_k \) and there is a directed edge \( (s_i, s_j) \) for all \( i < j \). As we mentioned, \textsc{Dir-Skew-MultiCut} has been considered earlier in \cite{gsv14} where it was shown to be in FPT parameterized by the solution size and is then used to prove that the well-known Directed Feedback Vertex set problem is in FPT. The demand pattern for \textsc{Skew-MultiCut} also arises in parameterized complexity of disjoint paths problem \cite{y98}.

Our results for \textsc{Skew-MultiCut} are special cases of more general results that examine the role that the demand graph \( H \) plays in the approximability of \textsc{MultiCut}. What structural aspects of \( H \) allow for better bounds than the worst-case results? For instance, can the constant factor approximation algorithms for \textsc{Multiway Cut} be understood in a more general setting? Some previous work has also examined the role that demand graph plays in \textsc{MultiCut}. Two examples are the original paper of Garg, Vazirani and Yannakakis \cite{gvy95} who showed that one can obtain an \( O(\log h) \)-approximation for \textsc{Undir-MulC} where \( h \) is the vertex cover size of the demand graph. This was generalized by Steurer and Vishnoi \cite{sv10} who showed that \( h \) can be chosen to be \( \min_{S,T} \max_{\emptyset \neq S \subseteq V} |S \cap T| \) where \( S \) is a vertex cover in \( H \) and \( T \) is an independent set in \( H \). Note that both these results are based on the flow-cut gap and yield only an \( O(\log k) \) upper bound for \textsc{Undir-Skew-MultiCut}.

We now describe our results for both \textsc{Undir-MulC} and \textsc{Dir-MulC} which yield, as corollaries, the results that we already mentioned and several others.

### 1.1 Overview of Results

We first discuss our result for \textsc{Undir-MulC}. We obtain a 2-approximation for a class of demand graphs. This class is inspired by the observation that the \textsc{Skew-MultiCut} demand graph does not contain a matching with two edges as an induced subgraph. More generally, a graph is said to be \( tK_2 \)-free for an integer \( t > 1 \) if it does not contain a matching of size \( t \) as an induced subgraph.

**Theorem 1.1.** There is a 2-approximation algorithm with running time \( \text{poly}(n, k^{O(1)}) \) on instances of \textsc{Undir-MulC} with supply graph \( G \) and \( tK_2 \)-free demand graph \( H \). Here \( n = |V(G)| \) and \( k = |V(H)| \).

Since demand graph in \textsc{Undir-Skew-MultiCut} instances are \( 2K_2 \)-free we obtain the following corollary.

**Corollary 1.1.** \textsc{Undir-Skew-MultiCut} admits a polynomial-time 2-approximation.

Note that a graph containing \( t - 1 \) parallel edges is \( tK_2 \)-free. Via known lower bounds \cite{gvy95}, \textsc{Distance-LP} has integrality gap \( \Omega(\log(t-1)) \) on \( tK_2 \)-free demand graphs. Thus, one cannot obtain a constant approximation for all fixed \( t \) using \textsc{Distance-LP}. Our algorithm relies on a different relaxation and a reduction to uniform metric labeling \cite{l10}. Also, as we noted earlier, assuming UGC, \textsc{Undir-MulC} is NP-hard to approximate within a constant factor \cite{all06}. Hence, assuming UGC, it is unlikely that one can obtain a fixed constant factor approximation for \textsc{Undir-MulC} with \( tK_2 \)-free demand graphs in \( \text{poly}(n, k, t) \) time.

We now turn our attention to \textsc{Dir-MulC}. As we mentioned earlier, \textsc{Dir-MulC} admits \( \tilde{O}(n^{11/23}) \) approximation \cite{all06} and is hard to approximate within a factor of \( \Omega(2^{\log^{-\epsilon} n}) \) assuming \( NP \neq ZPP \) \cite{all06}. A \( k \) approximation is trivial. It is natural to ask whether we can improve the bound of \( k \) when \( k \) is a fixed constant; however, it is known that the worst-case flow-cut gap in this case is \( k \) \cite{gsv14}, and hence one needs a different relaxation. What about the case of a specific demand graph \( H \) such as the one for \textsc{Dir-Skew-MultiCut}? To formalize this, for a fixed demand graph \( H \), we define the problem \textsc{Dir-MulC-H} as the special case of \textsc{Dir-MulC} where \( G \) is arbitrary but the demand graph is constrained to be \( H \). To be formal we need to define \( H \) as a “pattern” since we need to specify the nodes of \( G \) to which the nodes of \( H \) are mapped. However we avoid further notation and assume that \( V(H) \subseteq V(G) \).

We define \( \alpha_H \) to be the worst-case flow-cut gap over all instances with demand graph \( H \). We conjecture the following general hardness of approximation result.

**Conjecture 1.** For any fixed demand graph \( H \) and any fixed \( \epsilon > 0 \), unless \( P = NP \), there is no polynomial-time \((\alpha_H - \epsilon)\)-approximation for \textsc{Dir-MulC-H}.

In this paper we prove weaker forms of the conjecture, captured in the following two theorems:
Theorem 1.2. Assuming UGC, for any fixed directed bipartite graph $H$, and for any fixed $\varepsilon > 0$ there is no polynomial-time $(\alpha_H - \varepsilon)$ approximation for Dir-MulC-H.

Theorem 1.3. Assuming UGC, for any fixed directed graph $H$ on $k$ vertices and for any fixed $\varepsilon > 0$, there is no polynomial-time $\frac{\alpha_H}{2 \log k} - \varepsilon$ approximation for Dir-MulC-H.

Via known flow-cut gap results [20] and some standard reductions we obtain the following corollary.

Corollary 1.2. Assuming UGC, for any fixed $\varepsilon > 0$ the following hold.

- For any fixed $k$, if $H$ is a collection of $k$ disjoint directed edges then Dir-MulC-H is hard to approximate within a factor of $k - \varepsilon$.
- Separating $s$ from $t$ and $t$ from $s$ in a directed graph (Dir-Multiway-Cut with $2$ terminals) is hard to approximate within a factor of $2 - \varepsilon$.
- For any fixed $k$, Dir-Skew-MultiCut’s approximability coincides with its flow-cut gap.

In a recent independent work, Lee [16] has also shown that for fixed $k$, assuming UGC, Dir-MulC-H is $k - \varepsilon$ hard when $H$ is a collection of $k$ disjoint edges.

Our last result is on upper bounds for Dir-MulC. Corollary 1.2 shows that if $H$ contains, a matching of size $k$ as an induced subgraph then we cannot obtain a better than $k$ approximation. The following question arises naturally.

Question 2. Let $H$ be a fixed demand graph that does not contain a matching of size $k$ as an induced subgraph. Is there a $k$-approximation for Dir-MulC-H?

A positive answer to above question would imply a $2$-approximation for Dir-Skew-MultiCut. Since we are currently unable to improve the $O(\log k)$-approximation for Dir-Skew-MultiCut, we consider a relaxed version of the preceding question and give a positive answer. We say that a directed demand graph $H = (V,F)$ being a complete graph on an even number of nodes and remove edges from $H$ corresponding to a perfect matching $M$ on $V$ (if $uv \in M$ we remove $(u,v)$ and $(v,u)$ from $F$). We claim that the resulting graph $H'$ does not contain an induced $4$-matching-extension. What is the approximability of Dir-MulC-H? It is fair to say that previous work would not have found it easy to answer this question since $H'$ does not appear to have the same nice structure that the complete graph has. The theorem below shows that one can obtain a $3$ approximation for Dir-MulC-H.

Theorem 1.4. Consider Dir-MulC-H where $H$ does not contain an induced $k$-matching-extension. Then the flow-cut gap is at most $k-1$ and there is a polynomial-time rounding algorithm that achieves this upper bound.

The rounding scheme that proves the preceding theorem is built upon our recent insight for Dir-Multiway-Cut [7]. Interestingly the rounding scheme is itself oblivious to the demand graph $H$. It either provably obtains a $(k-1)$ approximation via the LP solution or provides a certificate that $H$ contains an induced $k$-matching-extension.

Techniques: At a high-level our results are based on a labeling view for Multicut. For undirected graphs we show that this yields Theorem 1.1. In directed graphs we show that a labeling based LP is no more powerful than Distance-LP which is stark contrast to the undirected graph setting. The labeling LP allows us to relate the hardness of Dir-MulC-H to the hardness of a min constraint satisfaction problems (Min CSPs) via a standard labeling LP for CSPs called Basic-LP. We crucially rely on a general hardness result for Min-\beta-CSP due to Ene, Vondrak and Wu [10] that generalized prior work of Manokaran et al. [17]. Finally, Theorem 1.4 builds upon our recent work on a simpler rounding scheme for Dir-Multiway-Cut [7].

Organization: Section 2 describes the factor 2-approximation for Undir-MulC with $K_2$-free demand graph. Section 3 describes the hardness of approximation results for Dir-MulC-H. Section 4 describes the $(k-1)$ approximation for Dir-MulC-H when $H$ does not contain an induced $k$-matching extension.
2 Approximating Undir-MulC with $tK_2$-free demand graph

In this section we obtain 2-approximation for $tK_2$-free demand graphs and prove Theorem 1.1.

**Theorem 1.1.** There is a 2-approximation algorithm with running time $\text{poly}(n, k^{O(1)})$ on instances of Undir-MulC with supply graph $G$ and $tK_2$-free demand graph $H$. Here $n = V(G)$ and $k = V(H)$.

Before we prove the theorem, we consider the Undir-MulC problem where the demand graph has $k$ vertices. Given supply graph $G = (V, E)$ let $S = \{s_1, \ldots, s_k\} \subset V$ be the terminals participating in the demand edges specified by $H$. A feasible solution $E' \subset E_G$ of the Undir-MulC instance will induce a partition over $S$ such that if $s_is_j$ is an edge in the demand graph $H$, then $s_i$ and $s_j$ belong to different components in $G - E'$. Note that two terminals that are not connected by a demand edge may be in the same connected component of $G - E'$. If $k$ is a fixed constant we can “guess” the partition of the terminals induced by an optimum solution. With the guess in place it is easy to see that the problem reduces to an instance of Undir-Multiway-Cut which admits a constant factor approximation. Thus, one can obtain a constant factor approximation for Undir-MulC in $2^{O(k \log k)} \text{poly}(n)$ time by trying all possible partitions of the terminals.

To prove Theorem 1.1 we use this idea of enumerating feasible partitions. However, $H$ is not necessarily of fixed size, and enumerating all possible partitions of the terminals is not feasible. Instead, we make use of the following theorem which bounds the number of maximal independent sets in a $tK_2$-free graph.

**Theorem 2.1.** (Balas and Yu [3]) Any $s$-vertex $tK_2$-free graph has at most $s^{O(t)}$ maximal independent sets and these can be found in $s^{O(t)}$ time.

We prove Theorem 1.1 by using the preceding theorem and reducing the Undir-MulC problem to the Uniform-MetricLabeling problem. We now describe the general MetricLabeling problem.

**MetricLabeling:** The input consists of an undirected edge-weighted graph $G = (V, E)$, a set of labels $L = \{1, \ldots, h\}$ and a metric $d(i,j)$, $i, j \in L$ defined over the labels. In addition, for each vertex $u \in V$ and label $i \in L$ there is a non-negative assignment cost $c(u,i)$. Given an assignment $f : V \rightarrow L$ of vertices to labels we define its cost as $\sum_{v \in V} c(u,f(u)) + \sum_{u \in E} w(uv)d(f(u), f(v))$. The goal is to find an assignment of minimum cost. The special case when the metric is uniform, that is $d(i,j) = 1$ for $i \neq j$, is referred to as Uniform-MetricLabeling.

**Theorem 2.2.** (Kleinberg and Tardos [13]) There is a 2-approximation for Uniform-MetricLabeling.

**Proof of Theorem 1.1.** Let the demand graph $H$ of the Undir-MulC instance be $tK_2$-free. Using Theorem 2.1 we can find all maximal independent sets in $H$. Let these independent sets be $I_1, \ldots, I_r$ where $r \leq |V_H|^{O(1)}$. Note that the independent sets are considered only in the demand graph. Note that instance also need to specify the mapping of vertices of $H$ to vertices of $G$. However, for ease of notation we will simply assume that $V_H \subset V_G$.

Consider the following instance of Uniform-MetricLabeling: The supply graph $G = (V, E)$ of the Undir-MulC instance is the input graph to the Uniform-MetricLabeling instance. The label set $L = \{1, 2, \ldots, r\}$, one for each maximal independent set in $H$. For each $u \in V_H$ let $c(u,i) = 0$ if $u \in I_i$ and $c(u,i) = \infty$ otherwise. And for each $u \in V \setminus V_H$, $c(u,i) = 0$ for all $i$.

We claim that the preceding reduction is approximation preserving. Assuming the claim, we can obtain the desired 2-approximation by solving the Uniform-MetricLabeling instance using Theorem 2.2. The size of the Uniform-MetricLabeling instance that is generated from the given Undir-MulC instance is $\text{poly}(n, |V_H|^{O(1)})$ which explains the running time. We now prove the claim.

Let $f : V \rightarrow L$ be an assignment of labels to the nodes whose cost is finite (such an assignment always exists since each terminal is in some independent set). Let $E' \subset E$ be the set of edges “cut” by this assignment; that is, $uv \in E'$ iff $f(u) \neq f(v)$. The cost of this assignment is equal to the weight of $E'$ since the metric is uniform and the labeling costs are 0 or $\infty$. We argue that $E'$ is a feasible solution for the Undir-MulC instance. Suppose not. Then there are terminals $u,v$ such that $uv$ is an edge in the demand graph $H$ and $u,v$ belong to the same connected component of $G - E'$. The label $j = f(u)$ corresponds to a maximal independent set $I_j$ in $H$ which means that $v \notin I_j$. Thus $f(v) \neq j$ since $c(v,j) = \infty$. Therefore, $u$ and $v$ are assigned different labels and cannot be in the same connected component.

Conversely, let $E' \subset E$ be a feasible solution for Undir-MulC instance and let $V_1, \ldots, V_s$ be vertex sets of the connected components of $G - E'$. Let $T_j$ be the terminals in $V_j$. Since, all pairs of terminals connected by an edge in $H$ are separated in $G - E'$, $T_j$ must be an
independent set in $H$. For each $T_j$, consider a maximal independent set in $H$ containing all the vertices of $T_j$; pick arbitrary one if more than one exists. Let this independent set be $I_i$. We construct a labeling $f$ by labeling all vertices of $V_j$ by label $i$. It is easy to see that all terminals are assigned a label corresponding to an independent set in $H$ containing that terminal. Hence, labeling cost is equal to zero. Also, all vertices corresponding to same connected component in $G - E'$ are assigned the same label. Hence, cost of the edges cut by the assignment $f$ is at most the cost of the edges in $E'$.

3 UGC-based hardness of approximation results for Dir-MulC

In this section we prove hardness of approximation for Dir-MulC-H, in particular Theorem 1.2 relating the hardness of approximation to the flow-cut gap. Recall that $\alpha_H$ is the worst-case flow-cut gap (equivalently, the integrality gap of the Distance-LP) for instances of Dir-MulC-H.

**Theorem 1.2.** Assuming UGC, for any fixed directed bipartite graph $H$, and for any fixed $\varepsilon > 0$ there is no polynomial-time $(\alpha_H - \varepsilon)$ approximation for Dir-MulC-H.

We prove the theorem via a reduction to Min-$\beta$-CSP and the hardness result of Ene, Vondrák and Wu [10]. We note that the result is technical and involves several steps. This is partly due to the fact that the theorem is establishing a meta-result. The theorem of [10] is in a similar vein. In particular [10] establishes that the hardness of Min-$\beta$-CSP depends on the integrality gap of a specific LP formulation Basic-LP (defined later). Our proof is based on establishing a correspondence between Dir-MulC-H and a specific constraint satisfaction problem Min-$\beta_H$-CSP where $\beta_H$ is constructed from $H$ (this is the heart of the reduction) and proving the following properties:

(I) Establish approximation equivalence between Dir-MulC-H and Min-$\beta_H$-CSP. That is, prove that each of them reduces to the other in an approximation preserving fashion.

(II) Prove that if the flow-cut gap for Dir-MulC-H (equivalently the integrality gap of Distance-LP) is $\alpha_H$ then the integrality gap of Basic-LP for Min-$\beta_H$-CSP is also $\alpha_H$.

From (I), we obtain that the hardness of approximation factor for Dir-MulC-H and Min-$\beta_H$-CSP coincide. From (II), we can apply the result in [10] which shows that, assuming UGC, the hardness of approximation for Min-$\beta_H$-CSP is the same as the integrality gap of Basic-LP. Putting together these two claims give us our desired result.

It is not straightforward to directly relate Distance-LP for Dir-MulC-H and Basic-LP for Min-$\beta_H$-CSP. Basic-LP appears to be stronger on first glance. In order to relate them we show that a seemingly strong LP for Dir-MulC we call LABEL-LP is in fact no stronger than Distance-LP. In fact this can be seen as the key technical fact unerlying the entire proof and is independently interesting since it is quite different from the undirected graph setting. It is much easier to relate LABEL-LP and Basic-LP. The rest of this section is organized as follows. In Section 3.1 we describe LABEL-LP and prove its equivalence with Distance-LP. In Section 3.2 we describe Min-$\beta$-CSP and Basic-LP and formally state the theorem of [10] that we rely on. We then subsequently describe our reduction from Dir-MulC-H to Min-$\beta_H$-CSP and complete the proof.

3.1 Label-LP and equivalence with Distance-LP for Dir-MulC

In Section 2 we saw that if demand graph $H$ has size $k$, then there is a labeling LP for MultiCut (the undirected problem) with size poly($2^k$, $n$) and integrality gap at most 2 which improves upon the integrality gap of Distance-LP which can be $\Omega(\log k)$. Here we describe a natural labeling LP for Dir-MulC (LABEL-LP), but in contrast to the undirected case, we show that it is not stronger than Distance-LP. We show this equivalence on an instance by instance basis. That is, for any Dir-MulC instance $I$, given a solution to Distance-LP, we can find a solution to LABEL-LP with same cost and vice versa.

Let the demand graph be $H$ with vertex set $V_H = \{s_1, \ldots, s_k\}$, and the supply graph be $G = (V_G, E)$ with $n$ vertices. We will assume here, for ease of notation, that $V_H \subset V_G$. Define a labeling set $L = \{0, 1\}^k$ which corresponds to all subsets of $V_H$. We interpret the labels in $L$ as $k$-length bit-vectors; if $\sigma \in L$ we use $\sigma[i]$ to denote the $i$'th bit of $\sigma$. For two labels $\sigma_1, \sigma_2 \in L$ we say $\sigma_1 \leq \sigma_2$ if $\forall i, \sigma_1[i] \leq \sigma_2[i]$. To motivate the formulation consider any set of edges $E' \subseteq E$ that can be cut. In $G' = G - E'$ we consider, for each $v \in V$, the reachability information from each of the terminals $s_1, s_2, \ldots, s_k$. For each $v$ this can be encoded by assigning a label $\sigma_v \in L$ where $\sigma_v[i] = 1$ iff $v$ is reachable from $s_i$ in $G'$.

The goal of the formulation to assign labels to vertices and to ensure that demand pairs are separated. An edge $e = (u, v)$ is cut if there is some $s_i$ such that $s_i$ can reach
u but s₁ cannot reach v. We add several constraints to ensure that the label assignment is consistent. The basic variables are \( z_{e,v} \) for each \( e \in E_G \) and \( \sigma \in L \) which indicate whether \( v \) is assigned the label \( \sigma \). We also a variable \( x_e \) for each edge \( e = (u,v) \in E_G \) that is derived from the label assignment variables. We start with the basic constraints involving these variables and then add additional variables that ensure consistency of the assignment.

- Each vertex is labelled by exactly one label. For \( v \in V_G \), \( \sum_{\sigma \in L} z_{v,\sigma} = 1 \).
- Vertex \( s_i \) is reachable from \( s_1 \). For \( s_i \in V_H \) and any \( \sigma \in L \) such that \( \sigma[i] = 0 \), \( z_{s_i,\sigma} = 0 \).
- Demand edges are separated. That is, if \( (s_i,s_j) \in E_H \), then \( s_j \) is not reachable from \( s_i \). That is, \( z_{s_j,\sigma} = 0 \) for any \( \sigma \) where \( \sigma[i] = 1 \) and \( (s_i,s_j) \in E_H \).

For each edge \( e = (u,v) \) we have variables of the form \( z_{e,u,\sigma} \) where the intention is that \( u \) is labeled \( \sigma_1 \) and \( v \) is labeled \( \sigma_2 \). To enforce consistency between edge assignment variables and vertex assignment variables we add the following set of constraints.

- For \( e = (u,v) \in E_G \), \( z_{u,\sigma} = \sum_{\sigma_2 \in L} z_{e,\sigma_1,\sigma_2} \) and \( z_{v,\sigma} = \sum_{\sigma_1 \in L} z_{e,\sigma_1,\sigma_2} \).

Finally, the auxiliary variable \( x_e \) indicates whether \( e \) is cut.

- For \( e = (u,v) \in E_G \), \( x_e = 1 \) if for some \( i \), \( u \) is reachable from \( s_i \) and \( v \) is not reachable from \( s_i \).

Then, \( x_e = 1 \) if \( z_{e,\sigma_1,\sigma_2} = 1 \) for \( \sigma_1 \neq \sigma_2 \). We thus set \( x_e = \sum_{\sigma_1,\sigma_2 \in L} z_{e,\sigma_1,\sigma_2} \).

It is not hard to show that if one constraint all the variables to be binary then the resulting integer program is a valid formation for Dir-MULC. Note that the number of variables is exponential in \( k = |V_H| \). Relaxing the integrality constraint on the variables, we get LABEL-LP as shown in Fig. 3.

**Theorem 3.1.** For any instance \( G, H \) of Dir-MULC-H, the optimum solution values for the formulations LABEL-LP and DISTANCE-LP are the same both in the fractional and integral settings.

**LABEL-LP**

\[
\begin{align*}
\min & \sum_{e \in E} w_e x_e \\
\text{s.t.} & \quad \sum_{\sigma \in L} z_{e,v,\sigma} = 1 \quad v \in V_G, \sigma \in L \\
& \quad z_{s_i,\sigma} = 0 \quad s_i \in V_H, \sigma \in L, \sigma[i] = 0 \\
& \quad z_{s_j,\sigma} = 0 \quad \sigma \in L, \sigma[i] = 1 \quad (s_i,s_j) \in E_H \\
& \quad \sum_{\sigma_2 \in L} z_{e,\sigma_1,\sigma_2} = z_{u,\sigma_1} \quad e = (u,v) \in E_G, \sigma_1 \in L \\
& \quad \sum_{\sigma_1 \in L} z_{e,\sigma_1,\sigma_2} = z_{v,\sigma_2} \quad e = (u,v) \in E_G, \sigma_2 \in L \\
& \quad \sum_{\sigma_1,\sigma_2 \in L} z_{e,\sigma_1,\sigma_2} = x_e \quad e \in E_G \\
& \quad \sum_{\sigma_1 \in L} z_{e,\sigma_1,\sigma_2} = 1 \quad v \in V_G, e \in E_G, \sigma_1, \sigma_2 \in L \\
& \quad z_{v,\sigma} \geq 0 \quad v \in V_G, e \in E_G, \sigma \in L \\
\end{align*}
\]

**Figure 3: LABEL-LP for Dir-MULC**

\( \Gamma_v = \{ v_\sigma \mid \sigma \in L \} \). We assign cost \( c(\sigma,\sigma') \) on the edge \( (u_\sigma, v_{\sigma'}) \). We assign a supply of \( z_{u,\sigma} \) on the vertex \( u_\sigma \) and a demand of \( z_{v,\sigma} \) on the vertex \( v_{\sigma'} \). The values \( z_{e,\sigma_1,\sigma_2} \) can be thought of as flow from \( u_{\sigma_1} \) to \( v_{\sigma_2} \) satisfying the following properties: (i) total flow out of \( u_{\sigma_1} \) must be equal to the supply \( z_{u,\sigma_1} \) \( (z_{u,\sigma_1} = \sum_{\sigma_2 \in L} z_{e,\sigma_1,\sigma_2}) \) (ii) total flow into \( v_{\sigma_2} \) must be equal to \( z_{v,\sigma_2} \) \( (z_{v,\sigma_2} = \sum_{\sigma_1 \in L} z_{e,\sigma_1,\sigma_2}) \) (iii) flow is non-negative \( (z_{e,\sigma_1,\sigma_2} \geq 0) \). The cost of the flow according to \( c \) is precisely \( x_e \) \( (= \sum_{\sigma_1,\sigma_2 \in L} z_{e,\sigma_1,\sigma_2}) \). In particular, given the values of the labeling variables \( z_{u,\sigma}, \sigma \in L \) and \( z_{v,\sigma'}, \sigma' \in L \) which can be thought of as two distributions on the labels, the smallest value of \( x_e \) that can be achieved is essentially the min-cost flow in \( B_{uv} \) with supplies and demands defined by the two distributions. In other words the other variables are completely determined by the distributions if one wants a minimum cost solution.

In the sequel we use \( \pi_u \) to denote the vector of assignment value \( z_{u,\sigma}, \sigma \in L \) and refer to \( \pi_u \) as the distribution corresponding to \( u \). We present the high-level reduction between the solutions of the two LP’s and refer to Section A for the full proof.

**From LABEL-LP to Distance-LP:** Let \( (x,z) \) be a feasible solution to LABEL-LP for an instance \( (G,H) \). This solution satisfies the following two conditions: (i) If \( (s_i,s_j) \in E_H \), then \( \sum_{\sigma \in \{0,1\}^{k:\sigma[i]=1}} z_{s_i,\sigma} = 1 \) and \( \sum_{\sigma \in \{0,1\}^{k:\sigma[i]=1}} z_{s_j,\sigma} = 0 \). (ii) For an edge \( e = (u,v) \in E_G \), and any terminal \( s_i, x_e \geq \sum_{\sigma \in \{0,1\}^{k:\sigma[i]=1}} z_{v,\sigma} - \sum_{\sigma \in \{0,1\}^{k:\sigma[i]=1}} z_{u,\sigma} \).

Suppose \( (s_i,s_j) \in E_H \) and \( s_i,a_1, \ldots, a_i, s_j \) is a path from \( s_i \) to \( s_j \) in \( G \). Then, plugging in the above inequalities for the edges of the path, we get \( \sum_{e \in P} x_e \geq \)}
1. Hence, \textbf{x} is a feasible solution to DISTANCE-LP and has same cost as \((x, z)\).

**From Distance-LP to Label-LP:** Let \(x\) be a feasible solution to DISTANCE-LP. We obtain a label assignment \(z'\) as follows. For a vertex \(u \in V_G\), let \(d(s_1, u) \leq d(s_2, u) \leq \cdots \leq d(s_k, u)\) (if not, rename the terminals accordingly). Here, \(d(u, v)\) denotes the shortest path distance from \(u\) to \(v\) as per lengths \(x_e\). For \(i \in [0, k]\), let \(\sigma_i = 10^k - i\). Then, set \(z'_u, \sigma_i = d(s_i, u)\) and for \(i \in [1, k-1]\), \(z'_u, \sigma_i = d(s_{i+1}, u) - d(s_i, u)\). For \(\sigma \notin \{\sigma_0, \ldots, \sigma_k\}\), \(z'_{u, \sigma} = 0\). Once the label assignment for vertices is defined, we obtain the values of other variables by considering each edge \(e = (u, v)\) and using the min-cost flow between \(z'_u\) and \(z'_v\), as described earlier. In Section \([3]\) we prove that this flow has cost \((z'_{u, v})\) at most \(\max_{i \in [1, k]} d(s_i, v) - d(s_i, u)\) which is upper bounded by \(x_e\). Hence, cost of solution \((z', z)\) to LABEL-LP is upper bounded by cost of \(x\).

### 3.2 Min-CSP and Basic-LP

Min-CSP refers to a minimization version of constraint satisfaction problems. We set up the formalism borrowed from \([16]\). Let \(L\) denote a finite set of labels. A real-valued function \(f : L^1 \rightarrow \mathbb{R}\) has *arity* \(i\). Let \(\Gamma = \{\psi \mid \psi : L^i \rightarrow [0, 1) \cup \{\infty\}, i \leq k\}\) be the set of functions defined on \(L\) with *arity atmost* \(k\) and range \([0, 1) \cup \{\infty\}\). Let \(\beta \subset \Gamma\) be a finite subset of \(\psi\). These functions are also referred to as *predicates*. \(k\) denotes the *arity* and \(L\) denotes the alphabet of \(\beta\). Each \(\beta\) induces an optimization problem Min-\(\beta\)-CSP.

#### Definition 3.1.

An instance of Min-\(\beta\)-CSP consists of the following:

- A vertex set \(V\) and a set of tuples \(T \subset \bigcup_{i=1}^{k} V^i\).
- A predicate \(\psi_i \in \beta\) for each tuple \(t \in T\) where cardinality of \(t\) matches the *arity* of \(\psi_i\).
- A non-negative weight function over the set of tuples, \(w : T \rightarrow \mathbb{R}^+\).

The goal is to find a label assignment \(\ell : V \rightarrow L\) to minimize \(\sum_{t=(v_{i_1}, \ldots, v_{i_j}) \in T} w_t \cdot \psi_i(\ell(v_{i_1}), \ldots, \ell(v_{i_j}))\).

Consider an integer programming formulation with following variables: for each vertex \(v \in V\) and label \(\sigma \in L\), we have a variable \(z_{v, \sigma}\) which is 1 if \(v\) is assigned label \(\sigma\). Also, for each tuple \(t = (v_{i_1}, \ldots, v_{i_j}) \in T\) and \(\alpha \in L^{|t|}\), we have a boolean variable \(z_{t, \alpha}\) which is 1 if \(v_{i_p}\) is labelled \(\alpha[p]\) for \(p \in [1, j]\). These variables satisfy following constraints:

- Each vertex receives unique label: \(\sum_{\sigma \in L} z_{v, \sigma} = 1\).
- Variables \(z_{v, \sigma}\) and \(z_{t, \alpha}\) are consistent. That is, if \(v \in t\) is assigned label \(\sigma\), then \(z_{t, \alpha}\) must be zero if \(\alpha\) does not assign label \(\sigma\) to \(v\). For every label \(\sigma \in L\), the following holds:

\[
\min \sum_{t \in T} w_t \sum_{\alpha \in L^{|t|}} z_{t, \alpha} \cdot \psi_i(\alpha)
\]

\[
\sum_{\alpha \in L} z_{t, \alpha} = 1 \quad v \in V
\]

\[
\sum_{\alpha \in L^{|t|}} z_{t, \alpha} = z_{v, \sigma} \quad t \in T, v = t[i], \sigma \in L
\]

\[
z_{v, \sigma}, z_{t, \alpha} \geq 0 \quad v \in V, \sigma \in L, t \in T, \alpha \in L^{|t|}
\]

\[
z_{v, \sigma}, z_{t, \alpha} \leq 1 \quad v \in V, \sigma \in L, t \in T, \alpha \in L^{|t|}
\]

Figure 4: Basic LP for Min-\(\beta\)-CSP

\(t \in T, v = t[i], \sigma \in L\), we have: \(z_{v, \sigma} = \sum_{\alpha \in L^{|t|}; \alpha[i] = \sigma} z_{t, \alpha}\).

The objective is minimize \(\sum_{t \in T} w_t \cdot \sum_{\alpha \in L^{|t|}} z_{t, \alpha} \cdot \psi_i(\alpha)\).

Basic-LP is the LP relaxation obtained by allowing the variables to take on values in \([0, 1]\) and is described in Fig 4. For instance \(I\), LP(\(I\)) and OPT(\(I\)) refer to the fractional and integral optimum values respectively.

**Definition 3.2.** For \(i \geq 2\), NAE\(_i\) : \(L^i \rightarrow \{0, 1\}\) be a predicate such that NAE\(_i\)\((\sigma_1, \ldots, \sigma_i) = 0\) if \(\sigma_1 = \sigma_2 = \cdots = \sigma_i = 1\) and 1 otherwise.

The following theorem shows that the hardness of Min-\(\beta_H\)-CSP coincides with the integrality gap of BASIC-LP if NAE\(_2\) is in \(\beta\).

**Theorem 3.2.** (Ene, Voron, Wu \([17]\)) Suppose we have a Min-\(\beta\)-CSP instance \(I = (V, T, \Psi, w)\) with fractional optimum (of Basic LP) \(LP(I) = c\), integral optimum \(OPT(I) = s\), and \(\beta\) contains the predicate NAE\(_2\). Then, assuming UGC, for any \(\epsilon\), for some \(\lambda > 0\), it is NP-hard to distinguish between instances of Min-\(\beta\)-CSP where the optimum value is at least \((s - \epsilon)\lambda\) and instances where the optimum value is less than \((c + \epsilon)\lambda\).

### 3.3 Dir-MulC-H and an equivalent Min-\(\beta\)-CSP Problem

In this section, we show that given a bipartite directed graph \(H = (S \cup T, E_H)\), we can construct a set of predicates \(\beta_H\) such that Dir-MulC-H is equivalent to Min-\(\beta_H\)-CSP. The notion of equivalence is as follows. We give a reduction from instances of Dir-MulC-H to instances of Min-\(\beta_H\)-CSP which preserves the cost of optimal integral solution and in addition also preserves the cost of optimum fractional solution to LABEL-LP and BASIC-LP. Similarly we give a reduction from Min-\(\beta_H\)-CSP to Dir-MulC-H which
The basic idea behind the construction of $\beta_H$ from $H$ is to simulate the constraints of LABEL-LP via the predicates of $\beta_H$. In addition to setting up $\beta_H$ correctly, we also need to preprocess the supply graph to prove the correctness of the reductions. Let the bipartite demand graph $H$ be $(S \cup T, E_H)$ with $S = \{s_1, \ldots, s_p\}$ and $T = \{t_1, \ldots, t_q\}$ as the bipartition. For $u \in S$ let $N_H^+(u) = \{v \in T \mid (u, v) \in E_H\}$ be the neighbors of $u$ in $H$. For $i \in [1, p]$, let $Y_i = \{j \in [1, p] \mid N_H^+(a_j) \subseteq N_H^+(a_i)\}$. That is, if $a_j \in Y_i$, the set of terminals that $a_j$ needs to be separated from is a subset of the terminals that $a_i$ needs to be separated from. For $j \in [1, q]$ let $Z_j = \{i \in [1, p] \mid a_i b_j \notin E_H\}$. That is, $Z_j$ is the set of all terminals in $S$ that do not need to be separated from $b_j$.

**Assumptions on supply graph:** We will assume that the supply graph $G$ in the instances of Dir-MulC-H satisfy the following properties.

- **Assumption I:** $G$ may contain undirected edges. The meaning of this is that a path may include this edge in either direction. A simple and well-known gadget shown in Fig 5 shows that this is without loss of generality.

- **Assumption II:** For $1 \leq j \leq q$ and $i \in Y_i$, there is an infinite weight edge from $a_i$ to $b_j$ in $G$. Moreover $b_j$ has no outgoing edge.

- **Assumption III:** For $1 \leq i \leq p$, and $j' \in Y_i$, there is an infinite weight edge from $a_i$ to $b_{j'}$ in $G$. Moreover $a_i$ has no other incoming edges.

The preceding assumptions are to make the construction of $\beta_H$ and the subsequent proof of equivalence with Dir-MulC-H somewhat more transparent and technically easier. Undirected edges allow us to use the NAE2 predicate in $\beta_H$. Assumption II and III simplify the reachability information of terminals that needs to be kept track of and this allows for a simpler label set definition and easier proof of equivalence.

DISTANCE-LP easily generalizes to handle undirected edges; in examining paths from $s_i$ to $t_i$ for a demand pair we allow an undirected edge to be used in both directions. A more technical part is to generalize LABEL-LP to handle undirected edges in the supply graph. For a directed edge $e$ recall that $x_e = \sum_{\sigma_1, \sigma_2 \in L, \sigma_1 \neq \sigma_2} z_{e, \sigma_1, \sigma_2}$. For an undirected edge $e$ we set $x_e = \sum_{\sigma_1, \sigma_2 \in L, \sigma_1 \neq \sigma_2} z_{e, \sigma_1, \sigma_2}$. See Section B for the justification of the assumptions.

**Constructing $\beta_H$ from $H$:** Next, we formally define $\beta_H$ for a bipartite graph $H = (S \cup T, E_H)$ where $S = \{a_1, \ldots, a_p\}$ and $T = \{b_1, \ldots, b_q\}$. Recall the definitions of $Y_i$ for $1 \leq i \leq p$ and $Z_j$ for $1 \leq j \leq q$ based on $E_H$. Observe that no vertex other than $b_j$ is reachable from $b_j$. And, since labels encode the reachability from terminals, we can ignore the reachability from $b_j$ and define $\beta_H$ with respect to terminal set $S$. For $\sigma \in \{0, 1\}^p$, let $J_\sigma = \{i \in [1, p] \mid \sigma[i] = 1\}$

- **Alphabet (Label Set)** $L = \{0, 1\}^p$. Labels encode the list of $a_i$’s from which a vertex is reachable.

- **For $i \in [1, p]$, a unary predicate $\psi_{a_i}$ encodes the correct label for $a_i$, and is defined as follows:** $\psi_{a_i}(\sigma) = 0$ if $J_\sigma = Y_i$, otherwise $\psi_{a_i}(\sigma) = \infty$.

- **For $i \in [1, q]$, predicate $\psi_{b_j}$ encodes the correct label for $b_j$.** $\psi_{b_j}(\sigma) = 0$ if $J_\sigma = Z_j$, otherwise $\psi_{b_j}(\sigma) = \infty$.

- **A binary predicate $C$ that encodes whether a directed edge is cut or not.** If $\sigma_1 \leq \sigma_2 C(\sigma_1, \sigma_2) = 0$, otherwise $C(\sigma_1, \sigma_2) = 1$.

- **A binary predicate NAE2 that encodes whether an undirected edge is cut or not.** If $\sigma_1 \leq \sigma_2$ NAE2(\sigma_1, \sigma_2) = 0, otherwise NAE2(\sigma_1, \sigma_2) = 1.

Thus $\beta_H = \{C, NAE2\} \cup \{\psi_{a_i} \mid i \in [1, p]\} \cup \{\psi_{b_j} \mid j \in [1, q]\}$. Min-$\beta_H$-CSP has label set $L$, predicate set $\beta_H$ and arity 2.

The main technical theorem we prove is the following. We remark that when we refer to Dir-MulC-H we are referring to the problem where the supply graph satisfies the assumptions I, II, III that we outlined previously.

**Theorem 3.3.** Let $H$ be a directed bipartite graph. There is a polynomial time reduction that given a Dir-MulC-H instance $I_M = (G = (V_G, E_G, w_G : E_G \to R^+), H = (S \cup T, E_H))$, outputs a Min-$\beta_H$-CSP instance $I_C = (V_C, T_C, \psi_{T_C} : T_C \to \beta_H, w_{T_C} : T_C \to R^+)$ such that the following holds: given a solution $(x, z)$ of the Label LP for $I_M$, we can construct a solution $z'$ of Basic LP for $I_C$ with cost at most that of $(x, z)$ and vice versa. More over, if $(x, z)$ is an integral solution, then $z'$ is also an integral solution and vice versa. A similar reduction exists from Min-$\beta_H$-CSP to Dir-MulC-H.

With the preceding theorem in place we can formally prove Theorem 1.2.

**Proof of Theorem 1.2:** Let $I_M$ be some fixed instance of Dir-MulC-H with flow-cut gap $\alpha_H$. From
Theorem 3.3 implies that Min-β-H-CSP reduces to Dir-MULC-H in an approximation preserving fashion. Thus, Dir-MULC-H is at least as hard to approximate as Min-β-H-CSP which implies that assuming UGC, the hardness of Dir-MULC-H is at least α_H - ε for any fixed ε > 0.

Basic-LP and Label-LP are almost identical except for the fact that Basic-LP is defined with label set \{0,1\}^k where k = p + q is the total number of terminals whereas Basic-LP is defined with label set \{0,1\}^p. However, since b_i’s do not have any outgoing edge, reachability from b_i is trivial. The formal proof of equivalence is long and somewhat tedious. We need to consider a reduction from Min-β-H-CSP to Dir-MULC-H and vice-versa. In each direction we need to establish the equivalence of the cost of Label-LP and Basic-LP for both integral and fractional settings. We will briefly sketch the reduction here. Full proofs can be found in Section 3.

Reduction from Min-β-H-CSP to Dir-MULC-H:

Given a Min-β-H-CSP instance I_C, equivalent Dir-MULC-H instance I_M is constructed as follows: (i) Vertex set of I_M is same as that of I_C. (ii) For i ∈ [1, p], name one of the vertex v ∈ V_C with constraint ψ_{a_i}(v) as vertex a_i and for rest of the vertices u ∈ V_C with ψ_{a_i}(u), add an undirected infinite weight edge between a_i and u. (iii) For constraint C(u, v), add a directed edge e_i = (u, v) and for constraint NAE_E(u, v), add an undirected edge e_i = u v. Add edges among a_i’s and b_j’s so as to satisfy Assumption II and III. Next, we show how to convert a solution for one LP to a solution to the other LP while preserving cost.

From Label-LP to Basic-LP: Let (x, z) be a feasible solution to Label-LP for I_M. Then, a feasible solution z’ to Basic-LP for I_C is a projection of z from label space \{0,1\}^p+q to label space \{0,1\}^p. Formally, z’ is defined as follows: For σ ∈ \{0,1\}^p, z’_{σ,0,σ} = \sum_{σ’ ∈ \{0,1\}^p} z_{σ,σ’,σ} (ii) For σ_1, σ_2 ∈ \{0,1\}^p, z’_{σ_1,σ_2} = \sum_{σ, σ’ ∈ \{0,1\}^p} z_{σ_1,σ_2,σ,σ’}. We can argue that z’ is at most the cost of solution (x, z).

From Basic-LP to Label-LP: Let z be a feasible solution to Basic-LP for I_C. Let σ_0 = 1, then a feasible solution (x’, z’) to Label-LP can be defined as an extension of z along σ_0. Formally, z’ is defined as follows: For σ ∈ \{0,1\}^p, σ’ ∈ \{0,1\}^p, v ∈ V_C, z’_{σ,0,σ'} = 1 if σ’ = σ_0 and 0 otherwise. Similarly, for σ_1, σ_2 ∈ \{0,1\}^p, σ’ ∈ \{0,1\}^p, z’_{σ_1,σ_2,0,σ'} = z_{σ_1,σ_2} if σ’ = 0 and 0 otherwise. We prove that x’ = ∑_{σ, σ_2 ∈ \{0,1\}^p} z’_{σ,σ,σ_2} · z_{σ_1,σ_2} = ∑_{σ, σ_2 ∈ \{0,1\}^p} z_{σ_1,σ_2,z_0}. Hence, cost of solution (x’, z’) is equal to cost of solution z.

3.4 Hardness for Non-bipartite Demand graphs

Here we prove Theorem 1.3 on the hardness of approximation of Dir-MULC-H when H is fixed and may not be bipartite. Let γ_H denote the hardness of approximation for Dir-MULC-H. Recall that α_H is the worst-case flow-cut gap for Dir-MULC-H.

Let the demand graph be H with 2^p vertices, V_H = \{s_σ | σ ∈ \{0,1\}^p\}. If number of vertices is not a power of two, then we can add dummy isolated vertices without changing the problem. We find r = 2p subgraphs H_1, ..., H_r such that H = H_1 ∪ ... ∪ H_r and

- Each H_i is a directed bipartite graph.
- α_H ≤ ∑_{i=1}^r γ_H_i.
- For 1 ≤ i ≤ r, there is an approximation preserving reduction from Dir-MULC-H_i to Dir-MULC-H. Hence, γ_H ≥ γ_H_i.

Since, H_i is bipartite, Theorem 1.2 implies, under UGC, that γ_H_i ≥ α_{H_i} - ε. Since, γ_H ≥ γ_H_i for all i ∈ [1, r], we have γ_H ≥ 1/r ∑_{i=1}^r γ_H_i. Therefore,

γ_H ≥ 1/r ∑_{i=1}^r γ_H_i ≥ 1/r (∑_{i=1}^r γ_H_i - ε) ≥ 1/r α_H - ε.

Since r = 2⌈log k⌉ where k = |V_H|, we obtain the proof of Theorem 1.3.

Next, we show how to construct H_i which satisfy the properties above. For each number j ∈ [1, p], define A_j = \{s_σ | σ ∈ \{0,1\}^p, σ(j) = 0\}, B_j = \{s_σ | σ ∈ \{0,1\}^p, σ(j) = 1\}. Let H_{2j-1} be the subgraph of H with vertex set V_H and edge set containing edges of H with head in B_j and tail in A_j. H_{2j} be the subgraph of H with vertex set V_H and edge set containing edges of H with head in A_j and tail in B_j.

V_{H_{2j-1}} = V_{H_{2j}} = V_H

E_{H_{2j-1}} = \{(s_{σ_1}, s_{σ_2}) ∈ E_H | s_{σ_1} ∈ A_j, s_{σ_2} ∈ B_j\}

E_{H_{2j}} = \{(s_{σ_1}, s_{σ_2}) ∈ E_H | s_{σ_1} ∈ B_j, s_{σ_2} ∈ A_j\}

By construction, it is clear that H_{2j-1}, H_{2j} are bipartite.
Lemma 3.3. \( H_i \) as defined above satisfy the following properties:

- \( E_H = \bigcup_{i=1}^p E_{H_i} \).
- \( \alpha_H \leq \sum_{i=1}^p \alpha_{H_i} \).
- For \( i \in [1, p] \), \( \gamma_H \geq \gamma_{H_i} \).

Proof. Let \( e = (s_{\sigma_1}, s_{\sigma_2}) \in E_H \). Since, there are no self-loops in \( H_i \), there exists \( j \in [1, p] \) such that either \( \sigma_1[j] = 0, \sigma_2[j] = 1 \) or \( \sigma_1[j] = 1, \sigma_2[j] = 0 \).

In the first case, \( e \in E_{H_{2j-1}} \), and in the second case \( e \in E_{H_{2j}} \).

- Given a Dir-MulC-H instance \( (G, H) \), idea is to solve \( (G_i, H_i) \) for \( i \in [1, p] \). Let \( I = (G, H) \) be a Dir-MulC-H instance. Let \( x \) be the optimal solution to Distance-LP on \( I \). Let \( I_i = (G_i, H_i) \) be the instance with the same supply graph \( G \) but demand graph \( H_i \). It is easy to see that \( x \) is a feasible fractional solution to \( I_i \), since \( H_i \) is a subgraph of \( H \). Since the worst-case integrality gap for Dir-MulC-H is \( \alpha_H \), there is a set \( E'_i \subseteq E_G \) such that \( w(E'_i) \leq \alpha_H w(x) \) and \( G - E'_i \) disconnects all demand pairs in \( H_i \). Clearly \( \bigcup_i E'_i \) is a feasible integral solution to \( (G, H) \) since \( H = \bigcup_i H_i \). The cost of \( \bigcup_i E'_i \) is at most \( \sum_i \alpha_{H_i} w(x) \). Since \( (G, H) \) was an arbitrary instance of Dir-MulC-H, this proves that \( \alpha_H \leq \sum_i \alpha_{H_i} \).

- We prove that there is an approximation preserving reduction from Dir-MulC-H to Dir-MulC-H which in turn proves that \( \gamma_H \geq \gamma_{H_i} \). Assume that \( i = 2j - 1 \) (case when \( i = 2j \) is similar). Let \( (G, H_i) \) be a Dir-MulC-H instance. \( G' \) is defined as follows:
  - \( V_{G'} = V_G \cup A'_j \cup B'_j \) where \( A'_j = \{ s'_\sigma \mid s_\sigma \in A_j \} \), \( B'_j = \{ s'_\sigma \mid s_\sigma \in B_j \} \).
  - \( G' \) contains all the edges of \( G \) and an infinite edge from \( s'_\sigma \) to \( s_\sigma \) for every \( s_\sigma \in A_j \) and infinite weight edge from \( s_\sigma \) to \( s'_\sigma \) for every \( s'_\sigma \in B_j \).

Let \( H' \) be a demand graph with vertex \( s_\sigma \) of \( H \) renamed as \( s'_\sigma \). Then, \( (G', H') \) is a Dir-MulC-H instance. Note that for \( s_\sigma \in A_j \), \( s'_\sigma \) in \( G' \) has no incoming edge and for \( s_\sigma \in B_j \), \( s'_\sigma \) in \( G' \) has no outgoing edge. Hence, for Dir-MulC instance \( (G', H') \), we only need to separate \( (s'_\sigma_1, s'_\sigma_2) \) if \( s_\sigma_1 \in A_j, s_\sigma_2 \in B_j \). Hence, Dir-MulC instances \( (G, H_i) \) and \( (G', H') \) are equivalent.

4 Approximating Dir-MulC

In this section, we prove Theorem 1.4 which improves the approximation ratio for Dir-MulC with restricted class of demand graphs. Recall that a directed demand graph \( H = (V, F) \) contains an induced \( k \)-matching extension if there are two subsets of \( V \), \( S = \{s_1, \ldots, s_k\} \) and \( T = \{t_1, \ldots, t_k\} \) satisfying the following properties:

(i) for \( 1 \leq i \leq k, (s_i, t_i) \in F \) and (ii) for \( i > j, (s_i, t_j) \notin F \).

Theorem 1.4. Consider Dir-MulC-H where \( H \) does not contain an induced \( k \)-matching-extension. Then the flow-cut gap is at most \( k - 1 \) and there is a polynomial-time rounding algorithm that achieves this upper bound.

Let \( G = (V, E) \) and \( H = (V, F) \) be the supply and demand graph for a Dir-MulC instance. We provide a generic randomized rounding algorithm that given a fractional solution \( x \) to LP [1] for an instance \( (G, H) \) of Dir-MulC returns a feasible solution; the rounding does not depend on \( H \). We can prove that the returned solution is a \((k-1)\)-approximation with respect to the fractional solution \( x \) or show that \( H \) contains an induced \( k \)-matching extension. This algorithm is inspired by our recent rounding scheme for Dir-Multiway-Cut [2].

Let \( x \) be a feasible solution to LP [1]. For \( u, v \in V \), define \( d(u, v) \) to be the shortest path length in \( G \) from vertex \( u \) to vertex \( v \) using lengths \( x_e \). We also define another parameter \( d_1(u, v) \) for each pair of vertices \( u, v \in V \). \( d_1(u, v) \) is the minimum non-negative number such that if we add an edge \( uv \) in \( G \) with \( x_{uv} = d_1(u, v) \) then \( u \) is still separated from all the vertices it has to be separated from. Formally, for \( u, v \in V \), \( d_1(u, v) := \max \{0, 1 - \min_{v' \in V \setminus \{u\}} d(u, v') \} \). If for some vertex \( u \), there is no demand edge leaving \( u \) in \( F \) then we define \( d_1(u, v) = 0 \) for all \( v \in V \). The following properties of \( d_1 \) are easy to verify.

Lemma 4.1. \( d_1(u, v) \) satisfies the following properties:

- \( \forall u \in V, d_1(u, u) = 0 \)
- \( \forall (u, v) \in F, \forall v' \in V, d_1(u, v') + d(v', v) \geq 1 \). Hence, \( \forall (u, v) \in F, d_1(u, v) \geq 1 \).
- \( \text{If } d_1(u, v') \neq 0, \text{ then there exists } (u, v) \in F \text{ such that } d_1(u, v') + d(v', v) = 1 \)
- \( \forall a, \forall (a, b) \in E, d_1(a, b) - d_1(a, u) \leq x_{ab} \)

Algorithm is a simple ball cut rounding around all the vertices as per \( d_1(u, v) \). We pick a number \( \theta \in (0, 1) \) uniformly at random. For all \( u \in V \), we consider \( \theta \) radius ball around \( u \) for all \( u \in V \). \( B_\theta = \{ v \in V \mid d_1(u, v) \leq \theta \} \). And then cut all the edges leaving the set \( B_\theta \): \( E' = \{ (v, v') \in E_G \mid v \in B_\theta, v' \notin B_\theta \} \).

Note that it is crucial that the same \( \theta \) is used for all \( u \).

Proving that \( E' \) is a feasible solution is easy. However, to bound the expected cost of the solution, we need the following lemma which shows that for any vertex \( v \),
Once we have this property, by linearity of expectation, edge $e$ the LP cost:

$$E_a(u,v) = \max(0,1 - \min_{v' \in V, v \neq v'} d(v,v'))$$

3: For all $u,v \in V$, compute $d_1(u,v) = \max(0,1 - \min_{v' \in V, v \neq v'} d(v,v'))$

4: Pick $\theta \in (0,1)$ uniformly at random

5: $B_u = \{ v \in V \mid d_1(u,v) \leq \theta \}$

6: $E' = \cup_{u \in V} \delta^+(B_u)$

7: Return $E'$

number of $u_i$ with different non-zero values of $d_1(u_i,v)$ is at most $k-1$.

**Lemma 4.2.** If for some $v \in V$ there exists $u_1, \ldots, u_k$ such that $0 \neq d_1(u_i,v) \neq d_1(u_j,v)$ for all $i \neq j$, then the demand graph $H$ contains an induced $k$-matching extension.

**Proof.** Rename the vertices $u_1, \ldots, u_k$ such that $d_1(u_1,v) > \cdots > d_1(u_k,v) > 0$. By Lemma 4.1, there exists $v_1', \ldots, v_k'$ such that $u_i v'_i \in E$ and $d_1(u_i,v_i) + d(v_i,v'_i) = 1$. Consider the subgraph of $H$ induced by the vertices $s_1, \ldots, s_k, t_1, \ldots, t_k$ where $s_i = u_i$, $t_i = v'_i$. Edge $(s_i,t_i) \in F$ as $(u_i,v'_i) \in F$. By construction $s_1, \ldots, s_k$ are distinct. We also argue that $t_1, \ldots, t_k$ are distinct. Suppose $t_i = t_j$, that is $v'_i = v'_j$ for $i < j$. Then, $d_1(u_j,v_j) + d(v_j,v'_j) < d_1(u_i,v_i) + d(v_i,v'_i) = d_1(u_i,v) + d(v_i,v'_i) = 1$. Since $u_j v'_j \in F$, by Lemma 4.1 $d_1(u_j,v) + d(v_j,v'_j) \geq 1$ which contradicts the inequality above.

For $i > j$, $d_1(s_i,v) + d(v,t_j) = d_1(u_j,v) + 1 - d_1(u_j,v) < 1$. By Lemma 4.1 $(s_i,t_j) \notin F$. Thus we have shown that $(s_i,t_i) \in F$ for $i \in [1,k]$ and $(s_i,t_j) \notin F$ for $i \geq j$. Thus $H$ contains an induced $k$-matching extension.

**Proof of Theorem 1.4.** We start by solving LP 1 and then perform the rounding scheme as per Algorithm 1.

By Lemma 4.1 for all $(u,v) \in F$, $d_1(u,v) \geq 1$ and since $\theta < 1$, we have $u \in B_u, v \notin B_u$. We remove all the edges going out of the set $B_u$ and hence, cut all the paths from $u$ to $v$. As argued above, for all $uv \in E_H$, $u \in B_u, v \notin B_u$ and we cut the edges going out of $B_u$. Hence, there is no path from $u$ to $v$ in $G - E'$ and $E'$ is a feasible DIR-MULC solution.

We claim that $\Pr[e \in E'] \leq (k-1)x_e$ for all $e \in E_G$. Once we have this property, by linearity of expectation, the expected cost of $E'$ can be bounded by $(k-1)$ times the LP cost: $E[\sum_{e \in E'} w_e] \leq (k-1) \sum_{e \in E_G} w_e x_e$.

Now we prove the preceding claim. Consider an edge $e = (a,b) \in E$. Edge $e \in E'$ only if for some $u \in V$, $e \in \delta^+(B_u)$ and this holds only if $\theta \in [d_1(u,a), d_1(u,b)]$. By Lemma 4.1, $d_1(u,b) \leq d_1(u,a) + x_e$. Hence, $e \in \delta^+(B_u)$, if $\theta \in [d_1(u,b) - x_{ab}, d_1(u,b)]$. Denote this interval by $I_u(e)$.

By Lemma 4.2 there are at most $k-1$ distinct elements in the set $\{d_1(u,b) \mid u \in V\}$. This implies that there are at most $k-1$ distinct intervals $I_u(e)$. In other words there exists $u_1, \ldots, u_r$, $r \leq k-1$ such that $\cup_{u \in V} I_u(e) = \cup_{i=1}^r I_{u_i}(e)$.

$$\Pr[(a,b) \in E'] \leq \Pr[\theta \in \cup_{i=1}^r I_{u_i}(e)] = \sum_{i=1}^r \Pr[\theta \in I_{u_i}(e)] \leq r x_e \leq (k-1)x_e.$$

Penultimate inequality follows from the fact that $I_{u_i}(e)$ has length $x_e$ and $\theta$ is chosen uniformly at random from $[0,1]$.

**5 Concluding Remarks and Open Problems**

Question 1 asks for the approximability and flow cut gap for UNDIR-SKEW-MULTICUT and DIR-SKEW-MULTICUT. In this paper, we proved that UNDIR-SKEW-MULTICUT admits a constant factor approximation. However, the question of determining the flow-cut gap is still open. Also, if this flow-cut gap turns out to be constant, we could ask the question about $tK2$-free demand graphs as well.

**Question 2.** Does there exists a function $f : N \rightarrow N$ such that the flow-cut gap for UNDIR-MULC with $tK2$-free demand graph is at most $f(t)$? If yes, what is the tightest possible bound on $f(t)$?

For DIR-SKEW-MULTICUT, we proved that flow-cut gap is a lower bound on hardness of approximation. However, determining the exact flow-cut gap for DIR-SKEW-MULTICUT is still an open question. If this flow-cut gap turns out to be constant, we could also ask about flow-cut gap when demand graph does not contain matching of size $k$ as an induced subgraph. This may also help answer Question 2 which asks for better approximation algorithms for such graphs.

**Question 3.** Does there exists a function $f : N \rightarrow N$ such that the flow-cut gap for UNDIR-MULC with $tK2$-free demand graph is at most $f(t)$? If yes, what is the tightest possible bound on $f(t)$?

For DIR-SKEW-MULTICUT, we proved that flow-cut gap is a lower bound on hardness of approximation. However, determining the exact flow-cut gap for DIR-SKEW-MULTICUT is still an open question. If this flow-cut gap turns out to be constant, we could also ask about flow-cut gap when demand graph does not contain matching of size $k$ as an induced subgraph. This may also help answer Question 2 which asks for better approximation algorithms for such graphs.

**Question 4.** Assuming $P \neq NP$, can we prove that DIR-MULC-D does not admit $k - \epsilon$ approximation algorithm if $H$ is a collection of $k$ disjoint directed edges?
[1] Ittai Abraham, Cyril Gavoille, Anupam Gupta, Ofer Neiman, and Kunal Talwar. Cops, robbers, and threatening skeletons: Padded decomposition for minor-free graphs. In Proceedings of the 46th Annual ACM Symposium on Theory of Computing, pages 79–88. ACM, 2014.

[2] Amit Agarwal, Noga Alon, and Moses S Charikar. Improved approximation for directed cut problems. In Proceedings of the thirty-ninth annual ACM symposium on Theory of computing, pages 671–680. ACM, 2007.

[3] Egon Balas and Chang Sung Yu. On graphs with polynomially solvable maximum-weight clique problem. Networks, 19(2):247–253, 1989.

[4] Chandra Chekuri, Sanjeev Khanna, Joseph Naor, and Pramod Viswanath. Delay-constrained unicast and the triangle-square cost. In Information Theory (ISIT), 2015 IEEE International Symposium on, pages 804–808. IEEE, 2015.

[5] Chandra Chekuri, Sudeep Kamath, Sreenar Kannan, and Pramod Viswanath. Approximate max-flow min-(multi) cut techniques. Discrete Algorithms, the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, pages 56(2):6, 2009.

[6] Jianer Chen, Liu Yang, Lu Songjian, O’sullivan Barry, and Igor Razgon. A fixed-parameter algorithm for the directed feedback vertex set problem. Journal of the ACM (JACM), 55(5), 2008.

[7] Michael Saks, Alex Samorodnitsky, and Leonid Zosin. A 2-approximation algorithm for the directed multiway cut problem. In Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016, pages 797–807, 2016.

[8] J Naor and L Zosin. A 2-approximation algorithm for the directed multiway cut problem. In 2013 IEEE 54th Annual Symposium on Foundations of Computer Science, pages 548–548. IEEE Computer Society, 1997.

[9] Ankit Sharma and Jan Vondrak. Multiway cut, pairwise realizable distributions, and descending thresholds. In Proceedings of the 46th Annual ACM Symposium on Theory of Computing, pages 724–733. ACM, 2014.

[10] David Steurer and Nisheeth Vishnoi. Connections between multi-cut and unique games. Technical Report TR09-125, 2009.

A Proof of Theorem 3.1

From Label-LP to Distance-LP: Let \((x, z)\) be a feasible solution to LABEL-LP for the given instance of \(G, H\). Consider a solution \(x'\) to DISTANCE-LP where we set \(x'_e = x_e\). We claim that \(x'\) is a feasible solution to DISTANCE-LP for \(G, H\). That is, for \((s_i, s_j) \in E_H\), and a path \(p\) from \(s_i\) to \(s_j\), we have \(\sum_{e \in p} x'_e \geq 1\).

Lemma A.1. For any edge \(e = (u, v) \in E_G\) and \(i \in \{1, \ldots, k\}\), \(x_e \geq \sum_{\sigma \in L, \sigma[i] = 1} z_{u, \sigma} - \sum_{\sigma \in L, \sigma[i] = 1} z_{v, \sigma} \).

Proof. Recall the interpretation of variables \(z_{e, \sigma_1, \sigma_2}\) as flow from set \(\Gamma_u = \{u_\sigma \mid \sigma \in L\}\) to \(\Gamma_v = \{v_\sigma \mid \sigma \in L\}\). Consider the following partition of \(\Gamma_u\) into \(\Gamma_u = \{u_\sigma \mid \sigma \in L, \sigma[i] = 1\}\) and \(\Gamma_u^2 = \{u_\sigma \mid \sigma \in L, \sigma[i] = 0\}\). Similarly, consider the partition of \(\Gamma_v\)
into $\Gamma_v^1$ and $\Gamma_v^2$. Amount of flow out of $\Gamma_u^1$ is equal to $\sum_{e \in L, \sigma[i]=1} z_{u,e}$, and amount of flow coming into $\Gamma_v^1$ is equal to $\sum_{e \in L, \sigma[i]=1} z_{v,e}$. Amount of flow from $\Gamma_u^1$ to $\Gamma_v^1$ is at most $\sum_{e \in L, \sigma[i]=1} z_{v,e}$. Hence, flow from $\Gamma_u^1$ to $\Gamma_v^2$ is at least $\sum_{e \in L, \sigma[i]=1} z_{u,e} - \sum_{e \in L, \sigma[i]=1} z_{v,e}$. For $u_{\tau_1} \in \Gamma_u^1$, $v_{\tau_2} \in \Gamma_v^2$, we have $\tau_1 \not\leq \tau_2$, and hence,

$$x'_e = x_e = \sum_{\sigma_1, \sigma_2 \in L: \sigma_1 \not\leq \sigma_2} z_{e,\sigma_1}\sigma_2 \geq \sum_{\sigma \in L: \sigma[i]=1} z_{u,\sigma} - \sum_{\sigma \in L: \sigma[i]=1} z_{v,\sigma}.$$  

Let $(s_i, s_j) \in E_H$. We prove that for any path $p$ from $s_i$ to $s_j$ in $G$ has $\sum_{e \in p} x'_e \geq 1$. Let the path $p$ be $s_i, a_1, \ldots, a_t, s_j$. Then, by Lemma A.1,

$$x_{s_i, a_1} + \sum_{t=1}^{t-1} x_{a_t, a_{t+1}} + x_{a_t, s_j} \geq \sum_{\sigma \in L: \sigma[i]=1} (z_{s_i, \sigma} - z_{a_1, \sigma}) + \sum_{t=1}^{t-1} (z_{a_t, \sigma} - z_{a_{t+1}, \sigma}) + (z_{a_t, \sigma} - z_{s_j, \sigma}) = \sum_{\sigma \in L: \sigma[i]=1} (z_{s_i, \sigma} - z_{s_j, \sigma})$$

LABEL-LP ensures that $z_{s_i, \sigma} = 0$ if $\sigma[i] = 0$ and $z_{s_j, \sigma} = 0$ if $\sigma[i] = 1$. Hence, $\sum_{\sigma \in L: \sigma[i]=1} z_{s_i, \sigma} = 1$ and $\sum_{\sigma \in L: \sigma[i]=1} z_{s_j, \sigma} = 0$. Hence the right hand side in the preceding inequality is 1.

**From Distance-LP to Label-LP:** Suppose $x$ is a feasible solution to DISTANCE-LP for the given instance $G, H$. We construct a solution $(x', z)$ for LABEL-LP such that $x'_e \leq x_e$ for all $e \in E_G$. The edge lengths given by $x$ induce shortest path distances in $G$ and we use $d(u, v)$ to denote this distance from $u$ to $v$. By adding dummy edges with zero cost as needed we can assume that $d(u, v) \leq 1$ for each vertex pair $(u, v)$. With this assumption in place we have that for any edge $e = (u, v)$ and any terminal $s_i$, $d(s_i, v) \leq d(s_i, u) + x_e$; hence $x_e \geq \max_{1 \leq i \leq k} (d(s_i, v) - d(s_i, u))$. We will in fact prove that $x'_e \leq \max_{1 \leq i \leq k} (d(s_i, v) - d(s_i, u))$.

We start by describing how to assign values to the variables $z_{u,\sigma}$. Recall that these induce values to the other variables if one is interested in a minimum cost solution. Let $d(u, v)$ denote the shortest distance from $u$ to $v$ in $G$ as per lengths $x_e$.

For a vertex $u$, consider the permutation $\pi^u : \{1, \ldots, k\} \to \{1, \ldots, k\}$ such that $d(s_{\pi^u(1)}, u) \leq \cdots \leq d(s_{\pi^u(k)}, u)$. In other words $\pi^u$ is an ordering of the terminals based on distance to $u$ (breaking ties arbitrarily). Define $\sigma^u_0, \ldots, \sigma^u_k$ as follows:

$$\sigma^u_i[j] = \begin{cases} 1 & j \in \{\pi^u(1), \ldots, \pi^u(i)\} \\ 0 & j \not\in \{\pi^u(1), \ldots, \pi^u(i)\} \end{cases}$$

In the assignment above it is useful to interpret $\sigma^u_i$ as a set of indices of the terminals. Hence $\sigma^u_0$ corresponds to $\emptyset$ and $\sigma^u_k$ to $\{\pi^u(1), \ldots, \pi^u(i)\}$. Thus, these sets form a chain with.

The assignment of values to the variables $z_{u,\sigma}$, $\sigma \in L$ is done as follows:

$$z_{u,\sigma} = \begin{cases} d(s_{\pi^u(1)}, u) & \sigma = \sigma^u_0 \\ d(s_{\pi^u(i+1)}, u) - d(s_{\pi^u(i)}, u) & \sigma = \sigma^u_i, i \in [1, k-1] \\ 1 - d(s_{\pi^u(k)}, u) & \sigma = \sigma^u_k \\ 0 & \text{otherwise} \end{cases}$$

**Lemma A.2.** $z_{u,\sigma}$ as defined above satisfy the following properties:

- $\forall u \in V_G, \sigma \in L, z_{u,\sigma} \geq 0$.
- $\forall u \in V_G, \sum_{\sigma \in L} z_{u,\sigma} = 1$.
- For $A \subseteq \{1, \ldots, k\}$, define $\sigma_A \in L$ as: $\sigma_A[i] = 1$ for $i \in A$ and 0 otherwise. Then,

$$\sum_{\sigma \geq \sigma_A} z_{u,\sigma} = 1 - \max_{i \in A} d(s_i, u)$$

- Terminals are labelled correctly. That is, for each $s_j$ and $\sigma \in L$, $z_{s_j, \sigma} = 0$ if $\sigma[j] = 0$.
- If $(s_i, s_j) \in E_H$, then $z_{s_j, \sigma} = 0$ for $\sigma \in L$ such that $\sigma[i] = 1$.

**Proof.** For $u \in V_G$, consider $\sigma^u_0, \sigma^u_1, \ldots, \sigma^u_k$ as defined above.

- $z_{u,\sigma} \geq 0$ is true by definition.
- By definition, $z_{u,\sigma} = 0$ if $\sigma \not\in \{\sigma^u_0, \ldots, \sigma^u_k\}$. Hence,

$$\sum_{\sigma \in L} z_{u,\sigma} = \sum_{i=0}^k z_{u,\sigma^u_i} = d(s_{\pi^u(1)}, u) + 1 - d(s_{\pi^u(k)}, u) + \sum_{i=1}^{k-1} d(s_{\pi^u(i+1)}, u) - d(s_{\pi^u(i)}, u) = 1$$

- Let $j = \arg \max_{i: \pi^u(i) \in A} d(s_{\pi^u(i)}, u)$. Then,
\[ \sigma_i^u, \ldots, \sigma_k^u \geq \sigma_A \text{ and } \sigma_i^v, \ldots, \sigma_j^v \not\geq \sigma_A. \] Hence,

\[
\sum_{\sigma \geq \sigma_A} z_{u, \sigma} = \sum_{i=0}^{k} z_{u, \sigma_i^u} = \sum_{i=j}^{k-1} d(s_{\sigma^v(i+1)}, u) - d(s_{\sigma^v(i)}, u) + 1 - d(s_{\sigma^v(k)}, u) = 1 - \max_{i \in A} d(s_{\sigma^v(i)}, u) = 1 - \max_{i \in A} d(s_i, u).
\]

- By definition of distance, \(d(s_j, s_j) = 0\). Consider \(A = \{j\}\). Applying the result from previous part, we get \(\sum_{\sigma \geq \sigma_A} z_{s_j, \sigma} = 1 - 0 = 1\). Hence, \(z_{s_j, \sigma} = 0\) if \(\sigma \not\geq \sigma_A\). Equivalently speaking, \(z_{s_j, \sigma} = 0\) if \(\sigma[j] = 0\).

- Let \((s_i, s_j) \in E_H\). Then, for the solution \(x\) to be feasible, we must have \(d(s_i, s_j) = 1\). Consider \(A = \{i\}\). Then, using result from previous part, we get \(\sum_{\sigma \geq \sigma_A} z_{s_j, \sigma} = 1 - 1 = 0\). Hence, \(z_{s_j, \sigma} = 0\) if \(\sigma \geq \sigma_A\). Equivalently speaking, \(z_{s_j, \sigma} = 0\) if \(\sigma[i] = 1\).

Consider an edge \(e = (u, v)\). Recall that once the distributions of \(\tilde{z}_u\) and \(\tilde{z}_v\) are fixed, \(x_e^i\) is simply the min-cost flow between these two distributions in the digraph \(B_{uv}\) with costs given by \(c\). Our goal is to show that this cost is at most \(\max\{0, \max_i (d(s_i, v) - d(s_i, u))\}\). Suppose we define a partial flow on zero-cost edges and have capacity left, we use \(\sigma_i^u\) if \(\sigma \not\geq \sigma_A\). Such \(\sigma_i^u\) is not saturated by the flow. If no such \(\ell\) exists then the greedy algorithm has sent a total flow of one unit on zero-cost edges and hence \(x_e^l = 0\). Thus, we can assume \(\ell\) exists. Moreover, in this case we can also assume that \(\ell < k\) for if \(\ell = k\) the greedy algorithm can send more flow since \(\sigma_i^u \leq \sigma_k^u\) for all \(i\).

Lemma A.3. The total flow sent by the greedy algorithm described is at least 1 - \(\max\{0, \max_i (d(s_i, v) - d(s_i, u))\}\).

Assuming the lemma we are done because the zero-cost flow is at least 1 - \(x_e\) and hence total cost of the flow is at most \(x_e\), thus proving \(x_e^i \leq x_e\) as desired. We now prove the lemma.

Consider the greedy flow. Let \(\ell\) be the maximum integer such that \(\nu_{\ell+1}^v = \sigma_{\ell}^v\). Such an \(\ell\) exists since \(\ell = 0\) is a candidate (corresponding to the empty set). Moreover, \(\ell < k\) since \(\sigma_k^v \not\geq \sigma_{\ell+1}^v\) since \(\ell < k\). Let \(\ell''\) be the minimum integer such that \(\sigma_{\ell''+1}^v \leq \sigma_{\ell''+1}^v\). \(\ell''\) exists by definition of \(\ell''\). We now claim several properties of the partial flow and justify them.

- \(\forall i \in [0, \ell'], j \in [\ell + 1, k], z_{e, \sigma_i^v} = 0\). This follows from the fact that the greedy algorithm did not saturate \(z_{e, \sigma_i^v}\).
- \(\forall i \in \ell' + 1, k, j \in [0, \ell' - 1], z_{e, \sigma_i^v} = 0\). From the definition of \(\ell', \ell''\), this is not a zero cost edge.
- \(\forall i \in [0, \ell'], \sum_{j=0}^{\ell} z_{e, \sigma_i^v} = z_{e, \sigma_i^v}\). From definition of \(\ell'\), for each \(i \leq \ell', \) there is a zero-cost edge from \(u_{\sigma_i^v}\) to \(\sigma_{\ell'}^v\). Since the greedy algorithm did not saturate \(\sigma_{\ell'}^v\), it means that \(u_{\sigma_i^v}\) is saturated and sends flow only to \(u_{\sigma_i^v}, \ldots, u_{\sigma_i^v}\).
- \(\forall j \in [\ell'', k], \sum_{i=\ell'+1}^{\ell} z_{e, \sigma_i^v} = z_{e, \sigma_j^v}\). By definition of \(\ell\), for \(j \geq \ell + 1\) we have the property that \(\sigma_j^v\) is saturated. As we argued above, for \(i \in [\ell'+1, k], j \in [0, \ell' - 1]\) we have \(z_{e, \sigma_i^v} = 0\). Hence, for \(j \geq \ell'' \geq \ell + 1\), we have \(\sum_{i=\ell'+1}^{\ell} z_{e, \sigma_i^v} = z_{e, \sigma_j^v}\).

From the preceding claim we see that the total value of the partial flow can be summed up as

\[
\sum_{\sigma_1, \sigma_2 \in \mathcal{L}} z_{e, \sigma_1 \sigma_2} = \sum_{i=0}^{\ell'} z_{u, \sigma_i^u} + \sum_{j=\ell'+1}^{k} z_{v, \sigma_j^v}.
\]
Moreover, by construction of $\tau_u$ and $\tau_v$,
\[
\sum_{i=0}^{k'} z_{u,\sigma_i} = d(s_{\pi_1(\ell+1)}, u).
\]
\[
\sum_{j=\ell'+1}^{k} z_{v,\sigma_j} = 1 - d(s_{\pi_1(\ell')}, v).
\]
Letting $h = \pi_{\ell+1} = \pi'_{\ell'}$, we see that from the preceding equalities that the total flow routed on the zero-cost edges is
\[
d(s_h, u) + 1 - d(s_h, v) = 1 - (d(s_h, v) - d(s_h, u)) \geq 1 - x_c.
\]
This finishes the proof.

**B Proof of Theorem 3.3**

The first two lemmas help establish that we can safely assume that the supply graph satisfies the assumptions I, II, and III. We omit the proof of the first lemma which involves tedious reworking of some of the details on equivalence of LABEL-LP and DISTANCE-LP.

**Lemma B.1.** For any instance $G, H$ of Dir-MULC-H where the supply graph has undirected edges, the optimum solution values for the formulations LABEL-LP and DISTANCE-LP are the same both in the fractional and integral settings.

Assuming the preceding lemma, following lemma is easy to prove:

**Lemma B.2.** For bipartite $H$, Dir-MULC-H with a general supply graph and Dir-MULC-H restricted to supply graphs satisfying Assumptions I, II and III are equivalent in terms of approximability and in terms of the integrality gap of DISTANCE-LP (equal to integrality gap of LABEL-LP).

**Proof.** We sketch the proof. Undirected edges can be handled by the gadget shown in Fig. 5. It is easy to see that given any instance of Dir-MULC-H with supply graph $G$ and bipartite demand graph $H$ we can first add dummy terminals to $G$ and assume that each terminal $a_i$ has only one outgoing infinite weight edge (to the original terminal) and each $b_j$ has only one incoming infinite weight edge. With this in place adding edges to satisfy Assumptions II and III can be seen to not affect the integral or fractional solutions to DISTANCE-LP.

We will assume for simplicity that all weights (for edges and constraints) are either 1 or $\infty$. Generic weights can be easily simulated by copies and the proofs make no essential use of weights other than that some are finite and others are infinite.

**B.1 Reduction from Min-$\beta_H$-CSP to Dir-MULC-H** Let the Min-$\beta$-CSP instance be $I_C = (V_C, T_C, \psi_{T_C} : T_C \rightarrow \beta_H, w_{T_C} : T_C \rightarrow R^+)$. We refer to tuple $t = (u)$ with $\psi_{T_C}(t) = \psi_{a_i}$ as constraint $\psi_{a_i}(u)$, $t = (u, v)$ with $\psi_{T_C}(t) = \psi_{b_j}$ as constraint $\psi_{b_j}(u, v)$. For each constraint $\psi_{a_i}$, $t = (u)$ with $\psi_{T_C}(t) = \psi_{b_j}$ as constraint $\psi_{b_j}(u, v)$.

We assume that for every $i \in [1, p]$, there is a constraint $\psi_{a_i}(u_i)$ for some vertex $u_i \in V_C$, and similarly for every $j \in [1, q]$ there is a constraint $\psi_{b_j}(v_j)$ for some vertex $v_j \in V_H$; moreover we will assume that $u_1, \ldots, u_p, v_1, \ldots, v_q$ are distinct vertices. One can ensure that this assumption holds by adding dummy vertices and dummy constraints with zero weight. We create an instance $I_M = (G = (V_G, E_G, w_G : E_G \rightarrow R^+)$. We now prove the equivalence of $I_C$ and $I_M$ from the point of view solutions to Basic-LP and Label-LP respectively.

Given two labels $\sigma$ and $\sigma'$ which can be interpreted as binary strings, we use the notation $\sigma \cdot \sigma'$ to denote the label obtained by concatenating $\sigma$ and $\sigma'$.

**From Label-LP to Basic-LP:** Suppose $(x, z)$ is a feasible solution to LABEL-LP for $I_M$. We construct a solution $z'$ to Basic-LP for $I_C$ in the following way. $z'$ is simply a projection of $z$ from label set $\{0, 1\}^p$ onto label set $\{0, 1\}^p$. Recall that in the instance $I_M$ the terminals $b_1, \ldots, b_q$ do not have any outgoing edges.
Hence, in the solution \((x, z)\) with label space \(\{0, 1\}^{p+q}\), which encodes reachability from both the \(a_i\)’s and the \(b_j\)’s, the information on reachability from the \(b_j\)’s does not play any essential role. We formalize this below.

- For \(v \in T_C, \sigma \in \{0, 1\}^p, z'_{v, \sigma} = \sum_{v' \in \{0, 1\}^q} z_{v', \sigma'}\).
- For unary constraint \(t = (v) \in T_C\) and \(\sigma \in \{0, 1\}^p, z'_{t, \sigma} = z'_{v, \sigma}\).
- For binary constraint \(t = (u, v) \in T_C, \sigma_1, \sigma_2 \in \{0, 1\}^p, z'_{t, \sigma_1, \sigma_2} = \sum_{\sigma' \in \{0, 1\}^q} \sum_{\sigma'' \in \{0, 1\}^q} z_{\sigma_1, \sigma_2, \sigma', \sigma''}\).

Note that if \((x, z)\) is an integral solution then \(z'\) as defined above is also an integral solution.

Feasibility of \(z'\) for Basic-LP is an “easy” consequence of the projection operation but we prove it formally.

**LEMMA B.3.** \(z'\) as defined above is a feasible solution to Basic-LP for instance \(I_C\).

**Proof.** From the definition of \(z'\), for each vertex \(v\),

\[
\sum_{\sigma \in \{0, 1\}^p} z'_{v, \sigma} = \sum_{\sigma \in \{0, 1\}^p} \sum_{\sigma' \in \{0, 1\}^q} z_{v, \sigma, \sigma'} = 1
\]

which proves that one set of constraints holds.

Next, we prove that for \(t \in T_C, v = t[i], \sigma \in L = \{0, 1\}^p\), the constraint \(z_{v, \sigma} - \sum_{\alpha \in L^{|\alpha|} : \alpha[i] = \sigma} z'_{t, \alpha} = 0\) holds. We consider unary and binary predicates separately.

- For \(t = (v) \) s.t. \(v = t[i], \sigma \in L = \{0, 1\}^p\),

\[
z'_{v, \sigma} - \sum_{\alpha \in L^{|\alpha|} : \alpha[i] = \sigma} z'_{t, \alpha} = z_{v, \sigma} - z'_{t, \sigma} = z'_{v, \sigma} - z'_{v, \sigma} = 0.
\]

- For \(t = (u, v) \in T_C, \sigma \in \{0, 1\}^p\)

\[
z'_{v, \sigma} - \sum_{\sigma_1 \in \{0, 1\}^p} z'_{t, \sigma_1} = z'_{v, \sigma} - \sum_{\sigma_1 \in \{0, 1\}^p} z_{v, \sigma_1, \sigma', \sigma''} = z'_{v, \sigma} - \sum_{\sigma' \in \{0, 1\}^q} \sum_{\sigma'' \in \{0, 1\}^q} z_{\sigma_1, \sigma', \sigma''} = 0.
\]

Similar argument holds for \(u\) as well. \(\square\)

**LEMMA B.4.** The cost of \(z'\) is at most \(\sum_{e \in E_G} w_e x_e\) which is the cost of \((x, z)\) to \(I_M\).

Before we prove Lemma B.4, we establish some properties satisfied by \((x, z)\).

**LEMMA B.5.** If the solution \((x, z)\) to Label-LP has finite cost, then the following conditions hold:

- For directed edge \(e = (u, v)\), and for \(i \in [1, p]\)

\[
x_e = \sum_{\sigma \in \{0, 1\}^p \cup \{\infty\}} z_{z, \sigma} - \sum_{\sigma \in \{0, 1\}^p \cup \{\infty\}} z_{z, \sigma} = \sum_{\sigma \in \{0, 1\}^p \cup \{\infty\}} z_{z, \sigma} - \sum_{\sigma \in \{0, 1\}^p \cup \{\infty\}} z_{z, \sigma}.
\]

Hence, if edge \(e\) has infinite weight \((w_e(c) = \infty)\), then

\[
\sum_{\sigma \in \{0, 1\}^p \cup \{\infty\}} z_{z, \sigma} \leq \sum_{\sigma \in \{0, 1\}^p \cup \{\infty\}} z_{z, \sigma}.
\]

- For \(i \in [1, p], \sigma \in \{0, 1\}^p, \sigma' \in \{0, 1\}^q\) s.t. \(J_\sigma \neq Y_i\), we have \(z_{a_i, \sigma, \sigma'} = 0\).

- For \(j \in [1, q]\), \(\sigma \in \{0, 1\}^p, \sigma' \in \{0, 1\}^q\) s.t. \(J_\sigma \neq Z_j\) we have \(z_{b_j, \sigma, \sigma'} = 0\).

- For an undirected edge \(e = uv \in E_G\) with \((w_e(c) = \infty)\), and \(\sigma_1, \sigma_2 \in \{0, 1\}^{p+q}, z_{u, \sigma_1, \sigma_2} = 0\) if \(\sigma_1 \neq \sigma_2\).

- For \(i \in [1, p], \sigma \in \{0, 1\}^p\), \(z_{u, \sigma, \sigma'} = 0\) and for \(i \in [1, p], \sigma' \in \{0, 1\}^q\), \(z_{u, \sigma, \sigma'} = 0\). Hence, for \(\sigma \in \{0, 1\}^p, z'_{u, \sigma, \sigma'} = 1\) if \(J_\sigma = Y_i\) and 0 otherwise.

- For \(j \in [1, q]\), \(\sigma \in \{0, 1\}^p, \sigma' \in \{0, 1\}^q\) s.t. \(J_\sigma \neq Z_j\) we have \(z_{b_j, \sigma, \sigma'} = 0\).

**Proof.** If \((x, z)\) has finite cost, then for an edge \(e\) with infinite weight \((w_G(c) = \infty)\), we must have \(x_e = 0\).

- Let \(e = (u, v)\) be a directed edge, and \(i \in [1, p]\)

\[
x_e = \sum_{\sigma_1, \sigma_2 \in \{0, 1\}^{p+q}, \sigma_1 \leq \sigma_2} z_{u, \sigma_1, \sigma_2} - \sum_{\sigma_1, \sigma_2 \in \{0, 1\}^{p+q}, \sigma_1 \leq \sigma_2} z_{u, \sigma_1, \sigma_2} = \sum_{\sigma_1, \sigma_2 \in \{0, 1\}^{p+q}, \sigma_1 \leq \sigma_2} z_{u, \sigma_1, \sigma_2} - \sum_{\sigma_1, \sigma_2 \in \{0, 1\}^{p+q}, \sigma_1 \leq \sigma_2} z_{v, \sigma_1, \sigma_2}.
\]

If edge \(e\) has infinite weight, then \(x_e = 0\) and

\[
\sum_{\sigma \in \{0, 1\}^{p+q}, \sigma[i] = 1} z_{u, \sigma} \leq \sum_{\sigma \in \{0, 1\}^{p+q}, \sigma[i] = 1} z_{v, \sigma}.
\]

- We prove the following two statements which in turn imply that for \(\sigma \in \{0, 1\}^p, \sigma' \in \{0, 1\}^q\), if \(J_\sigma \neq Y_i\), then \(z_{a_i, \sigma, \sigma'} = 0\).

\[
\forall j \in Y_i, \sum_{\sigma \in \{0, 1\}^{p+q}, \sigma[j] = 1} z_{a_i, \sigma, \sigma'} = 1,
\]

\[
\forall j \in [1, p] \setminus Y_i, \sum_{\sigma \in \{0, 1\}^{p+q}, \sigma[j] = 1} z_{a_i, \sigma, \sigma'} = 0.
\]
Let \( j \in Y_i \), then by construction of \( G \), there exists an infinite weight edge from \( a_j \) to \( a_i \). Using the result from previous part we get \( \sum_{\sigma \in \{0,1\}^p, \sigma' \in \{0,1\}^q: \sigma[j]=1} z_{a_i, \sigma} = \sum_{\sigma \in \{0,1\}^p, \sigma' \in \{0,1\}^q: \sigma[j]=1} z_{b_j, \sigma', \sigma} \geq 1 \). Label-LP enforces that term on the right side is lower bounded by 1 (\( a_j \) reachable from itself). Hence, term on the left side is lower bounded by 1. Since, it is also upper bounded by 1, it must be equal to 1.

Let \( j \in [1,p] \setminus Y_i \). By definition of \( Y_i \), we have \( N_H^+(a_j) \not\subseteq N_H^-(a_i) \). That is, there exists \( j' \in [1,q] \) such that \( a_j b_{j'} \in E_H \) and \( a_j b_{j'} \not\in E_H \). Since \( a_j b_{j'} \in E_H \), Label-LP enforces that
\[
\sum_{\sigma \in \{0,1\}^p, \sigma' \in \{0,1\}^q: \sigma[j]=1} z_{b_{j'}, \sigma', \sigma} = 0
\]
Also, we have \( a_j b_{j'} \not\in E_H \) and hence, there is an infinite weight edge from \( a_i \) to \( b_j \) in \( G \). Applying the result from previous part, we get \( \sum_{\sigma \in \{0,1\}^p, \sigma' \in \{0,1\}^q: \sigma[j]=1} z_{a, \sigma} = \sum_{\sigma \in \{0,1\}^p, \sigma' \in \{0,1\}^q: \sigma[j]=1} z_{b, \sigma', \sigma} = 0 \)

Next, to prove that \( z'_{a_j, \sigma} = 1 \) if \( J_\sigma = Y_i \) and 0 otherwise, we argue as follows:
\[
1 = \sum_{\sigma \in \{0,1\}^p, \sigma' \in \{0,1\}^q} z_{a_j, \sigma' \\
= \sum_{\sigma \in \{0,1\}^p: J_\sigma = Y_i, \sigma' \in \{0,1\}^q} z_{a_j, \sigma'} + \sum_{\sigma \in \{0,1\}^p: J_\sigma \neq Y_i, \sigma' \in \{0,1\}^q} z_{a_j, \sigma'} \\
= \sum_{\sigma \in \{0,1\}^p: J_\sigma = Y_i} z'_{a_j, \sigma}
\]

- For an undirected edge \( e = uv \), \( x_e = \sum_{\sigma \in \{0,1\}^p, \sigma' \in \{0,1\}^q} z_{u, \sigma} \cdot z_{v, \sigma'}. \) Since, weight of \( e \) is infinite, \( x_e \) must be 0. Hence, \( z_{u, \sigma} = 0 \) if \( \sigma \not\in \sigma_2 \). For \( \sigma_1 \in \{0,1\}^p+q \)
\[
z_{u, \sigma_1} = \sum_{\sigma_2 \in \{0,1\}^q \setminus J_\sigma} z_{u, \sigma_2} = z_{v, \sigma_1} = \sum_{\sigma_2 \in \{0,1\}^q \setminus J_\sigma} z_{v, \sigma_2}
\]

Let \( t = (u) \in T_C \) s.t. \( \psi_{t_C}(t) = \psi_{u}. \) If \( u = a_i \), then we have already proved that \( z'_{u, \sigma} = 1 \) if \( J_\sigma = Y_i \) and 0 otherwise. If \( u \neq a_i \), then there is an infinite weight undirected edge between \( u \) and \( a_i \) in \( G \). Hence, \( z'_{u, \sigma} = z'_{a_j, \sigma} \) for all \( \sigma \in \{0,1\}^p \) and the result follows.

\[
\square
\]

**Proof of Lemma B.4:** Next, we argue about the cost of the solution \( z' \). We assume here that \((x, z)\) has finite cost. For a constraint \( t \in T_C \), the cost according to \( z' \) is \( w_{t_C}(t) \sum_{\sigma \in L \setminus \{\alpha\}} z'_{\sigma, \alpha} \cdot \psi_{t}(\alpha) \). We consider four cases based on the type of \( t \).

- \( t \) corresponds to constraint of the form \( \psi_{u_1}(v) \). As argued in Lemma B.5, then \( z'_{u_1, \sigma} = z_{v, \sigma} = 0 \) if \( J_\sigma \not\subseteq Y_i \) and 1 if \( J_\sigma = Y_i \). On the other hand, \( \psi_{u_1}(\sigma) = 0 \) if \( J_\sigma = Y_i \) and \( \infty \) if \( J_\sigma \not\subseteq Y_i \). Hence, \( z'_{u_1, \sigma} \psi_{u_1}(\sigma) = 0 \) for all \( \sigma \). Therefore this constraint contributes zero to the cost.

- \( t \) corresponds to constraint of the form \( \psi_{b_j}(v) \). From Lemma B.3, \( z'_{b_j, \sigma} = z_{v, \sigma} = 0 \) if \( J_\sigma \neq Z_j \) and 1 if \( J_\sigma = Z_j \). And \( \psi_{b_j}(\sigma) = 0 \) if \( J_\sigma = Z_j \) and \( \infty \) if \( J_\sigma \neq Z_j \). Hence, \( z'_{b_j, \sigma} \psi_{b_j}(\sigma) = 0 \) for all \( \sigma \). Therefore the contribution of this constraint is zero.
• $t$ corresponds to constraint $C(u,v)$. This corresponds to a directed edge $e_t = (u,v)$ in $G$ and the cost paid by $(x, z)$ is $x_{e_t}$. The cost for $t$ in $z'$ is given by:

$$\sum_{\sigma_1, \sigma_2 \in \{0,1\}^p} z'_{t,\sigma_1,\sigma_2} \cdot C(\sigma_1, \sigma_2)$$

$$= \sum_{\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_1 \not\in \sigma_2} z'_{t,\sigma_1,\sigma_2}$$

$$= \sum_{\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_1 \not\in \sigma_2} \sum_{\sigma' \in \{0,1\}^q} z_{e_t,\sigma_1,\sigma_2}$$

$$\leq \sum_{\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_1 \not\in \sigma_2} \sum_{\sigma' \in \{0,1\}^q, \sigma_1 \not\in \sigma_2} z_{e_t,\sigma_1,\sigma_2}$$

$$= x_{e_t}.$$  

First equality follows from the fact that $C(\sigma_1, \sigma_2) = 0$ if $\sigma_1 \leq \sigma_2$ and 1 otherwise. Penultimate inequality follows because if $\sigma_1 \not\in \sigma_2$, then $\sigma_1 \cdot \sigma' \not\in \sigma_2 \cdot \sigma''$ for any $\sigma', \sigma'' \in \{0,1\}^q$.

• $t$ corresponds to constraint NAE2$(u,v)$. This corresponds to an undirected edge $e_t = uv$ in $G$ and the cost paid by $(x, z)$ is $x_{e_t}$. The cost for $t$ in $z'$ is given by:

$$\sum_{\sigma_1, \sigma_2 \in \{0,1\}^p} z'_{t,\sigma_1,\sigma_2} \cdot \text{NAE2}(\sigma_1, \sigma_2)$$

$$= \sum_{\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_1 \not\in \sigma_2} z'_{t,\sigma_1,\sigma_2}$$

$$= \sum_{\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_1 \not\in \sigma_2} \sum_{\sigma' \in \{0,1\}^q} z_{e_t,\sigma_1,\sigma_2}$$

$$\leq \sum_{\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_1 \not\in \sigma_2} \sum_{\sigma' \in \{0,1\}^q, \sigma_1 \not\in \sigma_2} z_{e_t,\sigma_1,\sigma_2}$$

$$= x_{e_t}.$$  

Combining the four cases, the total cost of the solution $z'$ is equal to the cost of the binary constraints each of which corresponds to an edge in $G$ with the same weight. From the above inequalities we see that the cost is at most $\sum_{e \in E_G} w_G(e) x_e$ which is the cost of $(x, z)$. □

**From Basic-LP to Label-LP:** Let $z$ be a Basic-LP solution to $I_G$. Let $x_0 = 1^q$. We define a solution $(x', z')$ to LABEL-LP for $I_M$ as follows:

• For $v \in V_C, \forall \sigma_1 \in \{0,1\}^p, \sigma_2 \in \{0,1\}^q$,

$$z'_{v,\sigma_1,\sigma_2} = \begin{cases} z_{v,\sigma_1} & \sigma_2 = \sigma_0 \\ 0 & \text{otherwise} \end{cases}$$

• For unary constraint $t = (u)$ s.t. $u \not\in \{a_1, \ldots, a_p, b_1, \ldots, b_q\}$ and $\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_3, \sigma_4 \in \{0,1\}^q$,

$$z'_{e_t,\sigma_1,\sigma_2,\sigma_3,\sigma_4} = \begin{cases} z_{u,\sigma_1} & \sigma_1 = \sigma_2, \sigma_3 = \sigma_4 = \sigma_0 \\ 0 & \text{otherwise} \end{cases}$$

• For binary constraint $t = (u,v) \in T_C$ such that $\psi_{T_C}(t) = C$ or NAE2, and $\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_3, \sigma_4 \in \{0,1\}^q$,

$$z'_{e_t,\sigma_1,\sigma_2,\sigma_3,\sigma_4} = \begin{cases} z_{e_t,\sigma_1,\sigma_2} & \sigma_3 = \sigma_4 = \sigma_0 \\ 0 & \text{otherwise} \end{cases}$$

• The edge variables $z'_e$ are induced by the $z'$ variables. We explicitly write them down. For directed edge $e \in E_G$, $z'_e = \sum_{\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_3, \sigma_4 \in \{0,1\}^q} z_{e_t,\sigma_1,\sigma_2,\sigma_3,\sigma_4}.$ For undirected edge $e \in E_G$, $z'_e = \sum_{\sigma_1, \sigma_2 \in \{0,1\}^p, \sigma_3, \sigma_4 \in \{0,1\}^q} z_{e_t,\sigma_1,\sigma_2,\sigma_3,\sigma_4}.$

It is easy to check that $(x', z')$ is integral if $z$ is integral.

**Lemma B.6.** $(x', z')$ is a feasible solution to LABEL-LP for $I_M$.

**Proof.** It is easy to check that all the variables are non-negative and upper bounded by 1.

We show that the other constraints are satisfied one at a time. Recall that LABEL-LP considered here has a constraint for undirected edges in addition to the constraints showed in Fig 3. The label set for LABEL-LP is $\{0,1\}^{p+q}$ which we can write as $\{\sigma_1, \sigma_2 | \sigma_1 \in \{0,1\}^p, \sigma_2 \in \{0,1\}^q\}$.

**Constraint 1:** For each $v \in V_C$, $\sum_{\sigma_1 \in \{0,1\}^p} z'_{v,\sigma_1,\sigma_0} = 1$

$$\sum_{\sigma_1 \in \{0,1\}^p} z'_{v,\sigma_1,\sigma_0} = \sum_{\sigma_1 \in \{0,1\}^p} z_{v,\sigma_1} = 1$$

**Constraint 2:** For $\sigma_1 \in \{0,1\}^p, \sigma_2 \in \{0,1\}^q, z'_{v,\sigma_1,\sigma_2} = 0$ if $\sigma_1[i] = 0$. And $z'_{v,\sigma_1,\sigma_2} = 0$ if $\sigma_2[j] = 0$. There is $t = (a_i) \in T_C$ such that $\psi_{T_C}(t) = \psi_{a_i}$. For $z$ to be a finite valued solution, we must have $z_{a_i,\sigma_1} = 0$. Since, $i \not\in Y$, we have that $z_{a_i,\sigma_1} = 0$ if $\sigma_1[i] = 0$. And hence, $z'_{v,\sigma_1,\sigma_2} = 0$ if $\sigma_1[i] = 0$. For $v \in V_C, z'_{v,\sigma_1,\sigma_2} = 0$ if $\sigma_2[j] = 0$. Hence, $z'_{v,\sigma_1,\sigma_2} = 0$ if $\sigma_2[j] = 0$. In particular, $z'_{a_i,\sigma_1,\sigma_2} = 0$ if $\sigma_2[j] = 0$.

**Constraint 3:** For $e = (u,v) \in E_G, \sigma_1 \in \{0,1\}^p, \sigma_3 \in \{0,1\}^q, z'_{e_t,\sigma_1,\sigma_2,\sigma_3,\sigma_4} = \sum_{\sigma_2 \in \{0,1\}^p} z_{e_t,\sigma_1,\sigma_2,\sigma_3,\sigma_4}.$ If $\sigma_3 \neq \sigma_0$, then all the terms are zero and hence, the equality holds. Else, $\sigma_3 = \sigma_0$ and there are two types of edges:
For $t = (u,v), e = e_t$, 
\begin{align*}
\sum_{\sigma_2 \in \{0,1\}^q, \sigma_4 \in \{0,1\}^q} z'_{e,\sigma_1,\sigma_2,\sigma_4} = \\
= \sum_{\sigma_2 \in \{0,1\}^q} z'_{e,\sigma_1,\sigma_0,\sigma_2,\sigma_4} = \\
= z_{u,\sigma_1} = z'_{u,\sigma_1,\sigma_0} = z'_{u,\sigma_1,\sigma_3}.
\end{align*}

For $t = (u,v), e = e_t$, 
\begin{align*}
\sum_{\sigma_2 \in \{0,1\}^q, \sigma_4 \in \{0,1\}^q} z'_{e,\sigma_1,\sigma_2,\sigma_4} = \\
= \sum_{\sigma_2 \in \{0,1\}^q} z'_{e,\sigma_1,\sigma_0,\sigma_2,\sigma_4} = \\
= \sum_{\sigma_2 \in \{0,1\}^q} z_{e,\sigma_1,\sigma_2} = \\
= z_{u,\sigma_1} = z'_{u,\sigma_1,\sigma_0} = z'_{u,\sigma_1,\sigma_3}.
\end{align*}

Constraint 4: For $e = (u,v) \in E_G, \sigma_2 \in \{0,1\}^q, \sigma_4 \in \{0,1\}^q$, \[ z'_{e,\sigma_1,\sigma_2,\sigma_4} = \sum_{\sigma_2 \in \{0,1\}^q, \sigma_4 \in \{0,1\}^q} z_{e,\sigma_1,\sigma_2,\sigma_4}. \] Proof is similar to the previous part.

Constraint 5: For directed edge $e$, \[ x'_e = \sum_{\sigma_1,\sigma_2 \in \{0,1\}^q, \sigma_3,\sigma_4 \in \{0,1\}^q, \sigma_3 \neq \sigma_2, \sigma_4} z'_{e,\sigma_1,\sigma_3,\sigma_4} = 0. \] This is true by definition of $x'_e$.

Constraint 6: For undirected edge $e$, \[ x'_e = \sum_{\sigma_1,\sigma_2 \in \{0,1\}^q, \sigma_3,\sigma_4 \in \{0,1\}^q, \sigma_3 \neq \sigma_2, \sigma_4} z'_{e,\sigma_1,\sigma_3,\sigma_4} = 0. \] This is true as well from the definition of $x'_e$.

\[ \square \]

**Lemma B.7.** The cost $(x', z')$ is upper bounded by the cost of $z$.

**Proof.** Recall that $\sigma_0 = 1^q$. We consider three cases based on the type of edge $e$

- $e = e_t = (u,v)$ for constraint $C(u,v)$. Then, \[ x'_{e_t} = \sum_{\sigma_1,\sigma_2 \in \{0,1\}^q, \sigma_3,\sigma_4 \in \{0,1\}^q, \sigma_3 \neq \sigma_2, \sigma_4} z'_{e_t,\sigma_1,\sigma_3,\sigma_4} = \sum_{\sigma_1,\sigma_2 \in \{0,1\}^q, \sigma_0 \neq \sigma_2} z'_{e_t,\sigma_1,\sigma_0,\sigma_2} = \sum_{\sigma_1,\sigma_2 \in \{0,1\}^q} z_{e_t,\sigma_1,\sigma_2}. \] For this case, $x'_{e_t} = x_{e_t}$.

- $e = e_t = (u,v)$ for constraint $NAE_2(u,v)$. Then, \[ x'_{e_t} = \sum_{\sigma_1,\sigma_2 \in \{0,1\}^q, \sigma_3,\sigma_4 \in \{0,1\}^q, \sigma_3 \neq \sigma_2, \sigma_4} z'_{e_t,\sigma_1,\sigma_3,\sigma_4} = \sum_{\sigma_1,\sigma_2 \in \{0,1\}^q, \sigma_0 \neq \sigma_2} z'_{e_t,\sigma_1,\sigma_0,\sigma_2} = \sum_{\sigma_1,\sigma_2 \in \{0,1\}^q} z_{e_t,\sigma_1,\sigma_2}. \]

Combining the above three facts we get the following. First, infinite any infinite weight edge $e$ in $G$ has $x'_e = 0$. For any finite weight edge $x'_e$ is the same as the fractional cost paid by the corresponding finite weight binary constraint in $I_C$. Hence, cost of $(x', z')$ is upper bounded by cost of $z$. \[ \square \]

**B.2 Reduction from Dir-MulC-H to Min-β-CSP** Let the Dir-MulC-H instance be $I_M = (G = (V_G, E_G, w_G : E_G \rightarrow R^+), (S \cup T, E_H))$. Recall that the supply graph satisfies assumptions I, II, and III. We reduce it an equivalent $\text{Min-β-CSP}$ instance $I_C = (V_C, T_C, \psi_{T_C} : T_C \rightarrow \beta_H, w_{T_C} : T_C \rightarrow R^+)$ as follows.

1. Vertex Set $V_C = V_G$.
2. $T_C, \psi_{T_C}, w_{T_C}$ are defined as follows:
   - For every $a_i \in S$, add a tuple $t = (a_i)$ in $T_C$ with $\psi_{T_C}(t) = \psi_{a_i}$ and $w_{T_C}(t) = 1$.
   - For every $b_j \in T$, add a tuple $t = (b_j)$ in $T_C$ with $\psi_{T_C}(t) = \psi_{b_j}$ and $w_{T_C}(t) = 1$.
   - For every directed edge $e = (u,v) \in E_G$, add a tuple $t = (u,v)$ in $T_C$ with $\psi_{T_C}(t) = C$ and $w_{T_C}(t) = w_G(e)$.
   - For every undirected edge $e = w \in E_G$, add a tuple $t = (u,v)$ in $T_C$ with $\psi_{T_C}(t) = \text{NAE}_2$ and $w_{T_C}(t) = w_G(e)$.

The proof of equivalence between LABEL-LP for $I_M$ and BASIC-LP for $I_C$ is essentially identical to the proof for the reduction in the other direction and hence we omit it.

This finishes the proof of Theorem 3.3.