No.435

A New Scaling Algorithm for the Maximum Mean Cut Problem

by

Kazuo Iwano Shinji Misono Shu Tezuka Satoru Fujishige

August 1990

A New Scaling Algorithm for the Maximum Mean Cut Problem

Kazuo Iwano* Shinji Misono* Shu Tezuka* Satoru Fujishige†

July 23, 1990

Abstract

In this paper, we present a new scaling algorithm for the maximum mean cut problem. The *mean* of a cut is defined by the cut capacity divided by the number of edges crossing the cut. The algorithm uses an approximate binary search and solves the circulation feasibility problem with relaxed capacity bounds. The maximum mean cut problem has recently been studied as a dual analogue of the minimum mean cycle problem in the framework of the minimum cost flow problem by Ervolina and McCormick. A graph G with lower and upper edge capacities is said to be *ϵ-feasible* if G has a feasible circulation when we relax the capacity bounds by ϵ; that is, we use $(lower(e)-\epsilon, upper(e)+\epsilon)$ bounds instead of $(lower(e), upper(e))$ bounds for each edge $e \in E$. During an approximate binary search we maintain two bounds, LB and UB, such that G is LB-infeasible and UB-feasible, and we reduce the interval size (LB, UB) by at least one third at each iteration. For a graph with n vertices, m edges, and integer capacities bounded by U, the running time of this algorithm is $O(\min\{m^2, mn+n^2\log n, mn\log\log n\}\log(nU))$. This time bound is better than or comparable to the time achieved by McCormick and Ervolina under the *similarity condition* (that is, $U = O(n^{O(1)})$). Our algorithm can be naturally used for the circulation feasibility problem, and thus provides a new scaling algorithm for the minimum cut problem.

1. Introduction

In this paper, we present a new scaling algorithm for the maximum mean cut problem for a graph G with n vertices, m edges, and two lower and upper integer capacity bounds (respectively denoted by $lower(e)$ and $upper(e)$ for an edge $e \in E$). The *mean* of a cut is defined by the cut capacity divided by the number of edges crossing the cut. The maximum mean cut problem has recently been studied as a dual analogue of the minimum mean cycle problem in the framework of the minimum cost flow problem [4,5,9,15]. A maximum mean

*Tokyo Research Laboratory, IBM Japan, 5-11 Sanbancho, Chiyoda-ku, Tokyo 102, Japan
†Institute of Socio-Economic Planning, University of Tsukuba, Tsukuba, Ibaraki 305, Japan

cut canceling algorithm was proposed by Ervolina and McCormick [4], and is related to Goldberg and Tarjan's minimum mean cycle canceling algorithm [7] for the minimum cost flow problem.

A graph G is said to be *ϵ-feasible* if G has a feasible circulation when we relax the capacity bounds by ϵ; that is, we use $(lower(e)-\epsilon, upper(e)+\epsilon)$ bounds instead of $(lower(e), upper(e))$ bounds for each edge $e \in E$. Let U be the maximum absolute value of capacities, and $MF(n, m, U)$ indicate the running time of a maximum flow algorithm for a graph with n vertices, m edges, and integer capacities whose absolute values are at most U.

Our first algorithm uses a binary search and solves the circulation feasibility problem with relaxed capacity bounds. During a binary search we maintain two bounds, LB and UB, such that G is LB-infeasible and UB-feasible, and we reduce the interval size (LB, UB) by a half at each iteration until the size becomes smaller than $1/m^2$. Since the interval size eventually becomes sufficiently small, a mean cut value in the interval is unique. Thus, the time-complexity of this algorithm is $O(MF(n, m, U)\log(nU))$.

We then propose a new scaling algorithm using an approximate binary search, which enables us to reduce the number of circulation feasibility problems solved in the algorithm from $O(\log(nU))$ to essentially one. The algorithm also maintains two bounds, LB and UB, like the first one. At each iteration, given a UB-feasible circulation, the algorithm either obtains an $(LB+2UB)/3$-feasible circulation or detects $(2LB+UB)/3$-infeasibility, while the first algorithm checks whether G is $(UB+LB)/2$ feasible at each iteration. In this way, we can reduce the size of an interval (LB, UB) by one third at each iteration, and can compute each iteration efficiently, as we will discuss later. The running time of the second algorithm is $O(\min\{m^2, MF(n, m, m^2)\}\log(nU))$, which is $O(\min\{m^2, mn + n^2\log n, mn\log\log n\}\log(nU))$ using Ahuja and Orlin's $O(mn + n^2\log U)$ time maximum flow algorithm [1] or Ahuja, Orlin, and Tarjan's $O(mn\log(\frac{n\log U}{m\log\log U} + 2))$ time algorithm [2]. This time bound is better than or comparable to McCormick and Ervolina's [15] $O(MF(n, m, U)\min\{m, \log(nU)\})$-time bound under the *similarity condition* (that is, $U = O(n^{O(1)})$). Our algorithm can be naturally used for the circulation feasibility problem, and thus provides a new scaling method for the minimum cut problem.

The remainder of the paper consists of four sections. In Section 2 we discuss previous work related to the maximum mean cut problem. In Section 3 we develop the first algorithm, which uses a binary search. In Section 4 we discuss the second algorithm, which uses an approximate binary search, and explore some practical considerations. In Section 5 we conclude the paper by stating the remaining open problems.

2 Previous work

The maximum mean cut problem has recently been studied as a dual concept of the minimum mean cycle problem in the framework of the minimum cost flow problem [4,5,9,15].

Ervolina and McCormick [4] first proposed a polynomial time maximum mean cut canceling algorithm, which can be regarded as a dual analogue of Goldberg and Tarjan's minimum mean cycle canceling algorithm [7]. Hassin [9] showed that a maximum mean cut canceling algorithm may take an exponential number of cancelations according to his definition of a *cut mean*, a cut value divided by its vertex cardinality. Note that Hassin's definition is different from that of Ervolina and McCormick [4], according to which a cut mean is a cut value divided by the number of edges crossing the cut.

Ervolina and McCormick's polynomial time algorithm for the maximum mean cut canceling algorithm is based on a polynomial time algorithm for the maximum mean cut problem [15]. Their polynomial time maximum mean cut algorithm uses the fact that the maximum mean cut value is identical to the minimum ϵ that G is ϵ-feasible [9,15]. Therefore, assuming that G is 0-infeasible, their proposed algorithm starts from $\epsilon = 0$ and repeatedly increases ϵ by solving the maximum flow problems at most $O(\min\{m, \log(nU)\})$ times until G becomes ϵ-feasible. They also suggested another approach that starts from any ϵ-feasible circulation and repeatedly decreases ϵ until G becomes ϵ-infeasible. However, they showed that the latter algorithm does not necessarily terminate in finite time. In our terms, McCormick and Ervolina's approach is based on successive improvements of only LB's or only UB's. In contrast, we will show that successive improvements of both LB and UB can lead us to an efficient algorithm.

When we regard the maximum mean cut problem as a dual analogue of the minimum mean cycle problem, we can associate algorithms in this paper with Orlin and Ahuja's scaling algorithms for the assignment and minimum cycle mean problems [18]. Orlin and Ahuja first devised a scaling algorithm for the assignment problem, and then applied it to the minimum mean cycle problem. They showed that an approximate binary search [19] reduces the number of computations of assignment problems from $O(\log(nU))$ to essentially one. In our algorithm, we can associate the relationship between the assignment problem and the minimum mean cycle problem with the relationship between the circulation feasibility problem and the maximum mean cut problem. Although McCormick and Ervolina suggested that a dual version of Karp's Theorem [13] for the minimum cycle mean may lead to a new algorithm for the maximum mean cut problem, such a theorem has not yet been formulated, as far as we know.

The circulation feasibility problem can be computed by solving the associated maximum flow problem [8,14]. Our algorithm consists of subroutines, each of which produces an $(LB + 2UB)/3$-feasible circulation given a UB-feasible circulation, if one exists. Although there are scaling algorithms for the maximum flow problems [1,2,3], none of them seems to give such a scaling algorithm. In this sense, our algorithm can provide a new scaling method for the circulation feasibility problem.

Herz [11] devised an algorithm for the circulation feasibility problem which repeatedly cancels a cycle and reduces ϵ by at least one as far as G is ϵ-feasible. Since each cycle cancelation takes $O(m)$-time, his algorithm runs in $O(mU)$ time, which is psuedo-polynomial. Our algorithm chooses an appropriate set of cycles to cancel and reduces the total number of cycle cancelations to the polynomial number in n, m, and $\log U$.

3 Generic Algorithm

In this section, we introduce basic terminology and relevant theorems, and then devise a generic algorithm for computing a maximum mean cut.

From now on, let $G = (V, E)$ be a connected directed graph with n vertices and m edges. Each edge $e \in E$ has lower and upper capacities (denoted by $lower(e)$ and $upper(e)$, respectively) such that $lower(e) \in Z \cup \{-\infty\}$, $upper(e) \in Z \cup \{+\infty\}$, and $lower(e) \leq upper(e)$. A cycle C is a sequence of vertices $u_0, u_1, \ldots, u_k$ such that $(u_j, u_{j+1}) \in E$ or $(u_{j+1}, u_j) \in E$ for each $j \in \{1, 2, \ldots, k\}$, where we interpret $k+1$ as 0. A cycle is said to be *elementary* if all vertices in the cycle are different. A subset $X \subseteq E$, such that $G' = (V, E - X)$ contains more components than G, is a *separating set* of G. A minimal separating set is a *cocycle* of G. A *cut* is an ordered pair of complementary non-empty subsets S and T of the vertex set V, denoted by $[S, T]$. For any cocycle X, there exists a cut $[S, T]$ such that $X = \{S, T\} (= (S, T) \cup (T, S))$, where (S, T) denotes a set of edges (u, v) with $u \in S$ and $v \in T$. The *value* of a cut $[S, T]$, denoted by $val(S, T)$, is defined as follows:

$$val(S, T) = \sum_{e \in (S,T)} lower(e) - \sum_{e \in (T,S)} upper(e). \tag{1}$$

An edge e *crosses a cut* $[S, T]$ when $e \in (S, T) \cup (T, S)$. We call a cut value divided by the number of the edges across the cut the *cut mean*, denoted by $mean(S, T)$; that is,

$$mean(S, T) = \frac{val(S, T)}{|\{S, T\}|}, \tag{2}$$

where $|A|$ denotes the size of a set A. A cut is called a *maximum mean cut* if its mean is the maximum among all cut means. We have the following lemma regarding two distinct cut means.

Lemma 3.1 *Let $[S,T]$ and $[S',T']$ be two cuts with different cut means. If G has integral capacities, then $|\, mean(S,T) - mean(S',T')\, | \geq 1/m^2$.*

Proof. Since all capacities are integral,

$$\left| \frac{val(S,T)}{|\{S,T\}|} - \frac{val(S',T')}{|\{S',T'\}|} \right| = \left| \frac{val(S,T)|\{S',T'\}| - val(S',T')|\{S,T\}|}{|\{S,T\}||\{S',T'\}|} \right| \geq 1/m^2.$$

□

From the following lemma about two successive terms of the Farey series of order m, the ascending series of irreducible fractions between 0 and 1 whose denominators do not exceed m, we know that the above lower bound $1/m^2$ is tight when constant factors are ignored.

Lemma 3.2 (Theorem 28 in [10]) *If h/k and h'/k' are two successive terms of the Farey series, then $kh' - hk' = 1$.* □

Denote the value of a maximum mean cut of G by $maxmean(G)$. We can now define the *maximum mean cut problem* as finding a maximum mean cut when given a digraph G as above. A *circulation* x is a real-valued function defined on E such that $\sum_{(u,v)\in E} x(u,v) = \sum_{(v,u)\in E} x(v,u)$ for every $v \in V$. That is, a circulation satisfies the *flow conservation rule* at each vertex. Let ϵ be a non-negative constant. We say that a circulation x is *ϵ-feasible* when x satisfies the following:

$$lower(e) - \epsilon \leq x(e) \leq upper(e) + \epsilon \tag{3}$$

for each edge $e \in E$. In particular, we call a 0-feasible circulation *feasible*. A digraph G with given lower and upper capacities is said to be *ϵ-feasible* if there exists an ϵ-feasible circulation. For the necessary and sufficient condition for the circulation feasibility, we have the following theorem:

Theorem 3.3 (Hoffman's Theorem, [12]) *Given a graph $G = (V,E)$ and two numbers $lower(e)$ and $upper(e)$ for each edge e, a feasible circulation x such that $lower(e) \leq x(e) \leq upper(e)$ for every $e \in E$ exists if and only if*

$$\sum_{e\in(S,T)} upper(e) - \sum_{e\in(T,S)} lower(e) \geq 0 \tag{4}$$

for every cut $[S,T]$. □

Note that Equation 4 can be expressed as $val_\epsilon(T,S) \equiv val(T,S) - \epsilon|\{S,T\}| \leq 0$ when we use relaxed bounds $(lower(e)-\epsilon, upper(e)+\epsilon)$.

For a digraph G and ϵ, we define a digraph $M(G,\epsilon,lower,upper) = (V_M, E_M)$ as follows. $V_M = V \cup \{s,t\}$ and $E_M = E \cup \{(s,u) \mid u \in V\} \cup \{(u,t) \mid u \in V\}$, and each edge e has a capacity $cap(e)$ as follows: for every $e \in E$,

$$cap(e) = (upper(e)+\epsilon) - (lower(e)-\epsilon)$$

and for every $u \in V$,

$$cap(s,u) = \sum_{(v,u)\in E} (lower(v,u)-\epsilon) \text{ and } cap(u,t) = \sum_{(u,v)\in E} (lower(u,v)-\epsilon).$$

The following lemma is well-known.

Lemma 3.4 *A graph G is ϵ-feasible if and only if a maximum flow in $M(G,\epsilon,lower,upper)$ from s to t saturates all outgoing edges from s.* □

Let $[S',T']$ be an arbitrary cut in $M(G,\epsilon,lower,upper)$ such that $s \in S'$ and $t \in T'$. Then we can naturally associate the cut $[S',T']$ with a cut $[S,T]$ in G such that $S = S'-\{s\}$ and $T = T'-\{t\}$. McCormick and Ervolina showed the following relationship between values of two associated cuts.

Lemma 3.5 ([15]) *Let $[S',T']$ be a cut in $M(G,0,lower,upper)$ and $[S,T]$ be the cut in G associated with $[S',T']$. Then we have $cap(S',T') = \sum_{e\in E} lower(e) - val(T,S)$.* □

We define $\epsilon^*(G)$ as the minimum value ϵ such that G is ϵ-feasible; that is,

$$\epsilon^*(G) = \min\{\epsilon \mid G \text{ is } \epsilon\text{-feasible}\}. \tag{5}$$

In this paper we assume that G is not feasible; that is, $\epsilon^*(G) > 0$. Let U be the maximum absolute values of edge capacities plus one, that is,

$$U = \max\{\max_{e\in E}|lower(e)|, \max_{e\in E}|upper(e)|\} + 1.$$

We now have the following lemma:

Procedure MAX-MEAN-CUT-I
begin
(1) $LB = 0, UB = U$;
(2) **do until** ($UB - LB < 1/m^2$);
(2.1) $\epsilon = (LB + UB)/2$;
(2.2) If FEASIBLE(G, ϵ)=*yes*
(2.3) **then** $UB = \epsilon$;
(2.4) else $LB = \epsilon$;
end
(3) $\epsilon^*(G) = mean(T_{LB}, S_{LB})$;
end

Figure 1: Algorithm MAX-MEAN-CUT-I

Lemma 3.6 *G is U-feasible.*

Proof. Since $lower(e) - U \leq 0 \leq upper(e) + U$ for each edge $e \in E$, a 0-circulation is U-feasible. □

McCormick and Ervolina [15] and Hassin [9] obtained the following theorem and lemma, which characterize a maximum mean cut.

Theorem 3.7 ([9,15]) *If G is not feasible, we have $\epsilon^*(G) = maxmean(G)$.* □

Corollary 3.8 ([15]) *A cut $[S,T]$ is a maximum mean cut if and only if G is $mean(S,T)$-feasible.* □

From Lemmas 3.1 and 3.6 and Theorem 3.7, we can naturally obtain an algorithm MAX-MEAN-CUT-I for computing a maximum mean cut, as shown in Figure 1. This algorithm maintains two bounds, UB and LB, such that G is UB-feasible and LB-infeasible. Note that initially G is U-feasible and 0-infeasible. The algorithm uses a binary search to reduce the size of an interval (LB, UB) by a half at each iteration. In each iteration FEASIBLE(G, ϵ) at Step (2.2), which is shown in Figure 2, checks whether G is ϵ-feasible by applying a maximum network flow algorithm to a graph $M(G, \epsilon, lower, upper)$. When the algorithm exits from the do-loop at Step (2) of MAX-MEAN-CUT-I, we have an ϵ-feasible circulation x, UB, and LB such that $UB - LB < 1/m^2$. Therefore, at this point the maximum cut mean $\epsilon^*(G)$ satisfies $| \epsilon - \epsilon^*(G) | < 1/m^2$. The number of iterations at Step (2) is

```
Procedure FEASIBLE(G, ε)
begin
(1)       Find a maximum flow in M(G, ε, lower, upper);
          Let f* be the obtained maximum flow.
(2)       If f* saturates all outgoing edges from s
(2.1)       then return yes; (G is ε-feasible).
(2.2)       else begin
                 Let [S'_LB, T'_LB] be a minimum cut in
                 M(G, ε, lower, upper);
                 return no and [S_LB, T_LB]; (G is ε-infeasible).
                 end
end
```

Figure 2: Algorithm FEASIBLE

$O(\log(m^2U)) = O(\log(nU))$. Finally, at Step (3), we can correctly compute $\epsilon^*(G)$, as shown in the next lemma:

Lemma 3.9 *A maximum mean cut in G is correctly computed at Step (3) in MAX-MEAN-CUT-I.*

Proof. Let LB and UB be the lower and upper bounds when algorithm MAX-MEAN-CUT-I exits from the do-loop at Step (2). Since G is UB-feasible and LB-infeasible, we have $LB < \epsilon^*(G) \leq UB$. Let $\epsilon_1 = mean(T_{LB}, S_{LB})$. From the lemma below, we have $LB < \epsilon_1$. Suppose G is ϵ_1-infeasible. We should have a relationship $LB < \epsilon_1 < \epsilon^*(G) \leq UB$. However, since $UB - LB < 1/m^2$, this relationship contradicts Lemma 3.1. Thus, G is ϵ_1-feasible, and from Lemma 3.1 $\epsilon_1 = \epsilon^*(G)$. □

Lemma 3.10 *Suppose G is ϵ_0-infeasible. Let $\epsilon_1 = mean(T_0, S_0)$, where $[S'_0, T'_0]$ is a minimum cut in $M(G, \epsilon_0, lower, upper)$. Then $\epsilon_0 < \epsilon_1$.*

Proof. Since G is ϵ_0-infeasible, from Hoffman's Theorem (Theorem 3.3), we have $val_{\epsilon_0}(T_1, S_1) > 0$ or $mean(T_1, S_1) > \epsilon_0$ for some cut $[S_1, T_1]$ in G. From Lemma 3.5, we have

$$cap(S', T') = \sum_{e \in E} lower(e) - val_{\epsilon_0}(T, S)$$

for any cut $[S', T']$ in $M(G, \epsilon_0, lower, upper)$. Therefore, the cut $[T_0, S_0]$ in G, which is associated with a minimum cut $[S'_0, T'_0]$ in $M(G, \epsilon_0, lower, upper)$, maximizes the value $val_{\epsilon_0}(T, S)$. Thus, we have $val_{\epsilon_0}(T_0, S_0) \geq val_{\epsilon_0}(T_1, S_1) > 0$; that is, $\epsilon_1 = mean(T_0, S_0)\epsilon_0$. □

From the above discussion, we obtained the following theorem:

Theorem 3.11 *The algorithm MAX-MEAN-CUT-I correctly computes the maximum mean cut in $O(MF(n, m, U) \log(nU))$ time.* □

In the next section, we will show that Step (2.2) can be computed more efficiently by computing the circulation feasibility essentially once instead of $O(\log(nU))$ times.

4 Efficient Implementation

In this section, we show that Step (2.2) in MAX-MEAN-CUT-I can be computed more efficiently. Although we check the ϵ-feasibility from the scratch $O(\log(nU))$ times in the algorithm MAX-MEAN-CUT-I, we may expect much faster implementation if we can check the ϵ-feasibility efficiently by using the UB-feasible circulation obtained in the previous iteration. Before introducing our implementation, we need Minty's Lemma:

Lemma 4.1 (Minty's Lemma, [8,16]) *Let $G = (V, E)$ be a graph whose edges are arbitrarily colored black, red, or yellow. Some edges may be uncolored. Assume that there is at least one edge, $e_0 \in E$, that is black. Then one of the following cases applies:*
(a) There exists an elementary cycle C, containing e_0, that does not contain an uncolored edge. In cycle C, all black edges have the same direction (that of e_0) and all yellow edges have a direction opposite to that of e_0. (Hereafter, we call this cycle a Minty-cycle.)
(b) There exists a cocycle which is not red (but may contain uncolored edges), with all black edges having the same direction (that of e_0) and all yellow edges having a direction opposite to that of e_0. (Hereafter, we call the cut associated with this cocycle a Minty-cut.) □

Suppose that we are given LB, UB, ϵ, δ, and x such that G is LB-infeasible and UB-feasible, x is a UB-feasible circulation, $\delta = (UB+2LB)/3$, and $\epsilon = 2(UB-LB)/3$. We now design an algorithm that obtains a $(\delta + \epsilon/2)$-feasible circulation x' or detects δ-infeasibility.

In order to use Minty's Lemma, we color edges in B black, edges in R red, and edges in Y yellow, where B, R, and Y are defined as follows: Let $B = B^+ \cup B^-$ such that

$$B^+ = \{e \in E \mid lower(e) - UB < x(e) \leq lower(e) - \delta - \epsilon/2\} \tag{6}$$

and

$$B^- = \{e \in E \mid lower(e) - \delta - \epsilon/2 < x(e) \leq lower(e) - \delta\}. \tag{7}$$

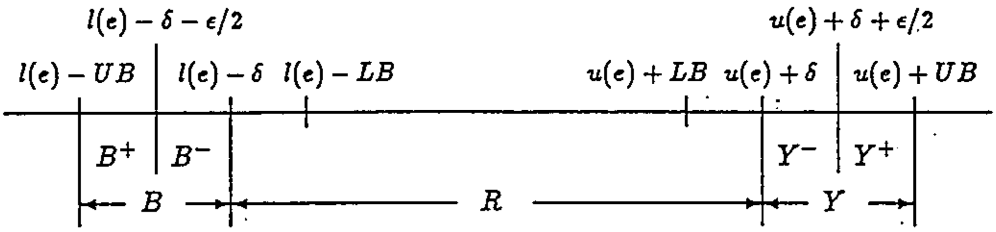

Figure 3: Coloring of edges. At the beginning of each iteration, G is UB-feasible and LB-infeasible. After augmenting the flow by $\epsilon/2$ units along with a Minty-cycle containing an edge $e \in B^+ \cup Y^+$, this edge will be moved to $B^- \cup Y^-$. In the figure, we use $l(e)$ and $u(e)$ instead of $lower(e)$ and $upper(e)$, respectively, for simplicity. Note that $\delta = (UB + 2LB)/3$ and $\epsilon = 2(UB - LB)/3$.

Let

$$R = \{e \in E \mid lower(e) - \delta < x(e) < upper(e) + \delta\}. \tag{8}$$

Let $Y = Y^- \cup Y^+$ such that

$$Y^- = \{e \in E \mid upper(e) + \delta \le x(e) < upper(e) + \delta + \epsilon/2\} \tag{9}$$

and

$$Y^+ = \{e \in E \mid upper(e) + \delta + \epsilon/2 \le x(e) < upper(e) + UB\}. \tag{10}$$

Figure 3 illustrates the way of edge coloring. If $B^+ \cup Y^+ = \emptyset$, a circulation x is already $(\delta + \epsilon/2)$-feasible. Otherwise, let $e_0 \in B^+ \cup Y^+$. We first try to find a Minty-cycle C_{e_0} containing e_0 and augment the flow along C_{e_0} by $\epsilon/2$ units so that $e_0 \notin B^+ \cup Y^+$ in the following way. Suppose that $e_0 \in B^+$. (If $e_0 \in Y^+$, we can define a new circulation analogously.) For an edge $e \in E$, we define $\chi(e) \in \{-1, 0, 1\}$ as follows: If $e \in C_{e_0}$ is in the same (*resp.* opposite) direction as that of e_0 in C_{e_0}, we define $\chi(e) = 1$ (*resp.* $\chi(e) = -1$). If $e \notin C_{e_0}$, then $\chi(e) = 0$. For each edge $e \in E$, we define a new circulation y by $y(e) = x(e) + \chi(e) \cdot \epsilon/2$. If there is no Minty-cycle containing e, we can conclude that G is δ-infeasible, as shown in the next lemma.

Lemma 4.2 *Let $e_0 \in B^+ \cup Y^+$. If there is no Minty-cycle containing e_0, then G is δ-infeasible.*

Proof. Suppose that $e_0 \in B^+$. (If $e_0 \in Y^+$, we can prove this lemma analogously.) From Minty's Lemma, there is a Minty-cut $[A, \overline{A}]$ that satisfies property (b) in the lemma. If an

```
Procedure MAX-MEAN-CUT-II
begin
(1)         LB = 0, UB = U, x = 0;
(2)         do until ( UB - LB < 1/m^2 );
(2.1)           δ = (UB + 2LB)/3;
                ε = 2(UB - LB)/3;
(2.2)           Define B+, B-, R, Y-, and Y+ as Equations 6-10.
(2.3)           do while ( B+ ∪ Y+ ≠ ∅ )
                  Let e0 ∈ B+ ∪ Y+.
(2.3.1)           if ∃ a Minty-cycle C_e0 such that e ∈ C_e0
(2.3.2)             then begin
                        Augment flow along this cycle by ε/2 units;
                        Update B, R, and Y;
                        end
(2.3.2)             else begin
                        Let [S_LB, T_LB] be a Minty-cut;
                        goto (2.4);
                        end
                end
(2.4)           if B+ ∪ Y+ = ∅
(2.4.1)           then UB = δ + ε/2;
(2.4.2)           else  LB = δ;
            end
(3)         ε*(G) = mean(T_LB, S_LB);
end
```

Figure 4: Algorithm MAX-MEAN-CUT-II

edge $e \in (A, \overline{A})$, then e is yellow and

$$upper(e) + \delta \leq x(e) < upper(e) + UB. \tag{11}$$

If an edge $e \in (\overline{A}, A)$, then e is black and

$$lower(e) - UB < x(e) \leq lower(e) - \delta. \tag{12}$$

Note that there are no red edges in $(A, \overline{A})$. Since $e_0 \in B^+$, we have

$$lower(e_0) - UB < x(e_0) \leq lower(e_0) - \delta - \epsilon/2. \tag{13}$$

Since x is a circulation,

$$\sum_{e \in (A,\overline{A})} x(e) - \sum_{e \in (\overline{A},A)} x(e) = 0. \tag{14}$$

From Equations 11, 12, 13, and 14, we have

$$\sum_{e\in(A,\overline{A})} (upper(e)+\delta) - \sum_{e\in(\overline{A},A)} (lower(e)-\delta) + \epsilon/2 \leq 0 \tag{15}$$

Therefore,

$$\sum_{e\in(A,\overline{A})} (upper(e)+\delta) - \sum_{e\in(\overline{A},A)} (lower(e)-\delta) \leq -\epsilon/2 < 0. \tag{16}$$

From Hoffman's theorem (Theorem 3.3) and Equation 16, G is not δ-feasible. □

From the above observation, we can obtain an algorithm MAX-MEAN-CUT-II, which is shown in Figure 4. Note that the algorithm correctly computes a maximum mean cut at Step (3) from Lemma 3.9.

In the following two lemmas, we discuss how efficiently we can compute Step (2.3). First, we can find a Minty-cycle for an edge $e \in B^+ \cup Y^+$ by a breadth-first search, which takes $O(m)$ time. Therefore, we have the following lemma:

Lemma 4.3 *We can compute Step (2.3) in MAX-MEAN-CUT-II in $O(m^2)$ time.*

Proof. Note that once an edge $e \in B^+ \cup Y^+$ moves out from $B^+ \cup Y^+$, it will never come back to this set in the same iteration. Therefore, the total number of augmentations at each iteration is at most m. □

In another way, we can compute Step (2.3) by solving a circulation feasibility problem, as shown in the next lemma:

Lemma 4.4 *We can compute Step (2.3) in MAX-MEAN-CUT-II by solving the following circulation feasibility problem defined on a digraph G. Let us define a lower (resp. upper) bound, denoted by $\alpha(e)$ (resp. $\beta(e)$), for each edge $e \in E$ as follows: Let $\gamma = \delta + \epsilon/2$.*

$$\alpha(e) = -\min\{m, f(2(x(e) - lower(e) + \gamma)/\epsilon)$$

and

$$\beta(e) = \min\{m, f(2(upper(e) + \gamma - x(e))/\epsilon),$$

where $f(k) = \lfloor k \rfloor$ if k is not an integer; otherwise, $f(k) = k - 1$.

Proof. Let us call the flow augmentation by $\epsilon/2$ units done at Step (2.3) in MAX-MEAN-CUT-II an $\epsilon/2$-*augmentation*. We now consider how many times we can (or have to) do $\epsilon/2$-augmentations through each edge. For example, an edge $e \in E$ can get $\epsilon/2$-augmentations until it belongs to Y^+. Therefore, the maximum number of $\epsilon/2$-augmentations for an edge $e \in E$ is

$$\beta(e) = \lfloor (upper(e) + \delta + \epsilon/2 - x(e))/(\epsilon/2) \rfloor.$$

However, the total amount of flow we move in Step (2.3) is at most $m\epsilon/2$. Therefore, without loss of generality, we can assume $\beta(e) = m$ when $\beta(e) \geq m$. We can similarly analyze the lower bound $\alpha(e)$ of the number of $\epsilon/2$-augmentations for $e \in E$. Therefore, we can replace Step (2.3) by the above-defined circulation feasibility problem. □

We can check the circulation feasibility of Lemma 4.4 by solving the maximum flow problem in a graph $M(G, 0, \alpha, \beta)$. Moreover, this maximum flow problem can be solved efficiently in $O(MF(n, m, m^2))$ time, since all the capacities are integers not greater than m^2. In this special case, Ahuja and Orlin's algorithm [1] runs in $O(mn + n^2 \log n)$, and that of Ahuja, Orlin, and Tarjan [2] in $O(mn \log(\frac{n \log n}{m \log \log n} + 2))$ or $O(mn \log \log n)$ time. Therefore, we have the following theorem:

Theorem 4.5 *Algorithm MAX-MEAN-CUT-II runs in*

$$O(\min\{m^2, mn + n^2 \log n, mn \log \log n\} \log(nU))$$

time. □

Corollary 4.6 *The circulation feasibility problem can be solved in the same time complexity.* □

For a practical improvement, we may reset UB to $\max\{\max\{lower(e) - x(e) \mid e \in E, x(e) \leq lower(e)\}, \max\{x(e) - upper(e) \mid e \in E, x(e) \geq upper(e)\}\}$ at the end of each iteration (Step (2.4.1) of MAX-MEAN-CUT-II). However, this modification does not improve the theoretical time complexity.

We can also generalize our algorithm by introducing an arbitrary scaling factor α so that $0 < \alpha < 1$, and at each iteration we check $((1 + \alpha)UB + (1 - \alpha)LB)/2$-feasibility or detect $(\alpha UB + (1 - \alpha)LB)$-infeasibility. The α-scaling method reduces the size of an interval (LB, UB) by at least $\min\{\alpha, (1 - \alpha)/2\}(UB - LB)$. Therefore, the choice of $\alpha = 1/3$ as in this paper gives the best improvement.

5 Conclusion

In this paper, we presented an efficient scaling algorithm for the maximum mean cut problem. Our algorithm uses an approximate binary search to find the maximum mean cut value. At each iteration, giving a UB-feasible circulation and keeping LB-infeasibility, we produce an $(LB+2UB)/3$-feasible circulation or detect $(2LB+UB)/3$-infeasibility by canceling at most m Minty-cycles or solving a maximum flow problem with integer capacities whose absolute values are bounded by m^2. In this way, we can attain the time-complexity $O(\min\{m^2, mn + n^2 \log n, mn \log\log n\} \log(nU))$, which is better than or comparable to the time achieved by McCormick and Ervolina [15] under the *similarity condition* (that is, $U = O(n^{O(1)})$).

Two questions remain open. First, can we improve the time-complexity of each iteration (Step (2.3) of MAX-MEAN-CUT-II) to $O(mn)$? Second, can we extend our algorithm to a strongly polynomial algorithm?

References

[1] R. K. Ahuja and J. B. Orlin, A Fast and Simple Algorithm for the Maximum Flow Problem, *Operations Research*, vol. 37, no. 5, pp. 748-759, 1990.

[2] R. K. Ahuja, J. B. Orlin, and R. E. Tarjan, Improved Time Bounds for the Maximum Flow Problem, *SIAM Journal on Computing*, vol. 18, pp. 939-954, 1989.

[3] J. Edmonds and R. M. Karp, Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems, *Journal of ACM*, 19, pp. 248-264, 1978.

[4] T. R. Ervolina and S. T. McCormick, A Strongly Polynomial Maximum Mean Cut Cancelling Algorithm for Minimum Cost Network Flow, UBC Faculty of Commerce Working Paper 90-MSC-009, May 1990.

[5] T. R. Ervolina and S. T. McCormick, A Strongly Polynomial Dual Cancel and Tighten Algorithm for Minimum Cost Network Flow, UBC Faculty of Commerce Working Paper 90-MSC-010, May 1990.

[6] A. V. Goldberg and R. E. Tarjan, A New Approach to the Maximum-Flow Problem, *Journal of ACM*, vol. 35, no. 4, pp. 921-940, 1988.

[7] A. V. Goldberg and R. E. Tarjan, Finding Minimum-Cost Circulations by Canceling Negative Cycles, *Journal of ACM*, vol. 36, no. 4, pp. 873-886, 1989.

[8] M. Gondran and M. Minoux, *Graphs and Algorithms*, (trans. S. Vajda), John Wiley & Sons Ltd., New York, N. Y., 1984.

[9] R. Hassin, Algorithms for the Minimum Cost Circulation Problem in Cycles, Based on Maximizing the Mean Improvement, February 1990.

[10] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers, Fifth Edition*, Oxford University Press, 1979.

[11] J. C. Herz, *Course de Théorie des graphes*, Faculté des Science de Lille, 1967.

[12] A. J. Hoffman, Some Recent Applications of the Theory of Linear Inequalities to Extremal Combinatorial Analysis. In "Combinatorial Analysis," *Proc. 10th Symp. on Appl. Math. Columbia Univ. 1958, Amer. Math. Soc. Providence*, R. I. (R. E. Bellman and M. Hall eds.) pp. 113-127, 1958.

[13] R. M. Karp, A Characterization of the Minimum Cycle Mean in a Digraph, *Discrete Mathematics*, vol. 23, pp. 309-311, 1978.

[14] E. L. Lawler, *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart and Winston, New York, 1978.

[15] S. T. McCormick and T. R. Ervolina, Computing Maximum Mean Cuts, UBC Faculty of Commerce Working Paper 90-MSC-011, May 1990.

[16] G. J. Minty, On the Axiomatic Foundations of the Theories of Directed Linear Graphs, Electrical Networks and Network Programming, *J. Math. Mech.* 15, pp. 485-520, 1966.

[17] J. B. Orlin, A Faster Strongly Polynomial Minimum Cost Flow Algorithm, *Proceedings of the 20th ACM Symposium on the Theory of Computing*, pp. 377-387, 1988.

[18] J. B. Orlin and R. K. Ahuja, New Scaling Algorithms for the Assignment and Minimum Cycle Mean Problems, Technical Report OR 178-88, Operations Research Center, MIT, May 1988.

[19] E. Zemel, A Linear Time Randomizing Algorithm for Searching Ranked Functions, *Algorithmica*, 2, pp. 81-90, 1987.