Technische Universität Berlin

FAKULTÄT II
MATHEMATIK UND
NATURWISSENSCHAFTEN
Institut für Mathematik

BENCHMARKS FOR STRICTLY FUNDAMENTAL CYCLE BASES

by

CHRISTIAN LIEBCHEN, GREGOR WÜNSCH,
EKKEHARD KÖHLER, ALEXANDER REICH,
AND ROMEO RIZZI

No. 003/2007

Benchmarks for Strictly Fundamental Cycle Bases

Christian Liebchen, Gregor Wünsch,
Ekkehard Köhler, Alexander Reich,
and Romeo Rizzi

5th February 2007

Abstract

In the MINIMUM STRICTLY FUNDAMENTAL CYCLE BASIS (MSFCB) problem one is looking for a spanning tree such that the sum of the lengths of its induced fundamental circuits is minimum.

We identify square planar grid graphs as being very challenging testbeds for the MSFCB. The best lower and upper bounds for this problem are due to Alon, Karp, Peleg, and West (1995) and to Amaldi et al. (2004).

We improve significantly their bounds, both empirically and asymptotically. Ideally, these new benchmarks will serve as a reference for the performance of any new heuristic for the MSFCB problem which will be designed only in the future.

1 Introduction

Consider the following problem. Given the $N \times N$ square planar grid graph $G_{N,N}$. Find a spanning tree T such that the sum of the lengths of its induced fundamental circuits is as small as possible. In Figure 1 we provide a very good solution for $G_{8,8}$. Is this optimal?

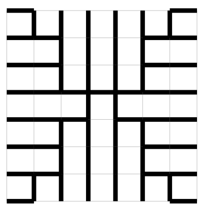


Figure 1: A very good SFCB of $G_{8,8}$. It costs 266. Can you give a cheaper one?

At first sight, this might appear being a kind of “toy problem.” Indeed, at the occasion of its annual web-based Christmas quiz (`www.mathe-kalender.de`), on December 18, 2006 the DFG Research Center MATHEON essentially asked the above question to more than 9 000 registered users (pupils, teachers, scientists, and others). Typically, each day about 1 500 users post their answers, and more than 60% of these answers are correct. In contrast, on Dec. 18, less than every *eighth* answer has been correct—a first indicator for the trickiness of this particular problem.

The fundamental circuits with respect to some spanning tree in a general graph form a strictly fundamental cycle basis, where we refer to Section 2 for any formal definition. We refer to the problem of finding a spanning tree whose fundamental circuits sum to a minimum value as the MINIMUM STRICTLY FUNDAMENTAL CYCLE BASIS (MSFCB) Problem. As a generalization, in the MINIMUM CYCLE BASIS (MCB) Problem one seeks for a general cycle basis of minimum length.

Applications. The MCB problem has many applications. These include biology and chemistry ([11]), traffic light planning ([15]), periodic railway timetabling ([17]), and electrical engineering ([6]). Typically, cycle bases are computed as a kind of preprocessing. Then, throughout the actual computations one ensures that a certain problem-specific property is true for the elements of the cycle basis in the graph of interest. By this, one can conclude that this property is actually true for *any* cycle in the graph, right as it is required by the practical application. In many cases, the shorter the used cycle basis, the shorter the time for the actual computations.

For some of these applications, due to structural reasons not all cycle bases are of use (e.g. traffic light planning and periodic railway timetabling), but strictly fundamental cycle bases—being the most specialized ones—always are. In other applications, such as electrical engineering, it is at least much more favorable to use strictly fundamental cycle bases, because of the numerical stability of the subsequent calculations ([3]). The practical relevance of the MSFCB problem is also reflected by numerous computational studies by different groups working in combinatorial optimization ([2, 7, 8, 9, 12, 16, 20]). We shortly overview these works and their findings.

Theory. As early as 1982, Deo et al. ([7]) proved the MSFCB problem to be NP-hard for general unweighted graphs. Yet, the many applications require solutions to be generated anyway. Hence, many heuristics were proposed and tested. However, for none of these heuristics neither any non-trivial approximation ratio nor any non-trivial bound on the absolute length of the resulting bases was proposed. The only statement into that direction is that Deo et al. ([7]) conjecture MSFCBs of unweighted graphs to have length $O(n^2)$.

The design of most of these heuristics has been led by the following observation: "A BFS produces spanning trees of short diameters. Thus, the BFS method on the average generates fundamental cycles of shorter total length (compared to some other approaches)." ([7]). In particular, these heuristics make local decisions that are mainly based on the degrees of the vertices, either in G or in some residual graph.

But also there can be applied totally different techniques. Actually, Elkin et al. ([10]) consider some average-stretch tree spanner problem. Profiting from the Unified Notation for Tree Spanner problems (UNTS, [19]) one can easily see that in the case of unweighted graphs their results apply immediately to the MSFCB problem. In particular, their recursive algorithm computes a SFCB of asymptotic length $O(m \cdot \log^2 n \log\log n)$. Let us emphasize that this is the first non-trivial theoretical guarantee on the quality of a solution to the SFCB problem, and it is obtained by a recursive approach. Moreover, for graphs with $|E| \in O(|V|^{2-\varepsilon})$, this result proves Deo's conjecture.

Why Planar Grids? In the absence of theoretical bounds for the many degree-based heuristics, their authors used empirical calculations to assess their quality. But to compare different heuristics empirically, these must be run on the very same input graphs.

But what are *good* such testbeds?

Liberti et al. ([16]) consider square planar grid graphs being "the most difficult testbeds for the MSFCB problem, both for heuristic and exact methods, due to the huge quantity of configurations having the same SFCB cost." In fact, also from a theoretical perspective this can be motivated in three ways. First, these graphs are almost regular because more than $n-4\sqrt{n}$ vertices have degree four—a nightmare for any degree-based heuristic. Second, within a fixed distance, the subgraphs around almost each vertex are isomorphic. Hence, any heuristic that bases its decisions on local configurations risks to perform poorly. Third, if G was a tree, then in the MSFCB problem no decisions are to be made and the problem clearly becomes trivial. An appropriate measure for the tree-alikeness of a graph is its tree-width ([4]). And with respect to that measure, grid graphs—having $\Theta(n)$ edges and tree-width $\sqrt{n}$—are prominent examples of being far away from being a tree ([22]). Thus the MSFCB problem is likely to keep its hardness. In addition, in several applications relevant instances are planar graphs, sometimes even grids (e.g. electrical engineering, traffic light scheduling).

Focusing on grid graphs could appear narrow. But it is commonly believed that these hold the key to better algorithms. Indeed, for square planar grid graphs Alon, Karp, Peleg, and West ([1]) design spanning trees of length $\frac{4}{3}n\log_2 n + O(n)$ ([14]). They prove these trees to be asymptotically optimal. Moreover, they conjecture their trees are "essentially optimal." This asymptotical upper bound lets us demonstrate how degree-based heuristics may fail. The degree-based $\mathcal{C}$-order heuristic can be implemented to compute "Machete"-trees (cf. [5], and Figure 1 for an example). These trees do not only minimize the diameter of trees in grids, but also the maximum stretch. At first sight these two parameters of a tree could appear being tightly related to its associated SFCB cost. It is again the UNTS ([19]) which makes it transparent that even for unweighted graphs no two of these three measures do actually coincide. As Machete-trees yield MSFCB objective values of $\Theta(n^{\frac{3}{2}})$, on grid graphs degree-based heuristics risk to fail drastically. This underlines grids being a relevant testbed.

As a matter of fact, Liberti et al. ([16]) select grid graphs as one of their testbeds. On the 50×50 grid they observe that their new $\mathcal{C}$-order heuristic attains an objective value of 46 452, compared to 48 254 of the NT heuristic. Unfortunately, from this isolated comparison it has to remain unclear whether these are good objective values at all. In fact, Amaldi et al. ([2]) also consider grid graphs in their computations. And they report a solution of objective value 23 026, that was obtained by local search techniques. This motivates the need for clear benchmark values for the MSFCB problem for the particularly challenging case of planar grid graphs—also for the future evaluation of new heuristics.

Of course, relevant benchmarks also include dual bounds. Since general cycle bases are a superset of strictly fundamental cycle bases, the value of an MCB clearly serves as a lower bound for the value of an MSFCB. On grid graphs, this yields a lower bound of $4 \cdot (\sqrt{n}-1)^2$. But exploiting the particular structure of grid graphs one can achieve asymptotically better lower bounds for the MSFCB problem. The first was given in [1] and it has value $\frac{\ln 2}{2048} n \log_2 n - O(n)$.

Contribution. The above discussion motivates a need for a collection of benchmark values for the MSFCB problem on square planar grid graphs. We provide two new families of lower bounds and two new families of upper bounds.

In Section 3 we sketch a proof of Köhler et al. ([14]) on how a new approach raises the asymptotical lower bound by Alon, Karp, Peleg, and West ([1]) to $\frac{1}{12}n\log_2 n - O(n)$, i.e. by a factor of more than 245. In addition, we identify $6n - 20\sqrt{n} + 22$ as a new lower bound. For $N \in \{3, \ldots, 61\}$ this constitutes the best known lower bound. It is a fact that *all* the primal solutions (upper bounds) that so far have been proposed in the literature are grids which fall within this range.

Finally, in Section 4 we introduce a new scheme for constructing very short strictly fundamental cycle bases—both empirically and asymptotically. We prove an upper bound on the length of their SFCB of $0.97n\log_2 n + O(n)$, hereby improving the objective value $\frac{4}{3}n\log_2 n + O(n)$ of the spanning trees due to Alon, Karp, Peleg, and West ([1]), which they assumed being essentially optimal. In our experiments we also compare their lengths to spanning trees that were obtained by using local search techniques ([2]). It turns out that our new trees improve the best solutions known so far for all $N \geq 20$. Interestingly, for $N = 10, 15, \ldots, 55$ they even constitute local optima with respect to the 2-neighborhood.

2 Preliminaries

We consider cycle bases of a 2-connected simple undirected unweighted[1] graph $G = (V, E)$. Define $n = |V|$, $m = |E|$, and $\nu = m - n + 1$, where ν is the *cyclomatic number* of G. Let C be a circuit (cf. [23, Ch. 3]) in G and denote by γ_C its $\{0,1\}$-incidence vector. The *cycle space* $\mathcal{C}$ of G is the following vector subspace over GF(2),

$$\mathcal{C} := \operatorname{span}\left(\{\gamma_C \,|\, C \text{ circuit in } G\}\right).$$

A *cycle basis* B of G is a set of ν circuits of G whose incidence vectors are a basis of $\mathcal{C}$. The *length* $\Phi(B)$ of a cycle basis of an unweighted graph is defined as $\Phi(B) = \sum_{C \in B} |C|$. A *minimum cycle basis* (MCB) of a graph G is a cycle basis of G of minimum length.

A set of circuits $\{C_1, \ldots, C_\nu\}$ such that

$$C_i \setminus (C_1 \cup \cdots \cup C_{i-1}) \neq \emptyset, \quad \forall i = 2, \ldots, \nu$$

is clearly a cycle basis. We call such a basis *weakly fundamental*. Notice that these were already considered by Whitney ([24]) in 1935.

Let T be some spanning tree of G. Depending on the context, we either regard T as a subgraph of G or as a set of edges $T \subset E$. For $e \in E \setminus T$, we denote by $C_T(e)$—or C_e for short—the *fundamental circuit* that e induces with respect to T, i.e. the unique circuit in $T \cup \{e\}$. To T there are associated ν fundamental circuits. These form a cycle basis which is called *strictly fundamental*. Here, we may write $\Phi(T)$ instead of $\Phi(B)$. A *minimum strictly fundamental cycle basis* (MSFCB) has minimum length among the

[1] Of course, minimum cycle basis problems are also investigated for weighted graphs. But as we aim to contribute to the particularly challenging case of planar unweighted grid graphs, we omit edge weights throughout our presentation.

set of strictly fundamental cycle bases. In the context of local search, for an arbitrary spanning tree T we define its 2-*neighborhood* as the set of spanning trees T' such that $|T \cap T'| \geq |T| - 2$.

In general, strictly fundamental cycle bases are a proper subset of weakly fundamental cycle bases, which in turn are a proper subset of general cycle bases of undirected graphs. Moreover, in general no two of the three corresponding minimization problems coincide ([18]).

With $N \in \mathbb{N}$, the planar *grid graph* $G_{N,N}$ is the graph on $V = \{1,\ldots,N\} \times \{1,\ldots,N\}$ with

$$E = \{\{(i,j),(i',j')\} : |i-i'| + |j-j'| = 1\} = \{\{u,v\} : ||u-v||_1 = 1\}.$$

In a graphical representation, e.g. in an embedding into $\mathbb{Z}^2$, the first index of a vertex represents its x-coordinate, the second index its y-coordinate. The graph $G_{N,N}$ has $n = N^2$ vertices and contains $m = 2 \cdot N \cdot (N-1)$ edges. Its cyclomatic number ν is $(N-1)^2$.

We denote the dual of an embedded planar graph G by G^*. The graph $(G_{N,N})^*$ is again the graph of a square $(N-1) \times (N-1)$ grid plus a further vertex F^∞, which corresponds to the outer face of the initial embedded planar graph. Recall from [23, Ch. 3] that the edge set of G can be identified with the edge set of G^*.

Now, consider a spanning tree T of $G_{N,N}$ and its dual counterpart, that we denote by T^*. In fact, T^* can be understood as the complement of T, as it contains the counterpart in G^* of each edge in $E(G_{N,N}) \setminus T$. The graph T^* is a spanning tree of G^*, although it is not necessarily connected when restricted to $G^* \setminus \{F^\infty\}$.

3 New Lower Bounds

Trivial lower bounds for the MSFCB problem are the length of a minimum weakly fundamental cycle basis, or even of an MCB ([24, 18]). Whereas the former is in general APX-hard to find ([21]), for the latter there are known polynomial-time algorithms (e.g. [13]). However, as a consequence of a result due to Alon et al. ([1]) one can conclude that these lower bounds risk to miss the optimum value of the MSFCB problem by a logarithmic-factor. And this is in particular the case for square planar grid graphs $G_{N,N}$. Here, the trivial lower bound is only $4 \cdot (\sqrt{n} - 1)^2$ whereas Alon et al. proved $\Phi(T) \geq \frac{\ln n}{2048} n - O(n)$ for all spanning trees T.

An alternative way to obtain lower bounds is to consider MIP formulations of the MSFCB problem (Liberti et al., [16]). Later, in [2] there was identified an improved MIP formulation, which turns out to be more efficient in their empirical computations. With this MIP formulation, we effected some spot tests on square planar grid graphs. Unfortunately, we had to observe that it is already *very* hard for standard MIP solvers—such as CPLEX© 10.1—to get beyond the trivial lower bound. This is why, in Sect. 3.1 we identified several classes of valid inequalities for these MIPs, some of them being defined even for general graphs. Indeed, these helped MIP solvers to detect better lower bounds.

Yet, in Corollary 6 we identify $6n - 20\sqrt{n} + 22$ being a lower bound for the MSFCB problem on a square planar grid, using purely combinatorial arguments. And then again, it got completely unpractical for the MIP—even in our refined version—to improve on this lower bound.

Of course, there exists some dimension N_0 such that our lower bound gets dominated by the asymptotically better lower bound by Alon et al. ([1]). But in the case of $N = 2^k + 2$, $k \in \mathbb{N}$ we also present one further lower bound function with

$$\Phi(T) \geq \frac{1}{12} n \log_2 n - O(n), \tag{1}$$

which is due to Köhler et al. ([14]). Clearly, it dominates Alon et al.'s lower bound. And it illustrates the predominance of the $6n - 20\sqrt{n} + 22$ bound over the trivial lower bound: The function that interpolates the asymptotically best lower bound (1) intersects with the trivial lower bound as early as $N_1 \approx 8.1$. But $6n - 20\sqrt{n} + 22$ intersects (1) only at $N_2 \approx 61.6$, i.e. in the case of more than 7 300 edges.

3.1 A MIP Formulation

In this section we qoute the MIP formulation by Amaldi et al. ([2]) for the MSFCB problem on general graphs. Moreover, we make it more efficient by identifying the first classes of valid inequalities. In particular, for planar grids two of these classes are even able to cut off *any* of the huge number of optimum solutions of the LP relaxation, hereby improving the lower bound of the root node in the Branch-and-Bound tree.

Let $G = (V, E)$ be a 2-connected graph with non-negative costs w_e on an edge $e \in E$. To ensure a spanning tree T to be computed, we resort on the following characterization: $|T| = |V| - 1$, and T is connected. We exploit the fact that T is connected, if and only if for each non-tree edge $e = \{i, j\} \in E \setminus T$ there exists a path in T between i and j.

Amaldi et al. introduce a binary variable z_{ij} for each edge $\{i, j\} \in E$, where $z_{ij} = 1$ iff $e \in T$. Of course, the correct cardinality of T is easy to state. Then, they are going to ensure the non-tree edge connectivity by introducing non-negative variables x, which are chosen to be well-suited to state the objective function of the MSFCB problem. The variables x can be understood as a multi-commodity flow in G, only using edges of T. For each edge $e = \{k, \ell\} \in E$, its endpoints are regarded as source and sink of a commodity for which one unit of flow is to be sent. As this flow may only use tree edges, the commodities in $E \setminus T$ guarantee T being connected. To state flow conservation, a directed graph D is derived from G, whose arc set consists of a pair of antiparallel arcs for each edge in $E(G)$, and the cost w_{ij} of an arc (i, j) is set equal to the cost of the edge $\{i, j\}$. Then, the variable $x_{ij}^{k\ell}$ encodes the directed flow through arc (i, j) for commodity $\{k, \ell\}$.

$$
\begin{aligned}
\min \sum_{\{k,\ell\}\in E} \sum_{(i,j)\in A} w_{ij} x_{ij}^{k\ell} \quad &+ \quad \sum_{\{i,j\}\in E} (1-2z_{ij})\, w_{ij} && \text{(2a)}\\
\sum_{j\in\delta(k)} (x_{kj}^{k\ell} - x_{jk}^{k\ell}) \quad &= \quad 1 \qquad \forall\{k,\ell\}\in E && \text{(2b)}\\
\sum_{j\in\delta(i)} (x_{ij}^{k\ell} - x_{ji}^{k\ell}) \quad &= \quad 0 \qquad \forall\{k,\ell\}\in E, \forall i\in V\setminus\{k,\ell\} && \text{(2c)}\\
x_{ij}^{k\ell} \quad &\leq \quad z_{ij} \qquad \forall\{k,\ell\}\in E, \forall\{i,j\}\in E && \text{(2d)}\\
x_{ji}^{k\ell} \quad &\leq \quad z_{ij} \qquad \forall\{k,\ell\}\in E, \forall\{i,j\}\in E && \text{(2e)}\\
\sum_{\{i,j\}\in E} z_{ij} \quad &= \quad n-1 && \text{(2f)}\\
x_{ij}^{k\ell} \geq 0 \qquad & \forall\{k,\ell\}\in E, \forall(i,j)\in A && \text{(2g)}\\
z_{ij} \in \{0,1\} \qquad & \forall\{i,j\}\in E. && \text{(2h)}
\end{aligned}
$$

In any integer feasible solution, within the first term of the objective function (2a) we find that

$$\sum_{a\in A} w_a x_a^e \geq d_T(e), \quad \text{for all } e\in E, \tag{3}$$

where in any optimum solution equality holds.

Although the MIP formulation (2) has been observed to behave better than other formulations ([16]), still there are some major shortcomings. First, the number of variables and constraints is large. For instance, there are $2\cdot m^2$ x-variables—in other words $\Theta(N^4)$. Already with this simple observation one might not expect too much for the solvability with, say, $N\geq 20$. But the second drawback is even worse. The LP relaxation has several trivial optimum solutions. For instance, take $z\equiv \frac{1}{2}+\frac{1}{2N}$. This particular choice admits the x-variables to sum up to $4\cdot(N-1)^2$, being the optimum value of the minimum weakly fundamental cycle basis problem on $G_{N,N}$. We will provide another set of optimum solutions of the LP relaxation in Example 2. Of course one can check this to be the optimum value of the LP relaxation by having a look at the dual problem. We conclude, adding valid inequalities to (2) will be key for its solvability.

Thus, in the remainder of this chapter we provide three classes of valid inequalities: two which are valid for general graphs and which are defined either in z-variables or in x-variables, and one class that exploits the particular structure of grid graphs and hereby can combine x- and z-variables.

Lemma 1. Consider the graph $G=(V,E)$ and an arbitrary proper subset H of E. Denote by $V(H)$ the vertices incident to edges in H. If $V(H)\subsetneq V$, then

$$\sum_{\substack{\{i,j\}\in(V(H),\overline{V(H)})\\ \{i,j\}\notin H,\ i,j\in V(H)}} z_{ij} \geq |V(H)| - \sum_{\{i,j\}\in H} z_{ij} \tag{4}$$

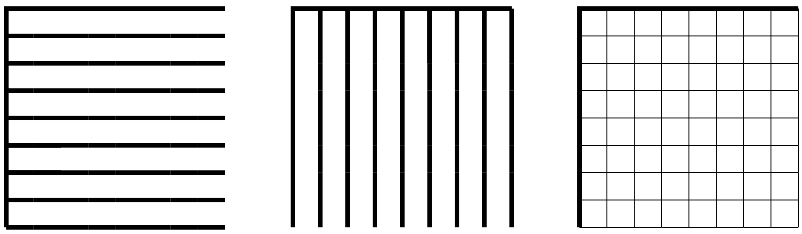
are valid for every integer feasible solution of the MIP.

Proof. In an integer feasible solution the edges with $z_{ij} = 1$ form a spanning tree T of G. Therefore, the right-hand side (RHS) of (4) equals the number of connected components of the graph with vertex set $V(H)$ and edge set H. As we assume $V \setminus V(H)$ to be nonempty, each component of $(V(H), H)$ must be reachable from the former vertex set. Hence, to ensure the connectivity of G via edges for which $z_{ij} = 1$, there must be at least as many such edges, as $(V(H), H)$ has connected components. □

On the one hand, we are aware of graphs in which there exist fractional vectors (x, z), which are feasible for the LP relaxation and whose objective value is strictly smaller than that of an integer optimum solution. Hence, these inequalities could appear to be reasonable candidates to add to the LP-relaxation of MIP (2).

On the other hand, it may happen that in an iterative cutting plane generation, one could only profit from inequalities of that type at rather late iterations.

Example 2. Consider the grid graph $G_{N,N}$. In Figures 2(a) and 2(b) we sketch two spanning trees T_1 and T_2 of $G_{N,N}$. Denote the corresponding solutions of the MIP (2) by (x_1, z_1) and (x_2, z_2), respectively.



(a) T_1 (b) T_2 (c) T_3

Figure 2: The Figures 2(a) and 2(b) depict the z vectors of the MIP solutions of two spanning trees T_1 and T_2 (in **bold**) for $G_{9,9}$. Figure 2(c) shows their convex combination $z_3 = \frac{1}{2}z_1 + \frac{1}{2}z_2$

Now, consider the convex combination of (x_1, z_1) and (x_2, z_2) which results in the following fractional vector $(x, z) := \frac{1}{2}(x_1, z_1) + \frac{1}{2}(x_2, z_2)$. Clearly, (x, z) is a feasible solution for the LP relaxation of MIP (2). But observe that there exists a vector (x', z) which is feasible for the LP, too, but in which

$$\sum_{\{k,\ell\} \in E} \sum_{(i,j) \in A} x'^{k\ell}_{ij} = \begin{cases} 1, & \text{if } z_{k\ell} = 1 \text{ and} \\ 2, & \text{otherwise.} \end{cases} \tag{5}$$

In particular, the objective value of (x', z) equals $4 \cdot (N-1)^2$. But this is just the optimum value of the LP relaxation of MIP (2). As z is the convex combination of two integer feasible points, any inequality that does not make use of any x-component, will never cut off (x', z), thus never increasing the LP value.

From the above example we conclude that for planar grids no valid inequality having non-zero coefficients *only* in z-components whatsoever, will ever be able to cut off one particular optimum solution of the LP relaxation, hereby never increasing the optimum value of any so refined LP.

This is why in the sequel we investigate valid inequalities which are either defined purely in terms of x-variables, or as a combination of x- and z-variables.

Lemma 3. Let $G=(V,E)$ be a 2-connected graph with a spanning tree T and consider a simple circuit C in G. Then,

$$\sum_{e \in C} d_T(e) \geq 2 \cdot (|C|-1). \tag{6}$$

Proof. Let T denote an arbitrary but fixed spanning tree of G. In the following we will prove the claim by induction over $|C \setminus T|$. Notice first that $|C \setminus T| = 0$ would imply $C \subseteq T$, which contradicts T being cycle-free.

Therefore, we select $|C \setminus T| = 1$ as the inductive base. In this case, the distances $d_T(e)$ in the tree are one for the $|C|-1$ tree edges, and $|C|-1$ for the unique non-tree edge. Hence, the claim holds.

In the inductive step, take a circuit C for which $|C \setminus T| = k \geq 2$. We will identify two circuits C_1 and C_2, each with $1 \leq |C_i \setminus T| < k$, $i=1,2$. Then, from (6) being true for C_1 and C_2 we argue that the claim is true for C.

Consider two vertices u and v within C which are connected through a path $P \subseteq T$ such that $V(P) \cap V(C) = \{u,v\}$, but $E(P) \cap E(C) = \emptyset$. Such a path exists because of $|C \setminus T| \geq 2$, otherwise T would not be connected. Denoting by P_1 and P_2 the two paths between u and v defined by C, $C_1 = P_1 \cup P$ and $C_2 = P_2 \cup P$ are simple circuits in G.

Now, as T is cycle-free, both P_1 and P_2 must include at least one non-tree edge, say e_{P_1} and e_{P_2}. Otherwise T would have contained a cycle. But then, C_1 contains at least one non-tree edge less than the circuit C, because it omits P_2, thus e_{P_2}, and the path P contains only tree-edges. The very same holds for C_2.

We may thus apply the inductive assumption to C_1 and C_2. Summing up (6) for these two circuits yields

$$\sum_{e \in C_1} d_T(e) + \sum_{e \in C_2} d_T(e) \geq 2 \cdot (|C_1| + |C_2| - 2).$$

By the construction of C_1 and C_2 it holds that $|C| = |C_1| + |C_2| - 2 \cdot |C_1 \cap C_2|$, and thus

$$\sum_{e \in C} d_T(e) + 2 \cdot \sum_{e \in C_1 \cap C_2} d_T(e) \geq 2 \cdot (|C|-1) + 4 \cdot |C_1 \cap C_2| - 2. \tag{7}$$

But since $C_1 \cap C_2 = P \subseteq T$ we have that $d_T(e) = 1$ for all $e \in C_1 \cap C_2$ and Equation (7) simplifies to

$$\sum_{e \in C} d_T(e) \geq 2 \cdot (|C|-1) + 2 \cdot |C_1 \cap C_2| - 2. \tag{8}$$

Finally, because of $|C_1 \cap C_2| \geq 1$, and thus $2 \cdot |C_1 \cap C_2| - 2 \geq 0$, Equation (8) implies (6) for the circuit C. $\square$

Corollary 4. Let $G=(V,E)$ be a 2-connected graph and consider a simple circuit C in G. Then,

$$\sum_{e \in C} \sum_{f \in A} x_f^e \geq 2 \cdot (|C|-1) \tag{9}$$

is a valid inequality for every integer feasible solution of the MIP (2).

Note that the above corollary holds for general graphs. Yet, we are interested most in investigating its effect on grid graphs.

3.2 Small Grids

Let C be some circuit in $G_{N,N}$. We denote by $\operatorname{diam}_H(C)$ the *horizontal diameter* of C, i.e. the difference between minimum and maximum x-coordinates in $\mathbb{Z}^2$ of vertices in C. Similarly, we denote the *vertical diameter* of C by $\operatorname{diam}_V(C)$. In particular,

$$|C| \geq 2 \cdot (\operatorname{diam}_H(C) + \operatorname{diam}_V(C)). \tag{10}$$

For $e \notin T$, we use $\operatorname{diam}_H(e) := \operatorname{diam}_H(C_T(e))$ as a short hand.

Let C be a circuit in $G_{N,N}$ and consider its enclosed finite region R. In $\mathring{C}$ we collect all the edges in $E(G_{N,N})$ that are incident with two faces of $G_{N,N}$ that have empty intersection with $\mathbb{R}^2 \setminus R$. In other words, $\mathring{C}$ refers to the edges *inside* C.

Proposition 5. Let $G_{M,N}$ be the $M \times N$ planar grid, T be a spanning tree in it, and C be a simple circuit in $G_{M,N}$. Then

$$\sum_{e \in (C \cup \mathring{C}) \setminus T} |C_T(e)| \geq 4 \cdot |C \setminus T| + 6 \cdot |\mathring{C} \setminus T|. \tag{11}$$

Proof. Using (10) it suffices to establish that

$$\sum_{e \in (C \cup \mathring{C}) \setminus T} 2 \cdot (\operatorname{diam}_H(e) + \operatorname{diam}_V(e)) \geq 4 \cdot |C \setminus T| + 6 \cdot |\mathring{C} \setminus T|. \tag{12}$$

We derive a lower bound for $\sum_{e \in (C \cup \mathring{C}) \setminus T} \operatorname{diam}_H(e) + \operatorname{diam}_V(e)$ by defining a function $d(e)$ such that

$$\operatorname{diam}_H(e) + \operatorname{diam}_V(e) \geq d(e), \quad \text{for all } e \in (C \cup \mathring{C}) \setminus T. \tag{13}$$

We define the function $d(e)$ as follows. By $e \notin T$ we already know that

$$\operatorname{diam}_H(e) \geq 1 \quad \text{and} \quad \operatorname{diam}_V(e) \geq 1. \tag{14}$$

To increase $d(e)$ beyond two, consider the spanning tree T^* in the dual graph $(G_{N,N})^*$ that corresponds to $E(G_{N,N}) \setminus T$. Take F^∞ as the root of T^*. Consider the two faces of $G_{N,N}$ that are incident with e. We refer to the one with the larger distance from F^∞ in T^* as $F(e)$.

For each edge $f \in (F(e) \setminus (C \cup \{e\})$, denote by $F(f) \neq F(e)$ the other face that f is incident with. Observe that $F(f) \neq F^\infty$ because of $f \notin C$. By the grid structure, each of these faces $F(f)$ is in a different direction with respect to $F(e)$, i.e. either north, east, south, or west. Now, if $f \notin T$, we know that $C_T(e)$ also must have $F(f)$ in its enclosed bounded region. This way, such an edge $f \in (F(e) \setminus (C \cup T \cup \{e\}))$ serves as a certificate that any lower bound on either $\operatorname{diam}_H(e)$ or $\operatorname{diam}_V(e)$, respectively, can be incremented. In total, we set

$$\operatorname{diam}_H(e) + \operatorname{diam}_V(e) \geq 2 + |F(e) \setminus (C \cup T \cup \{e\})| =: d(e), \tag{15}$$

which guarantees (13) still to be true.

When summing over all edges $e \in C \cup \mathring{C}$, we may rearrange the summation. To this end, observe that each edge $f \in \mathring{C} \setminus T$ has precisely one dual parent e among $(C \cup \mathring{C}) \setminus T$. Hence, it increments the lower bound on $\mathrm{diam}_H(e) + \mathrm{diam}_V(e)$ for precisely this one edge e. In other words, each edge $f \in \mathring{C} \setminus T$ counts three times: according to (14) it counts $d(f) = 2$ for its proper fundamental circuit $C_T(f)$, *plus* one increment for precisely its unique parent edge $e \in (C \cup \mathring{C}) \setminus T$. To summarize,

$$\sum_{e \in (C \cup \mathring{C}) \setminus T} d(e) \overset{(15)}{=} 2 \cdot |C \setminus T| \; + \; 3 \cdot |\mathring{C} \setminus T|. \tag{16}$$

Finally, we conclude that

$$\begin{aligned} \sum_{e \in (C \cup \mathring{C}) \setminus T} |C_T(e)| \quad & \overset{(10)}{\geq} \quad \sum_{e \in (C \cup \mathring{C}) \setminus T} 2 \cdot (\mathrm{diam}_H(e) + \mathrm{diam}_V(e)) \\ & \overset{(13)}{\geq} \quad \sum_{e \in (C \cup \mathring{C}) \setminus T} 2 \cdot d(e) \quad \overset{(16)}{\geq} \quad 4 \cdot |C \setminus T| \; + \; 6 \cdot |\mathring{C} \setminus T|. \end{aligned}$$

□

Corollary 6. Let $N \geq 3$ and $G_{N,N}$ be the $N \times N$ planar grid with $n = N^2$ vertices. Then for each spanning tree $T \subset E$

$$\Phi(T) = \sum_{e \in E \setminus T} |C_T(e)| \geq 6 \cdot n - 20\sqrt{n} + 22. \tag{17}$$

Proof. Simply take C as the circuit that contains precisely the edges that are incident with F^∞. Because of $E = C \dot{\cup} \mathring{C}$ we apply Proposition 5 to C. There, we minimize the RHS in (11) by maximizing $|C \setminus T|$. Now consider the four vertices which are not incident to any edge $\mathring{C}$. In any tree T, these must be incident with one edge $C \cap T$. As $N \geq 3$, we conclude that $|C \cap T| \geq 4$, thus $|C \setminus T| \leq 4\sqrt{n} - 8$. Finally, a simple calculation yields (17). □

Observe that $6n - 20\sqrt{n} + 22 \geq 4 \cdot n - 8\sqrt{n} + 4$, for all $N \geq 3$ and $n = N^2$.

As a special case, consider a spanning tree T in which for each edge $e \in E \setminus T$ its distance in T^* from F^∞ is at most two. Then, in particular in (15) equality holds. In the end, one can then argue that in (17) we will find equality, too. As for $N \in \{3, 4, 5\}$ there exist such spanning trees, we conclude that in these dimensions the bound in Corollary (6) is nothing but the optimum value of the MSFCB problem.

3.3 Large Grids

In this section we sketch that the strictly fundamental cycle basis B of any spanning tree T in the square $N \times N$ grid with $n = N^2 = (2^k + 2)^2$ vertices satisfies $\Phi(T) \geq \frac{1}{12} n \log_2 n - O(n)$. Hereby, our direct approach substantially improves the lower bound that has been obtained by Alon, Karp, Peleg, and West in [1, Thm. 6.6]—by a factor of more than 245.[2] Due to space limitations, unfortunately we cannot present any of our new proofs here. But we refer to [14] for the complete analysis.

In contrast to [1] we decided to tackle the lower bound problem from the perspective of the planar dual graph G^*. In particular, we exploit that for each spanning tree T, there is a one-to-one correspondence between its induced fundamental circuits in G, and its induced fundamental cuts in G^*. More precisely, if an edge $e \in E \setminus T$ induces a circuit in G, then its dual counterpart induces a cut in G^*—and both contain the very same edges.

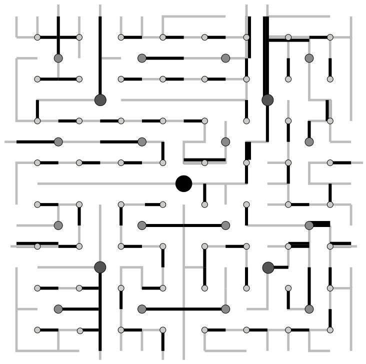


Figure 3: The dual tree T^* of a spanning tree T in $G_{18,18}$. In our lower bound for $\Phi(T)$ we only sum lower bounds on the fundamental circuits that are induced by the black edges

To detect sufficiently long circuits, according to Inequality (10) we resort on circuits that have large horizontal and/or vertical diameters. To obtain the claimed lower bound, it even turns out to be sufficient to either consider only the horizontal diameter $\mathrm{diam}_H(C)$ or its vertical diameter $\mathrm{diam}_V(C)$ of a circuit C.

We find circuits having large horizontal or vertical diameter by considering particular faces of $G_{N,N}$, or vertices of $(G_{N,N})^*$, respectively. The vertices that we use are the ones that are highlighted in Figure 3. In more detail, we organize these vertices in what we will call levels. We also assign a box to each such vertex. In a dual grid with $(N-1)^2 = \left(2^k+1\right) \cdot \left(2^k+1\right)$ vertices we establish k different levels of vertices as follows. The *level k* only contains the unique grid's center vertex. Its corresponding box equals $V(G_{N,N})$. The four quarters of the grid—which overlap on their borders—become the boxes of level $k-1$. By the particular choice of N, their center vertices are well-defined, and these constitute the vertices of level $k-1$. Recursively, each of these four quarters is again subdivided into four new quarters, which become the boxes of the next level, and their centers are the corresponding level-vertices. For a vertex u of level $\ell = 1, \ldots, k$ we call the set $\left\{v : d_{u,v} = 2^{\ell-1}\right\}$ the *border* of its *box* B_u.

For each such vertex v, we consider the subpath P of T^* that connects it with F^∞. But we only follow this path until the first edge e that is incident with the border of B_u. Assume w.l.o.g. that the edge e is "north" of v. In the primal grid, in fact e is a horizontal edge. Then we only consider the subsequence of vertical edges P_V (all being horizontal edges in $G_{N,N}$) of $P \subseteq T^*$ such that each edge is by one closer to e than its predecessor in P_V. In [14] we call this subpath the vertical *pseudo-path* of P. Then, for a vertex v of one particular level ℓ, $\ell \in \{1, \ldots, k\}$, we know that the $2^{\ell-1}$ edges of P_V induce fundamental circuits with respect to T of vertical diameters at least $1, 2, \ldots, 2^{\ell-1}$.

Now, we aim at summing the lower bounds on the diameters of the fundamental circuits for each occurrence of an edge on some pseudo-path. It is a simple observation that only pseudo-paths of different levels could share a, say, horizontal edge e of $G_{N,N}$, and thus potentially cause some double-counting (Lem. 3 in [14]). But one can further

[2] Let us mention that the authors of [1] state explicitly that they were not trying to "optimize constants."

observe that the lower bound on $\mathrm{diam}_V(e)$ is always larger for the pseudo-path that we defined for the vertex with the higher level index. Hence, to prevent us from any double-counting, for such an edge e we count as lower bound on $|C_T(e)|$ only the bound on $\mathrm{diam}_V(e)$ that we identify on the highest level. Doing so yields

Theorem 7 ([14])**.** Let $G_{N,N}$ be the planar grid graph with $n = N^2 = (2^k + 2)^2$ vertices. For every spanning tree T of $G_{N,N}$ there holds

$$\Phi(T) \geq \frac{1}{12} n \log_2 n - O(n). \tag{18}$$

4 New Upper Bounds

Alon et al. ([1]) provided spanning trees T_{AKPW} whose induced strictly fundamental cycle bases they showed to be bounded from above by $2n\log_2 n + o(n\log_2 n)$. An exact counting ([14]) revealed even that

$$\Phi(T_{\text{AKPW}}) \leq \frac{4}{3} n \log_2 n + O(n), \tag{19}$$

where $N = \sqrt{n}$, and $N = 2^k$ for some $k \in \mathbb{N}$. Although Alon, Karp, Peleg, and West ([1]) think of their trees as being essentially optimal, we are able to construct trees with an asymptotic coefficient for the $n\log_2 n$ term being strictly smaller than one. Moreover, we present trees which empirically perform very well already in small dimensions.

Fortunately, we are able to introduce a class of recursively defined trees that accommodates *both* goals. These spanning trees are the union of spanning trees in rectangular subgraphs of $G_{N,N}$, their building blocks. The trees differ in how their rectangular subgraphs—all respecting some arbitrary but fixed aspect ratio $\alpha \geq 1$—partition the faces of $G_{N,N}$. Hence, it remains to specify how to construct a spanning tree subject to a given parameter α for some grid $G_{M,N}$ having *aspect ratio* $\max\{\frac{M}{N}, \frac{N}{M}\} \approx \alpha$. This is done recursively. Assume w.l.o.g. that $M \geq N$. At the top-level of the recursion, we add to $T_\alpha(G_{M,N})$ the edges of the two longer borders of $G_{M,N}$ (here the horizontal ones), plus of one of its two other borders (cf. Figure 4). For the recursion, we partition

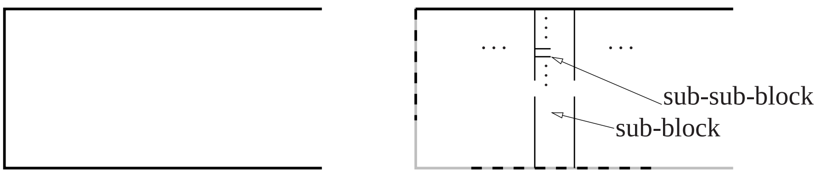


Figure 4: The shape of a block (left) and with a sketched interior recursively filled with smaller blocks (right), always keeping the aspect ratio.

the faces of $G_{M,N}$ into almost equally-sized rectangular subgraphs of aspect ratio again being close to α; only the faces of one horizontal path in $(G_{M,N})^*$, located almost in the middle of its two horizontal borders, are not contained in any of these rectangular subgraphs.

These trees are related to other families of trees as follows. In $G_{N,N}$, choosing $\alpha \geq \frac{N}{2} : 1$ there exists a partition of the grid such that we end with Machete-trees ([5], cf. Figure 1). Moreover, an aspect ratio of $\alpha = 1 : 1$ yields trees which can be obtained alternatively by a construction that is much similar to the one for T_{AKPW}.

According to the requirement—asymptotical or empirical quality—we will again subdivide our presentation into a part considering large grids and into a second part dealing with small grids. For both types of grids we will introduce trees with a block structure either having an aspect ratio of approximately $3:1$ or an aspect ratio of $2:1$, respectively. In addition, the trees differ in how the blocks are actually used to define a tree. Whereas on large grids it is sufficient to *cover* the grid with three (almost) equally-sized $3:1$ blocks, for small dimensions the grids are *tiled* with many $2:1$ blocks of many different sizes.

4.1 Large Grids

To achieve a good asymptotical upper bound we decided to construct trees out of the above described blocks with an aspect ratio of $3:1$. Unfortunately, it turns out to be tricky to subdivide or tile a square grid of arbitrary dimension with these particular blocks. Thus, we construct our trees bottom-up like. That means we take an atomic block of size 6×14 and arrange 32 copies of such a block to a new one having size 80×14. This procedure is then iterated providing spanning trees for dimensions

$$\left(\frac{1248}{496} \cdot 32^{k/2} + \frac{15}{31}\right) \times \left(\frac{419}{496} \cdot 32^{k/2} + \frac{30}{31}\right) \tag{20}$$

with k chosen integral and even. Finally, three copies of such a tree can be put onto each other and cover the entire square grid. Now, a detailed description of the construction of the tree and a precise analysis of it follows.

Construction of the tree. Whereas in Sec.4 we gave a brief top-down description of the tree that we consider we now introduce them bottom-up like, hereby having more control on the dimensions and, thus, by-passing rounding indispositions.

For every $k \in \mathbb{N}^+$ we construct recursively spanning trees as follows. For k equal to 1 consider the spanning tree T_1 as sketched in Figure 5. This tree is defined on a 6×14 grid and it has its *exit* on the lower horizontal border. The next tree, T_2, is constructed by arranging 32 copies of T_1. First, 16 copies are glued with this particular orientation side by side. Second, we mirror the other 16 copies of T_1 horizontally and place them such that their exits are opposite to the first 16 copies. At last, one vertical edge, which we will

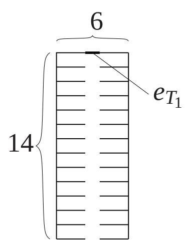


Figure 5: The spanning tree T_1 out of which all the trees T_k are constructed. T_1 has dimension 6×14, or side-lengths 5×13, respectively.

call e_{T_2} is added to connect the two so-constructed connected components.

The general rule here is to take the left vertical edge as connecting edge for the construction of the tree T_k with k even and the upper horizontal edge for the construction of the T_k with k odd. See Figure 5 for an example. By this construction, the tree T_2 is of dimension 81×28. In general, the tree T_k is constructed out of 32 copies of T_{k-1} and an additional connecting edge the very same way.

In order to finally state a spanning tree for a square grid and to prepare the analysis of the trees we introduce four sequences for the x and y length, w.r.t. edges, of T_k in dependence of k. As T_1 is a 6×14 grid tree we have $x_1 = 5$ and $y_1 = 13$. By construction, we get the following sequences taking the parity of k into account:

$$x_{2i} = 16 \cdot x_{2i-1} \qquad\qquad y_{2i} = 2 \cdot y_{2i-1} + 1 \tag{21}$$

for trees T_k with $k = 2i$ even. For odd $k = 2i+1$ the tree T_k has dimension

$$x_{2i+1} = 2 \cdot x_{2i-1} + 1 \qquad\qquad y_{2i+1} = 16 y_{2i-1}. \tag{22}$$

In the following we will only consider the spanning trees T_k for k even. Simple calculations transform (21) and the start values $x_2 = 80$ and $y_2 = 27$, respectively, into the explicit sequence of the lengths of T_k for even k:

$$x_k = \frac{78}{31} \cdot 32^{k/2} - \frac{16}{31} \qquad\qquad y_k = \frac{419}{496} \cdot 32^{k/2} - \frac{1}{31}. \tag{23}$$

If we now take a closer look at T_k, k even, we see that the ratio of its lengths is almost 3×1. In fact, the exact ratio x_k to y_k is always greater than 2.96 and converges to $\frac{1248}{419} \approx 2.978$.

Hence, if we take three copies of T_k and put them one upon another, then the resulting spanning tree, let us denote it by T_k^3, covers a grid of dimension

$$\left(\frac{1248}{496} \cdot 32^{k/2} + \frac{15}{31}\right) \times \left(\frac{1257}{496} \cdot 32^{k/2} + \frac{30}{31}\right).$$

We now claim that three times the size of T_k is an upper bound on the size of T_k^3 restricted to the square grid G with dimension

$$\left(\frac{1248}{496} \cdot 32^{k/2} + \frac{15}{31}\right) \times \left(\frac{1248}{496} \cdot 32^{k/2} + \frac{15}{31}\right).$$

So, how do we restricted T_k^3 to a square of the above size? Due to space limitations we will only give an idea of the procedure. Let us consider the boundary line L of G that, in a sense, cuts through the down-most copy T_k of T_k^3, cf. Figure 6. This T_k consists of several subtrees $T_{k-1}, T_{k-3}, \dots, T_1$. Those *odd* subtrees can have their exit pointing downwards, $\downarrow T_j$, or upwards,denoted by $\uparrow T_j$. If for a $j = 1, 3, \dots, k-3, k-1$ the boundary line L cuts through a subtree $\downarrow T_j$ we leave this part of tree unchanged and simply cut away what overhangs L. In the other case where the boundary line L cuts through a subtree $\uparrow T_j$ we cut away the overhanging parts as well, but—since we loose connectivity—we add an edge to $\uparrow T_j$ exactly where formerly the exit had been. If we do so for all $j = 1, 3, \dots, k-3, k-1$ we finally come up with a tree with less chords inducing lower length fundamental cycles than this down-most copy T_k of T_k^3 before.

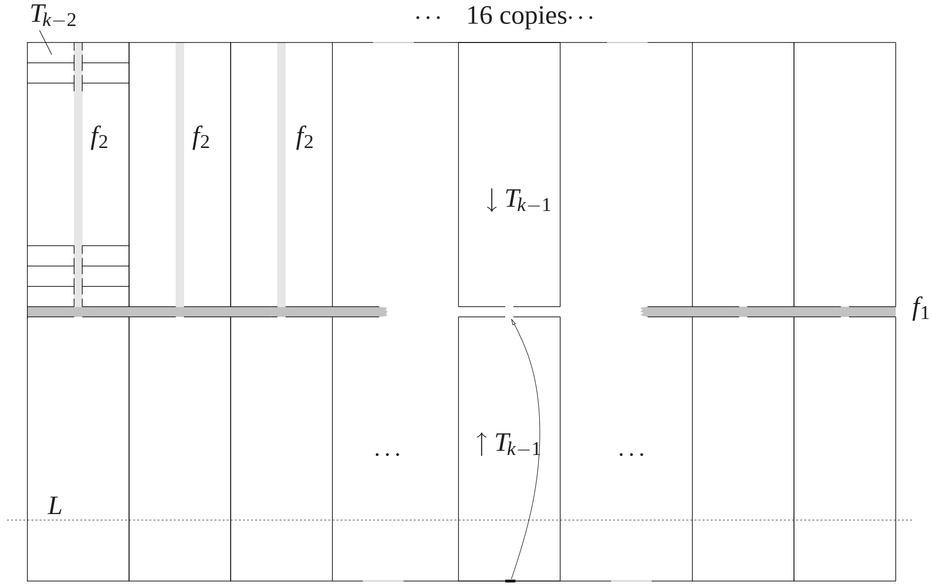


Figure 6: A schematic illustration of the tree T_k for an even $k \in \mathbb{N}^+$. Due to the construction rules T_k *consists of* $32^2 = 1024$ copies of T_{k-2} and different slots, i.e. one main slot f_1, dark-gray, and 32 subslots f_2, light-gray.

Analysis of the tree. So, for every $k \in \mathbb{N}^+$ let us denote with $\tilde{T}_k$ the spanning tree of the

$$\left(\frac{1248}{496} \cdot 32^{k/2} + \frac{15}{31}\right) \times \left(\frac{1248}{496} \cdot 32^{k/2} + \frac{15}{31}\right)$$

grid as described above. We are interested in an upper bound on the strictly fundamental cycle basis induced by $\tilde{T}_k$. As foreshadowed above we have

$$\Phi(\tilde{T}_k) \leq 3 \cdot \Phi(T_k). \tag{24}$$

In the following we develop a recursive formula for $\Phi(T_k)$. Because of the tree's special construction the following recursive formula holds,

$$\Phi(T_k) = 1024 \cdot \Phi(T_{k-2}) + f(T_k), \tag{25}$$

where $f(T_k)$ denotes the size of the fundamental cycles induced by edges that do not lie entirely within a copy of the smaller T_{k-2} tree. We call those areas slots. Then, $f(T_k)$ can be canonically subdivided into one main-slot and several sub-slots, cf. Figure 6. Obviously,

$$f(T_k) = f_1(T_k) + 32 \cdot f_2(T_k) \tag{26}$$

holds. Then, with the help of the sequences defining the lengths of the trees (Equations (21) and (22)) we straight-forward express f_1 and f_2 as

$$f_1(T_k) \le \sum_{i=1}^{\frac{1}{32}x_k+1} 2i + \sum_{j=1}^{15} 2\cdot\left(\sum_{i=1}^{\frac{1}{32}x_k+1}(2y_k+\frac{2j}{16}x_k+2i)\right) + \sum_{i=1}^{\frac{1}{32}x_k+1}(2x_k+2y_k+2i), \quad (27)$$

and

$$f_2(T_k) \le \sum_{i=1}^{\frac{1}{32}y_{k-1}+1} 2i \quad + \quad \sum_{j=1}^{15} 2\cdot\left(\sum_{i=1}^{\frac{1}{32}y_{k-1}+1}(2x_{k-1}+\frac{2j}{16}y_{k-1}+2i)\right) \quad (28)$$

$$+ \sum_{i=1}^{\frac{1}{32}y_{k-1}+1}(2x_{k-1}+2y_{k-1}+2i), \quad (29)$$

$$(30)$$

respectively. Further, plugging (27) and (28) into (26) and then (26) into (25) we yield the recursion:

$$\Phi(T_k) \le 1024\cdot\Phi(T_{k-2}) + \frac{20{,}323{,}353}{984{,}064}32^k + o(32^k). \quad (31)$$

Hereby, we omit the value for the recursion start T_2 because it is of no importance for the coefficient of the $n\cdot\log_2 n$ term.

After resolving (31) and applying the result to (24) one gets

$$\Phi(\tilde{T}_k) \le \frac{60{,}970{,}059}{1{,}968{,}128}\cdot 32^k\cdot k + o(32^k\cdot k).$$

Finally, making use of the special dimension, i.e.

$$\sqrt{n} = \frac{78}{31}\cdot 32^{k/2} - \frac{16}{31},$$

one can state the following upper bound:

$$\Phi(\tilde{T}_k) \le \frac{6,774,451}{6,922,240}\cdot n\cdot\log_2 n + o(n\cdot\log_2 n).$$

We summarize the section on upper bounds on large grids by stating the following lemma:

Lemma 8. Let $G_{N,N}$ denote the $N\times N$ square planar grid with $n = N^2$ vertices and with $N = \frac{78}{31}\cdot 32^{k/2} + \frac{15}{31}$ for some even integer k. Then the size of a minimum strictly fundamental cycle bases on $G_{N,N}$ can be bounded from above by

$$0.978\cdot n\cdot\log_2 n + O(n).$$

Once again remember that hereby the previously best asymptotical upper bound by Alon et al. ([1]) is enhanced by a factor of more than four third.

4.2 Small Grids

Unfortunately, the $3:1-$block structured trees, as described in the above section, are not perfectly suited for smaller dimensions. This is due to the fact that, although asymptotically $3:1$ turned out to be a very good aspect ratio for the blocks, it is not possible to decompose a square grid into such blocks without losing much of their advantage because of rounding "errors."

Therefore, for small grids, we chose a different block-structured graph. This time we demand a fix aspect ratio of $2:1$. Moreover, the $2:1-$blocks, in a sense, do not cover, but rather tile the square grid. The tiling procedure roughly goes as follows:

At first, two opposite $2:1-$blocks are put in the middle of the grid. See for example the bold line bordered vertical blocks with side lengths 8×15 in Figure 7. Then, horizontal $2:1-$blocks are added centrally aside such that rectangular subgrids in the four corners remain. Now in such a corner we always direct the next block such that its depth can be chosen as small as possible, always staying as close as possible to the target ratio $2:1$. During this procedure we do not pay attention to any rounding inaccuracies. In Figure 7 an example $2:1-$block structured tree for dimension $N=31$ is shown.

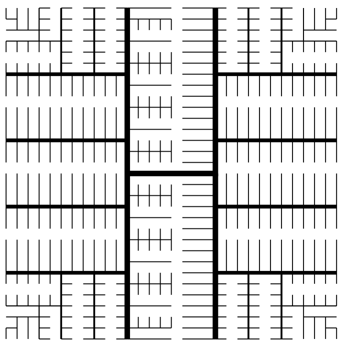


Figure 7: Notice the parquet-like structure of the tree with tiles having height-width ratio of 2 with small errors due to roundings. Inside, the blocks themselves are recursively filled with smaller blocks still maintaining the $2:1$ ratio.

5 Experimental Results

In this section we compare different spanning trees with respect to the length of the strictly fundamental cycle basis they induce. In addition to the degree-based tree-growing heuristics that we already referred to in the Introduction, local search techniques have also been considered. The most natural neighborhood that one can think of in the context of spanning trees, is the 2-neighborhood (cf. Section 2).

Amaldi et al. ([2]) reported the performance of several strategies for searching this type of neighborhood. In what they denote by local search (LS), the entire neighborhood is examined and they move to the tree with the best improvement. In a second deterministic strategy (ES), only those neighbors are tested, having a fixed ratio of branches according to a predefined order. To prevent LS to terminate too early in a too bad local optimum, Amaldi et al. ([2]) run metaheuristics such as variable-neighborhood search (VNS) and a tabu search (TS) on top of LS. In any of their computations, an adapted version of the tree-growing heuristic in [20] is used as the initial solution.

In our computations, we use the $2:1-$block-structured tree as initial solution. To improve them, in contrast to (LS) we do not examine the entire 2-neighborhood. Rather, whenever we identify a neighbor that improves the current solution, we greedily move to that neighbor. Of course, this method depends on the order the edges in the tree are checked. Empirical studies showed, however, that the influence of the edge-order is neglectable. For our computational studies we chose a random order of edges and ran our greedy-like approach—denoted by (GS)—ten times, considering best and average values of both length of the cycle basis and the running time of (GS).

Results. In Table 1 we compare the constructive heuristics, i.e. those that build up a tree without doing any subsequent local improvements. Moreover, we complement these values with information on lower bounds, once obtained by Corollary 6 and the trivial ones by a minimum weakly fundamental cycle basis (MWFCB). Notice that the latter were also used in the recent study of Amaldi et al. ([2]). In addition to these lower bound values, let us mention that for $N = 130$—being the dimension closest to 100 for which our asymptotic bound is defined exactly—the bound that we derived in Corollary 6 is only by about 3.5% weaker.

In our tables the *italic* numbers highlight the best known upper and lower bounds. For $N = 5$, these coincide, and we mark this in **boldface**. Observe that for any dimension $N \geq 10$, the new trees that we propose in Section 4 yield smaller SFCB values than any of the other constructive heuristics.

In Table 2 we compare the different local-search-type heuristics. For our greedy search (GS) we used a 3.2GHz Intel P4 computer (“A1”), running Linux and using LEDA©. Amaldi et al. used for their local search heuristics (LS) and (ES) also an Intel P4 computer running Linux, but with 2.66GHz (“A2”). Accordingly, the times stated in Table 2 refer to the particular architecture. The values for the metaheuristics (TS) and (VNS)—also quoted from [2]—each refer to 10 minute runs on the A2 environment. Much similar as in the purely constructive context, our new solutions improve the best known upper bounds for all dimensions $N \geq 20$.

As already mentioned before we ran our local search (GS) with a random order. In Table 2 the first column presents the value and running time for the best run out of 10 samples, and the deviations are very small. However, it has to be annotated that only for dimensions $N \in \{60, 80, 90, 100\}$ the start tree had *not* already been locally optimal.

6 Conclusions

Any serious summary of this paper has to remain a bit twofold. On the one hand, on square planar grid graphs—being a particularly challenging family of graphs for the

N	"2 : 1″" [1]	AKPW [5]	Machete [16]	C-Order [8, 16]	Deo's NT	Corollary 6	MWFCB
5	76	78	**72**	**72**	78	**72**	64
10	468	524	492	492	518	*422*	324
15	1 300	1 554	1 512	1 512	1 588	*1 072*	784
20	*2 550*	3 030	3 382	3 382	3 636	*2 022*	1 444
25	*4 368*	5 410	6 352	6 352	6 452	*3 272*	2 304
30	*6 656*	8 408	10 672	10 672	11 638	*4 822*	3 364
35	*9 592*	11 694	16 592	16 592	16 776	*6 672*	4 624
40	*13 162*	16 078	24 362	24 362	28 100	*8 822*	6 084
45	*17 236*	21 784	34 232	34 232	35 744	*11 272*	7 744
50	*21 920*	27 912	46 452	46 452	48 254	*14 022*	9 604
55	*27 356*	35 124	61 272	61 272	62 026	*17 072*	11 664
60	33 406	42 790	78 942	78 942	92 978	*20 422*	13 924
70	*47 300*	59 244	123 832	–	–	*28 022*	19 044
80	63 964	80 678	183 122	–	–	*36 822*	24 964
90	83 412	108 012	258 812	–	–	*46 822*	31 684
100	106 090	137 390	352 902	–	–	*58 022*	39 204

Table 1: Comparison of the cost of some selected trees, i.e. the length of the according strictly fundamental cycle bases. The rightmost column presents the previously best lower bound for small dimensions, obtained just by $4 \cdot (N-1)^2$. The penultimate column now states the consistently better lower bounds due to Corollary 6.

MSFCB Problem—we we have improved significantly the lower and upper bounds that were previously known for the MSFCB Problem.

On the other hand, just reconsider the rows with $N = 10$ in Tables 1 and 2. For this relatively small dimension, we simply believe the optimality gap that we are leaving here (about 10%) should better not the that big. Even worse for $N = 8$: Indeed, in Figure 8 we provide a spanning tree with smaller SFCB value than the one on the first page—and which only every *eighth* participant of MATHEON's 2006 christmas quiz has been aware of. Yet, we are not aware of any concise combinatorial proof for its optimality. Hence, further efforts are to be made.

Nevertheless, columns 4–6 of Table 1 illustrate impressively to what extent degree-based heuristics for the MSFCB problem are inferior to applying recursive approaches. In other words, for any heuristic for the MSFCB problem which will be designed only in the future, we strongly recommend to evaluate it also on planar square grid graphs (in contrast to what has been done in some studies in the past, e.g. [7, 8]), and there compare its performance to the *italic* values that we provide in Tables 1 and 2.

N	(GS)		(LS)		(ES)		(VNS)	(TS)
	cost	time	cost	time	cost	time	cost	cost
5	**72**	0:00:00	**72**	0:00:00	74	0:00:00	**72**	**72**
10	468	0:00:00	474	0:00:00	524	0:00:00	*466*	*466*
15	1 300	0:00:00	1 318	0:00:00	1 430	0:00:00	1 280	*1276*
20	*2 550*	0:00:00	2 608	0:00:03	3 186	0:00:00	2 572	2590
25	*4 368*	0:00:00	4 592	0:00:16	5 152	0:00:02	4 464	4430
30	*6 656*	0:00:01	6 956	0:00:47	8 488	0:00:03	6 900	6882
35	*9 592*	0:00:02	10 012	0:02:19	11 662	0:00:08	9 982	9964
40	*13 162*	0:00:07	13 548	0:06:34	15 924	0:00:26	13 524	13534
45	*17 236*	0:00:06	18 100	0:14:22	22 602	0:01:00	18 100	18100
50	*21 920*	0:00:09	23 026	0:31:04	33 274	0:01:10	23 026	23552
60	*33 374*	0:01:01	–	–	–	–	–	–
80	*63 810*	0:07:24	–	–	–	–	–	–
90	*83 222*	0:07:48	–	–	–	–	–	–
100	*105 766*	0:14:01	–	–	–	–	–	–

Table 2: An overview of the quality of five local search approaches. Missing values are marked with an "–" and running times are measured in *h*:*mm*:*ss*. The columns (LS)–(VNS) are cited from [2].

Acknowledgment

The authors would like to thank Janina Brenner for intensive discussions on the empirical lower bounds.

References

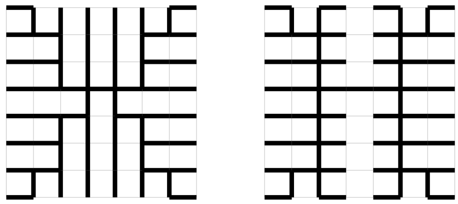
[1] N. Alon, R. M. Karp, D. Peleg, and D. B. West. A graph-theoretic game and its application to the k-server problem. *SIAM J. Comput.*, 24(1):78–100, 1995.

[2] E. Amaldi, L. Liberti, N. Maculan, and F. Maffioli. Efficient edge-swapping heuristics for finding minimum fundamental cycle bases. In C. C. Ribeiro and S. L. Martins, editors, *WEA*, volume 3059 of *Lecture Notes in Computer Science*, pages 14–29. Springer, 2004.

[3] S. Bächle and F. Ebert. Graph theoretical algorithms for index reduction in circuit simulation,. Preprint 245, DFG Research Center MATHEON, 2005.

[4] H. L. Bodlaender. A tourist guide through treewidth. *Acta Cybern.*, 11(1-2):1–22, 1993.

[5] P. Boksberger. Minimum stretch spanning trees. Diploma thesis, ETH Zürich, 2003.



(a) 266 (b) 262

Figure 8: Did you succeed in finding a better tree of cost only 262 like the one on the right hand side? Well, less than 12.5% of the participants of MATHEON's 2006 christmas quiz had been aware of it...

[6] B. Bollobás. *Modern Graph Theory*, volume 184 of *Graduate Texts in Mathematics*. Springer, 2002. 2nd printing.

[7] N. Deo, M. Krishnomoorthy, and G. Prabhu. Algorithms for generating fundamental cycles in a graph. *ACM Transactions on Mathematical Software*, 8(1):26–42, 1982.

[8] N. Deo, N. Kumar, and J. Parsons. Minimum-length fundamental-cycle set problem: A new heuristic and an simd implementation. Technical Report CS-TR-95-04, University of Central Florida, Orlando, 1995.

[9] V. Eiseler and D. Wasserrab. *Die Kreisbasenbibliothek CyBaL*. Technische Universität München, 2004. `http://www-m9.ma.tum.de/dm/cycles/cybal`, In German.

[10] M. Elkin, Y. Emek, D. A. Spielman, and S.-H. Teng. Lower-stretch spanning trees. In H. N. Gabow and R. Fagin, editors, *STOC*, pages 494–503. ACM, 2005.

[11] P. M. Gleiss. *Short Cycles*. Ph.D. thesis, Universität Wien, 2001.

[12] C. C. Gotlieb and D. G. Corneil. Algorithms for finding a fundamental set of cycles for an undirected linear graph. *Communications of the ACM*, 10(12):780–783, 1967.

[13] T. Kavitha, K. Mehlhorn, D. Michail, and K. E. Paluch. A faster algorithm for minimum cycle basis of graphs. In J. Diaz, J. Karhumäki, A. Lepistö, and D. Sanella, editors, *ICALP*, volume 3142 of *Lecture Notes in Computer Science*, pages 846–857. Springer, 2004.

[14] E. Köhler, C. Liebchen, R. Rizzi, and G. Wünsch. Reducing the optimality gap of strictly fundamental cycle bases in planar grids. Preprint 007/2006, TU Berlin, Mathematical Institute, 2006.

[15] E. Köhler, R. H. Möhring, and G. Wünsch. Minimizing total delay in fixed-time controlled traffic networks. In H. Fleuren, D. den Hertog, and P. Kort, editors, *Operations Research Proceedings 2004*, pages 192–199. Springer, 2005.

[16] L. Liberti, E. Amaldi, and F. M. N. Maculan. Mathematical models and a constructive heuristic for finding minimum fundamental cycle bases. *Yugoslav Journal of Operations Research*, 15(1):15–24, 2005.

[17] C. Liebchen. *Periodic Timetable Optimization in Public Transport*. dissertation.de – Verlag im Internet, 2006.

[18] C. Liebchen and R. Rizzi. Cycles bases of graphs. Technical Report 2005-018, TU Berlin, Mathematical Institute, 2005.

[19] C. Liebchen and G. Wünsch. The zoo of tree spanner problems. Technical Report 2006-013, TU Berlin, Mathematical Institute, 2006.

[20] K. Paton. An algorithm for finding a fundamental set of cycles of a graph. *Communications of the ACM*, 12(9):514–518, 1969.

[21] R. Rizzi. Manuscript, 2006.

[22] N. Robertson and P. D. Seymour. Graph minors. V. Excluding a planar graph. *J. Comb. Theory, Ser. B*, 41(1):92–114, 1986.

[23] A. Schrijver. *Combinatorial Optimization*, volume 24 of *Algorithms and Combinatorics*. Springer, 2003.

[24] H. Whitney. On the abstract properties of linear dependence. *American Journal of Mathematics*, 57:509–533, 1935.

Reports from the group

"Combinatorial Optimization and Graph Algorithms"

of the Department of Mathematics, TU Berlin

2006/32 *Romeo Rizzi and Christian Liebchen:* A New Bound on the Length of Minimum Cycle Bases

2006/24 *Christian Liebchen and Sebastian Stiller:* Delay Resistant Timetabling

2006/08 *Nicole Megow and Tjark Vredeveld:* Approximation Results for Preemptive Stochastic Online Scheduling

2006/07 *Ekkehard Köhler and Christian Liebchen and Romeo Rizzi and Gregor Wünsch:* Reducing the Optimality Gap of Strictly Fundamental Cycle Bases in Planar Grids

2006/05 *Georg Baier and Thomas Erlebach and Alexander Hall and Ekkehard Köhler and Heiko Schilling:* Length-Bounded Cuts and Flows

2005/30 *Ronald Koch and Martin Skutella and Ines Spenke :* Maximum k-Splittable Flows

2005/29 *Ronald Koch and Ines Spenke :* Complexity and Approximability of k-Splittable Flows

2005/28 *Stefan Heinz and Sven O. Krumke and Nicole Megow and Jörg Rambau and Andreas Tuchscherer and Tjark Vredeveld:* The Online Target Date Assignment Problem

2005/18 *Christian Liebchen and Romeo Rizzi:* Classes of Cycle Bases

2005/11 *Rolf H. Möhring and Heiko Schilling and Birk Schütz and Dorothea Wagner and Thomas Willhalm:* Partitioning Graphs to Speed Up Dijkstra's Algorithm.

2005/07 *Gabriele Di Stefano and Stefan Krause and Marco E. Lübbecke and Uwe T.Zimmermann:* On Minimum Monotone and Unimodal Partitions of Permutations

2005/06 *Christian Liebchen:* A Cut-based Heuristic to Produce Almost Feasible Periodic Railway Timetables

2005/03 *Nicole Megow, Marc Uetz, and Tjark Vredeveld:* Models and Algorithms for Stochastic Online Scheduling

2004/37 *Laura Heinrich-Litan and Marco E. Lübbecke:* Rectangle Covers Revisited Computationally

2004/35 *Alex Hall and Heiko Schilling:* Flows over Time: Towards a more Realistic and Computationally Tractable Model

2004/31 *Christian Liebchen and Romeo Rizzi:* A Greedy Approach to Compute a Minimum Cycle Bases of a Directed Graph

2004/27 *Ekkehard Köhler and Rolf H. Möhring and Gregor Wünsch:* Minimizing Total Delay in Fixed-Time Controlled Traffic Networks

2004/26 *Rolf H. Möhring and Ekkehard Köhler and Ewgenij Gawrilow and Björn Stenzel:* Conflict-free Real-time AGV Routing

2004/21 *Christian Liebchen and Mark Proksch and Frank H. Wagner:* Performance of Algorithms for Periodic Timetable Optimization

2004/20 *Christian Liebchen and Rolf H. Möhring:* The Modeling Power of the Periodic Event Scheduling Problem: Railway Timetables — and Beyond

2004/19 *Ronald Koch and Ines Spenke:* Complexity and Approximability of k-splittable flow problems

2004/18 *Nicole Megow, Marc Uetz, and Tjark Vredeveld:* Stochastic Online Scheduling on Parallel Machines

2004/09 *Marco E. Lübbecke and Uwe T. Zimmermann:* Shunting Minimal Rail Car Allocation

2004/08 *Marco E. Lübbecke and Jacques Desrosiers:* Selected Topics in Column Generation

2003/050 *Berit Johannes:* On the Complexity of Scheduling Unit-Time Jobs with OR-Precedence Constraints

2003/49 *Christian Liebchen and Rolf H. Möhring:* Information on MIPLIB's timetab-instances

2003/48 *Jacques Desrosiers and Marco E. Lübbecke:* A Primer in Column Generation

2003/47 *Thomas Erlebach, Vanessa Kääb, and Rolf H. Möhring:* Scheduling AND/OR-Networks on Identical Parallel Machines

2003/43 *Michael R. Bussieck, Thomas Lindner, and Marco E. Lübbecke:* A Fast Algorithm for Near Cost Optimal Line Plans

2003/42 *Marco E. Lübbecke:* Dual Variable Based Fathoming in Dynamic Programs for Column Generation

2003/37 *Sándor P. Fekete, Marco E. Lübbecke, and Henk Meijer:* Minimizing the Stabbing Number of Matchings, Trees, and Triangulations

2003/25 *Daniel Villeneuve, Jacques Desrosiers, Marco E. Lübbecke, and François Soumis:* On Compact Formulations for Integer Programs Solved by Column Generation

2003/24 *Alex Hall, Katharina Langkau, and Martin Skutella:* An FPTAS for Quickest Multicommodity Flows with Inflow-Dependent Transit Times

2003/23 *Sven O. Krumke, Nicole Megow, and Tjark Vredeveld:* How to Whack Moles

2003/22 *Nicole Megow and Andreas S. Schulz:* Scheduling to Minimize Average Completion Time Revisited: Deterministic On-Line Algorithms

2003/16 *Christian Liebchen:* Symmetry for Periodic Railway Timetables

2003/12 *Christian Liebchen:* Finding Short Integral Cycle Bases for Cyclic Timetabling

762/2002 *Ekkehard Köhler and Katharina Langkau and Martin Skutella:* Time-Expanded Graphs for Flow-Dependent Transit Times

761/2002 *Christian Liebchen and Leon Peeters:* On Cyclic Timetabling and Cycles in Graphs

752/2002 *Ekkehard Köhler and Rolf H. Möhring and Martin Skutella:* Traffic Networks and Flows Over Time

739/2002 *Georg Baier and Ekkehard Köhler and Martin Skutella:* On the k-splittable Flow Problem

736/2002 *Christian Liebchen and Rolf H. Möhring:* A Case Study in Periodic Timetabling

Reports may be requested from:

Sekretariat MA 6–1
Fakultt II – Institut fr Mathematik
TU Berlin
Straße des 17. Juni 136
D-10623 Berlin – Germany

e-mail: klink@math.TU-Berlin.DE

Reports are also available in various formats from

`http://www.math.tu-berlin.de/coga/publications/techreports/`

and via anonymous ftp as

`ftp://ftp.math.tu-berlin.de/pub/Preprints/combi/Report-`*number*`-`*year*`.ps`