

Pebble Game Algorithms and Sparse Graphs

Audrey Lee^{a,1} Ileana Streinu^{b,2}

^a*Computer Science Department, University of Massachusetts Amherst, MA, USA*

^b*Computer Science Department, Smith College, Northampton, MA, USA*

Abstract

A multi-graph G on n vertices is (k, ℓ) -sparse if every subset of $n' \leq n$ vertices spans at most $kn' - \ell$ edges. G is *tight* if, in addition, it has exactly $kn - \ell$ edges. For integer values k and $\ell \in [0, 2k)$, we characterize the (k, ℓ) -sparse graphs via a family of simple, elegant and efficient algorithms called the (k, ℓ) -pebble games.

Key words: sparse graph, pebble game, Henneberg sequence, matroid, circuit

1 Introduction

A multi-graph $G = (V, E)$ with $n = |V|$ vertices and $m = |E|$ edges is (k, ℓ) -sparse if every subset of $n' \leq n$ vertices spans at most $kn' - \ell$ edges. If, furthermore, $m = kn - \ell$, G is called *tight*. A (k, ℓ) -spanning graph is one containing a tight subgraph that spans the entire vertex set V . For brevity, we will refer to G as a *graph* instead of as a *multi-graph* (even though it may have loops and multiple edges) and will abbreviate (k, ℓ) -sparse as *sparse*.

Historical overview. Sparse graphs first appeared in Loréa (18), as examples of matroidal families. Classical results of Nash-Williams (20) and Tutte (28) identify the class of graphs decomposable into k edge-disjoint spanning trees with the (k, k) -tight graphs. Tay (26) relates them to generic body-and-bar rigidity in arbitrary dimensions. The $(2, 3)$ -tight graphs are the *generic*

Email addresses: alee@cs.umass.edu (Audrey Lee), streinu@cs.smith.edu (Ileana Streinu).

¹ Research supported by an NSF graduate research fellowship and by the NSF grant CCR-0310661 of the second author.

² Research partially funded by NSF grant CCR-0310661.

minimally rigid (or Laman) graphs for bar-and-joint frameworks in the plane (16), and the spanning ones correspond to those which are *rigid*.

A (k, a) -arborescence is defined as a graph where adding *any* a edges results in k edge-disjoint spanning trees. Results of Recski (22) and Lovasz and Yemini (19) identify Laman graphs with $(2, 1)$ -arborescences. For $\ell \in [k, 2k)$, this is extended by Haas (9) to an equivalence of (k, ℓ) -sparse graphs and $(k, \ell - k)$ -arborescences. Whiteley (29; 30) surveys several rigidity applications where sparse graphs appear, some having non-integer parameters associated to them. Frank, Szegő and Fekete (6; 24; 5) study inductive constructions for various subclasses of sparse graphs, motivated by the so-called Henneberg sequences appearing in Rigidity Theory (13), and Bereg (1) computes them with an $O(n^2)$ algorithm for the minimally rigid (Laman) case.

There exist many algorithms for decomposing a graph into edge-disjoint trees or forests (4; 7; 8; 25; 23). A variation on the $O(n^2)$ time matching-based algorithm of (11) for 2-dimensional rigidity became the simple and elegant *pebble game algorithm* of Jacobs and Hendrickson (14), further analyzed in (2). Practical applications in studies of protein flexibility led Jacobs et al. (15) to pebble game heuristics for special cases of three-dimensional rigidity. However, intriguingly, we have not found anywhere algorithms applicable to (k, a) -arborescences or to the *entire* class of (k, ℓ) -sparse graphs.

Our results. In this paper, we describe a *family of algorithms*, called the (k, ℓ) -pebble games, and prove that they recognize *exactly* the (k, ℓ) -sparse graphs, for the *entire range* $\ell \in [0, 2k)$.

In our terminology, Jacobs and Hendrickson's is a $(2, 3)$ -pebble game. We exhibit here the full extent to which their algorithm can be generalized, and characterize the recognized classes of *graphs*. We study the following fundamental problems.

- (1) **Decision:** is G a tight (or just sparse) graph?
- (2) **Spanning:** does G span a tight subgraph?
- (3) **Extraction:** extract a maximal sparse subgraph (ideally, spanning) from a given graph G .
- (4) **Optimization:** from a graph with *weighted* edges, extract a maximum weight sparse subgraph.
- (5) **Components:** given a non-spanning graph G , find its components (maximal tight induced subgraphs).

The pebble game algorithms run in time $O(n^2)$ using simple data structures and induce good algorithmic solutions for all the above problems. They exhibit the same complexity as Hendrickson's matching-based algorithm (10; 11) for 2-dimensional rigidity. For the *special case* of graphs decomposable into disjoint unions of spanning trees and pseudo-forests, corresponding to the range $\ell \in$

$[0, k]$ of (k, ℓ) -sparse graphs, we remark that there are $O(n^{3/2})$ algorithms due to Gabow and Westerman (7). But no better algorithms than the pebble games are known for the *entire* range of (k, ℓ) -sparse graphs.

2 Properties of Sparse Graphs

We start by showing why it is *natural* to restrict the range of the integer parameter ℓ to $[0, 2k)$. Then we identify a dual property related to a well-known theorem of Nash-Williams and Tutte (20; 28) on tree decompositions. Finally, we define *components* and give a detailed characterization of their main structural properties.

All graphs $G = (V, E)$ in this paper have $n = |V|$ vertices and $m = |E|$ edges. For subgraphs $E' \subset E$ induced on subsets $V' \subset V$, we use $n' = |V'|$ and $m' = |E'|$. The complete multi-graph on n vertices, with multiplicity a on loops and b on edges, is denoted by $K_n^{a,b}$, and the loopless version by K_n^b . The *degree* of a vertex is the number of incident edges, including loops. The parameters k and ℓ are integers.

Matroidal sparse graphs. The following Lemma justifies the choice of parameters and points to a small correction to the informal definition of sparse graphs we gave in the introduction: because for the range $\ell \in (k, 2k)$ and for $n' = 1$, $kn' - \ell$ becomes negative, we should require that every subset of $n' \leq n$ vertices spans at most $\max\{0, kn' - \ell\}$ edges.

Lemma 1 Properties of sparse graphs.

- (1) If $\ell \geq 2k$, the class of sparse graphs contains only the empty graph.
- (2) If $\ell < 0$, the union of two vertex disjoint sparse graphs may not be sparse.
- (3) **Loops and parallel edges** A sparse graph may contain at most $k - \ell$ loops per vertex. In particular, the sparse graphs are loopless when $\ell \geq k$. The multiplicity of parallel edges is at most $2k - \ell$.
- (4) **Single vertex graphs** In the upper range $\ell \in (k, 2k)$, there are no tight graphs on a single vertex.
- (5) **Small tight graphs** (Szegő (24)) If $\ell \in [\frac{3}{2}k, 2k)$ (called the Szegő range), there are no tight graphs on small sets of n vertices, for $n \in (2, \frac{\ell}{2k-\ell})$.
- (6) **Smallest tight graphs** When $\ell \in [\frac{3}{2}k, 2k)$, the smallest non-trivial tight sparse graphs have $\lceil \frac{\ell}{2k-\ell} \rceil$ vertices. For integer values of $\frac{\ell}{2k-\ell}$, there is only one tight graph on the minimum number of vertices: the complete multi-graph $K_{\frac{\ell}{2k-\ell}}^{2k-\ell}$; otherwise, there will be several.

Proof (1) For $\ell \geq 2k$, any subset of $n' = 2$ vertices would span at most $2k - \ell \leq 0$ edges. (2) If we take the vertex disjoint union of two tight sparse graphs on n_1 , resp. n_2 vertices, the union has $n = n_1 + n_2$ vertices and $k(n_1 + n_2) - 2l > kn - \ell$ edges, therefore it is not sparse. (3) Apply the sparsity condition $m' < kn' - \ell$ for $n' = 1$ and $n' = 2$. (4) Indeed, $kn - \ell < 0$ for $n = 1$, and the number of edges cannot be negative. (5) Assume $\ell \geq k$. A vertex may not span a negative number of edges, so $n \geq 2$. By part (3) above, a tight graph with $kn - \ell$ edges is a subgraph of the complete, loopless $(2k - \ell)$ -multi-graph $K_n^{2k-\ell}$; therefore $kn - \ell = m \leq (2k - \ell) \binom{n}{2}$. The inequality between the extremes leads to the condition $f(n) \geq 0$ for the quadratic function $f(n) = an^2 + bn + c$, with $a = 2k - \ell$, $b = \ell - 4k$ and $c = 2\ell$. The two roots of $f(n) = 0$ are $n_1 = 2$ and $n_2 = \frac{\ell}{2k - \ell}$. The *open* interval between the roots is non-trivial when it contains at least one integral value, i.e. when $n_2 \geq 3$. This happens exactly when $\ell > \frac{3}{2}k$. For values of n within this interval, all the subgraphs of $K_n^{2k-\ell}$ are (k, ℓ) -sparse, but none is tight. (6) Direct corollary of (5). ■

The range of values $\ell \in [0, k)$ is called the *lower range* and $\ell \in [k, 2k)$ is the *upper range*: the threshold case $\ell = k$ will occasionally be relevant for properties holding in either range (so we will specify when the lower and upper range intervals need to be taken as open or closed). The upper range is further subdivided into two, of which the *Szegő range* requires special care in applications such as Henneberg sequences. This phenomenon, of having to deal with special cases depending on the range of ℓ , is symptomatic for sparse graphs and impacts the choice of data structures for our algorithms. At the upper bound $\ell = 2k - 1$, the smallest tight graphs are complete graphs. For example, when $k = 3$ and $\ell = 5$, the smallest tight graph is K_5 . For other values of k and ℓ , there may be several *smallest* tight graphs. For example, when $k = 7$ and $\ell = 11$, there are 6 smallest tight graphs: all the multi-graphs on 4 vertices with a total of 17 edges and edge-multiplicity at most 3.

For values of the parameters k, ℓ and n in these ranges, we show now that the tight graphs form the set of *bases* of a matroid. The proof relies on a very simple property of *blocks* given below on page 6. White and Whiteley, in the appendix of (30), observed that the matroid *circuit* axioms are satisfied.

Theorem 2 (The (k, ℓ) -sparsity matroid) *Let n, k and ℓ satisfy: (1) $\ell \in [0, k]$ and $n \geq 1$; (2) $\ell \in (k, \frac{3}{2}k)$ and $n \geq 2$; (3) $\ell \in [\frac{3}{2}k, 2k)$ and $n = 2$ or $n \geq \frac{\ell}{2k - \ell}$. Then the collection of all the (k, ℓ) -tight graphs on n vertices, is the set of bases of a matroid whose ground set is the set of edges of the complete multi-graph on n vertices, with loop multiplicity $k - \ell$ and edge multiplicity $2k - \ell$.*

Proof We verify the three axioms of a basis system. **Equal cardinality** holds by definition. To prove **Non-emptiness**, we construct canonical tight graphs as follows. Let $V = \{1, \dots, n\}$. For $\ell \in [0, k)$, $n \geq 1$, place $k - \ell$ loops per

vertex; connect the vertices with ℓ trees (e.g. ℓ copies of the *same* tree). For $\ell \in [k, \frac{3}{2}k)$, $n \geq 2$, place $2k - \ell$ parallel edges between vertices 1 and 2. For each vertex $i > 2$, place $2k - \ell$ parallel edges between vertices i and 1, and $\ell - k < 2k - \ell$ edges between vertices i and 2. Finally, consider the case $\ell \in [\frac{3}{2}k, 2k)$. For $n = 2$, there is only one tight graph, the $(2k - \ell)$ -multi-edge. For $n \geq \frac{\ell}{2k - \ell}$, start with an arbitrary minimum-size tight graph on the set of vertices indexed from 1 to $\lceil \frac{\ell}{2k - \ell} \rceil$. For all vertices of larger index $i > \lceil \frac{\ell}{2k - \ell} \rceil$, place k edges between i and some of the vertices of index $\leq \lceil \frac{\ell}{2k - \ell} \rceil - 1$, saturating the multiplicity $2k - \ell$ of a vertex of index i before moving on to the next vertex of index $i + 1$.

To prove the **Basis exchange axiom**, let $G_j = (V, E_j)$, $j = 1, 2$, be two tight graphs and $e_2 \in E_2 \setminus E_1$. We must show that there exists an edge $e_1 \in E_1 \setminus E_2$ such that $(V, E_1 \setminus \{e_1\} \cup \{e_2\})$ is tight. Let $e_2 = uv$ (this includes the case $u = v$ when e_2 is a loop). Consider all the tight induced subgraphs (called *blocks*) $H_i = (V_i, E_i)$ of G_1 containing vertices u and v . Let $V' = \bigcap_i V_i$ and $H' = (V', E')$ be the subgraph of G_1 induced on V' . By Theorem 5(1) proved below in Section 2, H' is a block of G_1 . Not all the edges in H' are in G_2 , i.e. H' cannot be a block of G_2 , since V' also spans e_2 in G_2 and then the subgraph $E' \cup \{e_2\} \subset E_2$ would violate the sparsity of G_2 . Therefore, H' contains at least one edge $e_1 \in E_1 \setminus E_2$. We are done if we show that $H_3 = (V', E_1 \setminus \{e_1\} \cup \{e_2\})$ is sparse. Indeed, H' is the minimal subgraph of G_1 such that the addition of e_2 violates sparsity: any other subset would have been one of the V_i , and V' is contained in it. Since V' is contained in any subset on which sparsity was violated in $G_1 \cup \{e_2\}$, the removal of e_1 restores the counts. ■

In Theorem 2, the ground set $K_n^{k-\ell, 2k-\ell}$ was chosen to produce all the interesting bases. We may enlarge the ground set, by adding extra loops and parallel edges, or delete edges from it, by working with a subgraph of $K_n^{k-\ell, 2k-\ell}$, and we still obtain a matroid. In the first case, the bases will still be restricted to the number of edges required by the sparsity conditions; in the second case, the bases are maximal sparse subgraphs of G . This allows us later to refer to the *matroidal property of sparse graphs* as reason for the correctness of the *arbitrary order of edge insertion* in the pebble game algorithms, and of the greedy algorithm for the Optimization Problem (see (3), p.345 and (21)).

Partitioning. Nash-Williams (20) and Tutte (28) gave an alternative definition of (k, k) -tight graphs using vertex partitions and trees: a graph contains k edge-disjoint spanning trees if and only if every partitioning of the vertex set into p parts has at least $k(p - 1)$ edges between them. If, moreover, it has $kn - k$ edges, it is the edge-disjoint union of k spanning trees and a (k, k) -tight graph. We describe now a slight generalization of one direction of their criterion, for all (k, ℓ) -tight graphs.

Lemma 3 Let $G = (V, E)$ be a (k, ℓ) -tight graph and $P = \{V_1, \dots, V_p\}$ a partition of V . In the upper range $\ell \in (k, 2k)$, further assume that each $|V_i| \geq 2$. Then there are at least $\ell(p-1)$ edges between the partition sets V_i .

Proof Let E_i be the edge set induced by V_i in G and $n_i = |V_i|, m_i = |E_i|$. By sparsity and the assumption on the size of V_i , $m_i \leq kn_i - \ell, \forall i$ and $\sum_i m_i \leq \sum_i (kn_i - \ell) \leq kn - p\ell$. The number of edges between the partition sets is $m - \sum_i m_i \geq kn - \ell - (kn - p\ell) \geq \ell(p-1)$. ■

Lemma 4 Let $G = (V, E)$ be a tight graph. Then every vertex has degree $\geq k$. Moreover, if $\ell > 0$, then there is at least one edge between a vertex v and the rest of the vertices $V \setminus \{v\}$.

Proof If $v \in V$ had degree $d < k$, the induced subgraph on $V \setminus \{v\}$ would have $kn - \ell - d > kn - \ell - k = k(n-1) - \ell$ edges, contradicting the sparsity of G . This already implies the second part of the theorem for $\ell \in (k, 2k)$, because sparse graphs in this range have no loops. The other case $\ell \in (0, k]$ follows from Lemma 3. ■

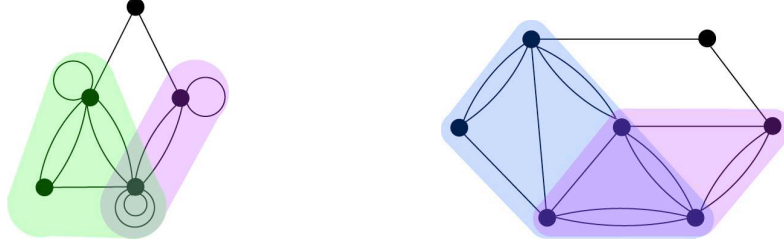
As a simple corollary, when $\ell > 0$, a tight graph is connected. We will make use of this small observation in Theorem 5 (4). Also, as a consequence of the theorem of Nash-Williams and Tutte we have that, for $\ell \in (0, k]$, a (k, ℓ) -tight graph contains ℓ edge-disjoint spanning trees.

Blocks, Components and Circuits. In a sparse graph, a subset of vertices $V' \subset V$ may span *exactly* $kn' - \ell$ edges, where $n' = |V'|$. In this case, the induced subgraph is called a *block*. A maximal block (with respect to the set of vertices) is called a *component*. We describe now basic properties of blocks and components.

We start with a decomposition theorem for a sparse graph into *components*, *free vertices* (not part of any component) and *free edges* (not spanned by any block, and hence component). In rigidity applications, the components correspond to rigid clusters. This decomposition will be used later in speeding up the pebble game. For this section, denote the range $\ell \in [0, k]$ as the *lower range* and $\ell \in (k, 2k)$ as the *upper range*.

Theorem 5 (Decomposition into Components) Let G be a sparse graph.

- (1) **Block intersection:** if two blocks intersect in at least: (a) one vertex, for the lower range [see Fig. 1a]; (b) two vertices, for the upper range [see Fig. 1b], then their intersection and union (with respect to the vertex sets) induce blocks.
- (2) **Component interaction:** sparse components are edge-disjoint. In the lower range, the components are vertex-disjoint [see Fig. 2a]. In the upper



(a) Lower range example ($k = 3, \ell = 1$): two blocks overlapping in one vertex. (b) Upper range example ($k = 3, \ell = 4$): two blocks overlapping in three vertices.

Fig. 1. Block intersection: two overlapping blocks whose union and intersection are also blocks.

range, they overlap in at most one vertex [see Fig. 2b].

(3) **Component connectivity:**

(a) When $\ell = 0$, there is at most one component, which may not be connected [see Fig. 2c].

(b) When $\ell > 0$, blocks (and therefore components) are connected [see Fig. 1, 2a, 2b, and 2d].

(4) **Decomposition:** G is decomposed into components, free vertices and free edges. More specifically:

(a) Lower range: a single vertex induces a block if and only if it has $k - \ell$ loops. In this case, if $\ell = 0$, the block may be a disconnected piece of a larger component, otherwise it is a component in itself [see Fig. 2c]. A vertex with fewer than $k - \ell$ loops in the lower range, or a vertex in the upper range is either free or part of a larger block (and hence component). When $\ell = 2k - 1$, there are no free vertices: each vertex is part of some block (and hence component), but it is never a block in itself.

(b) $\ell = k$: a single vertex is loop-free and is always a block. Thus, there are no free vertices, and V is partitioned into components (possibly connected by free edges)[see Fig. 2a].

(c) $\ell = 2k - 1$: there are no loops or parallel edges. A single vertex is free only when it is an isolated vertex of the graph. A single edge is always a block, thus there are no free edges, and E is partitioned into components [see Fig. 2b].

Proof (1) Let $B_i = (V_i, E_i), i = 1, 2$, be two blocks of a sparse graph $G = (V, E)$; they span $m_i = kn_i - \ell$ edges, $i = 1, 2$. Let G_\cap and G_\cup be the subgraphs of G induced on the intersection $V_1 \cap V_2$ (with n_\cap vertices and m_\cap edges), resp. union $V_1 \cup V_2$ (with n_\cup vertices and m_\cup edges), of their vertex sets. Then $m_\cup = m_1 + m_2 - m_\cap = (kn_1 - \ell) + (kn_2 - \ell) - m_\cap = k(n_1 + n_2) - 2\ell - m_\cap = k(n_\cap + n_\cup) - 2\ell - m_\cap = kn_\cup - \ell - (m_\cap - (kn_\cap - \ell))$. Since G is sparse, $m_\cup \leq kn_\cup - \ell$; thus, $m_\cap - (kn_\cap - \ell) \geq 0$, i.e., $m_\cap \geq kn_\cap - \ell$.

If $\ell \in [0, k]$, assume $n_\cap \geq 1$, and if $\ell \in (k, 2k)$, assume $n_\cap \geq 2$. Since G is

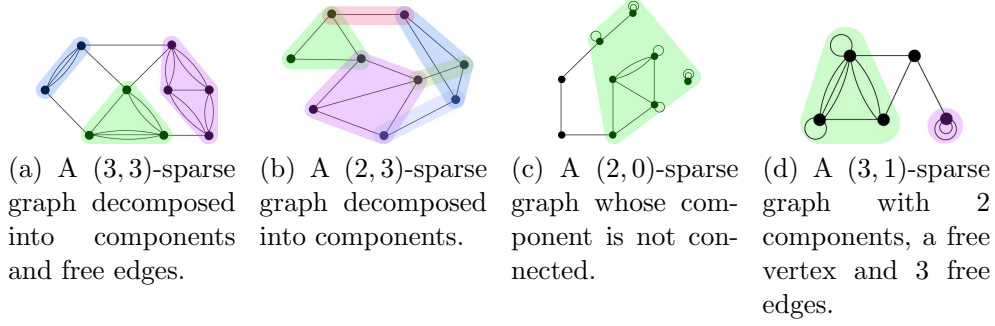


Fig. 2. Decomposition into components.

sparse, $m_\cap \leq kn_\cap - \ell$; therefore it follows that $m_\cap = kn_\cap - \ell$ and $m_\cup = kn_\cup - \ell$ and thus, both the induced intersection and union are blocks.

(2) Follows from the same calculations used in part (1).

(3) Lemma 4 implies that when $\ell > 0$, tight graphs are connected. For $(k,0)$ -sparse graphs, assume there exist several vertex-disjoint tight sparse subgraphs (blocks). A simple application of the sparsity counts shows that the union is also $(k,0)$ -tight.

(4) Take $n = 1$ and $n = 2$ in the definition of (k, ℓ) -sparsity, and analyze each case. ■

A reminder that, in matroid theory terminology, a set of elements of the ground set E of a matroid is *independent* if it is a subset of a basis. An element e of the ground set is *independent with respect to* a given independent set $I \subset E$ if $I \cup \{e\}$ is an independent set. Thus, sparse graphs are independent, and independent edges may be added to a sparse graph until it becomes tight. The obstructions to adding further edges in a sparse graph are the *blocks*, as stated in the following straightforward corollary to Theorem 2.

Corollary 6 *An edge is independent with respect to a sparse graph G if and only if its endpoints do not belong to some block of G .*

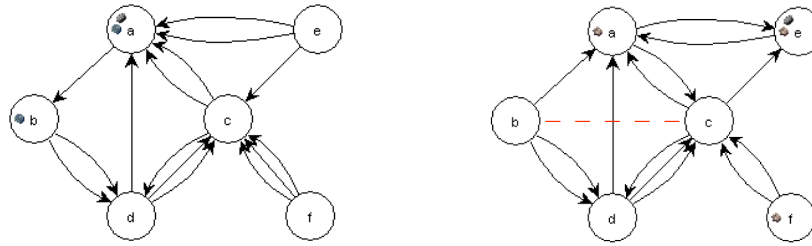
A minimal subset (of vertices and edges) violating sparsity is called a *circuit*. An edge which is not independent of a given sparse graph G violates the sparsity condition on some subset of vertices and induces a unique circuit, which can be identified using the criterion below.

Corollary 7 *Let G be a tight graph and let $e = uv$ be an edge not in G . The intersection of all the blocks containing u and v is a block H of G , called the minimal block spanning e . Furthermore, $H \cup \{e\}$ is a circuit in $G \cup \{e\}$.*

3 The basic (k, ℓ) -Pebble Game Algorithm

We turn now to the description of our generalized (k, ℓ) -pebble game for multi-graphs. Fig. 3 illustrates an example. We start with the simplest version, called the *Basic Pebble Game*. Later, we will extend it to a more efficient version which takes *components* into account. The correctness of the pebble game as a decision algorithm for sparse graphs is proven in the next section.

The algorithm depends on two parameters, k and ℓ : k is the *initial number of pebbles* on each vertex, and $\ell + 1$ is a lower bound on the total number of pebbles present at the two endpoints of an edge which is accepted during the execution of the algorithm.



(a) A *Well-constrained* $(3, 3)$ -pebble game output, with the final orientation and distribution of the remaining 3 pebbles on the input graph.

(b) An *Under-constrained* $(3, 3)$ -pebble game output: note the 4 remaining pebbles. If the dotted edge was part of the input, it could not be inserted: the pebble game would *fail*.

Fig. 3. Final state of the $(3, 3)$ -pebble game on two graphs.

The algorithm is built on top of a single-person *game*, played on a *board* consisting of a set of n nodes, initialized with k pebbles each. The player inserts edges between the nodes and orients them. The rules of the game indicate when an edge will be accepted (and therefore inserted) or rejected, and when the player can move pebbles and reorient already inserted edges. We give no rules for when this generic “game” should be stopped, nor do we specify what it means to win or to lose it: indeed, we do not analyze the game per se, but rather the algorithm built on top of it.

The algorithm takes a given graph as input, and considers its edges in an arbitrary order. It performs the moves of the game for the insertion or rejection of each edge. When all the edges have been considered, the algorithm ends with a classification of the input graph into one of four categories. The first two, *Well-constrained* and *Under-constrained*, correspond to *success* in accepting all the edges of the input graph; the other two, *Over-constrained* and *Other*, indicate the failure to fully accept the input graph. In Section 4 we prove that these categories correspond exactly to the input graph being *tight*, *sparse*,

spanning and neither sparse nor spanning. The algorithm is described in Fig. 4.

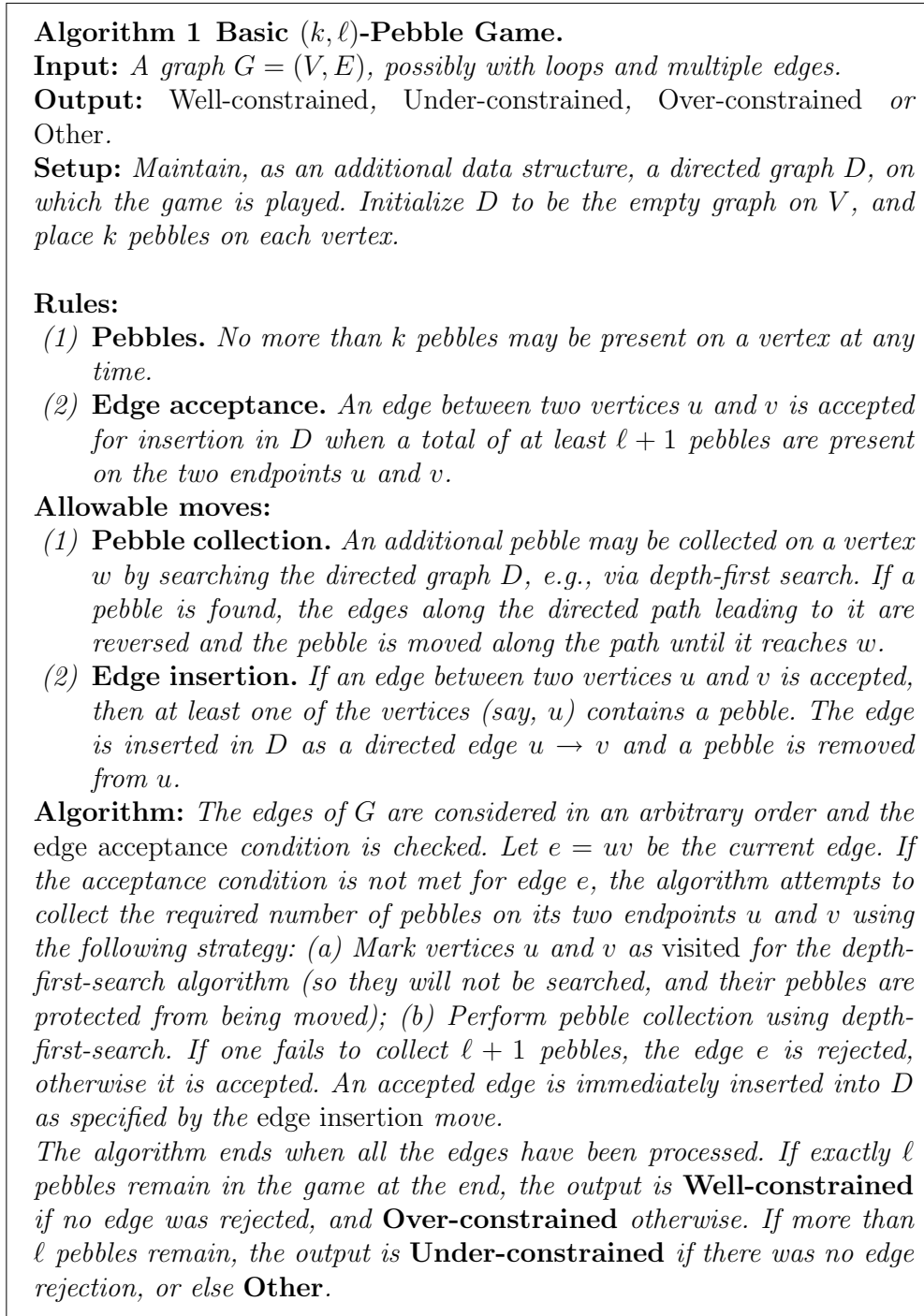


Fig. 4. Basic (k, ℓ) -pebble game algorithm.

Complexity analysis. Let m_a be the number of accepted edges in the final state of the game. Since each accepted edge requires the removal of one peb-

ble, $m_a = O(kn)$. The only data structure used by the pebble game is the additional digraph D , whose space complexity is $O(m_a + n) = O(kn)$. Each edge is considered exactly once and requires at most $\ell + 1$ depth-first searches through D , for a total of $O(\ell mn)$ time. For constant parameters k and ℓ , and dense input graphs with $O(n^2)$ edges, this algorithm has worst case $O(n^3)$ time and $O(n)$ space complexity. The time will be improved in Section 5.

4 Pebble Game Graphs coincide with Sparse Graphs

We are now ready to prove the main theoretical result of the paper, relating pebble games to sparse graphs.

Theorem 8 (Pebble Game Graphs and Sparse Graphs) *The class of Under-constrained pebble game graphs coincides with the class of sparse graphs, Well-constrained ones coincide with tight graphs, Over-constrained coincide with spanning ones and Other are neither sparse nor spanning.*

Corollary 9 *The basic Pebble Game solves the Decision, Extraction, Spanning and, with the slight modification of inserting the edges in sorted order of their weights, the Optimization problems for sparse graphs.*

The proof follows from the sequence of lemmas given below. For a vertex v in the directed graph D at some point in the execution of the pebble game algorithm, denote by $peb(v)$ the number of free pebbles on v , $span(v)$ the number of loops and by $out(v)$ its out-degree, i.e. the number of edges starting at v and ending at a different vertex (i.e. excluding loops). We extend these functions to vertex sets in a natural way: for $V' \subset V$, $peb(V') = \sum_{v \in V'} peb(v)$, $span(V')$ is the number of edges spanned by V' (including loops) and $out(V')$ is the number of edges starting at a vertex in V' and ending at a vertex in the complement $V \setminus V'$.

Lemma 10 (Invariants of the Pebble Game) *During the execution of the pebble game algorithm on a graph G with n vertices, for every vertex v and for every subset $V' \subset V$ on n' vertices, the following invariants are maintained on D . We assume that $n, n' \geq 1$ for $\ell \in [0, k]$ and $n, n' \geq 2$ for $\ell \in (k, 2k)$.*

- (1) $peb(v) + span(v) + out(v) = k$
- (2) $peb(V') + span(V') + out(V') = kn'$
- (3) $peb(V') + out(V') \geq \ell$. In particular, there are at least ℓ free pebbles in the digraph D .
- (4) $span(V') \leq kn' - \ell$

Proof (1) The invariant obviously holds when the game starts. When an edge is inserted into D and is oriented away from v , a pebble is removed from v ; this is true for loops as well, so the total sum is maintained. During a pebble search, if v lies along a path that is reversed to bring a pebble to the path's source, $out(v)$ remains unchanged. If v is the source of a path reversal, $out(v)$ is decreased by 1 and $peb(v)$ is increased by 1; if v is the target of a path reversal, $out(v)$ is increased by 1 and $peb(v)$ is decreased by 1. Hence the sum $peb(v) + span(v) + out(v)$ remains constant throughout the game.

(2) If m_1 is the number of non-loop edges spanned by V' , then $out(V') = \sum_{v \in V'} out(v) - m_1$ and $span(V') = m_1 + \sum_{v \in V'} span(v)$. Therefore $peb(V') + span(V') + out(V') = peb(V') + (m_1 + \sum_{v \in V'} span(v)) + ((\sum_{v \in V'} out(v)) - m_1) = \sum_{v \in V'} peb(v) + \sum_{v \in V'} span(v) + \sum_{v \in V'} out(v) = \sum_{v \in V'} (peb(v) + span(v) + out(v)) = kn'$ (by Invariant (1)).

(3) When the game starts, there are no outgoing edges from V' and $peb(V') = kn' \geq \ell$. Consider now the last time an edge incident to V' was inserted or reoriented by the pebble game algorithm. Four cases have to be analyzed: if the edge had both endpoints in V' , if it went between V' and $V \setminus V'$ and oriented away from or towards V' , or if it was an edge reorientation. In the first case, at least ℓ pebbles must be present on the endpoints of the edge after the insertion, so $peb(V') \geq \ell$. In the second case, the invariant was true before the insertion, and it is true after the insertion: if the edge was inserted away from V' , a pebble was consumed from V' but an outgoing edge was inserted; in the other case, the number of pebbles and outgoing edges was not modified. Finally, if the last move was an edge reorientation, it was either bringing in a pebble from the outside to the inside of V' , and decreasing by 1 the number of outgoing edges, or vice-versa, when it was decreasing the number of inside pebbles by one and increasing the number of outgoing edges.

(4) Straightforward, since $span(V') = kn' - (peb(V') + out(V'))$, and $peb(V') + out(V') \geq \ell$. ■

The following corollaries follow directly from Invariant 4.

Corollary 11 *For any subset $V' \subseteq V$, V' spans a block if and only if $peb(V') + out(V') = \ell$.*

Corollary 12 *Under-constrained pebble game graphs are sparse, Well-constrained ones are tight, Over-constrained ones are spanning.*

This completes the proof of one direction, characterizing the sparsity of the graphs classified by the algorithm. We move now to prove the other direction, that the algorithm classifies correctly sparse, tight and spanning graphs. Denote by $Reach(v)$ the *reachability region* of a vertex v (at some point during the execution of the algorithm): the set of vertices that can be reached via

directed paths from v in D . For example, in Figure 3b, $Reach(d) = \{a, c, d, e\}$.

Lemma 13 *If $e = uv$ is independent (but not yet inserted) in D , and strictly fewer than $\ell + 1$ pebbles are present on u and v , a pebble can be brought to one of u or v without changing the pebble count of the other vertex.*

Proof Let $V' = Reach(u) \cup Reach(v)$; e is independent, so $span(V') < k|V'| - \ell$. Since V' is a union of reachability regions, $out(V') = 0$. By Lemma 10, Invariant 3, $peb(V') > \ell$. By assumption, $peb(u) + peb(v) < \ell + 1$. Then there exists $w \in V'$ such that $w \neq u$ and $w \neq v$ with at least one free pebble. If $w \in Reach(u)$, bring the pebble from w to u . Otherwise, $w \in Reach(v)$; bring the pebble from w to v . ■

Lemma 14 *An edge is inserted by the pebble game if and only if it is independent in D .*

Proof Let $e = uv$ be an edge of G , not yet inserted into D , the current state of the directed graph pebble game data structure. By applying Lemma 13 repeatedly, it follows that $\ell + 1$ pebbles can be gathered on the endpoints u and v ; thus, e will be inserted into D by the pebble game. ■

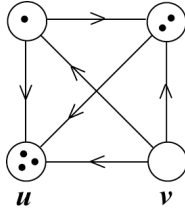


Fig. 5. A (3, 5)-pebble game where no edge parallel to uv can be inserted.

It is instructive to notice that it does not suffice to require that $\ell + 1$ pebbles be present in the reachability regions of u and v . In Fig. 5, an example of a (3, 5)-pebble game is shown. The reachability region for the pair u and v contains 6 pebbles, but not all can be collected on the two vertices u and v . No edge parallel to uv can be inserted. Note also that the reachability region of a vertex may change after a pebble move; the previous proof requires the independence of e at each application of Lemma 13.

Lemma 15 *The pebble game returns Under-constrained for sparse, but not tight graphs, Well-constrained for tight ones, Over-constrained for spanning graphs and Other for graphs that are neither spanning nor sparse.*

Proof Let G be a sparse graph with n vertices and m edges. Because sparse graphs form a matroid (Theorem 2), the order in which the edges are considered can be arbitrary. By Lemma 14, every independent edge is inserted by the pebble game. Thus, the pebble game is successful on sparse graphs.

If G is sparse, but not tight, $m < kn - \ell$. By Lemma 10, Invariant 3, the number of free pebbles in the final game graph must be $> \ell$ and the result is *Under-constrained*. If G is tight, $m = kn - \ell$ and the number of free pebbles is exactly ℓ ; the game returns *Well-constrained*. If G is spanning, it contains a tight subgraph which will be accepted, after which there won't be enough pebbles and the remaining edges will be rejected; the result is *Over-constrained* in this case. If G is neither spanning nor sparse, there must be $> \ell$ pebbles in the final game as well as at least one dependent (and thus rejected) edge. ■

Corollary 12 and Lemma 15 prove Theorem 8.

5 Component Pebble Games

The graph D maintained by the basic pebble game algorithm is sparse. It can therefore be decomposed into components. We now present a modification of the basic pebble game that maintains and uses these components to obtain an algorithm one order of magnitude faster.

Pebble Game with Components. Its input, output and additional directed graph D are the same as for the basic Pebble Game. We give first the overall structure of the algorithm in Figure 6; additional subroutines and some implementation details will be described next.

Algorithm 2 Component Pebble Game.

Input: A graph $G = (V, E)$, possibly with loops and multiple edges.

Output: Well-constrained, Under-constrained, Over-constrained or Other.

Method: Play the basic pebble game with the following modifications. Maintain components throughout the game. When considering edge $e = uv$, first check if u and v are in some common component or if $u = v$ (i.e., e is a loop) and $\ell \in [k, 2k)$. If so, **reject** and discard e . Otherwise, perform pebble searches to gather $\ell + 1$ pebbles on u and v , and insert edge e . Detect new component, if one is formed, and perform necessary component maintenance.

Fig. 6. Component pebble game algorithm.

If the endpoints of an edge do not belong to a component, the pebble searches are guaranteed to succeed, so the edge will be accepted. A newly inserted edge may be a free edge or lead to the creation of a new component (possibly by merging some already existing ones). We present next two different algorithms for computing the vertex set of the new component. Algorithm 3, shown in Figure 7, generalizes the approach of (14) and works similarly to breadth-first search on *in-coming edges* of an already detected block. An example is shown

in Figure 8; notice that vertex f , although it has an edge directed towards $\text{Reach}(a, c)$, contains a pebble and is not added to the component.

Algorithm 3 Component detection I

Input: Directed pebble game graph $D = (V, E')$, into which edge $e = uv$ has just been inserted. At least ℓ pebbles are present on u and v . If $\ell = 0$, the vertex set V_0 (which may be empty) of the single component of $D \setminus \{e\}$ is also given.

Output: The vertex set V' of the new component induced by e , if one was formed; \emptyset , otherwise.

Method:

- (1) If more than ℓ pebbles are present on u and v , **return** \emptyset : the new edge is free.
- (2) Otherwise, compute $\text{Reach}(u, v) = \text{Reach}(u) \cup \text{Reach}(v)$.
 - (a) If any $w \in \text{Reach}(u, v)$ has at least one free pebble, **return** \emptyset .
 - (b) Otherwise, initialize $V' = \text{Reach}(u, v)$. Initialize queue Q , and enqueue all vertices in $V \setminus V'$ with an edge into V' .
While Q has elements
 - (i) Dequeue vertex w from Q .
 - (ii) Compute $\text{Reach}(w)$.
 - (iii) If all vertices in $\text{Reach}(w)$ (other than u and v) have no free pebbles
 - (A) Set $V' = V' \cup \text{Reach}(w)$.
 - (B) Enqueue all vertices (that have not been previously enqueued) with an edge into $\text{Reach}(w)$.
- (3) If $\ell = 0$, merge into V' the vertices of the existing component of G (if it exists).
- (4) **Return** V' .

Fig. 7. Component detection algorithm I.

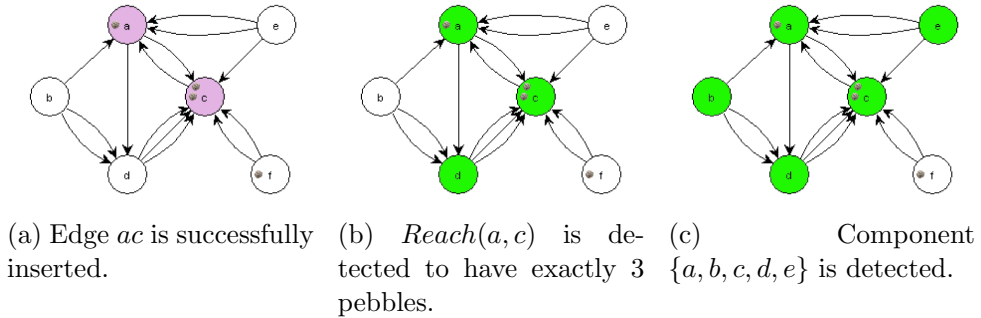


Fig. 8. (3,3) component detection after edge ac has just been inserted. First, $\text{Reach}(a, c) = \{a, c, d\}$ is detected as a block as it contains exactly 3 pebbles; then, component $\{a, b, c, d, e\}$ is detected.

Figure 9 describes the second component detection algorithm, which generalizes (2). It works by finding the *complement* of the vertex set of the newly

formed component, and does not require special treatment for $\ell = 0$.

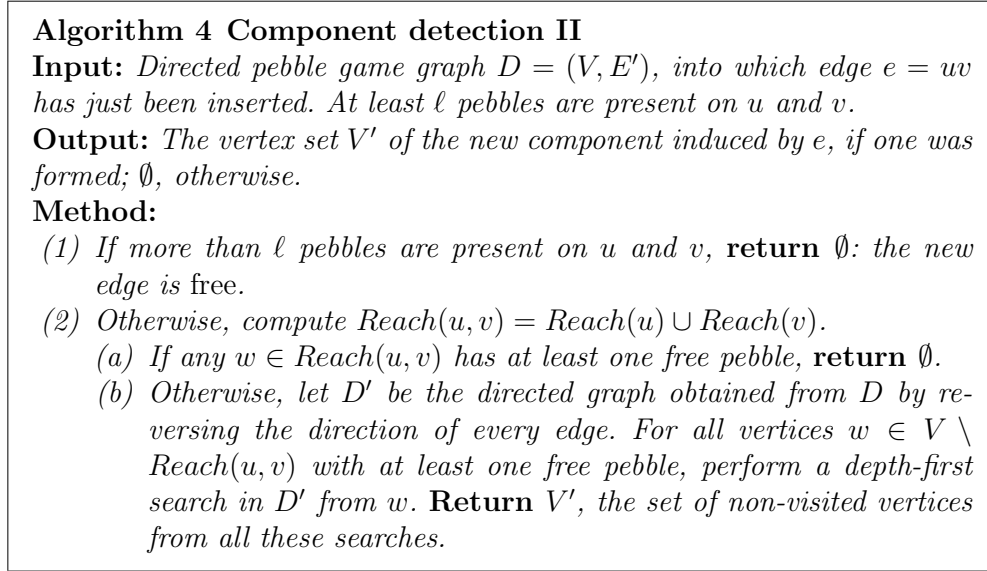


Fig. 9. Component detection algorithm II.

Component maintenance. Maintaining components requires additional bookkeeping. By Theorem 5, we must take the range of ℓ into account. When $\ell = 0$, there is at most one component, which is maintained by a simple *marking scheme*: a vertex is marked if and only if it lies in the component. When $\ell \in (0, k]$, the components are vertex disjoint. Their maintenance is accomplished with a simple *labeling scheme*: each vertex is labeled with an id of the component to which it belongs. In the upper range, when $\ell \in (k, 2k)$, components may overlap in a single vertex. We maintain a list of the components, represented by their vertex sets, as well as an $n \times n$ matrix. The matrix is used to provide constant time queries for whether two vertices belong to some common component; there is a 1 in entry $[i, j]$ if such a component exists and a 0 otherwise.

When a new component on V' has been detected, we must perform the necessary bookkeeping to update the data structures. When $\ell \in [0, k]$ (the *lower range*), we simply update the marks or labels of vertices in V' to record the newly detected component. For the *upper range*, we first mark all vertices in V' . Then, for each previous component V_i , all of whose vertices have been marked, delete V_i from the list of components and update the matrix.

We are now ready to state and prove that component pebble games correctly solve most of the fundamental problems presented in the Introduction: *Decision*, *Spanning*, *Extraction*, *Optimization* and *Components*.

Theorem 16 *The graphs recognized by component pebble games are the same as graphs recognized by basic pebble games, and components are correctly com-*

puted.

Proof The component pebble game differs from the basic pebble game by maintaining components and rejecting edges precisely when both endpoints lie in a component. Thus, by Corollary 6, the component pebble game accepts an edge if and only if it is independent.

To show that the component pebble games correctly maintain components, observe that Algorithm 3 detects a maximal connected subgraph (with respect to the vertices) with no outgoing edges in which exactly ℓ pebbles are present. By Lemma 10, Invariant 2, this subgraph must be a block. When $\ell > 0$, by Theorem 5(3b), components are connected; thus, Algorithm 3 detects a component. When $\ell = 0$, there may be at most one component by Theorem 5(3a) since the union of two blocks is a block, and the algorithm computes it.

The visited vertices in Algorithm 4 are those that can reach a pebble (in the original orientation) on a vertex other than u and v , and thus form the complement of the unique component containing e . ■

The time complexity on the algorithm is now $O(n^2)$ as dependent edges are rejected in constant time. Component detection and resulting updating of the data structures can be accomplished in linear time. More specific, implementation-related details on how to actually achieve this for the *upper range* are given in (17) using a similar data structure called *union pair-find*; while *union pair-find* maintains edge sets, the pebble game algorithm does not need to do it explicitly, and thus the associated implementation details for edge sets can be ignored.

Space complexity is linear for the *lower range* and $O(n^2)$ in the *upper range*, due to its additional matrix. An alternative solution to *union pair-find* presented in (17) uses only $O(n)$ space, though it requires the edges to be considered in a specific order and does not solve the Optimization problem.

6 Applications

Henneberg Sequences. Originating in (13) (see also (27)), these are inductive constructions for Laman graphs and other classes of rigid structures. We extend the concept to tight graphs: at the base case, start with a small tight graph; each inductive step would create a tight graph with an additional vertex by specifying b edges for removal before adding the new vertex of degree $k + b$. In addition, b can be chosen to be small $b \in [0, k]$.

We remind the reader of the **matroidal conditions** on tight graphs: (1)

$\ell \in [0, k]$ and $n \geq 1$, or (2) $\ell \in (k, \frac{3}{2}k)$ and $n \geq 2$, or (3) $\ell \in [\frac{3}{2}k, 2k)$ and $n \geq \lceil \frac{\ell}{2k-\ell} \rceil$. We refer to the smallest values of n as the **base-case conditions**; when n is strictly larger, we call them the **non-triviality conditions**. The following lemma is the key to proving the existence of a *Henneberg reduction*: given a tight graph, remove vertices one at a time until a base case is reached. This leads to an quadratic algorithm for computing the entire sequence.

Lemma 17 *Let v be a vertex of degree $k+b > k$ in a tight graph G . Then, after the removal of any edge $e = uv$, there exists a new edge whose insertion results in a tight graph. If $\ell \in [0, \frac{3}{2}k)$, this edge can be found among the neighbors of v ; otherwise, it is found in a larger set containing the neighbors of v whose size satisfies the **base-case conditions**.*

Proof Consider the sparse graph after the removal of e ; it is broken into components, free edges and free vertices. Let V' be the neighbors of v (but not v itself). If $\ell \in [\frac{3}{2}, 2k)$, add enough vertices to V' to satisfy the **base-case conditions**.

We claim that the vertex set V' cannot form a block; in fact, it cannot span more than $k|V'| - \ell - b$ edges. Indeed, suppose, for a contradiction, that V' spanned more than $k|V'| - \ell - b$ edges. Since the degree of v is $k + b$, the size of the induced set of edges in G on $V' \cup \{v\}$ is more than $k|V'| - \ell - b + k + b = (k|V'| - \ell) + k = k(|V'| + 1) - \ell = k|V' \cup \{v\}| - \ell$. This contradicts the sparsity of G .

Since V' does not form a block and its number of vertices satisfies the **base-case conditions**, it is not saturated with edges. Therefore, because of the matroidal property of base extension, there exists an edge not already spanned by V' , which can be added to restore tightness. ■

It is a simple exercise to show the existence of a vertex with bounded degree in $[k, 2k]$: indeed, the average degree in a sparse graph is at most $2k$, and each vertex v has degree at least k (or else sparsity would be violated on $V \setminus \{v\}$). We can then apply Lemma 17 $O(k)$ times repeatedly to compute a single Henneberg reduction step. The Henneberg sequence is obtained by iterating the Henneberg reduction step until we reach a base case.

This leads directly to an $O(n^2)$ algorithm for solving the Henneberg reduction problem by using the pebble game. Figure 10 describes one step of the algorithm. Each edge removal puts back one pebble and searches in a constant-size vertex subset for at most $O(k^2)$ possibilities of edge-insertion, taking a total of $O(n)$ time in the necessary pebble searches.

Circuits and redundancy. A graph G is said to be (k, ℓ) -*redundant* if it is spanning and the removal of *any* edge produces a graph which is still spanning. A *circuit* is a special type of redundant graph, where the removal of any edge

Algorithm 5 Henneberg reduction step.

Input: The directed graph produced by the pebble game, played on a tight graph G satisfying the **non-triviality conditions**.

Output: A Henneberg reduction step for G .

Method: Find a vertex of degree $k + b$, with $b \in [k, 2k)$. If $\ell = 0$, b may also be k . Compute the neighbor set V_v of v . If $\ell \in [\frac{3}{2}k, 2k)$, let V' be any set of size $\lceil \frac{\ell}{2k-\ell} \rceil$ that includes V_v , else $V' = V_v$.

Repeat b times:

Use the pebble game to find an edge with endpoints in V' which is not already spanned by V' and not in a component. Insert it.

Fig. 10. Henneberg reduction step algorithm.

produces a tight graph.

We can detect a circuit associated with a dependent edge $e = uv$ with respect to D during the pebble game by collecting ℓ pebbles on u and v and computing $Reach(u, v)$, which is done in linear time; the edges in D spanned by $Reach(u, v)$ along with e comprise the circuit.

To decide redundancy of the input graph G , simply detect circuits during the game and mark all the edges in circuits, as they are computed; if all edges are marked at the end of the game, the graph is redundant. If the graph is not redundant, unmarked edges are *bridges*; after their removal, the vertex sets of the sparsity components in the resulting graph correspond to the vertex sets of *redundant components*: induced subgraphs that are redundant. These algorithms run in $O(mn)$ time.

References

- [1] S. Bereg. Certifying and constructing minimally rigid graphs in the plane. In *Proc. 21th Annu. ACM Sympos. Comput. Geom.*, pp. 73–80, 2005.
- [2] A. Berg and T. Jordán. Algorithms for graph rigidity and scene analysis. In *Proc. 11th ESA*, LNCS vol. 2832, pp. 78–89. Springer, 2003.
- [3] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to Algorithms*. McGraw-Hill, 1990.
- [4] J. Edmonds. Edge-disjoint branchings. In B. Rustin, editor, *Combinatorial Algorithms*, pp. 91–96. Academic Press, New York, 1973.
- [5] Z. Fekete and L. Szegő. A note on $[k, l]$ -sparse graphs. TR 2005-05, Egerváry Research Group, Eötvös University, Budapest, Hungary, 2005.
- [6] A. Frank and L. Szegő. Constructive characterizations for packing and covering with trees. *Discrete Applied Mathematics*, 131(2):347–371, 2003.
- [7] H. Gabow and H. Westermann. Forests, frames, and games: algorithms for matroid sums and applications. In *Proc. STOC'88*, pp. 407–421, 1988.

- [8] H. N. Gabow. A matroid approach to finding edge connectivity and packing arborescences. In *Proc. STOC'91*, pp. 112–122, 1991.
- [9] R. Haas. Characterizations of arboricity of graphs. *Ars Combinatorica*, 63:129–137, 2002.
- [10] B. Hendrickson. *The molecule problem: determining conformation from pairwise distances*. PhD thesis, Cornell University, 1991.
- [11] B. Hendrickson. Conditions for unique graph realizations. *SIAM Journal of Computing*, 21(1):65–84, 1992.
- [12] B. Hendrickson. The molecule problem: Exploiting structure in global optimization. *SIAM Journal on Optimization*, 5(4):835–857, 1995.
- [13] E. L. Henneberg. *Die graphische Statik der starren Systeme*. Johnson Reprint 1968. Leipzig, 1911.
- [14] D. J. Jacobs and B. Hendrickson. An algorithm for two-dimensional rigidity percolation: the pebble game. *J. Comp. Phys.*, 137:346–365, 1997.
- [15] D. J. Jacobs, A. Rader, M. Thorpe, and L. A. Kuhn. Protein flexibility predictions using graph theory. *Proteins 44*, pp. 150 – 165, 2001.
- [16] G. Laman. On graphs and rigidity of plane skeletal structures. *Journal of Engineering Mathematics*, 4:331–340, 1970.
- [17] A. Lee, I. Streinu, and L. Theran. Finding and maintaining rigid components. In *Proc. CCCG'05*, Windsor, Canada, 2005.
- [18] M. Loréa. On matroidal families. *Discrete Math.*, 28:103–106, 1979.
- [19] L. Lovasz and Y. Yemini. On generic rigidity in the plane. *SIAM J. Algebraic and Discrete Methods*, 3(1):91–98, 1982.
- [20] C. S. A. Nash-Williams. Edge-disjoint spanning trees of finite graphs. *Journal London Mathematical Society*, 36:445–450, 1961.
- [21] J. G. Oxley. *Matroid Theory*. Oxford University Press, 1992.
- [22] A. Recski. A network theory approach to the rigidity of skeletal structures I. Modelling and interconnection. *Discr. Appl. Math.*, 7:313–324, 1984.
- [23] J. Roskind and R. E. Tarjan. A note on finding minimum-cost edge-disjoint spanning trees. *Math. Operat. Res.*, 10(4):701–708, 1985.
- [24] L. Szegő. On constructive characterizations of (k, l) -sparse graphs. TR 2003-10, Eötvös University, Budapest, Hungary, 2003.
- [25] R. E. Tarjan. Edge-disjoint spanning trees and depth-first search. *Acta Informatica*, 6:171–185, 1976.
- [26] T.-S. Tay. Rigidity of multi-graphs I. linking rigid bodies in n -space. *J. Comb. Theory Ser. B*, 36:95–112, 1984.
- [27] T.-S. Tay and W. Whiteley. Generating isostatic graphs. *Structural Topology*, 11:21–68, 1985.
- [28] W. T. Tutte. On the problem of decomposing a graph into n connected factors. *Journal London Math. Soc.*, 142:221–230, 1961.
- [29] W. Whiteley. The union of matroids and the rigidity of frameworks. *SIAM Journal Discrete Mathematics*, 1(2):237–255, May 1988.
- [30] W. Whiteley. Some matroids from discrete applied geometry. In J. O. J. Bonin and B. Servatius, eds., *Matroid Theory*, pp. 171–311. AMS 1996.