DISKRETE MATHEMATIK UND OPTIMIERUNG

Annabell Berger, Winfried Hochstättler:

Minconvex graph factors of prescribed size and a simpler reduction to weighted f-factors

FernUniversität in Hagen
Lehrgebiet Mathematik
Lehrstuhl für Diskrete Mathematik und Optimierung
D – 58084 Hagen

Minconvex graph factors of prescribed size and a simpler reduction to weighted f-factors

Annabell Berger
Winfried Hochstättler
Department of Mathematics and Computer Science
FernUniversität
Hagen, Germany

August 7, 2006

Abstract

Recently, András Sebő and Nicola Apollonio considered the following generalization of the problem to determine a matching of size k. Given a graph $G = (V, E)$ and an integer k, determine a subgraph $H = (V, F)$ of G that minimizes $\sum_{v \in V} d_H(v)^2$ or more general $\ell(d_H)$, where $\ell : \mathbb{N}^V \to \mathbb{R}$ is a separable convex function, and $|F| = k$. They proved, that the problem is polynomially solvable in two ways. First, it admits "augmenting pairs of alternating paths" which, according to their analysis, yields a complexity of $O(n^{5.5})$, if ℓ is a quadratic convex function. On the other hand they reduced the problem to a single minimum weight f-factor problem. Their reduction introduces quite a few new vertices and edges such that the resulting complexity is higher.

We present a simpler reduction of "minconvex factor of prescribed size" to a single minimum weighted f-factor problem resulting in an overall complexity of $O(m^2 \log(n))$.

1 Introduction

Starting from the problem to cover as many edges as possible with k vertices in a line graph, Nicola Apollonio and András Sebő [1, 2] considered the problem to find k edges in a given graph $G = (V, E)$ such that the sum of the squares of the degrees of the subgraph formed by these edges is minimized. More generally, they examine the problem to find $F \subseteq E$ in a given, not necessarily simple, graph $G = (V, E)$, such that $|F| = k$ and

$$\ell(d_F) := \sum_{v \in V} \ell_v(d_F(v))$$

is minimized, where ℓ is a discrete separable convex function on the degree sequence, i.e. for each $v \in V$ $\ell_v : \mathbb{N} \to \mathbb{R}$ is a discrete convex function.

1

They could show that the problem, while being NP-complete for general convex functions, is solvable in polynomial time in the separable case. In particular they prove that the global optimality criterion can be replaced by a local one, allowing for incremental search algorithms. This approach was generalized by Murota [6] to minimizing M-convex functions in the intersection of a hyperplane of constant component sum with a jump system. Apollonio and Sebő also presented a reduction to a weighted f-factor problem. Unfortunately, the number of vertices in this reduction increases considerably.

We present a simpler reduction which uses just one additional vertex. Together with the algorithm of Gabow [3] this yields an overall complexity of $O(m^2 \log n)$.

The paper is organized as follows: In Section 2 we provide necessary notation and some prerequisites. In Section 3 we review the main results of [1, 2] providing simplified proofs. In Section 4 we present our reduction and in Section 5 we conclude with algorithmic consequences.

2 Prerequisites

A function $k : \mathbb{N} \to \mathbb{R}$ is a *discrete convex function* if for all $i \in \mathbb{N}$

$$k(i+2) - k(i+1) \geq k(i+1) - k(i). \tag{1}$$

By induction this immediately implies

Proposition 2.1. *If k is a discrete convex function and $n_1, n_2, n_3 \in \mathbb{N}$. Then*

$$k(n_1 + n_2 + n_3) - k(n_1 + n_2) \geq k(n_1 + n_3) - k(n_1). \tag{2}$$

Proof.

$$
\begin{aligned}
k(n_1 + n_2 + n_3) - k(n_1 + n_2) &= \sum_{i=1}^{n_3} k(n_1 + n_2 + i) - k(n_1 + n_2 + i - 1) \\
&\geq \sum_{i=1}^{n_3} k(n_1 + i) - k(n_1 + i - 1) \\
&= k(n_1 + n_3) - k(n_1).
\end{aligned}
$$

$\square$

Let $G = (V, E)$ be a not necesserily simple graph. If $F \subseteq E$, by $d_F(v)$ we denote the degree of vertex v in the graph $H = (V, F)$ and by $d_F : V \to \mathbb{N}$ its *degree sequence*. Let 2^E denote the power set of E. A function $\ell : 2^E \to \mathbb{R}$ is called a *discrete separable convex function on the degree sequences of subgraphs of G* if there exist discrete convex functions ℓ_v for each $v \in V$ such that for all $F \subseteq E$ we have
$$\ell(F) = \sum_{v \in V} \ell_v(d_F(v)).$$

2

For an introduction to general discrete convex functions see [5].

If $f : V \to \mathbb{N}$ then an f-factor of G is a subgraph $H = (V, F)$ such that $d_F(v) = f(v)$ for all $v \in V$.

In this paper we consider the problem

Problem 2.2 (MCFS). *Given a graph $G = (V, E)$, an integer $k \geq 0$ and a discrete separable convex function on the degree sequences of subgraphs of G. Determine $F \subseteq E$ such that $|F| = k$ and $\ell(d_F)$ is minimized.*

If $F, F' \subseteq E$ then an F-F'-*alternating walk* is a walk which alternates between edges from $F \setminus F'$ and $F' \setminus F$. An F-$(E \setminus F)$ alternating path is simply called F-*alternating.* By $F \Delta F'$ we denote the symmetric difference

$$F \Delta F' = (F \cup F') \setminus (F \cap F').$$

3 Alternating walks

If ℓ is the sum of squares of the degrees and G has a matching $M \subseteq E$ of size k then clearly M is optimal for Problem 2.2 which, thus, generalizes matching (see e.g. [4]). Applying the established solution concept of alternating paths yields:

Proposition 3.1 ([1, 2]).   *1. If $F, F' \subseteq E$ and $|F| = |F'|$ then $F \Delta F'$ can be decomposed into pairwise edge disjoint F-F'-alternating walks such that in every vertex $v \in V$ exactly $|d_F(v) - d_{F'}(v)|$ of these walks end.*

   2. These walks can be grouped into even alternating walks $Q_j, j = 1, \ldots, t$ and pairs of odd alternating walks $P_i^+, P_i^-, i = 1, \ldots, s$ where $|P_i^+ \cap F| = |P_i^+ \setminus F| - 1$, but $|P_i^- \cap F| = |P_i^- \setminus F| + 1$ and

$$\forall 1 \leq i \leq s, 1 \leq j \leq t : |F \Delta Q_j| = |F| = |F \Delta P_i^+ \Delta P_i^-|.$$

Considering a pair af adjacent edges in two vertex disjoint triangles, it is immediate that modifying an edge set along even alternating walks may not be sufficient in order to optimize. Also note, that the walks may not be paths and their end vertices may coincide.

Since the alternating paths are pairwise edge disjoint their effect on the degree sequence of F decomposes, too. To be more precise, for an edge set $Q \subseteq E$ we denote by $\delta_F(Q) \in \mathbb{Z}^V$ its effect on the degree sequence of F, i.e.

$$\delta_F(Q) := d_{F \Delta Q} - d_F.$$

Then

$$d_{F'} = d_F + \delta_F(F' \Delta F) = d_F + \sum_{j=1}^{t} \delta_F(Q_j) + \sum_{i=s}^{t} \delta_F(P_i^+ \cup P_i^-).$$

The functions or – as we prefer to call them – vectors $\delta_F(Q_j)$ of even alternating paths are non-zero only on two vertices, namely equal to 1 on one end

vertex of the path and -1 on the other. The functions $\delta_F(P_i^+ \cup P_i^-)$, however, may have two 1's and two -1s, a 2 and two -1s, two 1s and a -2 or a 2 and a -2. As a consequence of Proposition 3.1 (2) all these vectors are *compatible*, i.e. if in a fixed coordinate some vector is negative (positive), then all other have to non-positive (non-negative) in that coordinate.

Compatible vectors behave nicely with respect to discrete separable convex functions.

Lemma 3.2 ([1, 2]). *Let $\lambda_1, \lambda_2 : V \to \mathbb{Z}$ be compatible vectors such that $d_F + \lambda_1, d_F + \lambda_2$ and $d_F + \lambda_1 + \lambda_2$ are non-negative and $\ell : 2^E \to \mathbb{R}$ be a discrete separable convex function. Then*

$$\ell(d_F + \lambda_1 + \lambda_2) - \ell(d_F) \geq (\ell(d_F + \lambda_1) - \ell(d_F)) + (\ell(d_F + \lambda_2) - \ell(d_F)). \quad (3)$$

Proof. By definition $\ell(d_F) = \sum_{v \in V} \ell_v(d_F)(v)$. Thus it suffices to show that for all $v \in V : \ell_v(d_F + \lambda_1 + \lambda_2)(v) \geq \ell_v(d_F + \lambda_1)(v) + \ell_v(d_F + \lambda_2)(v) - \ell_v(d_F)$. If $\lambda_1(v), \lambda_2(v) > 0$ then by Proposition 2.1

$$\ell_v(d_F + \lambda_1 + \lambda_2)(v) - \ell(d_F + \lambda_1)(v) \geq \ell_v(d_F + \lambda_2)(v) - \ell(d_F)(v).$$

If $\lambda_1(v), \lambda_2(v) < 0$ again Proposition 2.1 yields

$$\begin{aligned}
&\ell_v(d_F)(v) - \ell_v(d_F + \lambda_2)(v) \\
=\; &\ell_v(d_F + \lambda_1 + \lambda_2 + |\lambda_1| + |\lambda_2|)(v) - \ell_v(d_F + \lambda_1 + \lambda_2 + |\lambda_1|)(v) \\
\geq\; &\ell_v(d_F + \lambda_1)(v) - \ell_v(d_F + \lambda_1 + \lambda_2)(v)
\end{aligned}$$

implying the assertion. $\qquad\square$

Theorem 3.3 ([1, 2]). *Let (G, k, ℓ) be an instance of Problem 2.2 and $F \subseteq E$ satisfy $|F| = k$. Then F is an optimum solution of MCFS if and only if there exists neither an even alternating walk Q nor a pair of odd alternating walks P^+, P^i as in Proposition 3.1 such that*

$$\ell(d_{F \Delta Q}) < \ell(d_F) \text{ resp. } \ell(d_{F \Delta P^+ \Delta P^-}) < \ell(d_F).$$

Proof. If there exists such a alternating walks then clearly F is not optimal. On the other hand let $F' \subseteq E$ satisfy $|F'| = k$ and $\ell(d_{F'}) < \ell(d_F)$. Then Lemma 3.2 inductively implies

$$\begin{aligned}
0 \;>\; & \ell(d_{F'}) - \ell(d_F) \\
=\; & \ell\left(d_F + \sum_{j=1}^{t} \delta_F(Q_j) + \sum_{i=1}^{s} \delta_F(P_i^+ \cup P_i^-)\right) - \ell(d_F) \\
\geq\; & \sum_{i=1}^{s}(\ell(d_F + \delta_F(Q_j)) - \ell(d_F)) + \sum_{i=1}^{s}\left(\ell(d_F + \delta_F(P_i^+ \cup P_i^-) - \ell(F)\right).
\end{aligned}$$

Hence, at least one of the summands from the last expression must be negative, implying the assertion. $\qquad\square$

4   A Construction

In this section we reduce MCFS to a maximum weighted f-factor problem by adding an extra vertex that "attracts" edges of vertices of high degree. If a vertex has degree $d(v)$ then the first edge, if not included in F will save $\ell_v(d(v)) - \ell_v(d(v) - 1)$, the second $\ell_v(d(v) - 1) - \ell_v(d(v) - 2)$ and so on. If we connect v to a new vertex x with $d(v)$ parallel edges $e_{v,i}$, assigning to them the weights $\ell_v(i) - \ell_v(i-1)$ for $i = 1, \ldots, d(v)$, then a maximum weighted f-factor will prefer edges of larger index from these.

Hence, we construct an auxiliary graph $G' = (V \cup \{x\}, E \cup E^x)$ by adding a vertex x to G that is adjacent to each $v \in V$ by $d_G(v)$ parallel edges $e_{v,i}$ where $1 \leq i \leq d(v)$ and set

$$E^x := \bigcup_{v \in V} \{e_{v,i} \mid 1 \leq i \leq d(v)\}.$$

We define a weight function on the edgeset $E' := E \cup E^x$

$$w(e) := \begin{cases} 0 & \text{if} \quad e \in E \\ l(i) - l(i-1) & \text{if} \quad e = e_{v,i} \in E^x. \end{cases}$$

The function $f : V \cup \{x\}$ for our f-factor problem is defined as

$$f(v) := \begin{cases} 2|E| - 2k & \text{if} \quad v = x \\ d_G(v) & \text{if} \quad v \in V. \end{cases}$$

We say that an f-factor H' of G' is *proper*, if

$$\forall v \in V \; \forall 1 \leq i \leq \deg_G(v) - 1 : (e_{v,i} \in E(H') \Rightarrow e_{v,i+1} \in E(H')).$$

Thus, an f-factor is proper if and only if it prefers parallels of larger to those of smaller index. We have the following one-one-correspondence between edge sets $F \subseteq E$ of size k and proper f-factors of G'.

Lemma 4.1. *If H' is an f-factor of G' then $F = E \cap E(H')$ determines a subgraph $H = (V, F)$ of G such that $|F| = k$. If on the other hand $F \subseteq E$ determines a subgraph $H = (V, F)$ of G such that $|F| = k$, then*

$$E(H') := F \cup \bigcup_{v \in V} \{e_{v,i} \mid \deg_H(v) + 1 \leq i \leq \deg_G(v)\}$$

yields a proper f-factor H' of G'. If H' is proper we furthermore have

$$\ell(d_F) = \sum_{e \in E'} w(e) - \sum_{e \in E(H')} w(e) + \sum_{v \in V} \ell_v(0) = w(G') - w(H') + \sum_{v \in V} \ell_v(0).$$

Proof. If H' is an f-factor then $F = E \cap E(H')$ determines a subgraph $H = (V, F)$ of G such that

$$\sum_{v \in V} d_F(v) = \sum_{v \in V} d_G(v) - (2|E| - 2k) = 2k.$$

Hence $|F| = k$.

If, on the other hand, F is given and H' is as above, then $d_{H'}(x) = \sum_{v \in V}(d_G(v) - d_F(v)) = 2|E| - 2k$ and $\forall v \in V$ we have $d_{H'}(v) = d_F(v) + (d_G(v) - d_F(v)) = d_G(v)$. Hence H' is an f-factor of G', which is proper by construction. If H' is proper then furthermore

$$
\begin{aligned}
w(G') - w(H') &= \sum_{e \in E'} w(e) - \sum_{e \in E(H')} w(e) \\
&= \sum_{v \in V} \sum_{i=1}^{d_G(v)} (\ell_v(i) - \ell_v(i-1)) - \sum_{v \in V} \sum_{i=d_F(v)+1}^{d_G(v)} (\ell_v(i) - \ell_v(i-1)) \\
&= \sum_{v \in V} \sum_{i=1}^{d_F(v)} (\ell_v(i) - \ell_v(i-1)) = \ell(d_F) - \sum_{v \in V} \ell_v(0).
\end{aligned}
$$

$\square$

Theorem 4.2. *F determines a subgraph $H = (V, F)$ of G such that $|F| = k$ and*

$$\ell(d_F) := \sum_{v \in V} \ell_v(d_F(v))$$

is minimized if and only if H' is an f-factor of G' maximizing $w(F)$.

Proof. If F is any set of k edges of G then H' constructed as in Lemma 4.1 is a proper f-factor of G' of weight $w(G') - \ell(d_F) + \sum_{v \in V} \ell_v(0)$ which is of maximum weight if and only if $\ell(d_F)$ is a minimizer of MCFS.

If on the other hand H' is an f-factor such that $w(H')$ is maximized then since l is a discrete convex function and hence

$$\forall i \geq 1 \, \forall v \in V : \ell_v(i) - \ell_v(i-1) \leq \ell_v(i+1) - \ell_v(i),$$

eventually exchanging some parallel edges of equal weight we may assume that H' is proper and the claim follows. $\square$

5 Algorithmic consequences

The algorithm of Gabow [3] solves the weighted f-factor problem for $|V| = n, |E| = m$ in $O((\sum_{v \in V} f(v)) m \log n)$ (see e.g. [7]). It actually does not require the underlying graph to be simple, although we could eliminate parallels without increasing the order of its complexity by subdividing each parallel edge twice. Since

$$\sum_{v \in V \cup \{x\}} f(v) = 2|E| - 2k + \sum_{v \in V} d_G(v) = O(m)$$

Gabow's algorithm will solve our problem in $O(m^2 \log n)$.

References

[1] N. Apollonio and A. Sebő, *Minsquare factors and maxfix covers of graphs.*, in IPCO, G. L. Nemhauser and D. Bienstock, eds., vol. 3064 of Lecture Notes in Computer Science, Springer, 2004, pp. 388–400.

[2] N. Apollonio and A. Sebö, *Minconvex factors of prescribed size in graphs*, Technical Report 145, Les Cahiers du Laboratoire Leibniz, Grenoble, France, January 2006.

[3] H. N. Gabow, *An efficient reduction technique for degree-constrained subgraph and bidirected network flow problems*, in STOC, ACM, 1983, pp. 448–456.

[4] L. Lovasz and M. D. Plummer, *Matching Theory*, North-Holland Mathematics Studies : Annals of Discrete Mathematics, North-Holland, 1986. LOV l 86:1.

[5] K. Murota, *Discrete Convex Analysis: Monographs on Discrete Mathematics and Applications 10*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2003.

[6] ——, *M-convex functions on jump systems: A general framework for minsquare graph factor problem.*, SIAM J. Discrete Math., 20 (2006), pp. 213–226.

[7] A. Schrijver, *Combinatorial optimization : polyhedra and efficiency*, volume A, paths, flows, matchings, chapter 1-38, Springer, 2003. Schrijver.