# On Minimal Tree-Inducing Cycle-Cutsets and Their Use in a Cutset-Driven Local Search

**Alon Milchgrub**
School of Computer Science and Engineering
The Hebrew University of Jerusalem, Israel
alon.milchgrub@mail.huji.ac.il

**Rina Dechter**
Department of Computer Science
University of California Irvine
dechter@uci.edu

## Abstract

We prove that in grids of any size there exists a minimal cycle-cutset that its complement induces a single connected tree. More generally, any cycle-cutset in a grid can be transformed to a tree-inducing cycle-cutset, no bigger than the original one. We use this result to improve the known lower bounds on the size of a minimal cycle-cutset in some cases of grids, thus equating the lower bound to the known upper bound. In addition, we present a cycle-cutset driven stochastic local search algorithm in order to approximate the minimal energy of a sum of unary and binary potentials. We show that this method is on-par and even surpasses the state-of-the-art on some grid problems, when both are initialized by elementary means.

## 1 Introduction

A *cycle-cutset* in a graph $G = (V, E)$ is a subset $C$ of the vertices of $G$, such that the graph induced on $V' = V \backslash C$ is acyclic, i.e. a forest. *The minimal cycle-cutset problem* is finding a cycle-cutset of minimal cardinality. The problem is of interest to a wide variety of applications, including distributed computing and artificial intelligence in the context of Bayesian inference and constraint satisfaction. Due to the importance of the problem, it has been extensively studied, although the problem was proven to be NP-complete for general graphs. Some of the findings include polynomial algorithms solving the problem for some specific graph classes and lower and upper-bounds on the size of the minimal cycle-cutset. In particular, [1] has previously presented lower and upper bounds on the size of the minimal cycle-cutset of grids. It was shown that the minimal cycle-cutset of the $n \times m$ grid $M_{n,m}$ is of size at least

$$\frac{(m-1)(n-1)+1}{3}$$

and at most

$$\frac{mn}{3} + \frac{m+n}{6} + o(m,n)$$

The upper bound of [1] was later significantly improved by [2], who have shown an upper bound matching the lower bound of [1] in many cases and differing from the lower bound by at most 2, in other cases apart from when $m = 5$ and $n \geq 5$.

The rest of the paper is structured as follows. In Section 2, we present the main definitions used throughout the paper. In Section 3, we show that the forest induced by a minimal cycle-cutset in grids can always be transformed into a single connected component, i.e a single tree. We then use this result in Section 4 in order to improve the lower bound of [1] in some cases. Thus matching the lower bound with upper bound of [2] in these cases.

## 2 Preliminaries

**Definition 1** (Cycle cutset, Partition to trees)**.** Let $G = (V, E)$ be an undirected graph. A subset $C \subseteq V$ is a *cycle-cutset* in $G$ iff the graph $F$ induced by $V' = V \backslash C$ on $G$ is a forest. We define *the partition T of F to trees* as

$$T = \{t = (V_t, E_t) : \text{ t is a connected component of F}\}$$

i.e. $T$ is a set of (connected) trees, and $V'$ can be written as $V' = \biguplus_{t \in T} V_t$.

**Definition 2** (Tree-degree )**.** Let $c \in C$ be a vertex in cycle-cutset $C$ of $G$ and $t \in T$ a tree induced by $C$, and denote by $N(c)$ the neighbors of $c$ in the graph induced by $V' \cup \{c\}$ on $G$, we define *the tree-degree of c in t over* $C$ to be d, if $|N(c) \cap V_t| = d$ and for every other tree $t \neq t' \in T$ it holds that $|N(c) \cap V_{t'}| \leq 1$. In general, we say that the tree-degree of c is d if the condition above holds for some t, and that the tree-degree is undefined otherwise. If $d \geq 2$ we call $N(c) \cap V_t$ the *in-tree neighbors of c.*

**Definition 3** (Equivalent cutset vertices)**.** Let $c \in C$ be a cutset vertex and let $c' \in V$ be a vertex of the graph $G$. We say that $c'$ is *equivalent to c (under $C$),* if $C \backslash \{c\}$ is not a cutset, while $(C \backslash \{c\}) \cup \{c'\}$ is a cutset.

**Definition 4** (Induced degree)**.** Let $c \in C$ be a cutset vertex. The *induced degree of c under $C$* is the degree of $c$ in the graph induced by $V' \cup \{c\}$.

It can be shown that the following observations hold:

**Lemma 1.** ***a.*** *If c is cutset vertex of tree-degree 2, then every vertex along the path between its two in-tree neighbors is of equivalent to c.*

***b.*** *If c is a cutset vertex of tree-degree* 3*, then there exists a unique vertex c′ which is equivalent to c.*

***c.*** *A cutset induces a single tree iff the tree-degree of every cutset vertex is equal to its induced degree.*

**Definition 5** (Boundary of a tree)**.** Let $t \in T$ be a tree, we define *the boundary of t* to be all cutset vertices touching t and some other tree, and denote it B(t), i.e

$$B(t) = \{c \in C : N(c) \cap V_t \neq \emptyset, \, \exists t \neq t' \in T, \, s.t. \, N(c) \cap V_{t'} \neq \emptyset\}$$

## 3 Connectivity of the induced graphs in grids

**Theorem 1.** *Let $G$ be a grid graph and let C be a cutset such that the induced forest F is disconnected, i.e. $|T| \geq 2$, then there exists a a series of replacement moves, such that the resulting cutset $C'$ has no more elements than C, and the forest induced by $C'$ contains a single tree, i.e. $|C'| \leq |C|$ and the graph induced by $V \backslash C'$ is connected.*

**Corollary 1.** *In particular, it follow from Theorem 1 that there exists a minimal cutset that induces a connected tree.*

***Proof of Theorem 1****.* The proof follows immediately from the following Lemmas 2 and 3, and due to the space constraints we would not elaborate on the subject. □

**Lemma 2.** *Let G be a grid graph and let* **C** *be a cutset of G, that the induces a disconnected forest F, i.e. $|T| \geq 2$. Let $t \in T$ be a connected component of $F$. If there exists a vertex $c \in B(t)$ in the boundary of t with an tree-degree of 2, then c can be replaced with a vertex $c'$, that has a degree of 2 in the graph induced by $V' \cup \{c, c'\}$, thus reducing the number of connected components in the induced forest.*

***Proof of Lemma 2****.* Omitted due to space constraints □

**Lemma 3.** *Let G be a grid graph and let* **C** *be a cutset of G, that induces a disconnected forest F, i.e. $|T| \geq 2$. if there does not exist a tree $t \in T$ and a vertex $c \in B(t)$, having tree-degree (defined and) is equal or less than 2, then:*

***a.*** *for every tree $t \in T$, there exist at least 2 vertices in the boundary of t of tree-degree 3.*

***b.*** *There exists a series of replacement moves, such that the forest induced by the final cutset $C'$ is composed of less connected components than the original forest.*

Note that the conditions of Lemma 3 are the complementary of the conditions of Lemma 2 (ignoring cutset vertices of tree-degree less than 2, which can be trivially removed).

***Proof of Lemma 3***. The proof, which was omitted due to space constraints, makes use of the fact that a vertex in a grid can be equivalent to at most 2 cutset vertices of tree-degree 3, which follows from the following general claim. $\square$

**Lemma 4.** *Let $G = (V, E)$ is a planner graph, $v \in V'$ a vertex equivalent to cutset vertices $c_1, \dots, c_k \in C$, which are of tree-degree 3, and let $d$ be the degree of $v$ in the graph induced by $V' \cup \{c_i\}_{i=1}^{k}$, then $d \geq k+2$.*

*Proof of Lemma 4.* Omitted due to space constraints. $\square$

## 4 Improved lower bounds

We will use the results of the previous section in order to improve the known lower bound on the size of the minimal cutset of $M_{n,m}$. In particular, we will show that our lower bound is equal to the upper bound in these seleceted cases. In the following we denote by $olb_{n,m}$ the old lower bound of [1], i.e.

$$olb_{n,m} = \left\lceil \frac{(m-1)(n-1)+1}{3} \right\rceil$$

and by $nlb_{n,m}$ the new lower bound obtained by us.

**Lemma 5.** *Let $G := M_{n,m} = (V, E)$ be the $n \times m$ grid graph, and $C \subseteq V$ a cycle-cutset, such that the graph $T = (V_T, E_T)$ induced by $V' = V \backslash C$ is a single tree. Denote by $\alpha$ the number of cutset vertices which lie along the perimeter of the grid but not in its corners, i.e.*

$$\alpha = |C \cap (\{(1,j),(n,j) : 1 < j < m-1\} \cup \{(i,1),(i,m) : 1 < i < n-1\})|$$

*and by $\beta$ the number of cutset vertices which lie in the corner of the grid, i.e. $\beta = |C \cap \{(1,1),(1,m),(n,1),(n,m)\}|$. Denote by $n_C$ the cardinality of $C$, and by $p$ the number of connected components of the graph induced by $C$ (not $V \backslash C$). Then it holds that*

$$(n-1)(m-1) + \alpha + 2\beta \leq 2n_C + p \tag{1}$$

***Proof of Lemma 5***. Omitted due to space constraints. $\square$

We note that using the trivial facts that the number $p$ of connected components of $C$ is smaller than $|C| = n_C$, and that $\alpha + 2\beta \geq 1$, as there must be at least one cutset vertex along the perimeter of the grid, one receives from Lemma 5 that it holds that

$$(n-1)(m-1) + 1 \leq 3n_C$$

which is a restatement of the lower bound of [1].

Assume that $n \equiv r \mod 3$ and that $m \equiv s \mod 3$ ($0 \leq r, s \leq 2$), i.e. $n = 3q + r$ and $m = 3p + s$, and assume that w.l.o.g that $r \leq s$. Additionally, assume that $n_C = olb_{n,m}$, then it can be shown that

$$n_C = 3pq + p(r-1) + q(s-1) + \mathbb{1}[r \neq 0 \vee s \neq 2] \tag{2}$$

Plugging equation 2 in inequality 1 and rearranging, we get the following inequalities:

$$0 \leq n_C - p \leq 3 \cdot \mathbb{1}[r \neq 0 \vee s \neq 2] - (r-1) - \alpha - 2\beta \tag{3}$$

$$\alpha 2 + \beta \leq 3 \cdot \mathbb{1}[r \neq 0 \vee s \neq 2] - (r-1) \tag{4}$$

These inequalities are at the core of proving the improvements to the lower bounds of [1].

**Theorem 2.** *Let $m, n \geq 4$, such that $n \equiv 0 \mod 3$ and $m = 2 \mod 3$, and assume that at least one of $n$ and $m$ is even, then the size of the minimal cutset of the $n \times m$ grid $M_{n,m}$ (or the $m \times n$ grid $M_{m.n}$) is at least $olb_{n,m} + 1$, i.e. $nlb_{n,m} = olb_{n,m} + 1$.*

*Proof of Theorem 2 .* Omitted due to space constraints. □

Let $m, n$ be two integers at least one of which is even. Assume w.l.o.g. that $m$ is even and that $m \geq 6$. Using the upper bounds $ub_{n,m}$ of [2], we get that $nlb_{n,m} = ub_{n,m}$, if $n \geq 9$ and $n \equiv 0 \mod 3$ or if $n \geq 11$ and $n \equiv 2 \mod 3$, i.e. in every case in which [2] have shown an upper bound applicable in the conditions of Theorem 2, the upper bound is equal to the lower bound.

**Theorem 3.** *Let $m, n \geq 4$, such that both $n \equiv 0 \mod 3$ and $m = 0 \mod 3$ or both $n \equiv 2 \mod 3$ and $m = 2 \mod 3$, and assume that both $n$ and $m$ are even, then the size of the minimal cutset of the $n \times m$ grid $M_{n,m}$ (or the $m \times n$ grid $M_{m.n}$) is at least of size $olb_{n,m}+1$, i.e. $nlb_{n,m} = olb_{n,m}+1$.*

*Proof of Theorem 3.* Omitted due to space constraints. □

Let $m, n$ be two even integers, and assume that $m \geq 6$. Using the upper bounds $ub_{n,m}$ of [2], we get that $nlb_{n,m} = ub_{n,m}$, if $n \geq 9$ and $n \equiv 0 \mod 3$ or if $n \geq 11$ and $n \equiv 2 \mod 3$, i.e. in every case in which [2] have shown an upper bound applicable under the conditions of Theorem 3, the upper bound is equal to the lower bound.

## 5 Activate Cutset

We would like to use the notion of cycle-cutset in order to expand an optimal algorithm for energy minimization in trees presented in [3], thus receiving an approximate energy minimization algorithm that can handle arbitrary unary and binary potential over domains of finite size.

### 5.1 Problem definition

let $\bar{x} = x_1, \ldots, x_N$ be a set of variables over domain $\{1, \ldots, k\}$, for every $i \in \{1, \ldots, N\}$ let $\varphi_i : \{1, \ldots, k\} \to \mathbb{R}$ be an unary potential, and for every $1 \leq i < j \leq N$ let $\psi_{i,j} : \{1, \ldots, k\}^2 \to \mathbb{R}$ be a binary potential, then the problem of energy minimization is finding

$$\bar{x}^* = argmin_{\bar{x}} \sum_i \varphi_i(x_i) + \sum_{i<j} \psi_{i,j}(x_i, x_j)$$

Given an instance of the problem, it's "primal graph" is built by assigning a vertex to every variable $x_i$ is a vertex, and connecting two vertices $x_i$ and $x_i$ iff the matching potential $\psi_{i,j}$ is not identically zero.

### 5.2 Our approach

Although the problem is generally NP-hard, it is tractable in some cases, one of which is when the underling graph is cycle-free, i.e a forest. The algorithm for solving this case, presented in [3] as ACTIVATE, is essentially a form of belief propagation: After a root for every tree is selected, every vertex $v$ sends its parent $pa_v$ a message of the optimal assignment to the tree beneath $v$ given the value of $pa_v$. Once all the messages reach the root of the tree, the root selects an optimal assignment for itself and passes massages to its children to set their values accordingly. We will refer to this algorithm as the TREE-ALGORITHM.

If the graph does contain cycles, the cycles can be opened by finding a cycle-cutset $C$, setting the values of the variables of $C$ to some values, and using the exact TREE-ALGORITHM on the remaining (cycle-free) graph to solve the problem conditioned by the assignment to the cutset vertices. In our algorithm, named ACTIVATE-CUTSET, after the conditioned optimal solution is found, a new cutset is chosen and the process repeats. Cutset vertices either retain their values obtained by the exact TREE-ALGORITHM during a previous iteration (in which they were not a part of the cutset) or apply a local search algorithm in an attempt to improve the networks energy even further. The resulting algorithm is a local search algorithm, in which in every update step the entire complement of the cutset is updated optimally.

### 5.3 Cutset selection

The cutset selection algorithm used is based on the algorithm described in [4], where the cutset is built by randomly adding to it a vertex with probability based on its degree. Motivated by the ideas presented in [3] the probability a vertex is chosen in our implementation is governed not only by its degree, but by the number of iterations in which it has not changed its value and by the number of iterations it was not a part of the cycle-cutset as well.

In addition, in order to encourage the formation of forests in which the number of trees is (relatively) small and are composed mainly from a single big tree, the first stage of [4] is split to 2 stages: At first, the program tries to add to the cutset nodes which do not already have a node pointing to them as a parent, and only when no such nodes remains the program begins to add nodes which are pointed as the parent of another node, thus partitioning the forest created.

In order to reduce the cutset's size, after the forest is formed the partition of the forest into trees is explicitly found, and each cutset vertex of tree-degree less than 1 is removed from the cutset (and added to the forest nodes). Once all the redundant cutset nodes have been removed, the tree forming algorithm of [3] is run again (only on the forest nodes) in order to form a well directed forest.

### 5.4 Experiments

We have run ACTIVATE-CUTSET on the set of grid problems from problem set of the PASCAL2 Probabilistic Inference Challenge (PIC2011)[1]. The variables were initialized to the undefined value, thus effectively ignoring them until they attain a valid value, and in every update step only trees bigger than the 10% of the grid size were updated. This was done since some experiments suggested that initializing the variables randomly or updates small trees may have a tendency to draw the algorithm to high local minima. In addition, in order to avoid stagnation, if the assignment is not improved within 10 iterations, i.e. by 10 different cutsets, the algorithm is restarted by setting the cutset vertices to the undefined value. As a comparison, we have run $GLS^+$ [5], considered to be the state-of-the-art in energy minimization on the problem set. Each algorithm was run on each problem 10 times a bound amount of time. Since ACTIVATE-CUTSET was not implemented to employ a sophisticated initialization algorithm (such as the usage of mini-buckets by $GLS^+$), $GLS^+$ was set to a random initial assignment. Note that in the context of these problems an assignment obtaining **maximal** energy (sometimes referred to "Goodness") is required, and therefore higher values are better. Following are the energies obtained by each of the algorithm compared to the global maxima. For each problem instance Figure 1 depicts the average, minimal and maximal goodness obtained by ACTIVATE, in the four left bars, and the corresponding values obtained by $GLS^+$ in the right bars.

It can be seen in Figure 1 that for most problems ACTIVATE obtains a higher goodness than that obtained by $GLS^+$ at the same time. In general, we can see that ACTIVATE obtains higher energy faster.

## 6 Conclusion

In this work, we have established the basis to the notion of tree-inducing cycle-cutsets and the transformation of a general cutset to such a cutset. We have shown that in grids we can always transform a cycle-cutset to a tree-inducing cutset with no more vertices than the original one. These results lay the foundation to a more elaborate method of analyzing and bounding the size of minimal cutsets, thus allowing us to improve its lower bound in some cases. In other cases, a gap between the lower and the upper bounds remains, and more meticulous research should be undertaken in order to characterize better the classes in which the lower bound can be raised.

[1] http://www.cs.huji.ac.il/project/PASCAL/index.php

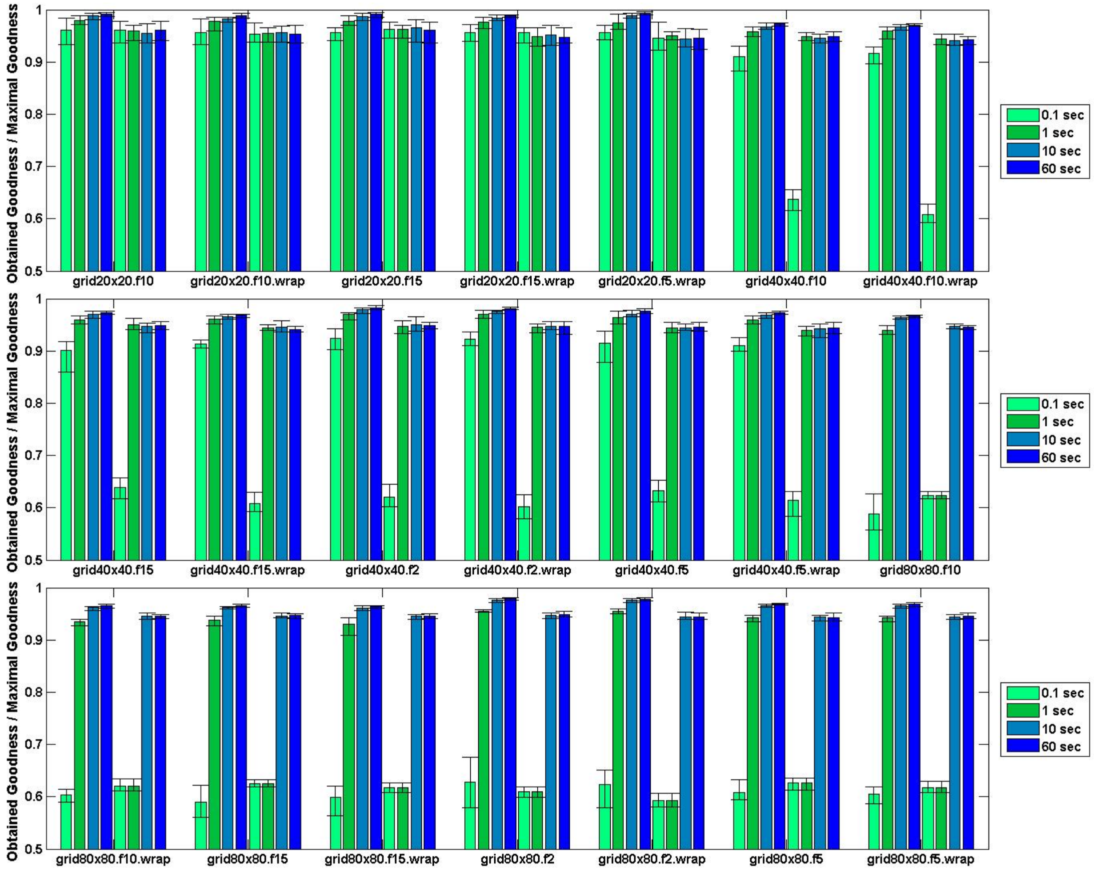


Figure 1: Average, minimal, and maximal ratio of obtained goodness to maximal goodness in grid problems of PIC2011 at several time bounds. In each set the 4 left bars correspond to ACTIVATE's results, and the 4 right bars correspond to the results of $GLS^{+}$.

In addition, we presented an algorithm which combines the notion of cycle-cutset with the well known Belief Propagation algorithm to obtain an approximate optimum of a sum of unary and binary potentials. This is done by the rather novel concept of traversal from one cutset to another and updating the induced forest, thus creating a local search algorithm, whose update phase spans over many variables. We have presented experiments indicating that this algorithm is on-par with the state-of-the-art in this domain (if not surpasses it) on some restricted problems of grids and when both algorithms use a elementary method of initialization. In this regard the algorithm should be further investigated in order to understand more fully the parameters governing its behavior. Additionally, the algorithm should be extended to handle potentials of higher arity than 2.

## References


[1] F.L. Luccio, “Almost exact minimum feedback vertex sets in meshes and butterflies”, *Information Processing Letters, Vol.* 66**,** (1998) 59–64.

[2] F. Madelaine and I. Stewart, “Improved upper and lower bounds on the feedback vertex numbers of grids and butterflies”, *Discrete Math.* , vol. 308, no. 18, (2008) 4144–4164.

[3] G. Pinkas and R. Dechter, "Improving Connectionist Energy Minimization", *Journal of Artificial Intelligence Research*, Vol. 3, (1995) 223-248.

[4] A. Becker, R. Bar-Yehuda, D. Geiger, “Randomized Algorithms for the Loop Cutset Problem”, *Journal of Artificial Intelligence Research*, Vol. 12, (2000) 219-234.

[5] F. Hutter, H.H. Hoos and T. Stützle, “Efficient stochastic local search for MPE solving”, Proc. of IJCAI-05, (2005) 169-174.