Bisections above Tight Lower Bounds

Matthias Mnich[1] and Rico Zenklusen[2,⋆]

[1] Cluster of Excellence, Saarbrücken, Germany
m.mnich@mmci.uni-saarland.de
[2] Massachusetts Institute of Technology, Cambridge, MA, USA
ricoz@math.mit.edu

Abstract. A bisection of a graph is a bipartition of its vertex set in which the number of vertices in the two parts differ by at most one, and the size of the bisection is the number of edges which go across the two parts.

Every graph with m edges has a bisection of size at least $\lceil m/2 \rceil$, and this bound is sharp for infinitely many graphs. Therefore, Gutin and Yeo considered the parameterized complexity of deciding whether an input graph with m edges has a bisection of size at least $\lceil m/2 \rceil + k$, where k is the parameter. They showed fixed-parameter tractability of this problem, and gave a kernel with $O(k^2)$ vertices.

Here, we improve the kernel size to $O(k)$ vertices. Under the Exponential Time Hypothesis, this result is best possible up to constant factors.

1 Introduction

A *bisection* of a graph is a bipartition of its vertex set in which the number of vertices in the two parts differ by at most one, and its *size* is the number of edges which go across the two parts. We are interested in finding bisections of maximum size in a given graph, which is known as the Max-Bisection problem. This problem is NP-hard by a simple reduction from the Max-Cut problem. On the other hand, there is a simple randomized polynomial-time procedure [1] that finds in any m-edge graph a bisection of size at least $\lceil m/2 \rceil$, and there are graphs (such as stars) for which this bound cannot be improved. Therefore, interest arose in the study of the problem Max-Bisection Above Tight Lower Bound (or Max-Bisection ATLB for short), where we seek a bisection of size at least $m/2 + k$ when given an m-edge graph G together with an integer $k \in \mathbb{N}$. The NP-hardness of Max-Bisection ATLB follows from the NP-hardness of Max-Bisection. On the positive side, Gutin and Yeo [1] showed that Max-Bisection ATLB is fixed-parameter tractable, that is, pairs (G, k) can be decided in time $f(k) \cdot n^{O(1)}$ for some function f dependent only on k, where n is the number of vertices of G. Fixed-parameter tractability directly implies the existence of a *kernelization* [2], which is a polynomial-time algorithm that efficiently compresses instances (G, k) to equivalent instances (G', k') (the

⋆ Supported by NSF grants CCF-1115849 and CCF-0829878, and by ONR grants N00014-11-1-0053 and N00014-09-1-0326.

M.C. Golumbic et al. (Eds.): WG 2012, LNCS 7551, pp. 184–193, 2012.

kernel) of size $|G'| + k' \leq g(k)$ for some function g dependent only on k. Gutin and Yeo's fixed-parameter tractability result [1] is based on proving a kernel with $O(k^2)$ vertices.

Here we improve their kernel as follows.

Theorem 1. Max-Bisection ATLB *admits a kernel with at most* $16k$ *vertices.*

We observe that the number of vertices in our kernel is asymptotically optimal, assuming the Exponential Time Hypothesis introduced by Impagliazzo et al.[3]. The hypothesis implies that a large family of NP-complete problems cannot be solved in subexponential time, including the Max-Cut problem. In the Max-Cut problem the goal is to find any bipartition of the vertices (not necessarily balanced as in the bisection case) that maximizes the number of edges crossing it. The Max-Cut problem on a graph $G = (V, E)$ can easily be reduced to the maximum bisection problem by adding $|V|$ additional isolated vertices to G to obtain G', and solving the bisection problem in G'. Clearly, the maximum bisection in G' induces a bipartition of the vertices in G that solves the Max-Cut problem. Now notice that a kernel for the Max-Bisection ATLB problem with $o(k)$ vertices, would imply that we could solve the Max-Bisection ATLB problem in $2^{o(k)}$ time by checking all bisections of the kernel. By the above relation to the Max-Cut problem this would yield a subexponential time algorithm for Max-Cut, contradicting the Exponential Time Hypothesis [4].

We remark that for the Max-Cut problem, deciding the existence of a cut with size $m/2+k$ is trivially fixed-parameter tractable, because any graph admits a cut of size $m/2 + \Omega(\sqrt{m})$ due to a classical result of Edwards [5,6]. In fact, the Max-Bisection problem forms an "extremal point" of a series of problems on *α-bisections*, which are cuts in which both sides of the bipartition have at least $(1/2-\alpha)n$ vertices. By a recent result of Lee et al. [7], for every $\alpha \in [0, 1/6]$ every n-vertex graph with m edges and no isolated vertices contains an α-bisection of size at least $m/2 + \alpha n$. Thus, deciding the existence of an α-bisection of size $m/2 + k$ is trivially fixed-parameter tractable for all $\alpha \in (0, 1/6]$. In this paper we prove an essentially optimal kernel for the extremal case $\alpha = 0$.

The problem Max-Bisection ATLB is an example of a so-called "above tight lower bound paramaterization", where the parameter k is chosen as the excess of the solution value of the given instance over a non-trivial tight lower bound on the solution value in arbitrary instances (here: $\lceil m/2 \rceil$). For many parameterizations above tight lower bound, it is often not even clear how to solve such problems in time m^k, let alone by a fixed-parameter algorithm in time $f(k) \cdot m^{O(1)}$ for some function f dependent only on k. By now, several techniques have been developed for fixed-parameter algorithms of above tight lower bound parameterizations of important computational problems, such as Max-r-Sat [8], Max-Lin2 [9], Permutation-CSPs [10], and Max-Cut [11]; see the survey by Gutin and Yeo [12]. Most of these techniques are based on probabilistic analysis of carefully chosen random variables, and they rarely yield kernels of linear size. Here, we introduce a new technique to establish a linear vertex-kernel for Max-Bisection ATLB, based on Edmonds-Gallai decompositions of graphs. We believe that this technique has the potential to find further applications in establishing linear kernels for problems parameterized above tight lower bound.

2 Preliminaries

Let G be a loopless undirected graph with vertex set $V(G)$ and edge set $E(G)$. We allow parallel edges. For each vertex $v \in V(G)$, let $N(v) = \{u \in V(G) \mid \{u, v\} \in E(G)\}$ be the set of neighbors of v. In particular, $v \notin N(v)$. For a vertex $v \in V(G)$ and a subset $U \subseteq V(G)$, we denote by $d_U(v) = |\{e \in E(G) \mid v \in e, e \setminus \{v\} \subseteq U\}|$ the *degree of v in U*, and write shorthand $d(v) = d_{V(G)}(v)$. Notice that if we have parallel edges adjacent to v, then $d(v) > |N(v)|$. Let $\sigma(G)$ denote the number of connected components of G. For sets $U, W \subseteq V(G)$, define $E(U, W) = \{\{u, w\} \in E(G) \mid u \in U, w \in W\}$ and $d(U, W) = |E(U, W)|$. For the special case of $U = \{u\}$ and $W = \{v\}$ being singleton sets, we use the shorthand $d(u, w) = d(\{u\}, \{w\})$. The subgraph of G induced by a subset $V' \subseteq V(G)$ is denoted by $G[V']$. For $v \in V$, we use the shorthand $G - v = G[V \setminus \{v\}]$.

For a graph G, a *matching* is a set M of pairwise non-adjacent edges; the vertices in $V(M) = \{v \in e \mid e \in M\}$ are *saturated by M* and vertices in $V(G) \setminus M$ are *unsaturated by M*. A matching is *perfect* for G if it saturates every vertex of G, and G is *factor-critical* if the graph $G - v$ admits a perfect matching for every $v \in V(G)$. Denote by $\nu(G)$ the cardinality of a maximum size matching in G.

For a graph G, an *Edmonds-Gallai decomposition* [13] is a tuple (X, Y, Z) such that $\{X, Y, Z\}$ forms a tripartion of $V(G)$, X is such that for every vertex $v \in X$ the size of a maximum cardinality matching in $G - v$ and G are the same, Y contains all neighbors of X in $V(G) \setminus X$, and $Z = V(G) \setminus (X \cup Y)$. Classical results on the Edmonds-Gallai decomposition imply that every connected component of $G[X]$ is factor-critical, every component of $G[Z]$ admits a perfect matching, and furthermore $\nu(G) = \frac{n - \sigma(G[X]) + |Y|}{2}$.

3 Proof of Theorem 1

In this section we prove our main result, Theorem 1.

Let G be a loopless undirected graph with $n = |V(G)|$ vertices and $m = |E(G)|$ edges, and let $k \geq 0$ be an integer. An important ingredient that we use to bound the size of the kernel is the following well-known fact, showing that large matchings lead to large bisections.

Proposition 1 ([14]). *Let G be a graph and M be a matching in G; then G has a bisection of size at least $\lceil m/2 \rceil + \lfloor |M|/2 \rfloor$ and such can be found in $O(m + n)$ time.*

Another important fact, that will prove to be useful in our reduction to obtain a small kernel, is that whenever there is a large set of vertices with the same neighbors (and same number of parallel edges to the neighbors), the problem can be reduced to a smaller one. This is a straightforward generalization of a reduction used by Gutin and Yeo [1].

Lemma 1 (straightforward extension of a result in [1]). *Let G be a loopless undirected graph with a set $I \subseteq V(G)$ of size $\lceil n/2 + j \rceil$ for some $j > 0$ such that $d(u, w) = d(v, w)$ for all $u, v \in I$ and $w \in V(G)$. Then G has a bisection of size $\lceil |E(G)|/2 \rceil + k$ if and only if the graph G' obtained from G by removing $2j$ arbitrary vertices of I has a bisection of size $\lceil E(G')/2 \rceil + k$.*

The above lemma can easily be seen to be true by observing that any balanced bipartition (V_1, V_2) of V must satisfy $|I \cap V_1| \geq j$ and $|I \cap V_2| \geq j$, due to the large size of I and the balancedness of the bipartition (V_1, V_2). One can then observe that any balanced bipartition (V_1, V_2) of V leading to a bisection of size at least $\frac{|E(G)|}{2} + k$ can be transformed into a bisection of G' of size $\frac{|E(G')|}{2} + k$ by removing any j vertices of $|I \cap V_i|$ from V_i, for $i = 1, 2$. Similarly, any bisection (V_1', V_2') of G' of size at least $\frac{|E(G')|}{2} + k$ can be completed to a bisection (V_1, V_2) of G of size $\frac{|E(G)|}{2} + k$ by adding any j vertices of $V(G) \setminus V(G')$ to V_1 to obtain V_1', and the remaining vertices of $V(G) \setminus V(G')$ to V_2 to obtain V_2'.

From now on, we assume that $\nu(G) < 2k$, since otherwise there is a bisection of size $m/2 + k$ due to Proposition 1, and such can be found efficiently through a simple randomized switching argument that can be derandomized through conditional expectations. (Details of such a derandomization in a similar setting are given by Ries and Zenklusen [15].) Furthermore, we assume that G does not contain any large set $I \subseteq V$ as defined in Lemma 1, for otherwise we could apply Lemma 1 to reduce the size of the graph.

Let (X, Y, Z) be a Gallai-Edmonds decomposition of G. As a reminder, $G[X]$ consists of factor-critical components, Y are all neighbors of X, and $G[Z]$ admits a perfect matching. Furthermore, $\nu(G) = \frac{n - \sigma(G[X]) + |Y|}{2}$.

We partition X into sets X_0, X_1, X_2, defined as

$$\begin{aligned} X_0 &= \{v \in X \mid d(v) = 0\}, \\ X_1 &= \{v \in X \mid d_X(v) = 0\} \setminus X_0, \\ X_2 &= \{v \in X \mid d_X(v) \geq 1\}. \end{aligned}$$

Hence, $G[X_2]$ contains all connected components of $G[X]$ with more than one vertex. Notice that since these components are factor-critical, each of them has size at least 3.

Lemma 2. *We have $|X_2|/3 + |Y| + |Z|/2 < 2k$.*

Proof. Consider a maximum matching $M \subseteq E(G)$ in G. It is a well-known property of the Gallai-Edmonds decomposition [13] that M saturates all vertices in $Y \cup Z$. More precisely, M can be partitioned into $M = M_X \uplus M_Y \uplus M_Z$, where M_X are the edges of M having both endpoints in X, M_Y are the edges of M having one endpoint in Y, and M_Z are the edges of M having both endpoints in Z. Furthermore, M_Z is a perfect matching in $G[Z]$, and all edges of M_Y connect a vertex of Y with one of X. Additionally, in each connected component $G[X']$ of $G[X]$, the edges of M_X with both endpoints in X' saturate all but one vertex of X'. Observe that X_2 are precisely those vertices in X that belong to connected components of $G[X]$ of size at least 3. Hence,

$$|M_X| \geq |X_2|/3,$$

because the number of edges $|M_X|$ is minimized when all vertices in X_2 are in connected components of size precisely 3. This is the case since in a connected component of $G[X]$ of size $2p + 1 \geq 3$ (whose size must be odd since $G[X]$ is factor-critical), the ratio between number of vertices and matching edges between them is $\frac{p}{2p+1}$, and this ratio is minimized for $p = 1$.

Furthermore, $|M_Y| = |Y|$ and $|M_Z| = |Z|/2$. Thus

$$|M| = |M_X| + |M_Y| + |M_Z| \geq |X_2|/3 + |Y| + |Z|/2.$$

Since $|M| < 2k$ (by assumption there is no matching of size $\geq 2k$), the result follows. $\square$

Our key technique to show that $|V(G)| \leq 16k$ is a generalized version of the randomized argument used to show that a large matching leads to a large bisection. We replace the role of a matching by what we call *switching units* on $V(G)$. A switching unit is a tuple (A, B) with $A, B \subseteq V(G)$, $A \cap B = \emptyset$ and $|A| = |B|$. We will construct a *switching family* $(A_i, B_i)_i$, which is a collection of mutually disjoint switching units, i.e., $(A_i \cup B_i) \cap (A_j \cup B_j) = \emptyset$ for $i \neq j$. Any switching family can be used to define a random bisection (V_1, V_2) of $V(G)$ by randomly and independently assigning the vertices of each switching unit (A_i, B_i) to the sets V_1, V_2 as follows: with probability $1/2$ assign the vertices of A_i to V_1 ($A_i \subseteq V_1$) and the vertices of B_i to V_2 ($B_i \subseteq V_2$), otherwise set $A_i \subseteq V_2$ and $B_i \subseteq V_1$. Furthermore, for all remaining vertices $\widetilde{V} = V(G) \setminus \bigcup_i (A_i \cup B_i)$, i.e., all vertices not part of any switching unit, we pick uniformly at random a bisection $(\widetilde{V}_1, \widetilde{V}_2)$ of $\widetilde{V}$, and assign all vertices in $\widetilde{V}_1$ to V_1 and all vertices of $\widetilde{V}_2$ to V_2. We call the thus obtained random bisection (V_1, V_2) a random bisection corresponding to the switching family $(A_i, B_i)_i$.

To compute the expected number of edges $\mathbb{E}[d(V_1, V_2)]$ in the random bisection (V_1, V_2) corresponding to $(A_i, B_i)_i$, we observe that all edges not having both endpoints in the same switching set $A_i \cup B_i$ are in the bisection (V_1, V_2) with probability at least $1/2$. (Notice that this probability can indeed be strictly larger than $1/2$, e.g., when considering a graph consisting of two vertices connected by a single edge, then this edge will be in the bisection with probability one.) It remains to consider for each switching unit (A_i, B_i) its contribution to $d(V_1, V_2)$. Notice that this contribution is deterministic and equals $d(A_i, B_i)$.

Since we are interested in how much $\mathbb{E}[d(V_1, V_2)]$ exceeds the tight lower bound $m/2$, we introduce the *excess* $\mathsf{ex}(A_i, B_i)$ of the switching unit (A_i, B_i) as follows:

$$\begin{aligned}\mathsf{ex}(A_i, B_i) &= 2d(A_i, B_i) - d(A_i \cup B_i, A_i \cup B_i) \\ &= d(A_i, B_i) - d(A_i, A_i) - d(B_i, B_i)\ .\end{aligned}$$

In words, the excess of (A_i, B_i) is the difference between the number of edges in $G[A_i \cup B_i]$ crossing the bisection (A_i, B_i) and those who do not. Using the notion of excess, we can express the expected number of edges in the random bisection (V_1, V_2) by

$$\mathbb{E}[d(V_1, V_2)] \geq \frac{m}{2} + \frac{1}{2}\sum_i \mathsf{ex}(A_i, B_i), \qquad (1)$$

because every edge is in the bisection with probability at least 1/2, except for edges e with both endpoints in one switching unit (A_i, B_i), in which case e is in the bisection if $e \in E(A_i, B_i)$ and e is not in the bisection if $e \in E(A_i, A_i) \cup E(B_i, B_i)$.

In the following, we describe a way to construct a switching family $(A_i, B_i)_i$ with a high total excess $\sum_i \mathsf{ex}(A_i, B_i)$. Let $Y = \{y_1, \ldots, y_\ell\}$. We start by constructing iteratively for each $i = 1, \ldots, \ell$ a switching unit (A_i, B_i), which might be chosen to be $(\emptyset, \emptyset)$. Assume that we already constructed switching units $(A_1, B_1), \ldots, (A_{i-1}, B_{i-1})$. Let

$$N_i = X_1 \setminus \bigcup_{j=1}^{i-1} (A_j \cup B_j),$$

where $N_1 = X_1$. Consider the partition of N_i into sets $N_i^0, N_i^1, \ldots$, where

$$N_i^j = \{v \in N_i \mid d(y_i, v) = j\}.$$

If $N_i = N_i^j$ for some $j \in \mathbb{Z}_+$, we set $A_i = B_i = \emptyset$. Otherwise, we start by assigning y_i to A_i and we choose any element $v \in N_i \setminus N_i^0$ that we assign to B_i. Then, as long as there is an unassigned pair $(u, v) \in N_i \times N_i$ with $u \in N_i^{j_1}$ and $v \in N_i^{j_2}$, where $j_1 < j_2$, we assign u to A_i and v to B_i. Clearly, at the end of this procedure, all elements in $N_i \setminus (A_i \cup B_i)$ belong to a single group N_i^j. The key observation is that for every pair $u \in N_i^{j_1}, v \in N_i^{j_2}$ with $j_1 < j_2$ that we add, $\mathsf{ex}(A_i, B_i)$ increases by at least one unit because v has at least one more edge adjacent to y_i than u has. Furthermore, also the assignment at the start of y_i to A_i and of an arbitrarily chosen vertex $v \in N_i \setminus N_i^0$ to B_i creates an excess of at least one unit. Thus, for each $i \in \{1, \ldots, \ell\}$, the switching unit (A_i, B_i) satisfies the following properties:

(a) $\mathsf{ex}(A_i, B_i) \geq |A_i \cup B_i|/2$, since any added pair of vertices increases the excess by at least one unit, and
(b) $d(y_i, v) = d(y_i, u)$ for any $u, v \in N_i$, since otherwise another pair of vertices could have been added to the switching unit (A_i, B_i).

The switching units (A_i, B_i) with $i \in \{1, \ldots, \ell\}$ are completed by adding switching units corresponding to a perfect matching M_Z of $G[Z]$ and a maximum matching M_{X_2} in $G[X_2]$, i.e., for each $\{u, v\} \in M_Z \cup M_{X_2}$, we construct a switching unit $(\{u\}, \{v\})$. These trivial switching units together with the ones constructed above complete the construction of our switching family, which we denote by (A_i, B_i).

We first provide a lower bound for $\mathbb{E}[d(V_1,V_2)]$. Let $\widetilde{X}_1 = X_1 \setminus \bigcup_{i=1}^{\ell}(A_i \cup B_i)$.

Lemma 3. *It holds that* $\mathbb{E}[d(V_1,V_2)] \geq \frac{m}{2} + \frac{1}{4}\left(|X_1 \setminus \widetilde{X}_1| + |Z|\right) + \frac{1}{6}|X_2|$.

Proof. Using (1), we have $\mathbb{E}[d(V_1,V_2)] \geq \frac{m}{2} + \frac{1}{2}\sum_i \mathsf{ex}(A_i,B_i)$. We recall that the sum $\sum_i \mathsf{ex}(A_i,B_i)$ is composed of three types of terms

$$\sum_i \mathsf{ex}(A_i,B_i) = \sum_{i=1}^{\ell} \mathsf{ex}(A_i,B_i) + \sum_{\{u,v\}\in M_Z} \mathsf{ex}(\{u\},\{v\}) + \sum_{\{u,v\}\in M_{X_2}} \mathsf{ex}(\{u\},\{v\}), \tag{2}$$

where (A_i,B_i) for $i \in \{1,\ldots,\ell\}$ are the iteratively constructed switching units, the second term corresponds to switching units stemming from a perfect matching M_Z in $G[Z]$, and the third term corresponds to switching units stemming from a maximum matching M_{X_2} in $G[X_2]$.

Consider the first term of (2). By property (a), we have $\mathsf{ex}(A_i,B_i) \geq \frac{|A_i \cup B_i|}{2}$ for $i \in \{1,\ldots,\ell\}$. Since $\bigcup_{i=1}^{\ell}(A_i \cup B_i)$ contains all edges of $X_1 \setminus \widetilde{X}_1$ (together with some additional vertices of Y), we obtain

$$\sum_{i=1}^{\ell} \mathsf{ex}(A_i,B_i) \geq \frac{1}{2}|X_1 \setminus \widetilde{X}_1| \ . \tag{3}$$

Now consider the second and third term of (2). Since M_Z is a perfect matching over $G[Z]$, we have $|M_Z| \geq \frac{|Z|}{2}$. Furthermore, since each connected component of $G[X_2]$ is factor-critical and has size at least 3, we have $|M_{X_2}| \geq \frac{|X_2|}{3}$. Notice that $\mathsf{ex}(\{u\},\{v\}) \geq 1$, and the inequality can be strict in case of parallel edges between u and v. Hence

$$\sum_{\{u,v\}\in M_Z} \mathsf{ex}(\{u\},\{v\}) \geq \frac{|Z|}{2},$$
$$\sum_{\{u,v\}\in M_{X_2}} \mathsf{ex}(\{u\},\{v\}) \geq \frac{|X_2|}{3} \ .$$

Combining the above inequalities with (3) and (2) and using $\mathbb{E}[d(V_1,V_2)] \geq \frac{m}{2} + \frac{1}{2}\sum_i \mathsf{ex}(A_i,B_i)$, the desired result is obtained. □

The next step is to show that not both $\widetilde{X}_1$ and X_0 can have a large size. For this we start with the following observation that follows immediately from property (b) of our iterative way to define the switching sets (A_i,B_i) for $i \in \{1,\ldots,\ell\}$.

Proposition 2. *All vertices in* $\widetilde{X}_1$ *have the same neighborhood structure, i.e., for any* $u,v \in \widetilde{X}_1$ *and* $w \in V(G)$, *we have* $d(u,w) = d(v,w)$.

Lemma 4. *If* $|\widetilde{X}_1| \geq 2k$ *and* $|X_0| \geq 2k-1$ *then* G *has a bisection of* G *of size at least* $m/2+k$, *and such can be found efficiently.*

Proof. Assume $|\widetilde{X}_1| \geq 2k$ and $|X_0| \geq 2k - 1$. Consider the following switching unit (A, B). Let $v \in N(u)$ for an arbitrary $u \in \widetilde{X}_1$. Notice that the choice of u does not matter due to Proposition 2, and $v \in Y$. Observe further that $N(u) \neq \emptyset$, since any element $u \in \widetilde{X}_1 \subseteq X_1$ has at least one neighbor in Y, as otherwise it would belong to X_0. Let $A = \{v\} \cup X_0'$, where $X_0' \subseteq X_0$ is any set with $|X_0'| = 2k - 1$, and let $B = \widetilde{X}_1'$ where $\widetilde{X}_1' \subseteq \widetilde{X}_1$ is any set with $|\widetilde{X}_1'| = 2k$. Notice that since all elements of B have the same neighborhood because $B \subseteq \widetilde{X}_1$, there is an edge between any vertex of B and v.

Instead of considering a random bisection using the switching units $(A_i, B_i)_i$, consider a random bisection (V_A, V_B) corresponding to the single switching unit (A, B). Notice that $G[A \cup B]$ is a bipartite graph with bipartition $\{A, B\}$. Hence, $\mathsf{ex}(A, B)$ is equal to the number of edges in $E(G)$ with both endpoints in $A \cup B$. Since each edge of B is connected to $v \in A$, we have $\mathsf{ex}(A, B) \geq |B| = 2k$. By (1) we thus obtain

$$\mathbb{E}[d(V_A, V_B)] \geq \frac{m}{2} + \frac{\mathsf{ex}(A, B)}{2} \geq \frac{m}{2} + k \ .$$

A bisection of size at least $m/2 + k$ can then be found efficiently through standard derandomization arguments using conditional expectations. □

We are now ready to combine all ingredients to obtain a kernel of size at most $16k$. To obtain the desired kernel, we first repeatedly apply Lemma 1 to reduce the given graph as long as the conditions of Lemma 1 are fulfilled, i.e., as long as there are large vertex sets with the same neighborhood structure. After that, if we can either apply Proposition 1, Lemma 4, or if the switching family $(A_i, B_i)_i$ leads to a random bisection (V_1, V_2) with $\mathbb{E}[d(V_1, V_2)] \geq \frac{m}{2} + k$, then we can obtain a large bisection with $\geq \frac{m}{2} + k$ edges. The remaining case, when none of these results leads to a large bisection, is covered by the following theorem.

Theorem 2. *Let G be a loopless graph on m edges and let $k \in \mathbb{N}$. Then either*

- *G has at most $16k$ vertices, or*
- *we can reduce G to a graph G' on $m' < m$ edges such that G has a bisection of size at least $m/2 + k$ if and only if G' has a bisection of size at least $m'/2 + k$, or*
- *G has a bisection of size at least $m/2 + k$ that we can find efficiently.*

Proof. First, suppose that G satisfies one of the following properties.

(i) If $\nu(G) \geq 2k$ then we obtain a bisection of G of size at least $\frac{m}{2} + k$ by Proposition 1.
(ii) If $|X_0| > \frac{|V(G)|}{2}$ then we can apply Lemma 1 to the vertices in X_0 to reduce the graph G, since all vertices in X_0 have the same neighborhood structure.
(iii) If $|\widetilde{X}_1| > \frac{|V(G)|}{2}$ then we can apply Lemma 1 to the vertices in $\widetilde{X}_1$, since all these vertices have the same neighborhood structure due to Proposition 2.
(iv) If $\min\{|X_0|, |\widetilde{X}_1|\} \geq 2k$ then Lemma 4 implies that there is a bisection of size at least $m/2 + k$ in G.

(v) If the random bisection (V_1, V_2) corresponding to the switching family $(A_i, B_i)_i$ satisfies $\mathbb{E}[d(V_1, V_2)] \geq \frac{m}{2} + k$ then a standard derandomization as given by Ries and Zenklusen [15] leads to a bisection of size at least $m/2 + k$ in G.

Second, suppose that G satisfies none of the conditions (i)–(v); we show that $|V(G)| \leq 16k$. By assumption, (v) does not hold, and so by Lemma 3 we have

$$\frac{m}{2} + k > \mathbb{E}[d(V_1, V_2)] \geq \frac{m}{2} + \frac{1}{4}(|X_1 \setminus \widetilde{X}_1| + |Z|) + \frac{1}{6}|X_2|,$$

implying

$$4k > |X_1 \setminus \widetilde{X}_1| + |Z| + \frac{2}{3}|X_2| \ . \tag{4}$$

Hence, we obtain

$$\begin{aligned} |V(G)| &= |X_0| + |X_1| + |X_2| + |Y| + |Z| \\ &= \underbrace{|X_1 \setminus \widetilde{X}_1| + |Z| + \frac{2}{3}|X_2|}_{<4k \text{ by (4)}} + \underbrace{\frac{1}{3}|X_2| + |Y|}_{<2k \text{ by Lemma 2}} + |X_0| + |\widetilde{X}_1| \\ &< 6k + |X_0| + |\widetilde{X}_1| \\ &= 6k + \underbrace{\min\{|X_0|, |\widetilde{X}_1|\}}_{<2k \text{ by (iv)}} + \underbrace{\max\{|X_0|, |\widetilde{X}_1|\}}_{\leq \frac{|V|}{2} \text{ by (ii) and (iii)}} \\ &< 8k + \frac{|V(G)|}{2} \ . \end{aligned}$$

Therefore, $|V(G)| < 16k$. □

Finally, as mentioned above, Theorem 1 is a direct consequence of Theorem 2.

4 Discussion

Our main result in this paper is a linear vertex-kernel for the Max-Bisection ATLB problem. Recently, Lee et al. [7] showed that for every $\alpha \in [0, 1/6]$, every n-vertex graph with m edges and no isolated vertices contains an α-bisection of size at least $m/2 + \alpha n$, where each side of the bipartition has at least $(1/2 - \alpha)n$ vertices. Thus, a natural problem to study is Max-α-Bisection ATLB for every $\alpha \in [0, 1/6]$, where we wish to decide the existence of an α-bisection of size at least $m/2 + \alpha n + k$ in a given n-vertex m-edge graph. We conjecture this problem to be fixed-parameter tractable and admit a polynomial-size kernel.

Acknowledgment. We are grateful to the referees for their helpful suggestions and comments.

References

1. Gutin, G., Yeo, A.: Note on maximal bisection above tight lower bound. Information Processing Letters 110, 966–969 (2010)
2. Downey, R.G., Fellows, M.R.: Parameterized complexity. Monographs in Computer Science (1999)
3. Impagliazzo, R., Paturi, R.: On the complexity of k-SAT. Journal of Computer and System Sciences 62, 367–375 (2001)
4. Cai, L., Juedes, D.: On the existence of subexponential parameterized algorithms. Journal of Computer and System Sciences 67, 789–807 (2003)
5. Edwards, C.S.: Some extremal properties of bipartite subgraphs. Canadian Journal of Mathematics 25, 475–485 (1973)
6. Edwards, C.S.: An improved lower bound for the number of edges in a largest bipartite subgraph. In: Recent Advances in Graph Theory (Proceedings of Second Czechoslovak Symposium, Prague, 1974), Prague, pp. 167–181 (1975)
7. Lee, C., Loh, P.S., Sudakov, B.: Bisections of graphs (2011), `http://arxiv.org/abs/1109.3180/v3`
8. Alon, N., Gutin, G., Kim, E.J., Szeider, S., Yeo, A.: Solving MAX-r-SAT above a tight lower bound. Algorithmica 61, 638–655 (2011)
9. Crowston, R., Fellows, M., Gutin, G., Jones, M., Rosamond, F., Thomassé, S., Yeo, A.: Simultaneously Satisfying Linear Equations Over $\mathbb{F}_2$: MaxLin2 and Max-r-Lin2 Parameterized Above Average. In: IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2011). Leibniz International Proceedings in Informatics (LIPIcs), vol. 13, pp. 229–240 (2011)
10. Gutin, G., van Iersel, L., Mnich, M., Yeo, A.: Every ternary permutation constraint satisfaction problem parameterized above average has a kernel with a quadratic number of variables. Journal of Computer and System Sciences 78, 151–163 (2012)
11. Crowston, R., Jones, M., Mnich, M.: Max-Cut Parameterized above the Edwards-Erdős Bound. In: Czumaj, A., Mehlhorn, K., Pitts, A., Wattenhofer, R. (eds.) ICALP 2012, part I. LNCS, vol. 7391, pp. 242–253. Springer, Heidelberg (2012)
12. Gutin, G., Yeo, A.: Constraint Satisfaction Problems Parameterized above or below Tight Bounds: A Survey. In: Bodlaender, H.L., Downey, R., Fomin, F.V., Marx, D. (eds.) Fellows Festschrift 2012. LNCS, vol. 7370, pp. 257–286. Springer, Heidelberg (2012)
13. Schrijver, A.: Combinatorial Optimization, Polyhedra and Efficiency. Springer (2003)
14. Haglin, D.J., Venkatesan, S.M.: Approximation and intractability results for the maximum cut problem and its variants. IEEE Transactions on Computing 40, 110–113 (1991)
15. Ries, B., Zenklusen, R.: A 2-approximation for the maximum satisfying bisection problem. European Journal of Operational Research 210, 169–175 (2011)