Electronic Colloquium on Computational Complexity, Report No. 66 (2005)

Narrow Proofs May Be Spacious: Separating Space and Width in Resolution

Jakob Nordström

Royal Institute of Technology (KTH)
SE-100 44 Stockholm, Sweden
jakobn@kth.se

June 4, 2005

Abstract

The width of a resolution proof is the maximal number of literals in any clause of the proof. The space of a proof is the maximal number of memory cells used if the proof is only allowed to resolve on clauses kept in memory. Both of these measures have previously been studied and related to the refutation size of unsatisfiable CNF formulas. Also, the resolution refutation space of a formula has been proven to be at least as large as the refutation width, but it has remained unknown whether space can be separated from width or the two measures coincide asymptotically. We prove that there is a family of k-CNF formulas for which the refutation width in resolution is constant but the refutation space is non-constant, thus solving an open problem mentioned in several previous papers.

1 Introduction

A proof system for a language L is a polynomial-time algorithm V such that for all $x \in \mathrm{L}$ there is a string π (a *proof*) for which $V(x, \pi) = 1$. For $x \notin \mathrm{L}$, it should hold for all strings π that $V(x, \pi) = 0$. The complexity of a proof system V is the smallest bounding function $g : \mathbb{N} \mapsto \mathbb{N}$ such that $x \in \mathrm{L}$ if and only if there is a proof π of size $|\pi| \leq g\big(|x|\big)$ for which $V\big(x, \pi\big) = 1$. If a proof system is of polynomial complexity, it is said to be polynomially bounded. A *propositional proof system* is a proof system for tautologies in propositional logic.

The central task of proof complexity is to construct and investigate the power of different propositional proof systems. This is done for at least two reasons.

The first reason is the connection to the question of P versus NP, which is recognized as a major open problem in computational complexity theory and mathematics. Since NP is exactly the set of languages with polynomially bounded proof systems, and since TAUTOLOGY can be seen to be the dual problem of SATISFIABILITY, we have the famous theorem of Cook and Reckhow [17] that $\mathsf{NP} = \mathsf{co\text{-}NP}$ if and only if there exists a polynomially bounded propositional proof system. Thus, if it could be shown that there are no polynomially bounded proof systems for propositional tautologies, $\mathsf{P} \neq \mathsf{NP}$ would follow as

ISSN 1433-8092

a corollary since P is closed under complement. One way of approaching this distant goal is to study stronger and stronger proof systems and try to prove superpolynomial lower bounds on proof size. However, despite the fact that the last decade has seen some impressive successes for a variety of propositional proof systems, it seems that we are still very far from fully understanding the reasoning power of even quite simple ones.

The second reason is that designing efficient algorithms for proving tautologies is a very important problem not only in theoretical computer science but also in applied research and in industry, for instance in the context of formal methods. All automated theorem provers, regardless of whether they actually produce a written proof or not, explicitly or implicitly define a system in which proofs are searched for and rules which determine what proofs in this system look like. Lower bounds on proofs in such proof systems give lower bounds on the running time of corresponding automated theorem provers. In the other direction, theoretical upper bounds on proof size in a system can give upper bounds on the running time of a proof search algorithm, provided that the algorithm can be shown to search for proofs in the system in an efficient manner.

Also, the field of proof complexity has rich connections to cryptography, artificial intelligence and mathematical logic. Some good surveys of proof complexity are [4, 6, 15, 34].

Any propositional logic formula can be converted to a formula in conjunctive normal form that is only linearly larger and is unsatisfiable if and only if the original formula is a tautology. Therefore, any sound and complete system which produces refutations of unsatisfiable formulas in conjunctive normal form can be considered as a general propositional proof system.

One such proof system, which is the focus of this paper, is *resolution*. The resolution proof system appeared in [11], and began to be studied in connection with automated theorem proving in [19, 20, 30]. Because of the simplicity of resolution—there is only one derivation rule—and because all lines in a proof are clauses, this system is well adapted to proof search algorithms. Many real-world automated theorem provers are based on resolution.

Being so simple and fundamental, resolution was a natural target to attack when trying to prove lower bounds in proof complexity. In this context, it is most straightforward to prove bounds on the *length* of proofs, i.e., the number of clauses, which is easily seen to be polynomially related to the proof size. In 1968, Tseitin [32] presented a superpolynomial lower bound on refutation length in resolution, but it was not until almost 20 years later that Haken [24] proved the first exponential lower bound, which has later been followed by many similar results, for instance in [5, 10, 16, 29, 33].

A second complexity measure for resolution refutations other than length is the minimal *width*, measured as the maximal size of a clause in the refutation. This measure was introduced by Ben-Sasson and Wigderson in [10], and was shown to be strongly correlated to proof length. Ben-Sasson and Wigderson proved that the width $W\bigl(F \vdash 0\bigr)$ of refuting a k-CNF formula F over n variables is bounded by the refutation length $L\bigl(F \vdash 0\bigr)$ by $W\bigl(F \vdash 0\bigr) = \mathrm{O}\bigl(\sqrt{n \log L(F \vdash 0)}\bigr)$, and used this to rederive and simplify many lower bounds on length by proving bounds on width.

The results on width lead to the question of whether other complexity measures could yield interesting insights as well. In [22, 31], Esteban and Torán initiated the investigation of proof *space* in resolution. Intuitively, the space of a resolution proof is the maximal number of clauses one needs to keep in memory while verifying the proof. A number of upper and lower bounds for proof space in resolution and other proof systems were subsequently presented in for instance [1, 8, 21, 23]. In several of these papers it was noted that the lower bounds on resolution refutation space for different formula families matched known lower bounds on refutation width. Atserias and Dalmau [3] showed that this was not a coincidence, but that the minimal refutation space $Sp\big(F \vdash 0\big)$ of any unsatisfiable k-CNF formula F is at least as large as the minimal refutation width $W\big(F \vdash 0\big)$ minus a constant term.

An immediate follow-up question to this is whether the lower bound on space in terms of width is asymptotically strict. That is, does there exist a family $\{F_n\}$ of k-CNF formulas such that $Sp\big(F_n \vdash 0\big) = \omega\big(W\big(F_n \vdash 0\big)\big)$ or does it always hold that $Sp\big(F_n \vdash 0\big) = \mathrm{O}\big(W\big(F_n \vdash 0\big)\big)$?

Another natural question concerns the relation between space and length. It is not too hard to see that upper bounds on width imply upper bounds on length, and as a consequence of the result in [3] this must be true for space with respect to length as well. In the other direction, we have the result from [10] that upper bounds on length imply upper bounds on width. Does a similiar bound hold for refutation space, or is there a family of k-CNF formulas $\{F_n\}$ that separates space from length in the sense that $Sp\big(F_n \vdash 0\big) = \omega\big(\sqrt{n \log L(F_n \vdash 0)}\big)$, for n the number of variables in F_n?

A plausible candidate for answering these two questions is the family of *pebbling contradictions* defined in terms of pebble games on directed acyclic graphs. Determining the space complexity of refuting pebbling contradictions in resolution has been mentioned as an interesting open problem in [7, 21, 23].

In this paper, we answer the first question above by separating space from width. This is done by proving a non-constant lower bound on space for pebbling contradictions on binary trees, thus at least partially solving the open problem about the space complexity of pebbling contradictions as well. More precisely, our results are as follows (formal definitions are given in Sections 2 and 4).

Theorem 1.1. *Let T_h denote the complete binary tree of height h and $Peb^d_{T_h}$ the pebbling contradiction of degree $d \geq 2$ defined on T_h. Then the space of refuting $Peb^d_{T_h}$ by resolution is bounded by $Sp\big(Peb^d_{T_h} \vdash 0\big) = \Omega\big(\sqrt{h}\big)$.*

Corollary 1.2. *There is a family $\{F_n\}_{n=1}^{\infty}$ of k-CNF formulas of size $\mathrm{O}(n)$ such that $W\big(F_n \vdash 0\big) = \mathrm{O}(1)$ but $Sp\big(F_n \vdash 0\big) = \Omega\big(\sqrt{\log n}\big)$.*

The organization of this paper is as follows. We start by presenting the resolution proof system in Section 2. Section 3 gives a short introduction to pebble games, and in Section 4 we review some previous results connecting resolution and pebbling. The lower bound on refutation space which separates space from width is then proven in three steps.

- First, we define a modified pebble game and establish a lower bound for this game in terms of the standard black-white pebble game (Sections 5 and 6).
- Next, we show that a resolution refutation of a pebbling contradiction induces a pebbling of the underlying graph in our modified pebble game (Section 7).

- Finally, we prove that if a set of clauses induces many pebbles, the set must be large. Since a resolution refutation induces a pebbling, and such a pebbling must contain many pebbles at some point, we deduce that the clause space of the resolution derivation must be large (Section 8).

We conclude in Section 9 by giving suggestions for further research.

2 The Resolution Proof System

A *literal* over a propositional logic variable x is either x itself or its negation $\overline{x}$. We define $\overline{\overline{x}} = x$. Two literals a and b are *strictly distinct* if $a \neq b$ and $a \neq \overline{b}$.

A *clause* $C = a_1 \vee \ldots \vee a_k$ is a set of literals. We say that C is *ordinary* if all literals are strictly distinct. All clauses are assumed ordinary unless otherwise stated. We say that C is a *subclause* of D if $C \subseteq D$. An ordinary clause containing at most k literals is called a *k-clause*.

A *CNF formula* $F = C_1 \wedge \ldots \wedge C_m$ is a set of clauses. A *k-CNF formula* is a CNF formula consisting of k-clauses.

In the following, we let A, B, C, D denote clauses, $\mathbb{C}, \mathbb{D}$ sets of clauses, x, y propositional variables, a, b, c literals, α, β truth value assignments and ν a truth value 0 or 1. We define

$$\alpha^{x=\nu}(y) = \begin{cases} \alpha(y) & \text{if } y \neq x, \\ \nu & \text{if } y = x. \end{cases}$$

We let $Vars(C)$ denote the set of variables and $Lit(C)$ the set of literals in a clause C. (Although the notation $Lit(C)$ is slightly redundant given the definition of a clause as a set of literals, we include it for clarity.) This notation is extended to sets of clauses by taking unions.

A *resolution derivation* $\pi : F \to A$ of a clause A from a CNF formula F is a sequence of clauses $\pi = \{D_1, \ldots, D_\tau\}$ such that $D_\tau = A$ and each line D_i, $1 \leq i \leq \tau$, is either one of the clauses in F (*axioms*) or is derived from clauses D_j, D_k in π with $j, k < i$ by the *resolution rule*

$$\frac{B \vee x \quad C \vee \overline{x}}{B \vee C}. \tag{1}$$

We refer to (1) as *resolution on the variable x* and $B \vee C$ as the *resolvent* of $B \vee x$ and $C \vee \overline{x}$ on x. A *resolution refutation* of a CNF formula F is a resolution derivation of the empty clause 0 (the clause with no literals) from F.

For F a formula and $\mathcal{G} = \{G_1, \ldots, G_n\}$ a set of formulas, we say that $\mathcal{G}$ *implies* F, denoted $\mathcal{G} \models F$, if every truth value assignment satisfying all formulas $G \in \mathcal{G}$ satisfies F as well.

Resolution is sound and implicationally complete. That is, if there is a resolution derivation $\pi : F \to A$ then $F \models A$, and if $F \models A$ then there is a resolution derivation $\pi : F \to A'$ for some $A' \subseteq A$. In particular, F is unsatisfiable if and only if there is a resolution refutation of F.

The graph G_π of a resolution derivation π is a directed acyclic graph (DAG) with the clauses of the derivation labelling the vertices and edges added from the assumption clauses to the resolvent for each application of the resolution rule. A resolution derivation π is *tree-like* if any clause in the derivation is used at most once as a premise in an application of the resolution rule, i.e., if G_π is a tree (we may make copies of the axioms in order to make G_π into a tree).

The *length* $L(F)$ of a CNF formula F is the number of clauses in it, and for π a resolution derivation $L(\pi)$ is the number of clauses in π. The length of refuting F by resolution is $L(F \vdash 0) = \min_{\pi: F \to 0} \{L(\pi)\}$. The length of refuting F by tree-like resolution $L_{\mathfrak{T}}(F \vdash 0)$ is defined by taking the minimum over all tree-like resolution refutations π_T of F.

The *width* $W(C)$ of a clause C is $|C|$. The width of a set of clauses $\mathbb{C}$ is $W(\mathbb{C}) = \max_{C \in \mathbb{C}}\{ W(C)\}$. The width of deriving a clause A from the formula F by resolution is $W(F \vdash A) = \min_{\pi: F \to A}\{ W(\pi)\}$.

If a resolution refutation has constant width, it must be of size polynomial in the number of variables. Conversely, if all refutations of a formula are very wide, it seems reasonable that any refutation of this formula must be very long as well. This intuition is made precise in the following theorem.

Theorem 2.1 ([10]). *The width of refuting a CNF formula F is bounded from above by*

$$W(F \vdash 0) \leq W(F) + \mathrm{O}\left(\sqrt{n \log L(F \vdash 0)}\right),$$

where n is the number of variables in F.

In [13], it was shown that this bound on width in terms of length is essentially optimal.

We next define the measure of *space*. Following the exposition in [22], a proof can be seen as a Turing machine computation, with a special read-only input tape from which the axioms can be downloaded and a working memory where all derivation steps are made. The clause space of a resolution proof is the maximal number of clauses that need to be kept in memory simultaneously during a verification of the proof. The variable space is the maximal "total" space needed, where also the width of the clauses is taken into account.

For the formal definition, it is convenient to use the following alternative definition of resolution introduced by [1].

Definition 2.2 (Resolution). A *clause configuration* $\mathbb{C}$ is a set of clauses. A sequence of clause configurations $\mathbb{C}_0, \ldots, \mathbb{C}_\tau$ is a *resolution derivation* from F if $\mathbb{C}_0 = \emptyset$ and for all $t \in [\tau]$, $\mathbb{C}_t$ is obtained from $\mathbb{C}_{t-1}$ by one of the following rules:

Axiom Download $\mathbb{C}_t = \mathbb{C}_{t-1} \cup \{C\}$ for some $C \in F$.

Erasure $\mathbb{C}_t = \mathbb{C}_{t-1} \setminus \{C\}$ for some $C \in \mathbb{C}_{t-1}$.

Inference $\mathbb{C}_t = \mathbb{C}_{t-1} \cup \{D\}$ for some $D \notin \mathbb{C}_{t-1}$ inferred by resolution from $C_1, C_2 \in \mathbb{C}_{t-1}$.

A derivation $\pi : F \to A$ of a clause A from F is a resolution derivation $\mathbb{C}_0, \ldots, \mathbb{C}_\tau$ such that $\mathbb{C}_\tau = \{A\}$. A *resolution refutation* of F is a derivation of 0 from F.

Definition 2.3 (Clause space [1, 7]). The *clause space* of a resolution derivation $\mathbb{C}_0, \ldots, \mathbb{C}_\tau$ is $\max_{0 \leq i \leq \tau}\{|\mathbb{C}_i|\}$. The clause space of deriving A from F is $Sp(F \vdash A) = \min_{\pi: F \to A}\{Sp(\pi)\}$. $Sp(F \vdash 0)$ is the minimal clause space of any resolution refutation of F.

Definition 2.4 (Variable space [1]). The *variable space* of a configuration $\mathbb{C}$ is $VarSp(\mathbb{C}) = \sum_{C \in \mathbb{C}} W(C)$. The variable space of a resolution derivation $\mathbb{C}_0, \ldots, \mathbb{C}_\tau$ is $\max_{0 \le i \le s}\{ VarSp(\mathbb{C}_i)\}$, and $VarSp(F \vdash 0)$ is the minimal variable space of any resolution refutation of F.

Restricting the resolution derivations to tree-like resolution, we get the measures $Sp_{\mathfrak{T}}(F \vdash 0)$ and $VarSp_{\mathfrak{T}}(F \vdash 0)$ in analogy with $L_{\mathfrak{T}}(F \vdash 0)$ defined above.

All contradictory CNF formulas can be refuted in clause space linear in the formula size. More precisely:

Theorem 2.5 ([22]). *Any unsatisfiable CNF formula F on n variables can be refuted in clause space $n+2$.*

Theorem 2.6 ([22]). *Any unsatisfiable CNF formula F with m clauses can be refuted in clause space $m+1$, i.e., $Sp(F \vdash 0) \le L(F) + 1$.*

Thus the interesting question is which formulas demand this much space, and which formulas can be refuted in for instance logarithmic or even constant space. It has been shown that there are polynomial-size formulas that meet the upper bounds of Theorems 2.5 and 2.6 up to a multiplicative constant.

Theorem 2.7 ([1, 31]). *There is a polynomial-size family $\{F_n\}_{n=1}^{\infty}$ of unsatisfiable 3-CNF formulas such that $Sp(F \vdash 0) = \Omega(L(F)) = \Omega(|Vars(F)|)$.*

Lower bounds on clause space have been presented for a number of different CNF formula families [1, 8, 31]. As was mentioned above, in these papers it was observed that the lower bounds on refutation space coincided with the lower bounds on refutation width. This lead to the conjecture that the width measure is a lower bound for the clause space measure, a conjecture that was proven true in [3].

Theorem 2.8 ([3]). *Let F be an arbitrary unsatisfiable CNF formula. Then it holds that $Sp(F \vdash 0) \ge W(F \vdash 0) - W(F)$.*

In fact, if one does the calculations in the proof of Theorem 2.8 carefully, one can sharpen the inequality to $Sp(F \vdash 0) - 3 \ge W(F \vdash 0) - W(F)$. In other words, the extra clause space exceeding the minimum 3 needed for any resolution derivation is bounded from below by the extra width exceeding the minimum $W(F)$. An immediate corollary of this theorem is that for polynomial-size k-CNF formulas, constant clause space implies polynomial proof length.

A very natural question, which has remained open, is what holds in the other direction. Do the space and width measures coincide asymptotically or is there a formula family separating space from width? We remark that in order for this question to be interesting, we should restrict our attention to families of k-CNF formulas. Any resolution refutation of an unsatisfiable CNF formula F with *minimum* clause width k can be shown to require clause space at least $k+2$ [22], so it is easy to find CNF formulas $\{F_n\}$ of growing width such that $W(F_n \vdash 0) - W(F_n) = \mathrm{O}(1)$ but $Sp(F_n \vdash 0) = \Omega(n)$.

In this paper, we settle the open question of the relationship between space and width by proving that there is a family of k-CNF formulas $\{F_n\}$ such that $W(F_n \vdash 0) = \mathrm{O}(1)$ but $Sp(F_n \vdash 0) = \omega(1)$.

3 Pebble Games

Pebble games were devised for studying programming languages and compiler construction, but have found a variety of applications in computational complexity theory. In connection with resolution, pebble games have been employed both to analyze resolution derivations with respect to how much memory they consume (using the original definition of space in [22]) and to construct CNF formulas which are hard for different variants of resolution in various respects (see for example [2, 9, 12, 14]). An excellent survey of pebbling up to 1980 is [28].

The black pebbling price of a DAG G captures the memory space, i.e., the number of registers, required to perform the deterministic computation described by G. The space of a non-deterministic computation is measured by the black-white pebbling price of G. We say that vertices of G with indegree 0 are *sources* and vertices with outdegree 0 *targets*.

Definition 3.1 (Pebble game). Suppose that G is a DAG with sources S and a unique target z. The *black-white pebble game* on G is the following 1-player game. At any point in the game, there are black and white pebbles placed on some vertices of G, at most one pebble per vertex. A *pebble configuration* is a pair of subsets $\mathbb{P} = (B, W)$ of $V(G)$, comprising the black- and white-pebbled vertices. The rules of the game are as follows:

1. If all immediate predecessors of an empty vertex v have pebbles on them, a black pebble may be placed on v.
2. A black pebble may be removed from any vertex at any time.
3. A white pebble may be placed on any empty vertex at any time.
4. If all immediate predecessors of a white-pebbled vertex v have pebbles on them, the white pebble on v may be removed.

A *legal black-white pebbling reaching* (B, W) in G is a sequence of configurations $\mathcal{P} = \{\mathbb{P}_0, \ldots, \mathbb{P}_\tau\}$ such that $\mathbb{P}_0 = (\emptyset, \emptyset)$, $\mathbb{P}_\tau = (B, W)$, and for all $t \in [\tau]$, $\mathbb{P}_t$ follows from $\mathbb{P}_{t-1}$ by one of the rules above.

The *cost* of a pebbling configuration $\mathbb{P} = (B, W)$ is $\mathrm{cost}(\mathbb{P}) = |B \cup W|$ and the cost of a of a legal pebbling $\mathcal{P} = \{\mathbb{P}_0, \ldots, \mathbb{P}_\tau\}$ is $\max_{t \in [\tau]}\{\mathrm{cost}(\mathbb{P}_t)\}$. The *black-white pebbling price* of (B, W), denoted $\mathit{BW\text{-}Peb}(B, W)$, is the minimal cost of any legal pebbling reaching (B, W).

A *legal black-white pebbling of* G is a pebbling reaching $(\{z\}, \emptyset)$, and the *black-white pebbling price* of G, denoted $\mathit{BW\text{-}Peb}(G)$, is the minimal cost of any legal pebbling of G.

A *legal black pebbling of* G is a pebbling reaching $(\{z\}, \emptyset)$ using only black pebbles, i.e., $W_t = \emptyset$ for all t, and the *(black) pebbling price* of G, denoted $\mathit{Peb}(G)$, is the minimal cost of any legal black pebbling of G.

A black-white pebbling *visiting* z is a pebbling such that $\mathbb{P}_0 = \mathbb{P}_\tau = (\emptyset, \emptyset)$ and there exists a $t \in [\tau]$ such that $z \in B_t \cup W_t$. The minimum cost of such a pebbling is denoted $\mathit{BW\text{-}Peb}^{\emptyset}(G)$.

It is easy to see that $\mathit{BW\text{-}Peb}^{\emptyset}(G) \leq \mathit{BW\text{-}Peb}(G) \leq \mathit{BW\text{-}Peb}^{\emptyset}(G) + 1$.

We think of the moves in a pebbling as occurring at integral time intervals $t = 1, 2, \ldots$ and talk about the pebbling move "at time t" (which is the move resulting in configuration $\mathbb{P}_t$) or the moves "during the time interval $[t_1, t_2]$".

In this paper we will consider pebblings of complete binary trees. We let T denote a complete binary tree considered as a DAG with edges directed towards the root. We write T_h when we want to specify that the height of the tree is h. We use z to denote the unique target vertex of T, i.e., the root.

The black pebbling price of T_h can be established by induction over the tree height. We omit the easy proof.

Theorem 3.2. *$Peb(T_h) = h + 2$.*

General bounds for the black-white pebbling price of trees of any arity was presented in [27]. We give a simplified proof with tight bounds for the case of complete binary trees.

Theorem 3.3. *$BW\text{-}Peb(T_h) = \lfloor \frac{h}{2} \rfloor + 3$ and $BW\text{-}Peb^{\emptyset}(T_h) = \lfloor \frac{h-1}{2} \rfloor + 3$.*

The proof is facilitated the following proposition, which is an immediate consequence of Definition 3.1.

Proposition 3.4 ([18]). *If $\mathcal{P}$ is a black-white pebbling of a DAG G visiting the target, then one can get a dual pebbling $\overline{\mathcal{P}}$ of G by reversing the sequence of moves and switching the colours of the pebbles.*

Proof of Theorem 3.3. In all of this proof, we let z_1, z_2 denote the immediate predecessors of the root z of the tree.

We first show that $BW\text{-}Peb^{\emptyset}(T_{h+2}) \geq BW\text{-}Peb^{\emptyset}(T_h) + 1$. Suppose not, and let $\mathcal{P}$ be a pebbling in cost $K = BW\text{-}Peb^{\emptyset}(T_h)$ for T_{h+2} making the minimum number of pebbling moves. Let T_h^i, $i \in [4]$, be the four disjoint subtrees of height h in T_{h+2}. It is easy to see that $\mathcal{P}$ restricted to $V(T_h^i)$ yields a legal pebbling of T_h^i visiting its root. It follows that there must exist distinct times t_i, $i \in [4]$, when T_h^i contains K pebbles and the rest of T_{h+2} is empty. Number the subtrees so that $t_1 < t_2 < t_3 < t_4$.

Suppose that the root z of T_{h+2} has been pebbled before time t_3. Then we can get a shorter pebbling of T_{h+2} by completing the subpebbling of T_h^3 but ignoring pebbles moves outside T_h^3 after time t_3.

Consequently, z must be pebbled for the first time after t_3. But at time t_3 the rest of the tree is empty, so in that case we can get a shorter legal pebbling by ignoring all moves outside T_h^3 before time t_3 and performing all moves in $\mathcal{P}$ after time t_3. Contradiction. Thus $BW\text{-}Peb^{\emptyset}(T_{h+2}) \geq BW\text{-}Peb^{\emptyset}(T_h) + 1$.

Next, it is easy to see that $BW\text{-}Peb^{\emptyset}(T_{h+1}) \leq BW\text{-}Peb(T_h)$. First black-pebble z_1 with using a pebbling $\mathcal{P}$ in cost $BW\text{-}Peb(T_h)$. Place white pebbles on z and z_2, and then remove the pebbles from z_1 and z. Finally, use the dual pebbling $\overline{\mathcal{P}}$ to get the white pebble off z_2 in the same cost $BW\text{-}Peb(T_h)$.

Since $BW\text{-}Peb(T_1) = BW\text{-}Peb^{\emptyset}(T_1) = 3$, it follows that $BW\text{-}Peb^{\emptyset}(T_h) \geq \lfloor \frac{h-1}{2} \rfloor + 3$ and $BW\text{-}Peb(T_h) \geq \lfloor \frac{(h+1)-1}{2} \rfloor + 3 = \lfloor \frac{h}{2} \rfloor + 3$. It remains to demonstrate that there are pebblings meeting these lower bounds. We construct such pebblings inductively.

Suppose for h odd that $BW\text{-}Peb(T_h) = BW\text{-}Peb^{\emptyset}(T_h) = \lfloor \frac{h-1}{2} \rfloor + 3 = \lfloor \frac{h}{2} \rfloor + 3$. Using the same pebbling as above for T_{h+1}, it is easy to see that $BW\text{-}Peb^{\emptyset}(T_{h+1}) = \lfloor \frac{h}{2} \rfloor + 3$, and since the pebbling cost cannot increase by more than one when the height is increased by one we get $BW\text{-}Peb^{\emptyset}(T_{h+2}) = \lfloor \frac{h}{2} \rfloor + 4 = \lfloor \frac{h+1}{2} \rfloor + 3$. In the same way we get $BW\text{-}Peb(T_{h+1}) = \lfloor \frac{h+1}{2} \rfloor + 3$.

To pebble T_{h+2} in cost $\lfloor \frac{h+1}{2} \rfloor + 3$ leaving a pebble on z, first black-pebble the root z_1 of the subtree T^1_{h+1} in cost $\lfloor \frac{h+1}{2} \rfloor + 3$. Leaving the pebble on z_1, make a pebbling visiting the root z_2 of T^2_{h+1} in cost $\lfloor \frac{h}{2} \rfloor + 3 = \lfloor \frac{h+1}{2} \rfloor + 2$ using the pebbling for T^2_{h+1} constructed inductively. In this pebbling there is a time t when z_2 is pebbled and T^2_{h+1} contains at most $\lfloor \frac{h+1}{2} \rfloor + 1$ pebbles. At this time t, place a black pebble on z and remove the black pebble on z_1 without exceeding the total limit of $\lfloor \frac{h+1}{2} \rfloor + 3$ pebbles on T_{h+2}. Then finish the pebbling of T^2_{h+1}. The theorem follows. □

4 Resolution and Pebbling Contradictions

A *pebbling contradiction* defined on a DAG G encodes the pebble game on G by defining the sources to be true and the targets false, and specifying that truth propagates through the graph according to the pebbling rules.

Definition 4.1 (Pebbling contradiction [10]). Let G be a DAG with sources S and a unique target z and with all vertices of G having indegree 0 or 2, and let $d \in \mathbb{N}^+$. Associate d distinct variables $x(v)_1, \ldots, x(v)_d$ with every vertex $v \in V(G)$. The dth degree *pebbling contradiction* on G, denoted Peb^d_G, is the conjunction of the following clauses:

- $\bigvee_{i=1}^{d} x(s)_i$ for all $s \in S$ (*source axioms*),
- $\overline{x(z)}_i$ for all $i \in [d]$ (*target axioms*),
- $\overline{x(u_1)}_i \vee \overline{x(u_2)}_j \vee x(v)_1 \vee \ldots \vee x(v)_d$ for all $v \in V(G) \setminus S$, where u_1, u_2 are the two predecessors of v, and all $i, j \in [d]$ (*pebbling axioms*).

The formula Peb^d_G is a $(2{+}d)$-CNF formula with $\mathrm{O}(d^2 \cdot |V(G)|)$ clauses over $d \cdot |V(G)|$ variables. See Figure 1 for an example pebbling contradiction.

It is easy to see that pebbling contradictions are unsatisfiable. Peb^d_G can be refuted in resolution by deriving $\bigvee_{i=1}^{d} x(v)_i$ for all $v \in V(G)$ inductively in topological order and then resolve with the target axioms $\overline{x(z)}_i$, $i \in [d]$. This proves the next theorem.

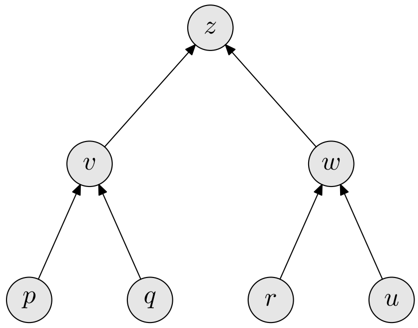
Theorem 4.2 ([9]). *For any DAG G with all vertices having indegree 0 or 2, there is a resolution refutation $\pi : Peb^d_G \to 0$ with $L(\pi) = \mathrm{O}(d^2 \cdot |V(G)|)$ and $W(\pi) = 2d$.*

Tree-like resolution is good at refuting pebbling contradictions Peb^1_G but is bad at refuting Peb^d_G for $d \geq 2$.

Theorem 4.3 ([7]). *For any DAG G with all vertices having indegree 0 or 2, there is a tree-like resolution refutation π of Peb^1_G such that $L(\pi) = \mathrm{O}(|V(G)|)$ and $Sp(\pi) = \mathrm{O}(1)$.*

Theorem 4.4 ([9]). *For any DAG G with all vertices having indegree 0 or 2, $L_{\mathfrak{T}}(Peb^2_G \vdash 0) = 2^{\Omega(Peb(G))}$.*

For refutation clause space, the upper bound $Sp(Peb^d_G \vdash 0) = Peb(G) + C$, where C is a constant independent of d, is fairly obvious: Just take an optimal



$$
\begin{aligned}
& \big(x(p)_1 \lor x(p)_2\big) \\
\land\ & \big(x(q)_1 \lor x(q)_2\big) \\
\land\ & \big(x(r)_1 \lor x(r)_2\big) \\
\land\ & \big(x(u)_1 \lor x(u)_2\big) \\
\land\ & \big(\overline{x(p)}_1 \lor \overline{x(q)}_1 \lor x(v)_1 \lor x(v)_2\big) \\
\land\ & \big(\overline{x(p)}_1 \lor \overline{x(q)}_2 \lor x(v)_1 \lor x(v)_2\big) \\
\land\ & \big(\overline{x(p)}_2 \lor \overline{x(q)}_1 \lor x(v)_1 \lor x(v)_2\big) \\
\land\ & \big(\overline{x(p)}_2 \lor \overline{x(q)}_2 \lor x(v)_1 \lor x(v)_2\big) \\
\land\ & \big(\overline{x(r)}_1 \lor \overline{x(u)}_1 \lor x(w)_1 \lor x(w)_2\big) \\
\land\ & \big(\overline{x(r)}_1 \lor \overline{x(u)}_2 \lor x(w)_1 \lor x(w)_2\big) \\
\land\ & \big(\overline{x(r)}_2 \lor \overline{x(u)}_1 \lor x(w)_1 \lor x(w)_2\big) \\
\land\ & \big(\overline{x(r)}_2 \lor \overline{x(u)}_2 \lor x(w)_1 \lor x(w)_2\big) \\
\land\ & \big(\overline{x(v)}_1 \lor \overline{x(w)}_1 \lor x(z)_1 \lor x(z)_2\big) \\
\land\ & \big(\overline{x(v)}_1 \lor \overline{x(w)}_2 \lor x(z)_1 \lor x(z)_2\big) \\
\land\ & \big(\overline{x(v)}_2 \lor \overline{x(w)}_1 \lor x(z)_1 \lor x(z)_2\big) \\
\land\ & \big(\overline{x(v)}_2 \lor \overline{x(w)}_2 \lor x(z)_1 \lor x(z)_2\big) \\
\land\ & \overline{x(z)}_1 \\
\land\ & \overline{x(z)}_2
\end{aligned}
$$

Figure 1: The pebbling contradiction $Peb^2_{T_2}$ for the complete binary tree of height 2.

black pebbling and derive $x(v)_1 \lor \ldots \lor x(v)_d$ when vertex v is pebbled. This is not quite an optimal strategy with respect to clause space, however. We can do at least a little bit better.

Theorem 4.5 ([23]). $Sp\big(Peb^2_{T_h} \vdash 0\big) \le \left\lceil \frac{2h+1}{3} \right\rceil + 3 \approx \frac{2}{3} Peb(G)$.

It is not known if this bound is tight, since no corresponding lower bound on $Sp\big(Peb^d_G \vdash 0\big)$ has been shown for pebbling degree $d \ge 2$ in unrestricted resolution (in terms of the pebbling price or otherwise). The only previously known lower bound on the refutation clause space of pebbling contradictions is a bound $Sp_{\mathfrak{T}}\big(Peb^d_{T_h} \vdash 0\big) = \Omega(h)$ for the special case of tree-like resolution [23]. Unfortunately, this does not tell us anything about unrestricted resolution. For tree-like resolution, if the only way of deriving D is from clauses C_1, C_2 such that $Sp_{\mathfrak{T}}\big(F \vdash C_i\big) \ge s$, then $Sp_{\mathfrak{T}}\big(F \vdash D\big) \ge s + 1$ since one of the clauses C_i must be kept in memory while deriving the other clause. This seems to be very different from how unrestricted resolution works with respect to space.

However, the resolution refutation of $Peb^d_{T_h}$ in the proof of Theorem 4.5 in [23] is structurally quite similar to the optimal black-white pebbling of T_h presented in [27], and it is hard to see how any resolution refutation could do better than this. This raises the suspicion that the black-white pebbling price

$BW\text{-}Peb(T_h) \approx h/2$ might be a lower bound for $Sp\left(Peb^d_{T_h} \vdash 0\right)$, and in general that $Sp\left(Peb^d_G \vdash 0\right) \geq BW\text{-}Peb(G)$ for any DAG G and $d \geq 2$.

This suspicion is somewhat strengthened by the fact that for variable space, we do have a lower bound for unrestricted resolution.[1]

Theorem 4.6 ([7]). *For any $d \in \mathbb{N}^+$, $VarSp\left(Peb^d_G \vdash 0\right) \geq BW\text{-}Peb(G)$.*

If the refutation clause space of pebbling contradictions would be constant, Theorem 4.6 would imply that as G gets larger, the clauses in memory get wider, and thus weaker. Still it would somehow be possible to derive a contradiction from a constant number of these clauses of unbounded width. This appears counterintuitive.

On the other hand, for $d = 1$ refutations of Peb^1_G in constant space have exactly these "counterintuitive" properties. The resolution refutation of Peb^1_G in [7] is constructed by first downloading the pebbling axiom for the target z and then propagating falsity downwards by resolving with pebbling axioms for vertices $v \in V(G) \setminus S$ in reverse topological order. This finally yields a clause $\bigvee_{v \in S} \overline{x(v)}_1 \vee x(z)_1$ of width $|S| + 1$, which can be eliminated by resolving one by one with the source axioms $x(v)_1$ for all $v \in S$ and then with the target axiom $\overline{x(z)}_1$ to yield the empty clause 0.

If we want to establish a non-constant lower bound on $Sp\left(Peb^d_G \vdash 0\right)$ for $d \geq 2$, we have to pin down why this case is different. Intuitively, the difference is that with only one variable per vertex, a single CNF clause $\overline{x(v_1)} \vee \ldots \vee \overline{x(v_m)}$ can express the disjunction of the falsity of an arbitrary number of vertices $v_1, \ldots, v_m$, but for $d = 2$, the straightforward way of expressing that both variables $x(v_i)_1$ and $x(v_i)_2$ are false for at least one out of m vertices requires 2^m CNF clauses.

A resolution proof refutes a pebbling contradiction by deriving $x(v)_i$ and $\overline{x(v)}_i$ for some variable $x(v)_i$ and then resolving to get 0, or, in other words, by proving that some vertex v is both true and false. Arguing very informally, if we let black pebbles in a DAG G correspond to the conjunction of truth $\bigvee_{i=1}^{d} x(v)_i$ for all black-pebbled vertices v, and white pebbles in G correspond to the disjunction of falsity $\bigwedge_{i=1}^{d} \overline{x(w)}_i$ for all white-pebbled vertices w, the clauses in a pebbling contradiction encode that truth propagates "upwards" and falsity "downwards" in Peb^d_G exactly in accordance with the rules of the black-white pebble game on G. In view of this, is does not seem too far-fetched that a resolution refutation should somehow have to mimic a pebbling of the DAG on which the formula is based.

If we could make the connection between resolution and pebbling by associating truth with black pebbles and falsity with white pebbles, for $d \geq 2$ we would expect that such a connection should yield a lower bound on the refutation space of a pebbling contradiction in terms of the pebbling price of the underlying graph. This is the guiding intuition behind our proof.

5 Modifying the Black-White Pebble Game

To prove a lower bound on the refutation clause space of pebbling contradictions, we want to interpret resolution derivations in terms of pebble placements on the

[1] To be precise, the result in [7] is for $d = 1$, but the proof generalizes easily to any $d \in \mathbb{N}^+$.

corresponding graph. The translation from sets of clauses to sets of pebbles, which is presented in Section 7, follows the ideas sketched at the end of the previous section, but the problem is that the pebble configurations induced by a resolution derivation using this translation do not obey the rules of the black-white pebble game. Therefore, we need to introduce new rules for the pebble game and define a slightly altered cost function.

In this section, we present the modified pebble game used for analyzing resolution derivations. Assuming a technical lemma, we then show that for binary trees we get essentially the same bound on pebbling price in this new pebble game as for the black-white pebble game of Definition 3.1. The rather lengthy proof of the key lemma is given in the next section.

We define our adapted pebble game in two steps. Our first modification is that in the context of resolution, it appears that a more natural rule for white pebble removal is that a white pebble can be removed from a vertex when a black pebble is placed on this same vertex. This does not really change anything.

Definition 5.1 (S-pebble game). Suppose that G is a DAG with sources S and a single target z. The *superpositioned black-white pebble game*, or *S-pebble game*, is as in Definition 3.1, except that a vertex may have both a black and a white pebble on it, and the pebbling rules are (1)–(3) in Definition 3.1 and (4') below instead of rule (4) in Definition 3.1.

4'. A white pebble on v can be removed only if there is a black pebble on v.

Lemma 5.2. *Suppose that* $\mathcal{S} = \{\mathbb{S}_0, \ldots, \mathbb{S}_\tau\}$ *is an S-pebbling of a DAG* G. *Then there is an ordinary black-white pebbling* $\mathcal{P} = \{\mathbb{P}_0, \ldots, \mathbb{P}_\tau\}$ *such that* $W_t \subseteq W'_t$ *and* $B_t \subseteq B'_t$ *for* $\mathbb{P}_t = (B_t, W_t)$ *and* $\mathbb{S}_t = (B'_t, W'_t)$. *In particular,* $\mathrm{cost}(\mathcal{P}) \leq \mathrm{cost}(\mathcal{S})$.

The proof is an easy induction over $\mathcal{S}$, yielding a pebbling $\mathcal{P}$ as stated in the lemma. To avoid being overly formalistic, we ignore the fact that the inductive construction might yield "idle moves" $\mathbb{P}_t = \mathbb{P}_{t+1}$ and moves simultaneously removing a white pebble and placing a black one in $\mathcal{P}$. It is clear that this is not a problem.

Our second, and far more substantial, modification of the pebble game is motivated by the fact that when analyzing resolution derivations, we are forced to deal with "backward" pebbling moves and even "illegal erasures" of white pebbles. In order to prove lower bounds for a pebble game allowing for such moves, we have to keep track exactly on which white pebbles have been used to get a black pebble on a vertex. Loosely put, removing a white pebble from a vertex v without placing a black pebble on the same vertex should be in order, provided that all black pebbles placed on vertices above v in the DAG with the help of the white pebble on v are removed as well.

We need some notation and terminology to define and analyze our new pebble game. Recall that T denotes a complete binary tree with root z. We use p, q, r, u, v, w to denote arbitrary vertices in $V(T)$ and U, V, W to denote arbitrary subsets of vertices in $V(T)$.

For v a vertex of T, we let T^v denote the vertices in the complete binary subtree of T rooted in v, and $T^v_* = T^v \setminus \{v\}$ the vertices in T^v without its root v. We let P^v denote the vertices in the unique path from v to the root z of T and $P^v_* = P^v \setminus \{v\}$ the path without v (see Figure 2).

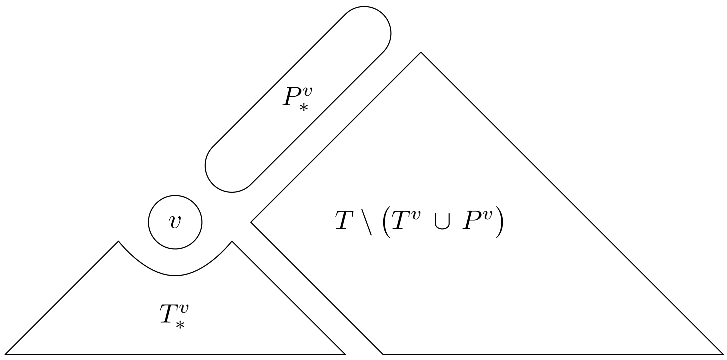


Figure 2: Referencing sets of vertices of a tree T relative to a vertex $v \in V(T)$.

We say that the vertex v is *below* u if $v \in T_*^u$ and *above* u if $v \in P_*^u$. We say that u and v are *unrelated* if $v \notin T^u \cup P^u$. We let $succ(v)$ denote the immediate successor of v and $pred(v)$ the immediate predecessors. For a leaf v we have $pred(v) = \emptyset$ and for the root z we have $succ(z) = \emptyset$. If $succ(u) = succ(v)$ for $u \neq v$, u and v are *siblings*, and we write $v = sibl(u)$. We blur the distinction somewhat between a tree T and the vertices in $V(T)$ and write for instance $T \setminus (T^v \cup P^v)$ instead of $V(T) \setminus (T^v \cup P^v)$ to denote all vertices in the tree unrelated to v.

The following definition extends the relations "below" and "above" from vertices to sets.

Definition 5.3. For sets of vertices V, W in a binary tree, we say that W is *above* V if there is no $w \in W$ such that $P_*^w \cap V \neq \emptyset$, and that W is a *roof over* V if in addition for each $v \in V$ there is a $w \in P^v \cap W$. The set W is *below* the vertex u if $W \subseteq T_*^u$. If $P_*^w \cap W = \emptyset$ for all $w \in W$, the vertex set W is *simple*.

Next, we present the concept used to keep track of for each black pebble which white pebbles (if any) this black pebble is dependent on. It might be easier to parse this rather technical definition by first studying Figure 3 and the explanations in Example 5.5.

Definition 5.4 (Subconfiguration). If v is a vertex of T and $W \subseteq T_*^v$ is a simple set below v, we say that $v\langle W\rangle$ is a *subconfiguration* with a black pebble on v supported by white pebbles on $w \in W$. The black pebble on v in $v\langle W\rangle$ is said to be *dependent on* the white pebbles in W. We refer to $v\langle\emptyset\rangle$ as an *independent black pebble*.

We define the *cover* of $v\langle W\rangle$ to be $cover(v\langle W\rangle) = T^v \setminus \bigcup_{w \in W} T^w$. The *boundary* of $v\langle W\rangle$ is $\partial v\langle W\rangle = \{v\} \cup W$. The *interior* of $v\langle W\rangle$ is $int(v\langle W\rangle) = cover(v\langle W\rangle) \setminus \partial v\langle W\rangle$ and the *closure* is $cl(v\langle W\rangle) = cover(v\langle W\rangle) \cup \partial v\langle W\rangle$.

If $cover(v\langle V\rangle) \subseteq cover(u\langle U\rangle)$, we say that $v\langle V\rangle$ is *covered by* $u\langle U\rangle$ and write $v\langle V\rangle \preceq u\langle U\rangle$. If $v\langle V\rangle \preceq u\langle U\rangle$ and $v\langle V\rangle \neq u\langle U\rangle$, we write $v\langle V\rangle \prec u\langle U\rangle$. If $cover(v\langle V\rangle) \cap cover(u\langle U\rangle) = \emptyset$, the subconfigurations $v\langle V\rangle$ and $u\langle U\rangle$ are *non-overlapping*. If $cl(v\langle V\rangle) \cap cl(u\langle U\rangle) = \emptyset$, $v\langle V\rangle$ and $u\langle U\rangle$ are *non-touching*.

When we specify the set W of white-pebbled vertices in $v\langle W\rangle$ by enumerating the members of W, we abuse notation somewhat by omitting the curly brackets inside $\langle$ and $\rangle$ around this set.

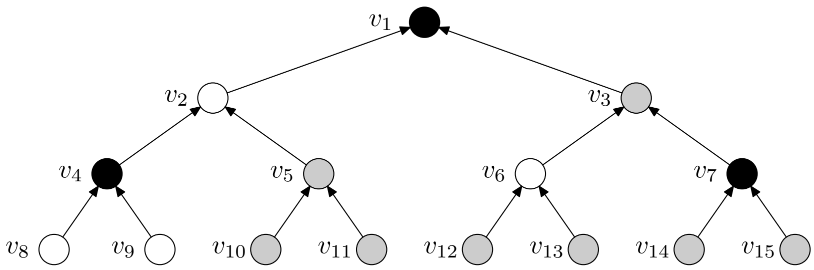


Figure 3: The subconfigurations $v_1\langle v_2, v_6\rangle$, $v_4\langle v_8, v_9\rangle$ and $v_7\langle\emptyset\rangle$.

Example 5.5. Consider the subconfigurations in Figure 3. For $v_1\langle v_2, v_6\rangle$ we have

$$\begin{aligned} \mathit{cover}(v_1\langle v_2, v_6\rangle) &= \{v_1, v_3, v_7, v_{14}, v_{15}\}, \\ \partial v_1\langle v_2, v_6\rangle &= \{v_1, v_2, v_6\}, \\ \mathit{int}(v_1\langle v_2, v_6\rangle) &= \{v_3, v_7, v_{14}, v_{15}\}, \\ \mathit{cl}(v_1\langle v_2, v_6\rangle) &= \{v_1, v_2, v_3, v_6, v_7, v_{14}, v_{15}\}. \end{aligned}$$

Since $\mathit{cl}(v_4\langle v_8, v_9\rangle) = \{v_4, v_8, v_9\}$, the subconfigurations $v_1\langle v_2, v_6\rangle$ and $v_4\langle v_8, v_9\rangle$ are non-touching. For $v_7\langle\emptyset\rangle$ we have $\mathit{cover}(v_7\langle\emptyset\rangle) = \{v_7, v_{14}, v_{15}\}$, so $v_7\langle\emptyset\rangle$ and $v_1\langle v_2, v_6\rangle$ are overlapping, or more precisely it holds that $v_7\langle\emptyset\rangle \prec v_1\langle v_2, v_6\rangle$.

Note that $\preceq$ is an order relation and that the minimal elements are subconfigurations $v\langle \mathit{pred}(v)\rangle$. We will use the following characterization of $\preceq$ repeatedly.

Observation 5.6. *$v\langle V\rangle \preceq u\langle U\rangle$ if and only if $v \in T^u$, $P^v \cap U = \emptyset$ and V is a simple roof below v over $U \cap T^v$.*

Proof. By Definition 5.4, U and V are simple sets below u and v, respectively, and $v\langle V\rangle \preceq u\langle U\rangle$ if and only if $\mathit{cover}(v\langle V\rangle) = T^v \setminus \bigcup_{w\in V} T^w \subseteq T^u \setminus \bigcup_{w \in U} T^u = \mathit{cover}(u\langle U\rangle)$.

($\Rightarrow$) Suppose that $\mathit{cover}(v\langle V\rangle) \subseteq \mathit{cover}(u\langle U\rangle)$. Since $v \in \mathit{cover}(v\langle V\rangle) \subseteq \mathit{cover}(u\langle U\rangle)$, we have $v \in T^u$ but $v \notin \bigcup_{w\in U} T^w$, and this second condition is equivalent to $P^v \cap U = \emptyset$. If V is not a roof over $U \cap T^v$, there is a $w \in U \cap T^v$ such that $P^w \cap V = \emptyset$. For such a w we would have $w \in \mathit{cover}(v\langle V\rangle)$ but $w \notin \mathit{cover}(u\langle U\rangle)$, which contradicts $\mathit{cover}(v\langle V\rangle) \subseteq \mathit{cover}(u\langle U\rangle)$.

($\Leftarrow$) Suppose that $v \in T^u$ and $P^v \cap U = \emptyset$ for V a simple roof below v over $U \cap T^v$, but that $\mathit{cover}(v\langle V\rangle) \not\subseteq \mathit{cover}(u\langle U\rangle)$. By assumption $T^v \subseteq T^u$ and $v \notin \bigcup_{w\in U} T^w$, so $v \in \mathit{cover}(u\langle U\rangle)$. Thus there must exist a $v' \in T^v_*$ such that $v' \in \mathit{cover}(v\langle V\rangle) \setminus \mathit{cover}(u\langle U\rangle)$ and $\mathit{succ}(v') \in \mathit{cover}(u\langle U\rangle)$. This implies that $v' \in U \cap T^v$, but the fact that $v' \in \mathit{cover}(v\langle V\rangle)$ shows that $P^{v'} \cap V = \emptyset$. That is, V is not a roof over $U \cap T^v$. Contradiction. ☐

Our modified black-white pebble game is defined in terms of subconfigurations of black and white vertices.

Definition 5.7 (Labelled black-white pebble game). For T a binary tree with root z, a *labelled black-white pebbling*, or *L-pebbling*, of T is a sequence $\mathcal{L} = \{\mathbb{L}_0 = \{\emptyset\}, \mathbb{L}_1, \ldots, \mathbb{L}_\tau\}$ of sets of subconfigurations $\mathbb{L}_t$ such that $\mathbb{L}_{t+1}$ is obtained from $\mathbb{L}_t$ by one of the following rules:

Introduction $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \{v\langle pred(v)\rangle\}$ for $v\langle pred(v)\rangle \notin \mathbb{L}_t$.

Merger $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \{u\langle (U \cup V) \setminus \{v\}\rangle\}$ for $u\langle U\rangle$, $v\langle V\rangle \in \mathbb{L}_t$ such that $v \in U$.

Reversal $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \{v\langle V\rangle\}$ if $v\langle V\rangle \prec u\langle U\rangle$ for some $u\langle U\rangle \in \mathbb{L}_t$.

Erasure $\mathbb{L}_{t+1} = \mathbb{L}_t \setminus \{v\langle V\rangle\}$ for $v\langle V\rangle \in \mathbb{L}_t$.

A *legal L-pebbling of* T is an L-pebbling $\mathcal{L}$ such that $\mathbb{L}_\tau = \{z\langle\emptyset\rangle\}$.

We write $u\langle U\rangle = \text{merge}(v\langle V\rangle, w\langle W\rangle)$ if $u\langle U\rangle = v\langle (V \cup W) \setminus \{w\}\rangle$ for $w \in V$, and refer to this as a *merger on* w. We let $Bl(\mathbb{L}_t) = \{v \in V(T) \mid \exists\, v\langle W\rangle \in \mathbb{L}_t\}$ denote the black pebbles and $Wh(\mathbb{L}_t) = \{w \in V(T) \mid \exists v\langle W\rangle \in \mathbb{L}_t \text{ s.t. } w \in W\}$ the white pebbles in $\mathbb{L}_t$.

In the L-pebble game, one can remove a white pebble without placing a black pebble on the same vertex, but if so the rule for erasure makes sure that any black pebble dependent on the removed white pebble is removed as well. A normal removal of a white pebble from w according to rule (4') of the S-pebble game corresponds to merging $v\langle V\rangle$ and $w\langle W\rangle$ into $v\langle (V \cup W) \setminus \{w\}\rangle$ and then erasing $v\langle V\rangle$ and $w\langle W\rangle$. Note that if $u\langle U\rangle = \text{merge}(v\langle V\rangle, w\langle W\rangle)$, then $cover(u\langle U\rangle) = cover(v\langle V\rangle) \dot\cup cover(w\langle W\rangle)$, where $\dot\cup$ denotes disjoint union.

The "backward" pebbling moves mentioned in the beginning of this section are moves according to the reversal rule. The intuition for $cover(v\langle W\rangle)$ is that this is the set of vertices already taken care of by $v\langle W\rangle$, in the sense that if the rest of the pebbling is performed optimally no black pebbles will be placed on $cover(v\langle W\rangle)$. For the relation $\preceq$, the intuition is that if $v\langle V\rangle \preceq u\langle U\rangle$, any legal pebbling reaching $u\langle U\rangle$ should be at least as expensive as an optimal pebbling reaching $v\langle V\rangle$. Arguing informally, it seems plausible that making reversals can only weaken pebbling configurations, and so the rule for reversal should not affect the pebbling cost.

We have not yet defined L-pebbling cost, however. It turns out that to make the proof of our lower bound for refutation space go through, we cannot define configuration cost as $|Bl(\mathbb{L}_t) \cup Wh(\mathbb{L}_t)|$. We have to be slightly more careful.

Definition 5.8 (L-pebbling price). The *cost* of a set of subconfigurations $\mathbb{L}$ is defined as $\text{cost}(\mathbb{L}) = \max_B\{|B \cup Wh(\mathbb{L})|\}$ where $B \subseteq Bl(\mathbb{L})$ ranges over all subsets such that if $u, v \in B$, $u \neq v$, either u and v are unrelated or, assuming that $u \in P^v_*$, there is a white pebble $w \in Wh(\mathbb{L}) \cap (P^v \setminus P^u)$ in between u and v. We say that such a set $B \subseteq Bl(\mathbb{L})$ is an *admissible choice* for $\mathbb{L}$.

The cost of an L-pebbling $\mathcal{L} = \{\mathbb{L}_0, \ldots, \mathbb{L}_\tau\}$ is $\max_{t \in [\tau]}\{\text{cost}(\mathbb{L}_t)\}$. The *L-pebbling price* $\text{L-Peb}(v\langle W\rangle)$ of a subconfiguration $v\langle W\rangle$ is the minimum cost of any L-pebbling such that $\mathbb{L}_\tau = \{v\langle W\rangle\}$, and the L-pebbling price of T is $\text{L-Peb}(T) = \text{L-Peb}(z\langle\emptyset\rangle)$.

Although the restriction on $B \subseteq Bl(\mathbb{L}_t)$ in Definition 5.8 might seem very technical, it can be given some intuitive motivation. At all times when there are two black pebbles on the same path, except for when the immediate successor of a black-pebbled vertex has just been black-pebbled, we would expect there to be a white pebble in between them supporting the uppermost black pebble. Otherwise it seems that one of these black pebbles must be redundant. For example, the black pebble on v_7 in Figure 3 appears redundant in view of the black pebble on v_1.

We want to prove that the L-pebbling price of a binary tree T is asymptotically at least as large as the black-white pebbling price $\mathit{BW\text{-}Peb}(T)$. The main obstacle in the proof is how to handle the reversal moves. It might seem intuitively clear that an optimal L-pebbling strategy does not need any reversal moves, but unfortunately, the proof of this turns out to be rather involved.

What we need to get rid of reversal moves is Lemma 5.9 stated below. We spend the rest of this section demonstrating how the desired lower bound $\text{L-Peb}(T) = \Omega\bigl(\mathit{BW\text{-}Peb}(T)\bigr)$ follows from this assumption, postponing a proof of Lemma 5.9 until the next section.

Lemma 5.9. *Suppose that $\mathcal{L}$ is an L-pebbling of a complete binary tree T. Then there is an L-pebbling $\mathcal{L}'$ of T without reversals such that* $\operatorname{cost}(\mathcal{L}') = \mathrm{O}\bigl(\operatorname{cost}(\mathcal{L})\bigr)$.

It is not too hard to see that taking a legal reversal-free L-pebbling $\mathcal{L} = \bigl\{\mathbb{L}_0, \ldots, \mathbb{L}_\tau\bigr\}$ of T and looking at $\bigl\{\mathit{Bl}\bigl(\mathbb{L}_t\bigr), \mathit{Wh}\bigl(\mathbb{L}_t\bigr)\bigr\}$ for $1 \leq t \leq \tau$, we get a legal S-pebbling of T in at most the same cost. We prove this formally in the next two lemmas.

Lemma 5.10. *Suppose that $\mathcal{L}$ is a reversal-free L-pebbling of T. Then there is a reversal-free L-pebbling $\mathcal{L}'$ of T with* $\operatorname{cost}(\mathcal{L}') \leq \operatorname{cost}(\mathcal{L})$ *such that every $v\langle V\rangle$ in $\mathcal{L}'$ occurs during one contiguous time interval, and every $v\langle V\rangle$ in $\mathcal{L}'$ except $z\langle\emptyset\rangle$ is used in exactly one merger, after which it is erased.*

Proof. We construct $\mathcal{L}'$ by backward induction over $\mathcal{L} = \bigl\{\mathbb{L}_0, \ldots, \mathbb{L}_\tau\bigr\}$. Let $\mathbb{L}'_\tau = \mathbb{L}_\tau = \bigl\{z\langle\emptyset\rangle\bigr\}$. Our induction hypothesis is that $\mathbb{L}'_t \subseteq \mathbb{L}_t$ for $\mathbb{L}'_t$ consisting of non-overlapping subconfigurations. The backward induction step from $t+1$ to t is a case analysis over the moves $\mathbb{L}_t \rightsquigarrow \mathbb{L}_{t+1}$ in $\mathcal{L}$.

Introduction $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \bigl\{v\langle \mathit{pred}(v)\rangle\bigr\}$: Set $\mathbb{L}'_t = \mathbb{L}'_{t+1} \setminus \bigl\{v\langle \mathit{pred}(v)\rangle\bigr\}$. Note that we might have $\mathbb{L}'_t = \mathbb{L}'_{t+1}$ if $v\langle \mathit{pred}(v)\rangle \notin \mathbb{L}'_{t+1}$. In any case, the induction hypothesis holds for $\mathbb{L}'_t$.

Merger $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \bigl\{v\langle(V \cup W) \setminus \{w\}\rangle\bigr\}$: If $v\langle(V \cup W) \setminus \{w\}\rangle \notin \mathbb{L}'_{t+1}$, set $\mathbb{L}'_t = \mathbb{L}'_{t+1}$. The induction hypothesis trivially remains true. Otherwise, set $\mathbb{L}'_t = \bigl(\mathbb{L}'_{t+1} \cup \bigl\{v\langle V\rangle, w\langle W\rangle\bigr\}\bigr) \setminus \bigl\{v\langle(V \cup W) \setminus \{w\}\rangle\bigr\}$. By the induction hypothesis we have $v\langle V\rangle, w\langle W\rangle \notin \mathbb{L}'_{t+1}$, since $\mathbb{L}'_{t+1}$ is non-overlapping and $v\langle V\rangle$ and $w\langle W\rangle$ are covered by $\operatorname{merge}(v\langle V\rangle, w\langle W\rangle)$ by Definitions 5.4 and 5.7. For the same reason $\mathbb{L}'_t$ must be non-overlapping. We can get from $\mathbb{L}'_t$ to $\mathbb{L}'_{t+1}$ in three steps $\mathbb{L}'_{t+1/3} = \mathbb{L}'_t \cup \bigl\{v\langle(V \cup W) \setminus \{w\}\rangle\bigr\}$, $\mathbb{L}'_{t+2/3} = \mathbb{L}'_{t+1/3} \setminus \bigl\{v\langle V\rangle\bigr\}$, $\mathbb{L}'_{t+1} = \mathbb{L}'_{t+2/3} \setminus \bigl\{w\langle W\rangle\bigr\}$ by first merging $v\langle V\rangle$ and $w\langle W\rangle$, then erasing $v\langle V\rangle$ and finally erasing $w\langle W\rangle$.

Erasure $\mathbb{L}_{t+1} = \mathbb{L}_t \setminus \bigl\{v\langle V\rangle\bigr\}$: All erasure moves in $\mathcal{L}'$ are taken care of in connection with mergers, so set $\mathbb{L}_t = \mathbb{L}'_{t+1}$.

We claim that all moves in $\mathcal{L}'$ constructed in this way are legal. If $u\langle U\rangle \in \mathbb{L}'_t$, then $u\langle U\rangle \in \mathbb{L}_t$ and for $u\langle U\rangle \neq u\langle \mathit{pred}(u)\rangle$ we know that this subconfiguration must have been derived at a time $t' \leq t$ in $\mathcal{L}$ by a merger of $v\langle V\rangle, w\langle W\rangle \prec u\langle U\rangle$. Thus the backward construction of $\mathcal{L}'$ will yield a correct derivation of $u\langle U\rangle$.

Also, any subconfiguration $v\langle V\rangle$ occurs only in one merger, after which it is immediately erased. At all times $t' > t$ after which $v\langle V\rangle$ was erased from $\mathcal{L}'$ directly after the first merger move, there is a $u\langle U\rangle \succ v\langle V\rangle$ in $\mathbb{L}'_{t'}$ and $\mathbb{L}'_{t'}$ is non-overlapping so $v\langle V\rangle$ never appears again (this can easily be formalized by a forward induction argument).

Finally, by construction $\mathbb{L}'_t \subseteq \mathbb{L}_t$, and for the merger moves in $\mathcal{L}'$ we have $\mathbb{L}'_{t+1/3}, \mathbb{L}'_{t+2/3} \subseteq \mathbb{L}_{t+1}$. This shows that for all $\mathbb{L}' \in \mathcal{L}'$ and all admissible choices $B \subseteq Bl(\mathbb{L}')$ according to Definition 5.8 such that $\mathrm{cost}(\mathbb{L}') = |B \cup Wh(\mathbb{L}')|$, the set B is also an admissible choice for a corresponding $\mathbb{L} \in \mathcal{L}$, so $\mathrm{cost}(\mathbb{L}') \leq \mathrm{cost}(\mathbb{L})$, and it follows that $\mathrm{cost}(\mathcal{L}') \leq \mathrm{cost}(\mathcal{L})$. □

Lemma 5.11. *Suppose that $\mathcal{L}$ is a reversal-free L-pebbling of a complete binary tree T. Then there is an S-pebbling $\mathcal{S}$ of T such that* $\mathrm{cost}(\mathcal{S}) \leq \mathrm{cost}(\mathcal{L})$.

Proof. By Lemma 5.10, without loss of generality we can assume that each $v\langle V\rangle$ is erased from $\mathcal{L}$ precisely after it has been used in a merger, and that $v\langle V\rangle$ is erased before $w\langle W\rangle$ when both subconfigurations are eliminated after a move $v\langle (V \cup W) \setminus \{w\}\rangle = \mathrm{merge}(v\langle V\rangle, w\langle W\rangle)$, so that the white pebble on w is removed before the black pebble on w. Also, by the construction in Lemma 5.10 it holds that if the subconfigurations $v\langle V\rangle, u\langle U\rangle \in \mathbb{L}_t$ are distinct but overlapping, then they are ordered, say $v\langle V\rangle \prec u\langle U\rangle$, and $u\langle U\rangle$ has just been derived in one step from $v\langle V\rangle$ by a merger. Because of this "non-overlappingness" property, we have $\mathrm{cost}(\mathbb{L}_t) = |Bl(\mathbb{L}_t) \cup Wh(\mathbb{L}_t)|$ for all $\mathbb{L}_t \in \mathcal{L}$.

It follows that we are done if we can construct an S-pebbling $\mathcal{S}$ with moves matching the moves in $\mathcal{L}$ exactly. For such an $\mathcal{S}$ it must hold that $\mathrm{cost}(\mathcal{S}) \leq \mathrm{cost}(\mathcal{L})$, since there will be no pebbles that Definition 5.8 does not charge $\mathcal{L}$ but that $\mathcal{S}$ has to pay for.

Let $\mathbb{S}_0 = (\emptyset, \emptyset)$ and construct $\mathbb{S}_{t+1}$ inductively by looking at the moves in $\mathbb{L}_t \rightsquigarrow \mathbb{L}_{t+1}$.

Introduction $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \{v\langle pred(v)\rangle\}$: Place white pebbles on $pred(v)$ and then a black pebble on v in $\mathcal{S}$.

Merger $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \{v\langle (V \cup W) \setminus \{w\}\rangle\}$ for $v\langle V\rangle$, $w\langle W\rangle \in \mathbb{L}_t$: No change of pebbles in $\mathcal{S}$, but note that if $v\langle V\rangle$ and $w\langle W\rangle$ are now removed, the change in pebbles on T in $\mathcal{L}$ is exactly the same as after an application of rule (4') on v.

Erasure $\mathbb{L}_{t+1} = \mathbb{L}_t \setminus \{v\langle V\rangle\}$: This is the only nontrivial case. In general, an erasure move in an L-pebbling can remove an arbitrary number of white pebbles without any black pebbles being even close to these white pebbles, and there is no way we can match such a move in an S-pebbling. But since we can assume that $\mathcal{L}$ is an L-pebbling as described in Lemma 5.10, we know that $v\langle V\rangle$ has just been used in a merger. It follows that the only pebble that disappears when going from $\{Bl(\mathbb{L}_t), Wh(\mathbb{L}_t)\}$ to $\{Bl(\mathbb{L}_{t+1}), Wh(\mathbb{L}_{t+1})\}$ is either the black pebble on v, which is always a legal removal, or some white pebble on $w \in V$ which has just been eliminated in the merger move by a black pebble, and this is a legal removal according to rule (4').

We see that $\mathcal{S}$ generated in this way is a legal S-pebbling, if we modify each introduction step into three pebble placement moves. □

Putting it all together, we get that the L-pebbling price is asymptotically at least as large as the black-white pebbling price.

Theorem 5.12. *For T a complete binary tree,* $\text{L-Peb}(T) = \Omega\big(\textit{BW-Peb}(T)\big)$.

Proof. Let $\mathcal{L}$ be an arbitrary L-pebbling of T. Assuming Lemma 5.9, there exists a legal L-pebbling $\mathcal{L}'$ of T without reversal moves with $\text{cost}(\mathcal{L}') = \text{O}\big(\text{cost}(\mathcal{L})\big)$. By Lemma 5.11, we can construct an S-pebbling $\mathcal{S}$ of T for which $\text{cost}(\mathcal{S}) \leq \text{cost}(\mathcal{L}')$. Finally, using Lemma 5.2 we get a plain old black-white pebbling $\mathcal{P}$ of T such that $\text{cost}(\mathcal{P}) \leq \text{cost}(\mathcal{S})$. The theorem follows. □

6 Getting Rid of Reversal Pebbling Moves

We now prove Lemma 5.9, i.e., that the reversal rule can be eliminated from the L-pebble game without increasing the pebbling price by more than a constant factor. This provides the missing link in the proof of Theorem 5.12.

Although this section is *very* technical, the structure of the argument is quite straightforward. Before plunging into the proof, we try to give an informal overview of where we are going.

1. First we show that without loss of generality one can assume that an optimal L-pebbling $\mathcal{L}$ is *non-overlapping*, by which we mean that all subconfigurations in $\mathbb{L}_t \in \mathcal{L}$ are non-overlapping with exception for those involved in the current merger or reversal move (Definition 6.6 and Lemma 6.9).

2. Then we observe that if we restrict an L-pebbling to a subset of the vertices in T in the natural way, we get a valid L-pebbling on this subset. We refer to this restriction operation as *projection* (Definition 6.7 and Proposition 6.10).

3. This leads to the idea of trying to get rid of reversals in the following way: When the cover of a set of subconfigurations $\mathbb{L}$ shrinks as the result of a reversal move, we can eliminate this reversal by projecting the L-pebbling moves made so far on what remains *after* the reversal move. If we do this by forward induction for all reversal moves in $\mathcal{L}$, we get a reversal-free L-pebbling $\mathcal{L}'$.

4. The problem is that these projection operations do not preserve pebbling cost—the pebbling $\mathcal{L}$ may contain reversal moves such that the projected pebbling $\mathcal{L}'$ becomes more expensive than $\mathcal{L}$. We identify which kind of reversals in $\mathcal{L}$ spoil our construction of a reversal-free and cheap pebbling $\mathcal{L}'$ by projection. Allowing some temporary wishful thinking, we then establish that if such *wasteful* reversals could somehow be avoided, the construction sketched above would work (Definition 6.12, Lemma 6.13 and Corollary 6.14).

5. Finally, we prove that wasteful reversals can be eliminated. If a pebbling $\mathcal{L}$ makes a wasteful reversal, we can replace such a move by a stronger, non-wasteful reversal without increasing the total pebbling cost by more than a constant factor (Lemma 6.19).

Summing this up, Lemma 5.9 follows.

The rest of this section contains the formal proof of the lemma. Although the technical machinery might appear cumbersome, we believe that the proof should be easier to follow if the reader tries to digest what the definitions mean and what is proven about them simply by drawing a binary tree of suitable height and working out small examples in this binary tree while reading.

Below, we assume without loss of generality that no obviously redundant pebbling moves are performed, in the sense that if a subconfiguration $v\langle V\rangle$ is derived at time t, then this subconfiguration is not just thrown away but is used at some time $t' > t$ further on in the pebbling before being erased. We state this formally.

Observation 6.1. *Let $\mathcal{L} = \{\mathbb{L}_0, \ldots, \mathbb{L}_\tau\}$ be an arbitrary L-pebbling. Then there is a pebbling $\mathcal{L}' = \{\mathbb{L}'_0, \ldots, \mathbb{L}'_{\tau'}\}$ such that $\mathrm{cost}(\mathcal{L}') \leq \mathrm{cost}(\mathcal{L})$, $\mathbb{L}'_{\tau'} = \mathbb{L}_\tau$ and if $v\langle V\rangle \in \mathbb{L}'_t \setminus \mathbb{L}'_{t-1}$, then $v\langle V\rangle$ is used in a merger or reversal move before being erased from $\mathcal{L}'$ at some time $t' > t$.*

Proof sketch. $\mathcal{L}'$ can be constructed by backward induction over $\mathcal{L}$ in the same manner as in the proof of Lemma 5.10 on page 16. □

We start by extending Definition 5.4 to sets of subconfigurations, or *L-configurations*.

Definition 6.2 (L-configuration). An *L-configuration* is a set of subconfigurations $\mathbb{L} = \{v_i\langle V_i\rangle \mid i \in [m]\}$.

We define $\mathit{cover}(\mathbb{L}) = \bigcup_{v_i\langle V_i\rangle\in\mathbb{L}} \mathit{cover}(v_i\langle V_i\rangle)$. We say that $\mathbb{L}_1$ is *covered by* $\mathbb{L}_2$ and write $\mathbb{L}_1 \preceq \mathbb{L}_2$ if $\mathit{cover}(\mathbb{L}_1) \subseteq \mathit{cover}(\mathbb{L}_2)$. If $\mathit{cover}(\mathbb{L}_1) = \mathit{cover}(\mathbb{L}_2)$, we say that $\mathbb{L}_1$ and $\mathbb{L}_2$ *coincide* and write $\mathbb{L}_1 \sim \mathbb{L}_2$. $\mathbb{L}$ is *non-overlapping* if all distinct $v\langle V\rangle, u\langle U\rangle \in \mathbb{L}$ are pairwise non-overlapping and *non-touching* if all distinct $v\langle V\rangle, u\langle U\rangle \in \mathbb{L}$ are pairwise non-touching. Two L-configurations $\mathbb{L}_1, \mathbb{L}_2$ are *mutually non-overlapping* or *mutually non-touching* if all $v\langle V\rangle \in \mathbb{L}_1$ and $u\langle U\rangle \in \mathbb{L}_2$ are pairwise non-overlapping or non-touching, respectively.

For an arbitrary set of vertices $V \subseteq V(T)$, we define the *canonical representation* $\mathrm{canon}(V)$ of V to be the unique non-touching $\mathbb{L}'$ such that $\mathit{cover}(\mathbb{L}') = V$. For $\mathbb{L}$ an arbitrary L-configuration, we define $\mathrm{canon}(\mathbb{L})$ to be the canonical representation of $\mathit{cover}(\mathbb{L})$. For $\mathbb{L}$ with $\mathrm{canon}(\mathbb{L}) = \mathbb{L}'$, the *boundary* of $\mathbb{L}$ is defined as $\partial\mathbb{L} = \bigcup_{v\langle V\rangle\in\mathbb{L}'} \partial v\langle V\rangle$, the *interior* is $\mathit{int}(\mathbb{L}) = \bigcup_{v\langle V\rangle\in\mathbb{L}'} \mathit{int}(v\langle V\rangle)$ and the *closure* is $\mathit{cl}(\mathbb{L}) = \bigcup_{v\langle V\rangle\in\mathbb{L}'} \mathit{cl}(v\langle V\rangle)$.

Example 6.3. Returning to Figure 3 on page 14, if we look at the L-configuration $\mathbb{L} = \{v_1\langle v_2, v_6\rangle,\ v_4\langle v_8, v_9\rangle,\ v_7\langle\emptyset\rangle\}$ we have $\mathit{cover}(\mathbb{L}) = \{v_1, v_3, v_4, v_7, v_{14}, v_{15}\}$. Since $v_7\langle\emptyset\rangle$ is covered by $v_1\langle v_2, v_6\rangle$ and the subconfigurations $v_1\langle v_2, v_6\rangle$ and $v_4\langle v_8, v_9\rangle$ are non-touching, we get the canonical representation simpy by leaving out $v_7\langle\emptyset\rangle$, i.e., $\mathrm{canon}(\mathbb{L}) = \{v_1\langle v_2, v_6\rangle, v_4\langle v_8, v_9\rangle\}$. Using this canonical representation of $\mathbb{L}$, we see that

$$\begin{aligned}
\partial\mathbb{L} &= \{v_1, v_2, v_4, v_6, v_8, v_9\},\\
\mathit{int}(\mathbb{L}) &= \{v_3, v_7, v_{14}, v_{15}\},\\
\mathit{cl}(\mathbb{L}) &= \{v_1, v_2, v_3, v_4, v_6, v_7, v_8, v_9, v_{14}, v_{15}\}.
\end{aligned}$$

The L-configuration $\mathbb{L}$ is overlapping because of $v_7\langle\emptyset\rangle$ and $v_1\langle v_2, v_6\rangle$, but for instance $\mathbb{L}_1 = \{v_1\langle v_2, v_6\rangle,\ v_7\langle\emptyset\rangle\}$ and $\mathbb{L}_2 = \{v_4\langle v_8, v_9\rangle\}$ are mutually non-touching.

An alternative constructive definition of canonical representation is given in the following observation. We leave it to the reader to verify that the two descriptions of canonical representation are indeed equivalent.

Observation 6.4. *The canonical representation of V can be constructed as follows: for each $v \in V$ such that $succ(v) \notin V$ or $v = z$, add the subconfiguration $v\langle W\rangle$, where $W \subseteq T_*^v$ is the maximal set such that for all $w \in W$ it holds that $P_*^w \setminus P_*^v \subseteq V$ but $w \notin V$.*

We allow a mild abuse of notation by omitting curly brackets around singleton L-configurations, writing for instance $v\langle V\rangle \preceq \mathbb{L}$, $u\langle U\rangle = \mathbb{L}$ and $w\langle W\rangle \cup \mathbb{L}$ instead of $\{v\langle V\rangle\} \preceq \mathbb{L}$, $\{u\langle U\rangle\} = \mathbb{L}$ and $\{w\langle W\rangle\} \cup \mathbb{L}$.

As a final preliminary before moving on to part 1 in the proof outline above, we collect some properties of the L-pebbling cost function of Definition 5.8.

Proposition 6.5. *Suppose that $\mathbb{L}, \mathbb{L}_1, \mathbb{L}_2$ are arbitrary L-configurations.*

1. *If $\mathbb{L}_1 \subseteq \mathbb{L}_2$ then $\mathrm{cost}(\mathbb{L}_1) \leq \mathrm{cost}(\mathbb{L}_2)$.*
2. $\frac{1}{2}(\mathrm{cost}(\mathbb{L}_1) + \mathrm{cost}(\mathbb{L}_2)) \leq \mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2) \leq 2\,(\mathrm{cost}(\mathbb{L}_1) + \mathrm{cost}(\mathbb{L}_2))$*.*
3. *If $\mathbb{L}$ is non-overlapping, $\mathrm{cost}(\mathbb{L}) = \left|Bl(\mathbb{L}) \cup Wh(\mathbb{L})\right|$, and if in addition $\mathbb{L}$ is non-touching, $\mathrm{cost}(\mathbb{L}) = \left|Bl(\mathbb{L})\right| + \left|Wh(\mathbb{L})\right| = \left|\partial\mathbb{L}\right|$.*
4. *If $\mathbb{L}_1$ and $\mathbb{L}_2$ are mutually non-touching, $\mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2) = \mathrm{cost}(\mathbb{L}_1) + \mathrm{cost}(\mathbb{L}_2)$.*
5. *For all $\mathbb{L}$, $\mathrm{cost}(\mathbb{L}) \geq \left|\partial\mathbb{L}\right|$. If $\mathbb{L}$ is non-touching equality holds, and if $\mathbb{L}$ is non-overlapping but touching strict inequality holds.*
6. *If $\mathbb{L}'_i = \mathrm{canon}(\mathbb{L}_i)$ for $i = 1, 2$, then $\mathrm{cost}(\mathbb{L}'_1 \cup \mathbb{L}'_2) \leq \mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2)$.*
7. *If $\mathbb{L}' = \mathrm{canon}(\mathbb{L})$, then $\mathrm{cost}(\mathbb{L} \cup \mathbb{L}') = \mathrm{cost}(\mathbb{L})$, and there is an L-pebbling from $\mathbb{L}$ to $\mathbb{L}'$ which does not cost more than $\mathbb{L}$.*

The reason we need Proposition 6.5 is that the cost function in the L-pebble game does not charge for all pebbles in $Bl(\mathbb{L}) \cup Wh(\mathbb{L})$. Because of this, the pebbling cost might change in unintuitive ways when L-configurations are modified. Informally, what Proposition 6.5 says is that the changes in cost cannot be *too* unintuitive. All claims in the proposition follow from Definitions 5.7, 5.8 and 6.2. Before proving the proposition, we try to explain what the different parts are used for.

Part 1 is the fundamental observation that the pebbling cost can never decrease when new subconfigurations are added. The L-pebbling cost function is not an additive measure, but part 2 says that there are linear upper and lower bounds on $\mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2)$ in terms of $\mathrm{cost}(\mathbb{L}_1)$ and $\mathrm{cost}(\mathbb{L}_2)$.

For non-overlapping and non-touching L-configurations, L-pebbling cost behaves as ordinary ordinary black-white pebbling cost, in that the cost of a union of subconfigurations is the sum of the individual costs adjusted for overlapping pebbles. This is parts 3 and 4.

Parts 5 and 6 tell us that for any given set of vertices, the cheapest way of covering this set is to use canonical L-configurations, and if $\mathbb{L}$ is not canonical, by part 7 it does not cost anything extra to make $\mathbb{L}$ canonical.

Proof of Proposition 6.5. Part 1 is obvious. If $\mathbb{L}_1 \subseteq \mathbb{L}_2$, any admissible choice for $B \subseteq Bl(\mathbb{L}_1)$ in the sense of Definition 5.8 is also an admissible choice for the black pebbles in $\mathbb{L}_2$. The first inequality in part 2 follows from part 1, since $\mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2) \geq \max\{\mathrm{cost}(\mathbb{L}_1), \mathrm{cost}(\mathbb{L}_2)\} \geq \frac{1}{2}(\mathrm{cost}(\mathbb{L}_1) + \mathrm{cost}(\mathbb{L}_2))$.

The second inequality in part 2 is more complicated. Let $B \subseteq Bl(\mathbb{L}_1 \cup \mathbb{L}_2)$ be an admissible choice such that $\mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2) = |B \cup Wh(\mathbb{L}_1 \cup \mathbb{L}_2)|$. Set $B_i = B \cap Bl(\mathbb{L}_i)$ for $i = 1, 2$. If B_1 and B_2 were admissible choices for $\mathbb{L}_1$ and $\mathbb{L}_2$ respectively, we would have $\mathrm{cost}(\mathbb{L}_1) + \mathrm{cost}(\mathbb{L}_2) \geq \mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2)$, but this need not be the case in general. All pairs of unrelated vertices $u, v \in B$ are still admissible choices for $\mathbb{L}_1$ and $\mathbb{L}_2$, but there may exist $u, v \in B_1$ with $u \in P^v_*$ such that there is a $w \in Wh(\mathbb{L}_2)$ in between u and v but $(P^v \setminus P^u) \cap Wh(\mathbb{L}_1) = \emptyset$. If so, we cannot charge for both u and v in $\mathbb{L}_1$.

Consider a *closest* pair of such vertices $u, v \in B_1$, i.e., such that $u \in P^v_*$ and $(P^v \setminus P^u) \cap Wh(\mathbb{L}_2) \neq \emptyset$ but $(P^v \setminus P^u) \cap Wh(\mathbb{L}_1) = \emptyset$, and in addition there is no black pebble in $\mathbb{L}_1$ in between u and v, or formally $(P^v_* \setminus P^u) \cap Bl(\mathbb{L}_1) = \emptyset$. Mark u to be removed from B_1 and associate u with a vertex $w_u \in (P^v \setminus P^u) \cap Wh(\mathbb{L}_2)$ which made u admissible for $\mathbb{L}_1 \cup \mathbb{L}_2$. Going through all closest pairs $u, v \in B_1$, $u \in P^v_*$, in this way and throwing away all vertices u marked for removal, we get an admissible choice $B'_1 \subseteq B_1$ for $\mathbb{L}_1$, and each u thrown away is associated with a distinct $w_u \in Wh(\mathbb{L}_2) \setminus Wh(\mathbb{L}_1)$. Then we do the same procedure for $\mathbb{L}_2$. We see that for each black pebble u eliminated in B_1 and B_2, there is a unique white pebble $w_u \in Wh(\mathbb{L}_1 \cup \mathbb{L}_2)$ that contributes to the cost of $\mathbb{L}_1$ or $\mathbb{L}_2$. Hence, $\mathrm{cost}(\mathbb{L}_1) + \mathrm{cost}(\mathbb{L}_2) \geq \frac{1}{2}\mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2)$.

For part 3, if $\mathbb{L}$ is non-overlapping it is easy to verify that $B = Bl(\mathbb{L})$ is an admissible choice in Definition 5.8, and if in addition $\mathbb{L}$ is non-touching it holds that $\partial\mathbb{L} = Bl(\mathbb{L}) \,\dot{\cup}\, Wh(\mathbb{L})$.

In part 4, the inequality $\mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2) \geq \mathrm{cost}(\mathbb{L}_1) + \mathrm{cost}(\mathbb{L}_2)$ is immediate from the mutual non-touchingness, and equality follows from the fact that $Wh(\mathbb{L}_2)$ cannot help us choose a larger $B \subseteq Bl(\mathbb{L}_1)$ and vice versa.

For $\mathbb{L}' = \mathrm{canon}(\mathbb{L})$ it holds that $Bl(\mathbb{L}') \subseteq Bl(\mathbb{L})$ and $Wh(\mathbb{L}') \subseteq Wh(\mathbb{L})$, and $Bl(\mathbb{L}')$ is always an admissible choice for $\mathbb{L}$ in Definition 5.8. If $\mathbb{L}$ is non-overlapping but touching the inclusions above are strict. Using part 3, we get part 5. Part 6 follows in the same way by observing that $Bl(\mathbb{L}'_1 \cup \mathbb{L}'_2) \subseteq Bl(\mathbb{L}_1 \cup \mathbb{L}_2)$ and $Wh(\mathbb{L}'_1 \cup \mathbb{L}'_2) \subseteq Wh(\mathbb{L}_1 \cup \mathbb{L}_2)$.

For part 7, $Bl(\mathbb{L} \cup \mathbb{L}') = Bl(\mathbb{L})$ and $Wh(\mathbb{L} \cup \mathbb{L}') = Wh(\mathbb{L})$, which shows that the cost is the same. To get the statement about pebbling, suppose for $v \in T^u_*$ that $v\langle V\rangle, u\langle U\rangle$ are overlapping and that $v\langle V\rangle \not\preceq u\langle U\rangle$. Then we can derive $w\langle W\rangle$ such that $\mathit{cover}(w\langle W\rangle) = \mathit{cover}(v\langle V\rangle) \cup \mathit{cover}(u\langle U\rangle)$ and substitute it for $v\langle V\rangle, u\langle U\rangle$ at no extra cost by first deriving $u_i\langle V \cap T^{u_i}_*\rangle$ for all $u_i \in U \cap \mathit{int}(v\langle V\rangle)$ from $v\langle V\rangle$ by reversals, and then merging all $u_i\langle V \cap T^{u_i}_*\rangle$ in turn with $u\langle U\rangle$. ☐

We define non-overlapping pebblings as L-pebblings where each introduction is immediately followed by a merger when possible, each merger is immediately followed by the erasure of the operands, and all reversals from a subconfiguration $u\langle U\rangle$ are performed in sequence after which $u\langle U\rangle$ is erased. We refer to these merger-and-erasures and reversals-and-erasure moves as *expansions* and *implosions*, respectively.

Definition 6.6 (Non-overlapping pebbling). A *non-overlapping L-pebbling* $\mathcal{L}$ is a sequence of the following types of moves.

Introduction $\mathbb{L}_{t+1} = \mathbb{L}_t \cup v\langle pred(v)\rangle$, for $v\langle pred(v)\rangle \not\preceq \mathbb{L}_t$ and $\mathbb{L}_t$ non-touching.

Expansion $\mathbb{L}_{t+3} = \big(\mathbb{L}_t \cup \text{merge}(u\langle U\rangle, v\langle V\rangle)\big) \setminus \big\{u\langle U\rangle, v\langle V\rangle\big\}$ for $u\langle U\rangle$, $v\langle V\rangle \in \mathbb{L}_t$ and $\mathbb{L}_t$ non-overlapping.

Implosion $\mathbb{L}_{t+m+1} = \big(\mathbb{L}_t \setminus u\langle U\rangle\big) \cup \mathbb{M}$ for $\mathbb{L}_t$ and $\mathbb{M} = \big\{v_i\langle V_i\rangle \mid i \in [m]\big\}$ non-touching, and $\mathbb{M} \preceq u\langle U\rangle \in \mathbb{L}_t$.

We say that $u\langle U\rangle \rightsquigarrow \mathbb{M}$ is a *nontrivial implosion* if $\mathbb{M} \prec u\langle U\rangle$.

Note that after introduction and expansion the resulting L-configuration is non-overlapping, and after implosion $\mathbb{L}_{t+m+1}$ is non-touching.

We want to prove that without loss of generality we can assume L-pebblings to be non-overlapping. The notation in the proof of this fact is simplified by introducing *projections*.

Definition 6.7 (Projection). Let $u\langle U\rangle, v\langle V\rangle$ be arbitrary subconfigurations, $\mathbb{L}$ an arbitrary L-configuration, and $\mathbb{M}$ an arbitrary non-touching L-configuration.

If $u\langle U\rangle$ and $v\langle V\rangle$ are overlapping, the *projection* of $u\langle U\rangle$ on $v\langle V\rangle$ is defined as $\text{proj}_{v\langle V\rangle}(u\langle U\rangle) = \text{canon}(cover(u\langle U\rangle) \cap cover(v\langle V\rangle))$, i.e., the unique subconfiguration $w\langle W\rangle$ such that $cover(w\langle W\rangle) = cover(u\langle U\rangle) \cap cover(v\langle V\rangle)$. If $u\langle U\rangle$ and $v\langle V\rangle$ are non-overlapping, we define $\text{proj}_{v\langle V\rangle}(u\langle U\rangle) = \emptyset$.

The projection of $u\langle U\rangle$ on $\mathbb{M}$ is $\text{proj}_{\mathbb{M}}(u\langle U\rangle) = \bigcup_{v\langle V\rangle \in \mathbb{M}} \text{proj}_{v\langle V\rangle}(u\langle U\rangle)$, and $\text{proj}_{\mathbb{M}}(\mathbb{L}) = \bigcup_{u\langle U\rangle \in \mathbb{L}} \text{proj}_{\mathbb{M}}(u\langle U\rangle)$.

In order to grasp this definition, it might be helpful to study the example in Figure 4. Here and in the following, we adopt the convention that projections resulting in $\emptyset$ are implicitly eliminated from all L-configurations.

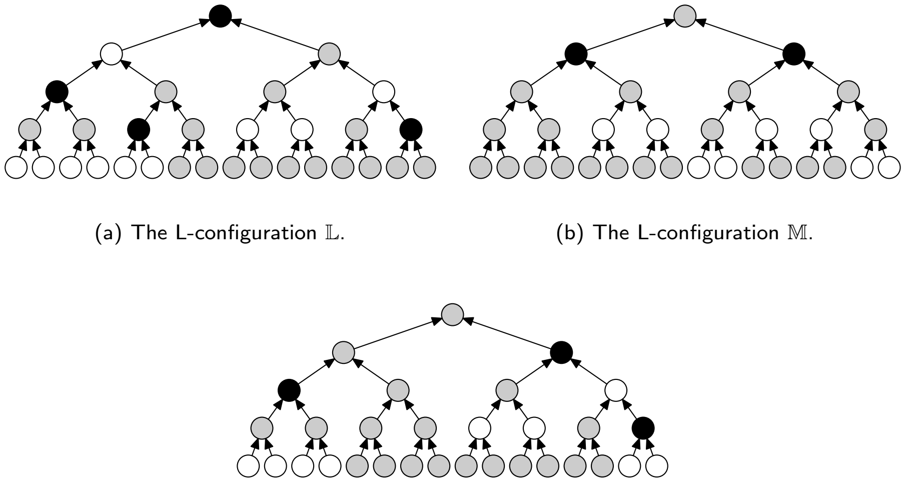
The next observation says that any L-configuration $\mathbb{L}$ can be written as a disjoint union of the sets of subconfigurations of $\mathbb{L}$ covered by each subconfiguration in $\text{canon}(\mathbb{L})$, and that the cost of $\mathbb{L}$ is the sum of the costs of the sets in this disjoint union. The proof is immediate from Definition 6.7 and Proposition 6.5, parts 4 and 5.

Observation 6.8. *Let* $\mathbb{L}' = \text{canon}(\mathbb{L})$. *Then it holds that* $\mathbb{L}$ *is a disjoint union of the sets* $\text{proj}_{v\langle V\rangle}(\mathbb{L}) = \big\{\text{proj}_{v\langle V\rangle}(u\langle U\rangle) \mid u\langle U\rangle \in \mathbb{L},\ u\langle U\rangle \preceq v\langle V\rangle\big\}$ *for all* $v\langle V\rangle \in \mathbb{L}'$. *Also,* $\text{cost}(\mathbb{L}) = \sum_{v\langle V\rangle \in \mathbb{L}'} \text{cost}(\text{proj}_{v\langle V\rangle}(\mathbb{L}))$, *and for all* $v\langle V\rangle \in \mathbb{L}'$ *it holds that* $\text{cost}(v\langle V\rangle) \leq \text{cost}(\text{proj}_{v\langle V\rangle}(\mathbb{L}))$.

Using Proposition 6.5, Definition 6.7 and Observation 6.8, we can prove that for every overlapping L-pebbling we can find a non-overlapping pebbling which is at least as good and at least as cheap.

Lemma 6.9. *Suppose that* $\mathcal{L}$ *is an arbitrary legal L-pebbling of* T. *Then there is a non-overlapping L-pebbling* $\mathcal{L}'$ *of* T *such that* $\text{cost}(\mathcal{L}') \leq \text{cost}(\mathcal{L})$.

Proof. Given $\mathcal{L} = \big\{\mathbb{L}_0, \ldots, \mathbb{L}_\tau\big\}$, we create the "backbone" $\mathcal{L}' = \big\{\mathbb{L}'_0, \ldots, \mathbb{L}'_\tau\big\}$ of a non-overlapping pebbling by setting $\mathbb{L}'_t = \text{canon}(\mathbb{L}_t)$. By Proposition 6.5, part 5, $\text{cost}(\mathbb{L}'_t) \leq \text{cost}(\mathbb{L}_t)$, so we are done if we can fill in the holes in the



(a) The L-configuration $\mathbb{L}$.

(b) The L-configuration $\mathbb{M}$.

(c) The projection $\mathrm{proj}_{\mathbb{M}}(\mathbb{L})$.

Figure 4: Example L-configurations $\mathbb{L}$ and $\mathbb{M}$ and projected L-configuration $\mathrm{proj}_{\mathbb{M}}(\mathbb{L})$.

transitions $\mathbb{L}'_t \rightsquigarrow \mathbb{L}'_{t+1}$ in cost $\max\{\mathrm{cost}(\mathbb{L}_t), \mathrm{cost}(\mathbb{L}_{t+1})\}$ using the non-overlapping moves of Definition 6.6. This is basically just an exercise in applying Proposition 6.5. Consider the moves $\mathbb{L}_t \rightsquigarrow \mathbb{L}_{t+1}$ in $\mathcal{L}$.

Introduction $\mathbb{L}_{t+1} = \mathbb{L}_t \cup v\langle \mathit{pred}(v)\rangle$: If $v\langle \mathit{pred}(v)\rangle \preceq \mathbb{L}'_t$, set $\mathbb{L}'_{t+1} = \mathbb{L}'_t$. Otherwise, introduce $v\langle \mathit{pred}(v)\rangle$ and canonize by expanding (at most three times) to get $\mathbb{L}'_{t+1} = \mathrm{canon}(\mathbb{L}_{t+1})$ in cost at most $\mathrm{cost}(\mathbb{L}_{t+1})$ by parts 6 and 7 of Proposition 6.5.

Merger $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \mathrm{merge}(u\langle U\rangle, v\langle V\rangle)$ for $u\langle U\rangle, v\langle V\rangle \in \mathbb{L}_t$: $\mathbb{L}_{t+1} \sim \mathbb{L}_t$, so set $\mathbb{L}'_{t+1} = \mathbb{L}'_t = \mathrm{canon}(\mathbb{L}_{t+1})$.

Reversal $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \{v\langle V\rangle\}$ for $v\langle V\rangle \prec u\langle U\rangle \in \mathbb{L}_t$: $\mathbb{L}_{t+1} \sim \mathbb{L}_t$, so set $\mathbb{L}'_{t+1} = \mathbb{L}'_t = \mathrm{canon}(\mathbb{L}_{t+1})$.

Erasure $\mathbb{L}_{t+1} = \mathbb{L}_t \setminus \{v\langle V\rangle\}$ for $v\langle V\rangle \in \mathbb{L}_t$: If $v\langle V\rangle \preceq \mathbb{L}_{t+1}$ we have $\mathbb{L}_{t+1} \sim \mathbb{L}_t$ and can set $\mathbb{L}'_{t+1} = \mathbb{L}'_t$, so assume that $v\langle V\rangle \not\preceq \mathbb{L}_{t+1}$.

Since $\mathbb{L}'_t$ is non-touching, there is a $u\langle U\rangle \in \mathbb{L}'_t$ such that $v\langle V\rangle \preceq u\langle U\rangle$. It follows from Observation 6.8 that for $w\langle W\rangle \in \mathbb{L}'_t$, $w\langle W\rangle \neq u\langle U\rangle$, we have $\mathrm{proj}_{w\langle W\rangle}(\mathbb{L}_{t+1}) = \mathrm{proj}_{w\langle W\rangle}(\mathbb{L}_t)$. Thus, letting $\mathbb{L}^u_i = \mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_i)$ for $i = t, t+1$, by Proposition 6.5, part 4, it is sufficient to show that we can implode $u\langle U\rangle = \mathrm{canon}(\mathbb{L}^u_t) = \mathrm{canon}(\mathbb{L}^u_{t+1} \cup v\langle V\rangle)$ into $\mathbb{M} = \mathrm{canon}(\mathbb{L}^u_{t+1})$ in cost at most $\max\{\mathrm{cost}(\mathbb{L}^u_{t+1} \cup v\langle V\rangle), \mathrm{cost}(\mathbb{L}^u_{t+1})\} = \mathrm{cost}(\mathbb{L}^u_{t+1} \cup v\langle V\rangle)$. By part 1 of the same proposition, it is enough to check that $\mathrm{cost}(\mathbb{M} \cup u\langle U\rangle) \leq \mathrm{cost}(\mathbb{L}^u_{t+1} \cup v\langle V\rangle)$. But this is just part 6 of the proposition.

Eliminating "idle moves" $\mathbb{L}'_{t+1} = \mathbb{L}'_t$, we see that we get a non-overlapping pebbling in accordance with Definition 6.6. $\square$

Lemma 6.9 tells us that as far as pebbling cost is concerned, without loss of generality we may assume that an L-pebbling $\mathcal{L}$ that reaches $z\langle\emptyset\rangle$ is non-overlapping. This completes part 1 in the proof of Lemma 5.9 sketched at the beginning of this section.

If $\mathcal{L} = \{\mathbb{L}_0, \ldots, \mathbb{L}_\tau\}$ is a non-overlapping pebbling ending in an implosion $u\langle U\rangle \rightsquigarrow \mathbb{M}$, it seems natural to try to replace the moves in $\mathcal{L}$ leading to $u\langle U\rangle$ by a reversal-free pebbling reaching $\mathbb{M} \preceq u\langle U\rangle$. Since $u\langle U\rangle$ and $\mathbb{L}_{\tau-1} \setminus u\langle U\rangle$ are mutually non-touching by definition, this substitution should not affect the cost of the pebbling outside $cl(u\langle U\rangle)$. Intuitively, one natural candidate for such a substitution is the projection of $\mathcal{L}$ on $\mathbb{M}$. We next show that projecting any L-pebbling on any non-touching L-configuration $\mathbb{M}$, we get a legal L-pebbling inside $cl(\mathbb{M})$, modulo some insignificant technical details easily taken care of. This is part 2 in our proof outline.

Proposition 6.10. *For $\mathcal{L} = \{\mathbb{L}_0, \ldots, \mathbb{L}_\tau\}$ an arbitrary L-pebbling and $\mathbb{M}$ a non-touching L-configuration, let $\mathrm{proj}_{\mathbb{M}}(\mathcal{L}) = \{\mathbb{L}'_0, \ldots, \mathbb{L}'_\tau\}$ for $\mathbb{L}'_t = \mathrm{proj}_{\mathbb{M}}(\mathbb{L}_t)$. Then $\mathrm{proj}_{\mathbb{M}}(\mathcal{L})$ is a legal L-pebbling if we eliminate idle moves $\mathbb{L}'_{t+1} = \mathbb{L}'_t$ and take care of that one reversal or erasure $\mathbb{L}_t \rightsquigarrow \mathbb{L}_{t+1}$ in $\mathcal{L}$ may correspond to a sequence of reversals or erasures respectively in $\mathrm{proj}_{\mathbb{M}}(\mathcal{L})$. Legalizing $\mathrm{proj}_{\mathbb{M}}(\mathcal{L})$ by performing these moves one by one does not affect the pebbling cost. Also, if $\mathcal{L}$ does not contain any reversals, then neither does $\mathrm{proj}_{\mathbb{M}}(\mathcal{L})$.*

Proof. By case analysis over the pebbling moves $\mathbb{L}_t \rightsquigarrow \mathbb{L}_{t+1}$ in $\mathcal{L}$.

Introduction If $v\langle pred(v)\rangle \not\preceq \mathbb{M}_t$ the projection does not change, and otherwise adding $v\langle pred(v)\rangle = \mathrm{proj}_{\mathbb{M}}(v\langle pred(v)\rangle)$ is a legal introduction move.

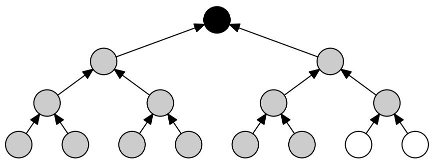
Merger Suppose that $\mathbb{L}_{t+1} = \mathbb{L}_t \cup \{u\langle (U \cup V) \setminus \{v\}\rangle\}$ for $u\langle U\rangle, v\langle V\rangle \in \mathbb{L}_t$ such that $v \in U$. For all $w\langle W\rangle \in \mathbb{M}$ such that $v \notin int(w\langle W\rangle)$, it is easy to verify that $u\langle (U \cup V) \setminus \{v\}\rangle$ projects the same subconfigurations on $w\langle W\rangle$ as do $u\langle U\rangle$ and $v\langle V\rangle$ together. Suppose that $v \in int(w\langle W\rangle)$. Since $\mathbb{M}$ is non-touching there is at most one such $w\langle W\rangle \in \mathbb{M}$, and $\mathrm{proj}_{w\langle W\rangle}(u\langle (U \cup V) \setminus \{v\}\rangle)$ can be verified to be a legal merger of $\mathrm{proj}_{w\langle W\rangle}(u\langle U\rangle)$ and $\mathrm{proj}_{w\langle W\rangle}(v\langle V\rangle)$.

Reversal If $v\langle V\rangle \preceq u\langle U\rangle$ it holds that $\mathrm{proj}_{\mathbb{M}}(v\langle V\rangle) \preceq \mathrm{proj}_{\mathbb{M}}(u\langle U\rangle)$, so adding $\mathrm{proj}_{\mathbb{M}}(v\langle V\rangle)$ is a sequence of legal reversals. As this sequence of reversals is performed, the pebbling cost increases monotonously by part 1 of Proposition 6.5.

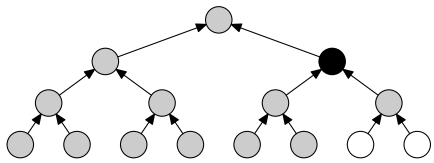
Erasure If $\mathbb{L}_{t+1} = \mathbb{L}_t \setminus \{v\langle V\rangle\}$ for $v\langle V\rangle \in \mathbb{L}_t$, removing $\mathrm{proj}_{\mathbb{M}}(v\langle V\rangle)$ from $\mathbb{L}'_t$ is a sequence of legal erasures. As this sequence of erasures is performed, the pebbling cost decreases monotonously by part 1 of Proposition 6.5.

We see that the cost of this pebbling is $\max_{t\in[\tau]}\{\mathrm{proj}_{\mathbb{M}}(\mathbb{L}_t)\}$, and if $\mathcal{L}$ is reversal-free then so is $\mathrm{proj}_{\mathbb{M}}(\mathcal{L})$, since every move in $\mathcal{L}$ is matched by the same kind of moves in $\mathrm{proj}_{\mathbb{M}}(\mathcal{L})$. □

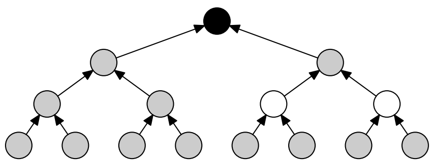
In view of Proposition 6.10, the transformation from a non-overlapping pebbling $\mathcal{L}$ to a reversal-free pebbling $\mathcal{L}'$ seems obvious: by forward induction over the moves in $\mathcal{L}$, replace each implosion $u\langle U\rangle \rightsquigarrow \mathbb{M}$ at time t by a local projection of $\{\mathbb{L}_0, \ldots, \mathbb{L}_t\}$ on $\mathbb{M}$. Since by induction there are no reversals before time t, the projection must be a reversal-free pebbling inside $cl(\mathbb{M})$. Doing this for all



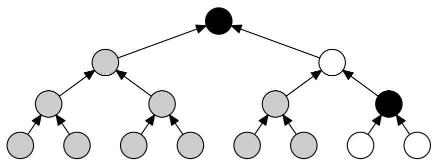
(a) The subconfiguration $u\langle U\rangle$.



(b) Wastefully lowered black pebble.



(c) Superfluous white pebble.



(d) Wasteful "split" of $u\langle U\rangle$.

Figure 5: A subconfiguration $u\langle U\rangle$ and three wasteful implosions of $u\langle U\rangle$.

implosions, we get a globally reversal-free pebbling $\mathcal{L}'$ ending in $z\langle\emptyset\rangle$. This is the transformation described in part 3 of our road map for the proof of Lemma 5.9.

There is only one problem. Although $\mathrm{proj}_{\mathbb{M}}(\mathcal{L})$ is a legal L-pebbling, it is not true in general that $\mathrm{cost}(\mathrm{proj}_{\mathbb{M}}(\mathcal{L})) \leq \mathrm{cost}(\mathcal{L})$. For instance, if $v\langle V\rangle \preceq u\langle\emptyset\rangle$ for $V \neq \emptyset$, then $\mathrm{proj}_{v\langle V\rangle}(u\langle\emptyset\rangle) = v\langle V\rangle$ and hence $\mathrm{cost}(\mathrm{proj}_{v\langle V\rangle}(u\langle\emptyset\rangle)) = 1 + |V| > \mathrm{cost}(u\langle\emptyset\rangle) = 1$. Looking at this counterexample, however, it seems clear that having gotten as far as $u\langle\emptyset\rangle$, reversing to the weaker and more expensive configuration $v\langle V\rangle$ is a blind alley that no optimal pebbling would go into. What we want to do next is to define formally which reversals are *wasteful* in this sense, and to prove that for pebblings avoiding such wasteful reversals, projection does not increase the pebbling cost.

Since the definition of wastefulness turns out to be quite technical, we first try to give some more intuition for which kind of reversals we disapprove of.

Example 6.11. Consider the subconfiguration $u\langle U\rangle$ in Figure 5(a).

1. If $v \in T^u_*$, the reversal $u\langle U\rangle \rightsquigarrow v\langle T^v_* \cap U\rangle$ is acceptable only if $T^v_* \cap U \subsetneq U$, i.e., if we get rid of white pebbles by lowering the black pebble from u to v. The reversal in Figure 5(b) does not satisfy this.

2. For V a simple roof below u over U, we approve of $u\langle U\rangle \rightsquigarrow u\langle V\rangle$ only if for all $w \in V$ it holds that $T^w \cap U \neq \emptyset$. Otherwise, unnecessary white pebbles have been introduced, as in Figure 5(c).

3. If $u\langle U\rangle$ is imploded into non-touching $\{v_1\langle V_1\rangle, v_2\langle V_2\rangle\}$ such that, say, $v_2 \in T^{v_1}_*$, it should not be the case that $v_1\big\langle\big(V_1 \setminus P^{v_2}\big) \cup V_2\big\rangle \preceq u\langle U\rangle$, for if so we could have reversed to this stronger subconfiguration instead of $\{v_1\langle V_1\rangle, v_2\langle V_2\rangle\}$ at no extra cost. The implosion in Figure 5(d) violates this condition.

The reversals from $u\langle U\rangle$ in figures 5(b), 5(c) and 5(d) are all examples of wasteful implosions for which our reversal-free pebbling $\mathcal{L}'$ constructed by projection may become more expensive than $\mathcal{L}$. Looking at these examples, it is easy to believe that such moves are nonoptimal and that it ought to be possible to eliminate them. The formal definition of wastefulness is as follows.

Definition 6.12 (Wasteful implosion). For a non-touching L-configuration $\mathbb{M} = \{v_i\langle V_i\rangle\} \preceq u\langle U\rangle$, the implosion $u\langle U\rangle \rightsquigarrow \mathbb{M}$ is *non-wasteful* if

1. for every $v \in \mathit{Bl}(\mathbb{M}) \setminus \{u\}$ there is a $w \in U \cap T_*^{\mathit{succ}(v)}$ such that it holds for the path $p_v = P^w \setminus P_*^{\mathit{succ}(v)}$ that $p_v \cap \big(\mathit{Bl}(\mathbb{M}) \cup \mathit{Wh}(\mathbb{M})\big) = \emptyset$,
2. for every $v \in \mathit{Wh}(\mathbb{M})$ there is a $w \in U \cap T^v$ such that it holds for the path $p_v = P^w \setminus P_*^v$ that $p_v \cap \mathit{Bl}(\mathbb{M}) = \emptyset$,
3. the paths from $\big(\mathit{Bl}(\mathbb{M}) \cup \mathit{Wh}(\mathbb{M})\big) \setminus \{u\}$ to $\mathit{Wh}\big(u\langle U\rangle\big) = U$ can all be chosen pairwise disjoint, i.e., such that $p_v \cap p_{v'} = \emptyset$ if $v \neq v$.

If $u\langle U\rangle \rightsquigarrow \mathbb{M}$ is not a non-wasteful implosion it is said to be *wasteful*.

Definition 6.12 identifies the offending reversal moves for which our projective construction of a reversal-free but cheap pebbling fails. Continuing according to part 4 in our proof plan, we show that for pebblings without such wasteful moves the projective construction works. This is the next lemma.

Lemma 6.13. *Suppose that* $\mathcal{L} = \big\{\mathbb{L}_0, \ldots, \mathbb{L}_{\tau-2}, \mathbb{L}_{\tau-1} = u\langle U\rangle \rightsquigarrow \mathbb{M}\big\}$ *is a pebbling without reversals except for a final non-wasteful implosion* $u\langle U\rangle \rightsquigarrow \mathbb{M}$*. Then* $\mathrm{cost}(\mathrm{proj}_{\mathbb{M}}(\mathbb{L}_t)) \leq \mathrm{cost}(\mathbb{L}_t)$ *for all* $t > \tau$*, and* $\mathrm{cost}(\mathrm{proj}_{\mathbb{M}}(\mathcal{L})) \leq \mathrm{cost}(\mathcal{L})$*.*

Proof. Let $\mathbb{L}'_t = \mathrm{proj}_{\mathbb{M}}(\mathbb{L}_t)$ for all $t < \tau$. By Proposition 6.10, it suffices to show $\mathrm{cost}(\mathbb{L}'_t) \leq \mathrm{cost}(\mathbb{L}_t)$ to get $\mathrm{cost}(\mathrm{proj}_{\mathbb{M}}(\mathcal{L})) \leq \mathrm{cost}(\mathcal{L})$.

By the proof of Lemma 5.10, $\mathit{cover}(\mathbb{L}_t)$ grows monotonously with t in a non-redundant reversal-free pebbling, so in particular $\mathbb{L}_t \preceq u\langle U\rangle$ for all t. By the same lemma, $\mathcal{L}$ is non-overlapping. This means that we can confine ourselves to looking at non-overlapping L-configurations $\mathbb{L}_t$, since the overlapping L-configurations during expansion moves do not affect the pebbling cost. Using Proposition 6.5, part 3, we have that $\mathrm{cost}(\mathbb{L}_t) = \big|\mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t)\big|$. Hence, to prove $\mathrm{cost}(\mathbb{L}'_t) \leq \mathrm{cost}(\mathbb{L}_t)$ it is enough to find for each $v \in \mathit{Bl}(\mathbb{L}'_t) \cup \mathit{Wh}(\mathbb{L}'_t)$ an associated $v_L \in \mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t)$ such that $v_L \neq w_L$ if $v \neq w$.

If $v \in \big(\mathit{Bl}(\mathbb{L}'_t) \cup \mathit{Wh}(\mathbb{L}'_t)\big) \cap \big(\mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t)\big)$, an obvious choice is $v_L = v$. Suppose therefore that $v \in \big(\mathit{Bl}(\mathbb{L}'_t) \cup \mathit{Wh}(\mathbb{L}'_t)\big) \setminus \big(\mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t)\big)$. Then $v \in \partial\mathbb{M}$, since it is easy to check that $v \in \mathit{int}(\mathbb{M})$ implies $v \in \mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t)$. Also, there is a subconfiguration $w_v\langle W_v\rangle \in \mathbb{L}_t$ such that $v \in \mathit{int}\big(w_v\langle W_v\rangle\big)$, namely the $w_v\langle W_v\rangle$ projecting the pebble on v. Lastly, note that if $v \in \big(\mathit{Bl}(\mathbb{L}'_t) \cup \mathit{Wh}(\mathbb{L}'_t)\big) \cap \partial\mathbb{M}$, it is a routine matter to verify that v has the same colour in $\mathbb{L}'_t$ and $\mathbb{M}$, i.e., either $v \in \mathit{Bl}(\mathbb{L}'_t) \cap \mathit{Bl}(\mathbb{M})$ or $v \in \mathit{Wh}(\mathbb{L}'_t) \cap \mathit{Wh}(\mathbb{M})$. We choose $v_L \in \mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t)$ for such vertices v by first associating a unique $v_u \in U = \mathit{Wh}\big(u\langle U\rangle\big)$ to v as follows.

1. If $v \in \mathit{Bl}(\mathbb{L}'_t) \cap \mathit{Bl}(\mathbb{M})$, pick a vertex $v_u \in \big(U \cap T_*^{\mathit{succ}(v)}\big) \setminus T^v$ and a path $p_v = P^{v_u} \setminus P_*^{\mathit{succ}(v)}$ to v_u such that $p_v \cap \big(\mathit{Bl}(\mathbb{M}) \cup \mathit{Wh}(\mathbb{M})\big) = \emptyset$ as guaranteed by Definition 6.12. For the subconfiguration $w_v\langle W_v\rangle \in \mathbb{L}_t$ projecting the black pebble on v, we must have $\mathit{succ}(v) \in \mathit{cover}\big(w_v\langle W_v\rangle\big)$ since $v \in \mathit{int}\big(w_v\langle W_v\rangle\big)$, and consequently $p_v \cap \mathit{cover}\big(w_v\langle W_v\rangle\big) \neq \emptyset$.

2. If $v \in \mathit{Wh}(\mathbb{L}'_t) \cap \mathit{Wh}(\mathbb{M})$, pick $v_u \in U \cap T^v$ and $p_v = P^{v_u} \setminus P^v_*$ such that $p_v \cap \mathit{Bl}(\mathbb{M}) = \emptyset$ as guaranteed by the definition. For $w_v\langle W_v\rangle \in \mathbb{L}_t$ projecting the white pebble on v, we have $v \in \mathit{int}(w_v\langle W_v\rangle) \subseteq \mathit{cover}(w_v\langle W_v\rangle)$, so $p_v \cap \mathit{cover}(w_v\langle W_v\rangle) \neq \emptyset$.

According to Definition 6.12 all paths p_v can be chosen disjoint, since by assumption $u\langle U\rangle \rightsquigarrow \mathbb{M}$ is a non-wasteful implosion.

We now use the paths to the associated $v_u \in U$ to choose a distinct $v_L \in \mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t)$ for all $v \in (\mathit{Bl}(\mathbb{L}'_t) \cup \mathit{Wh}(\mathbb{L}'_t)) \setminus (\mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t))$. By construction, $\mathbb{L}'_t \preceq \mathbb{L}_t \preceq u\langle U\rangle$, and in particular $w_v\langle W_v\rangle \preceq u\langle U\rangle$ for all $w_v\langle W_v\rangle$ found above. Note that $p_v \cap \mathit{cover}(w_v\langle W_v\rangle) \neq \emptyset$ and that $p_v \not\subseteq \mathit{cover}(u\langle U\rangle)$ since the lowest vertex in p_v is a white pebble of $u\langle U\rangle$. This implies that $W_v \cap p_v \neq \emptyset$, for otherwise $p_v \subseteq \mathit{cover}(w_v\langle W_v\rangle)$ which yields the contradiction $w_v\langle W_v\rangle \not\preceq u\langle U\rangle$. Thus we can choose $v_L \in \mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t)$ to be the vertex in $W_v \cap p_v$. Since all paths p_v are disjoint, all such v_L are distinct. They must also be distinct from the $v \in (\mathit{Bl}(\mathbb{L}'_t) \cup \mathit{Wh}(\mathbb{L}'_t)) \cap (\mathit{Bl}(\mathbb{L}_t) \cup \mathit{Wh}(\mathbb{L}_t))$, since $\mathbb{L}_t$ is non-overlapping. It follows that $\mathrm{cost}(\mathbb{L}'_t) \leq \mathrm{cost}(\mathbb{L}_t)$. ☐

We can use Lemma 6.13 to eliminate non-wasteful implosions one by one. If $\mathcal{L} = \{\mathbb{L}_0, \ldots, \mathbb{L}_\tau \rightsquigarrow (\mathbb{L}_\tau \setminus u\langle U\rangle) \cup \mathbb{M}\}$ is a non-overlapping reversal-free pebbling except for a final non-wasteful implosion $u\langle U\rangle) \rightsquigarrow \mathbb{M}$, then by definition $\mathbb{L}_\tau = \{v_i\langle V_i\rangle\}$ is non-touching, and using Observation 6.8 each $\mathbb{L}_t$ can be written as a non-touching union of $\mathbb{L}_t^{v_i} = \mathrm{proj}_{v_i\langle V_i\rangle}(\mathbb{L}_t)$ such that $\mathrm{cost}(\mathbb{L}_t) = \sum_{v_i\langle V_i\rangle \in \mathbb{M}} \mathrm{cost}(\mathbb{L}_t^{v_i})$. For all $v_i\langle V_i\rangle \in \mathbb{L}_t$, $\mathcal{L}^{v_i} = \{\mathbb{L}_0^{v_i}, \ldots, \mathbb{L}_{\tau-1}^{v_i}, \{v_i\langle V_i\rangle\}\}$ can be seen to be pairwise non-touching pebblings without reversals. It follows by Lemma 6.13 that we can locally replace $\mathbb{L}_t^u$ for the imploded subconfiguration $u\langle U\rangle$ by $\mathrm{proj}_{\mathbb{M}}(\mathbb{L}_t^u)$ without increasing the global pebbling cost. Doing this by forward induction for all implosions in turn, we get Corollary 6.14.

Corollary 6.14. *Let $\mathcal{L} = \{\mathbb{L}_0, \ldots, \mathbb{L}_\tau = \{z\langle\emptyset\rangle\}\}$ be a non-overlapping L-pebbling of T without wasteful implosions. Then there is an L-pebbling $\mathcal{L}'$ of T without any reversal moves such that* $\mathrm{cost}(\mathcal{L}') \leq \mathrm{cost}(\mathcal{L})$.

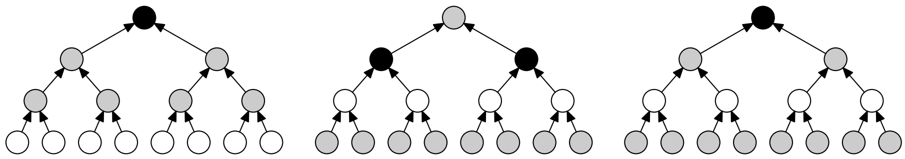
This concludes part 4 in the proof outline on page 18.

All that remains is to show that in an arbitrary non-overlapping L-pebbling we can always replace wasteful implosions by non-wasteful ones without increasing the pebbling cost by more than a constant factor. It will take a couple of technical lemmas before we get there, but the intuition from Example 6.11 is clear: if $\mathbb{L}_t \rightsquigarrow \mathbb{L}_{t+m+1}$ is a wasteful implosion, we should be able to match this move with a non-wasteful implosion $\mathbb{L}'_t \rightsquigarrow \mathbb{L}'_{t+m+1}$ instead, where $\mathbb{L}'_i \succeq \mathbb{L}_i$ and $\mathrm{cost}(\mathbb{L}'_i) \leq \mathrm{cost}(\mathbb{L}_i)$ for $i = t, t+m+1$. The only thing that complicates the matter is that we may have to pay extra for the transitional L-configurations during the implosion $\mathbb{L}'_t \rightsquigarrow \mathbb{L}'_{t+m+1}$ because of overlapping subconfigurations.

The cornerstone of our proof is the fact that for every wasteful implosion $u\langle U\rangle \rightsquigarrow \mathbb{L}$, there is a non-wasteful implosion to $\mathbb{M} \succ \mathbb{L}$ with $\mathrm{cost}(\mathbb{M}) \leq \mathrm{cost}(\mathbb{L})$.

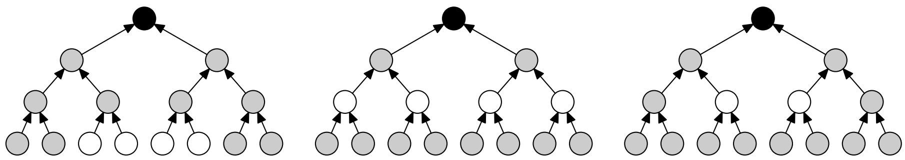
Lemma 6.15. *If $u\langle U\rangle \rightsquigarrow \mathbb{L}$ is a wasteful implosion, then there is a non-touching $\mathbb{M}$ such that $u\langle U\rangle \succeq \mathbb{M} \succ \mathbb{L}$, $\mathrm{cost}(\mathbb{M}) \leq \min\{\mathrm{cost}(u\langle U\rangle), \mathrm{cost}(\mathbb{L})\}$ and $u\langle U\rangle \rightsquigarrow \mathbb{M}$ is a non-wasteful implosion.*

Proof. If $u\langle U\rangle \rightsquigarrow \mathbb{M}$ is a non-wasteful implosion, it holds that $\mathrm{cost}(\mathbb{M}) = |\mathit{Bl}(\mathbb{M})| + |\mathit{Wh}(\mathbb{M})| \leq \mathrm{cost}(u\langle U\rangle) = 1 + |U|$, since by Definition 6.12 every $v \in (\mathit{Bl}(\mathbb{M}) \cup \mathit{Wh}(\mathbb{M})) \setminus \{u\}$ can be associated with a distinct $w \in U$.



(a) The subconfiguration. (b) Wasteful implosion $\mathbb{L}$. (c) Non-wasteful $\mathbb{M} \succ \mathbb{L}$.

Figure 6: Wasteful and corresponding non-wasteful implosion according to condition 1.



(a) The subconfiguration. (b) Wasteful implosion $\mathbb{L}$. (c) Non-wasteful $\mathbb{M} \succ \mathbb{L}$.

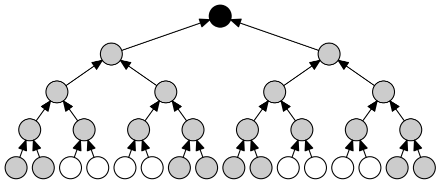
Figure 7: Wasteful and corresponding non-wasteful implosion according to condition 2.

We demonstrate that if $u\langle U\rangle \rightsquigarrow \mathbb{L}$ is a wasteful implosion, we can find an $\mathbb{M}$ such that $u\langle U\rangle \succeq \mathbb{M} \succ \mathbb{L}$ and $\mathrm{cost}(\mathbb{M}) \leq \mathrm{cost}(\mathbb{L})$. If $u\langle U\rangle \rightsquigarrow \mathbb{M}$ is a wasteful implosion, we repeat this construction. Sooner or later the process must terminate for some $\mathbb{M} \preceq u\langle U\rangle$ such that $u\langle U\rangle \rightsquigarrow \mathbb{M}$ is non-wasteful—if nothing else, by definition the trivial implosion $u\langle U\rangle \rightsquigarrow u\langle U\rangle$ is.

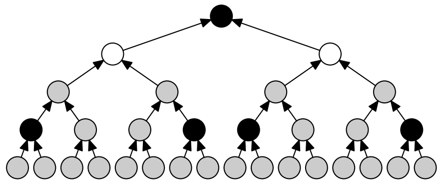
According to Definition 6.12, the configuration $\mathbb{L} = \{v_i\langle V_i\rangle\}$ can be wasteful with respect to $u\langle U\rangle$ in three ways.

1. Some black pebble $v \in \mathit{Bl}(\mathbb{L}) \setminus \{u\}$ lacks a path. If $\mathit{succ}(v) \in \mathit{Wh}(\mathbb{L})$ we must have $\mathit{succ}(v) \in \mathit{cover}(u\langle U\rangle)$, so we can add $\mathrm{canon}(\{\mathit{succ}(v)\}) = \mathit{succ}(v)\langle v, \mathit{sibl}(v)\rangle$ to $\mathbb{L}$ and set $\mathbb{M} = \mathrm{canon}(\mathbb{L} \cup \mathit{succ}(v)\langle v, \mathit{sibl}(v)\rangle) \succ \mathbb{L}$ with $\mathrm{cost}(\mathbb{M}) \leq \mathrm{cost}(\mathbb{L}) + |\{\mathit{sibl}(v)\}| - |\{v, \mathit{succ}(v)\}| < \mathrm{cost}(\mathbb{L})$. Otherwise, all paths from $\mathit{succ}(v)$ downwards in $T^{\mathit{sibl}(v)}$ are either blocked by $r_1, \ldots, r_m \in \mathit{Bl}(\mathbb{L}) \cap T^{\mathit{sibl}(v)}$ or reach sources in $T^{\mathit{sibl}(v)}$ without passing pebbled vertices (we can have $m = 0$). From this we can conclude that $V = T^{\mathit{succ}(v)} \setminus \left(T^v \cup \bigcup_{i\in[m]} T^{r_i}\right) \subseteq \mathit{cover}(u\langle U\rangle)$, so we can add $\mathrm{canon}(V) = \mathit{succ}(v)\langle v, r_1, \ldots, r_m\rangle \preceq u\langle U\rangle$ to $\mathbb{L}$, which increases the cost by 1 for $\mathit{succ}(v)$. Setting $\mathbb{M} = \mathrm{canon}(\mathbb{L} \cup \mathit{succ}(v)\langle v, r_1, \ldots, r_m\rangle) \succ \mathbb{L}$ removes the pebbles from v and $r_1, \ldots, r_m$ and decreases the cost by at least 1, so $\mathrm{cost}(\mathbb{M}) \leq \mathrm{cost}(\mathbb{L})$. See Figure 6 for a simple example.

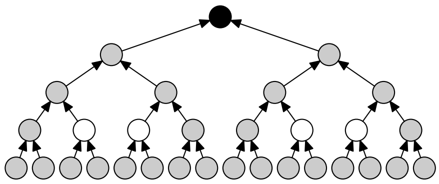
2. There is a white pebble $w \in \mathit{Wh}(\mathbb{L})$ such that all paths downwards in T^w either are blocked by $r_1, \ldots, r_m \in \mathit{Bl}(\mathbb{L}) \cap T^w_*$ or reach sources in T^w without passing pebbled vertices. If so, we have $V = T^w \setminus \bigcup_{i\in[m]} T^{r_i} \subseteq \mathit{cover}(u\langle U\rangle)$, and we can add $\mathrm{canon}(V) = w\langle r_1, \ldots, r_m\rangle \preceq u\langle U\rangle$ to $\mathbb{L}$ at no extra cost and set $\mathbb{M} = \mathrm{canon}(\mathbb{L} \cup w\langle r_1, \ldots, r_m\rangle) \succ \mathbb{L}$. Here we get a strict inequality $\mathrm{cost}(\mathbb{M}) < \mathrm{cost}(\mathbb{L})$ since the canonization eliminates at least the pebbles on w. This case is illustrated in Figure 7.



(a) The subconfiguration $u\langle U\rangle$.



(b) Wasteful implosion $\mathbb{L}$ of $u\langle U\rangle$.



(c) Non-wasteful implosion $u\langle U\rangle \rightsquigarrow \mathbb{M} \succ \mathbb{L}$.

Figure 8: Wasteful and corresponding non-wasteful implosion according to condition 3.

3. There are paths for all $v \in \big(Bl(\mathbb{L}) \cup Wh(\mathbb{L})\big) \setminus \{u\}$ to $w \in U$, but they cannot be chosen disjoint. Start picking disjoint paths bottom-up so that when we choose a path for a white pebble $v \in Wh(\mathbb{L})$ we have already determined paths for all $w \in \big(Bl(\mathbb{L}) \cup Wh(\mathbb{L})\big) \cap T^v_*$, and when we choose a path for a black pebble $v \in Bl(\mathbb{L})$ we have already determined paths for all $w \in \big(Bl(\mathbb{L}) \cup Wh(\mathbb{L})\big) \cap T^{sibl(v)}_*$. Note that for black pebbles, the vertex $sibl(v)$ itself cannot be black-pebbled, for if so there would be no path for v and we would have case 1. For the same reason, $succ(v)$ is not white-pebbled, and then $sibl(v)$ cannot be white-pebbled either since $\mathbb{L}$ is non-touching.

   At some point we reach a v such that no matter how we choose the paths below, we cannot choose a disjoint path for v. Consider the colour of v.

   (a) v is black. There are white pebbles in $U \cap T^{sibl(v)}$ reachable from v, but they are all blocked by already chosen paths from $r_1, \ldots, r_m \in Bl(\mathbb{L}) \cap T^{sibl(v)}$. This means that $\{succ(r_i) \mid i \in [m]\} \subseteq cover(u\langle U\rangle)$, so we can add $\mathrm{canon}(\{succ(r_i) \mid i \in [m]\}) = \big\{r_i\langle pred(r_i)\rangle \mid i \in [m]\big\}$ to $\mathbb{L}$ at an additional cost $2m$. Reasoning in the same way, we can also include $succ(v)\langle v, succ(r_1), \ldots, succ(r_m)\rangle$ at a further cost of 1 for the unpebbled vertex $succ(v)$. When we canonize, the pebbles on $v, r_1, \ldots, r_m, succ(r_1), \ldots, succ(r_m)$ all disappear and the cost decreases by $2m+1$, resulting in $\mathbb{M} \succ \mathbb{L}$ with $\mathrm{cost}(\mathbb{M}) \leq \mathrm{cost}(\mathbb{L})$.

   (b) v is white. The construction is analogous. Let the blocking black pebbles be $r_1, \ldots, r_m \in Bl(\mathbb{L}) \cap T^v_*$. Again we can add $r_i\langle pred(r_i)\rangle$ for all $i \in [m]$ at an extra cost $2m$. Since $succ(r_1), \ldots, succ(r_m)$ blocks all paths from v we have $T^v \setminus \big(T^{succ(r_1)} \cup \ldots \cup T^{succ(r_m)}\big) \subseteq cover(u\langle U\rangle)$, so $v\langle succ(r_1), \ldots, succ(r_m)\rangle$ can be added as well at no additional cost. Canonizing decreases the cost by $2m+1$, which yields $\mathbb{M} \succ \mathbb{L}$ with $\mathrm{cost}(\mathbb{M}) < \mathrm{cost}(\mathbb{L})$. The transition from Figure 8(b) to Figure 8(c) is accomplished by applying this construction twice.

In all cases we can find $\mathbb{M}$ with $\mathrm{cost}(\mathbb{M}) \leq \mathrm{cost}(\mathbb{L})$ and $\mathbb{M} \succ \mathbb{L}$. The lemma follows. $\square$

The following transitivity property of non-wasteful implosions is immediate from Definition 6.12.

Observation 6.16. *If $u\langle U\rangle \rightsquigarrow \{v_i\langle V_i\rangle \mid i \in [m]\}$ and $v_i\langle V_i\rangle \rightsquigarrow \mathbb{M}_i$ for $i \in [m]$ are all non-wasteful implosions, then $u\langle U\rangle \rightsquigarrow \{\mathbb{M}_i \mid i \in [m]\}$ is a non-wasteful implosion.*

It follows from Observation 6.16, that if $u\langle U\rangle \rightsquigarrow \mathbb{L}$ is a wasteful implosion and $u\langle U\rangle \rightsquigarrow \mathbb{M} \succ \mathbb{L}$ is a corresponding non-wasteful implosion for $\mathbb{M}$ minimal, then all nontrivial "local implosions" from subconfigurations in $\mathbb{M}$ to subconfigurations in $\mathbb{L}$ are wasteful.

Lemma 6.17. *Suppose that $u\langle U\rangle \rightsquigarrow \mathbb{L}$ is a wasteful implosion and let $\mathbb{M} \succ \mathbb{L}$ be minimal such that $u\langle U\rangle \rightsquigarrow \mathbb{M}$ is non-wasteful. Then for each $v\langle V\rangle \in \mathbb{M}$ and each non-touching $\mathbb{L}'$ such that $\mathbb{M} \succ \mathbb{L}' \succeq \mathbb{L}$, either $\mathrm{proj}_{v\langle V\rangle}(\mathbb{L}') = v\langle V\rangle$ or $v\langle V\rangle \rightsquigarrow \mathrm{proj}_{v\langle V\rangle}(\mathbb{L}')$ is a wasteful implosion. In particular, for each $v\langle V\rangle \in \mathbb{M}$ it holds that $\mathrm{cost}(v\langle V\rangle) \leq \mathrm{cost}(\mathrm{proj}_{v\langle V\rangle}(\mathbb{L}'))$.*

Proof. Suppose that there are $v\langle V\rangle \in \mathbb{M}$ and $\mathbb{L}'$ such that $\mathrm{proj}_{v\langle V\rangle}(\mathbb{L}') \prec v\langle V\rangle$ and $v\langle V\rangle \rightsquigarrow \mathrm{proj}_{v\langle V\rangle}(\mathbb{L}')$ is a non-wasteful implosion. Then by transitivity $\mathbb{M}' = \big(\mathbb{M} \cup \mathrm{proj}_{v\langle V\rangle}(\mathbb{L}')\big) \setminus v\langle V\rangle \prec \mathbb{M}$ is a non-wasteful implosion of $u\langle U\rangle$. This contradicts the minimality of $\mathbb{M}$.

If $v\langle V\rangle \rightsquigarrow \mathrm{proj}_{v\langle V\rangle}(\mathbb{L}')$ is a wasteful implosion, Lemma 6.15 says that there is a non-wasteful implosion to $\mathbb{M}' \succ \mathrm{proj}_{v\langle V\rangle}(\mathbb{L}')$ such that $\mathrm{cost}(\mathbb{M}') \leq \mathrm{cost}(\mathrm{proj}_{v\langle V\rangle}(\mathbb{L}'))$. But we have just proven that this non-wasteful $\mathbb{M}'$ must be identical with $v\langle V\rangle$, so $\mathrm{cost}(v\langle V\rangle) \leq \mathrm{cost}(\mathrm{proj}_{v\langle V\rangle}(\mathbb{L}'))$. $\square$

Very roughly, the next lemma says that wasteful implosions are preserved under mergers.

Lemma 6.18. *Suppose for $i = 1, 2$ that $u_i\langle U_i\rangle \succeq \mathbb{L}_i$ and $\mathrm{cost}(u_i\langle U_i\rangle) \leq \mathrm{cost}(\mathbb{L}_i)$ for $\mathbb{L}_i$ non-overlapping, and that $u_1\langle U_1\rangle$ and $u_2\langle U_2\rangle$ are mutually non-overlapping with $u_2 \in U_1$. Then $\mathrm{cost}(\mathrm{merge}(u_1\langle U_1\rangle, u_2\langle U_2\rangle)) \leq \mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2)$.*

Proof. The L-configurations $\mathbb{L}_1$ and $\mathbb{L}_2$ must be mutually non-overlapping since they are covered by $u_1\langle U_1\rangle$ and $u_2\langle U_2\rangle$, respectively. By part 3 of Proposition 6.5, we have $\mathrm{cost}(\mathbb{L}_1 \cup \mathbb{L}_2) = \big|\mathit{Bl}\big(\mathbb{L}_1\big) \cup \mathit{Wh}\big(\mathbb{L}_1\big) \cup \mathit{Bl}\big(\mathbb{L}_2\big) \cup \mathit{Wh}\big(\mathbb{L}_2\big)\big|$. The only way this could be less than $\mathrm{cost}(\mathrm{merge}(u_1\langle U_1\rangle, u_2\langle U_2\rangle)) = \mathrm{cost}(u_1\langle U_1\rangle) + \mathrm{cost}(u_2\langle U_2\rangle) - 1 \leq \mathrm{cost}(\mathbb{L}_1) + \mathrm{cost}(\mathbb{L}_2) - 1$ is if there are at least two vertices in $\bigcap_{i=1,2}\big(\mathit{Bl}\big(\mathbb{L}_i\big) \cup \mathit{Wh}\big(\mathbb{L}_i\big)\big)$. But $\mathit{Bl}\big(\mathbb{L}_i\big) \cup \mathit{Wh}\big(\mathbb{L}_i\big) \subseteq \mathit{cl}(\mathbb{L}_i) \subseteq \mathit{cl}(u_i\langle U_i\rangle)$ since $\mathbb{L}_i \preceq u_i\langle U_i\rangle$ by the assumptions of the lemma, and also by assumption $\mathit{cl}(u_1\langle U_1\rangle) \cap \mathit{cl}(u_1\langle U_1\rangle) = \{u_2\}$, so this is impossible. $\square$

Combining Lemmas 6.17 and 6.18, we can provide the fifth and final component in the proof of Lemma 5.9, namely that any non-overlapping L-pebbling $\mathcal{L}$ can be transformed into a pebbling $\mathcal{L}'$ without wasteful implosions such that $\mathcal{L}'$ has asymptotically the same cost as $\mathcal{L}$.

Lemma 6.19. *Suppose that $\mathcal{L}$ is a non-overlapping L-pebbling of T. Then there is a non-overlapping pebbling $\mathcal{L}'$ of T without wasteful implosions such that* $\mathrm{cost}(\mathcal{L}') = \mathrm{O}\big(\mathrm{cost}(\mathcal{L})\big)$.

Proof. Given a non-overlapping L-pebbling $\mathcal{L}$, we build a non-overlapping pebbling $\mathcal{L}'$ such that if we let $\mathbb{L}_i \in \mathcal{L}$ denote the starting configuration of the ith move according to the rules in Definition 6.6, there is a corresponding $\mathbb{L}'_i \in \mathcal{L}'$ such that the following invariants hold.

1. $\mathbb{L}'_i$ is non-touching.
2. $\mathbb{L}'_i \succeq \mathbb{L}_i$.
3. For all $u\langle U\rangle \in \mathbb{L}'_i$, it holds that $\mathrm{cost}(u\langle U\rangle) \leq \mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_i))$.
4. The cost of the L-pebbling transition from $\mathbb{L}'_{i-1}$ to $\mathbb{L}'_i$ in $\mathcal{L}'$ does not exceed $4 \cdot \max\big\{\mathrm{cost}(\mathbb{L}_{i-1}), \mathrm{cost}(\mathbb{L}_i)\big\}$.

To see that the lemma follows from this, note that invariants 1 and 2 imply that for every $v\langle V\rangle \in \mathbb{L}_i$ there is a $u\langle U\rangle \in \mathbb{L}'_i$ such that $v\langle V\rangle \preceq u\langle U\rangle$. Then plugging invariant 3 into Proposition 6.5, part 4, we get $\mathrm{cost}(\mathbb{L}'_i) = \sum_{u\langle U\rangle\in\mathbb{L}'_i} \mathrm{cost}(u\langle U\rangle) \leq \sum_{u\langle U\rangle\in\mathbb{L}'_i} \mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_i)) = \mathrm{cost}(\mathbb{L}_i)$. Using invariant 4 to bound the cost of the pebbling transitions $\mathbb{L}'_{i-1} \rightsquigarrow \mathbb{L}'_i$, we get the desired result $\mathrm{cost}(\mathcal{L}') = \mathrm{O}\big(\mathrm{cost}(\mathcal{L})\big)$.

The construction is by forward induction over the moves in $\mathcal{L}$. Assume that the invariants hold for $\mathbb{L}_t$ and $\mathbb{L}'_t$.

Introduction $\mathbb{L}_{t+1} = \mathbb{L}_t \cup v\langle \mathit{pred}(v)\rangle$: If $v\langle \mathit{pred}(v)\rangle \leq \mathbb{L}'_t$ we set $\mathbb{L}'_{t+1} = \mathbb{L}'_t$. For the subconfiguration $u\langle U\rangle \in \mathbb{L}'_t$ such that $v\langle \mathit{pred}(v)\rangle \preceq u\langle U\rangle$, we have $\mathrm{cost}(u\langle U\rangle) \leq \mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_t)) \leq \mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_t \cup v\langle \mathit{pred}(v)\rangle))$, and for $u'\langle U'\rangle \in \mathbb{L}'_t$ distinct from $u\langle U\rangle$ nothing changes. All invariants stay true.

If $v\langle \mathit{pred}(v)\rangle \not\leq \mathbb{L}'_t$, we introduce $v\langle \mathit{pred}(v)\rangle$ and expand to get $\mathbb{L}'_{t+1} = \mathrm{canon}(\mathbb{L}'_t \cup v\langle \mathit{pred}(v)\rangle)$. Invariants 1 and 2 obviously hold. We claim that invariant 3 holds with respect to $\mathbb{L}_{t+1}$ for all L-configurations $\mathbb{L}'$ in the transition $\mathbb{L}'_t \rightsquigarrow \mathbb{L}'_{t+1}$ upto and including $\mathbb{L}'_{t+1} = \mathrm{canon}(\mathbb{L}'_t \cup v\langle \mathit{pred}(v)\rangle)$. This claim yields invariants 3 and 4 for $\mathbb{L}'_{t+1}$.

To prove the claim, observe that invariant 3 holds for $\mathbb{L}'_t \cup v\langle \mathit{pred}(v)\rangle$ with respect to $\mathbb{L}_{t+1} = \mathbb{L}_t \cup v\langle \mathit{pred}(v)\rangle$ by the induction hypothesis and the fact that $\mathrm{proj}_{v\langle \mathit{pred}(v)\rangle}(\mathbb{L}_t \cup v\langle \mathit{pred}(v)\rangle) = v\langle \mathit{pred}(v)\rangle$. Since $\mathbb{L}'_{t+1}$ is obtained by repeated merging of non-overlapping subconfigurations from $\mathbb{L}'_t \cup v\langle \mathit{pred}(v)\rangle$, and since by induction over each such merger these subconfigurations meet the conditions in Lemma 6.18, the claim follows.

Expansion $\mathbb{L}_{t+3} = \big(\mathbb{L}_t \cup \mathrm{merge}(v_1\langle V_1\rangle, v_2\langle V_2\rangle)\big) \setminus \big\{v_1\langle V_1\rangle, v_2\langle V_2\rangle\big\}$: By induction $\mathbb{L}'_t \succeq \mathbb{L}_t \sim \mathbb{L}_{t+3}$, so there is a $u\langle U\rangle \in \mathbb{L}'_t$ such that $v_1\langle V_1\rangle, v_1\langle V_1\rangle \preceq u\langle U\rangle$. For $u'\langle U'\rangle \in \mathbb{L}'_t$ distinct from $u\langle U\rangle$ there are no changes, and if $\mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_{t+3})) \geq \mathrm{cost}(u\langle U\rangle)$ nothing needs to be done and we can set $\mathbb{L}'_{t+3} = \mathbb{L}'_t$.

It can be the case, however, that the expansion within $\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_{t+3})$ decreased the cost so that $u\langle U\rangle$ is now too expensive. If so, we implode

$u\langle U\rangle$ to a minimal non-wasteful L-configuration $\mathbb{M} \succeq \mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_{t+3})$ and set $\mathbb{L}'_{t+3} = \big(\mathbb{L}'_t \setminus u\langle U\rangle\big) \cup \mathbb{M}$.

Invariants 1 and 2 are immediate. Invariant 3 follows from Lemma 6.17 since $\mathbb{M}$ is chosen minimal. Thus, $\mathrm{cost}(\mathbb{M}) \leq \mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_{t+3}))$, and by the induction hypothesis we know that $\mathrm{cost}(u\langle U\rangle) \leq \mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_t))$. Using parts 1 and 2 of Proposition 6.5, we see that the implosion sequence $\mathbb{L}'_t \rightsquigarrow \mathbb{L}'_{t+3}$ causes an extra cost of at most

$$\begin{aligned}\mathrm{cost}(u\langle U\rangle \cup \mathbb{M}) &\leq 2 \cdot (\mathrm{cost}(u\langle U\rangle) + \mathrm{cost}(\mathbb{M})) \\ &\leq 2 \cdot (\mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_t)) + \mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_{t+3}))) \\ &\leq 4 \cdot \max_{i\in\{t,t+3\}} \big\{\mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_i))\big\},\end{aligned}$$

which yields invariant 4.

Implosion $\mathbb{L}_{t+m+1} = \big(\mathbb{L}_t \cup \mathbb{M}\big) \setminus v\langle V\rangle$ for $\mathbb{M} = \big\{v_i\langle V_i\rangle \mid i \in [m]\big\}$: This case is completely analogous to the expansion case. Again $v\langle V\rangle$ is covered by some $u\langle U\rangle \in \mathbb{L}'_t$, and if $\mathrm{cost}(u\langle U\rangle) > \mathrm{cost}(\mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_{t+m+1}))$ we implode $u\langle U\rangle$ to a minimal non-wasteful $\mathbb{M} \succeq \mathrm{proj}_{u\langle U\rangle}(\mathbb{L}_{t+m+1})$ and set $\mathbb{L}'_{t+m+1} = \big(\mathbb{L}'_t \setminus u\langle U\rangle\big) \cup \mathbb{M}$. Using Lemma 6.17 and Proposition 6.5, we get invariants 1-4.

Going through the moves in $\mathcal{L} = \big\{\mathbb{L}_0, \ldots, \mathbb{L}_\tau\big\}$, this construction yields a L-pebbling $\mathcal{L}' = \big\{\mathbb{L}'_0, \ldots, \mathbb{L}'_{\tau'}\big\}$ without wasteful implosions such that $\mathbb{L}'_{\tau'} \succeq \mathbb{L}_\tau$ and $\mathrm{cost}(\mathcal{L}') \leq 4 \cdot \mathrm{cost}(\mathcal{L})$. □

Thereby, the proof of Lemma 5.9 outlined in the beginning of this section is complete. We repeat the proof in condensed form for completeness.

Proof of Lemma 5.9. Let $\mathcal{L}$ be an arbitrary L-pebbling of T. By Observation 6.1, we can assume $\mathcal{L}$ to be non-redundant. Using Lemma 6.9, we get a non-overlapping pebbling $\mathcal{L}'$ with $\mathrm{cost}(\mathcal{L}') \leq \mathrm{cost}(\mathcal{L})$. If $\mathcal{L}'$ contains wasteful implosions, Lemma 6.19 yields a non-wasteful pebbling $\mathcal{L}''$ in $\mathrm{cost}(\mathcal{L}'') = \mathrm{O}\big(\mathrm{cost}(\mathcal{L}')\big)$. Finally, Corollary 6.14 transforms $\mathcal{L}''$ into a reversal-free L-pebbling $\mathcal{L}'''$ of T such that $\mathrm{cost}(\mathcal{L}''') \leq \mathrm{cost}(\mathcal{L}'') = \mathrm{O}\big(\mathrm{cost}(\mathcal{L})\big)$. The lemma follows. □

7 Resolution Derivations Induce Labelled Pebblings

In this section, we shift our focus to resolution and show that clause configurations can be interpreted in terms of labelled pebble configurations in such a way that resolution derivations induce legal L-pebblings. We first give some technical preliminaries. Then we try to explain the intuition for how sets of clauses are translated into sets of pebbles. Finally, we state the formal definitions and prove the correspondence between resolution derivations and L-pebblings.

We start with the technicalities. For simplicity, in the following we will write $v_1, \ldots, v_d$ instead of $x(v)_1, \ldots, x(v)_d$ for the d variables associated with the vertex v in a dth degree pebbling contradiction.

Definition 7.1. Assume that G is a DAG with a unique target z and all vertices having indegree 0 or 2. Then we define ${}^*Peb^d_G = Peb^d_G \setminus \big\{\overline{z}_1, \ldots, \overline{z}_d\big\}$ to be the pebbling contradiction with target axioms removed.

We observe that without loss of generality we may study the formula $*Peb^d_G$ and prove lower bounds for $Sp\big(*Peb^d_G \vdash \bigvee_{l=1}^{d} z_l\big)$ instead of $Sp\big(Peb^d_G \vdash 0\big)$.

Observation 7.2. *For any DAG G with a unique target z and all vertices having indegree 0 or 2, it holds that $Sp\big(Peb^d_G \vdash 0\big) = Sp\big(*Peb^d_G \vdash \bigvee_{l=1}^{d} z_l\big)$.*

Proof. For any resolution derivation $\pi^* : *Peb^d_G \to \bigvee_{l=1}^{d} z_l$, we can get a resolution refutation of Peb^d_G from π^* in the same space by resolving $\bigvee_{l=1}^{d} z_l$ with all $\overline{z}_l$, $l = 1, \ldots, d$, in space 3. In the other direction, for $\pi : Peb^d_G \to 0$ we can extract a derivation of $\bigvee_{l=1}^{d} z_l$ in at most the same space by simply omitting all downloads of and resolution steps on $\overline{z}_l$ in π, leaving the literals z_l in the clauses. Instead of the final empty clause 0 we get some clause $D \subseteq \bigvee_{l=1}^{d} z_l$, and since $*Peb^d_T \nvDash D \subsetneq \bigvee_{l=1}^{d} z_l$ and resolution is sound, we have $D = \bigvee_{l=1}^{d} z_l$. ☐

The following easy lemma will be used repeatedly.

Lemma 7.3. *Suppose that C, D are clauses and $\mathbb{C}, \mathbb{D}$ sets of clauses.*

1. $\mathbb{C} \cup \{C\} \vDash D$ *if and only if* $\mathbb{C} \vDash \overline{a} \vee D$ *for all* $a \in Lit(C)$.
2. $\mathbb{C} \cup \mathbb{D} \vDash D$ *for* $\mathbb{D} = \{D_1, \ldots, D_m\}$ *if and only if* $\mathbb{C} \vDash \bigvee_{i \in [m]} \overline{a}_i \vee D$ *for all choices of literals* $(a_1, \ldots, a_m) \in Lit(D_1) \times \cdots \times Lit(D_m)$.

Proof. For part 1, assume that $\mathbb{C} \cup \{C\} \vDash D$ and consider an assignment α such that $\alpha(\mathbb{C}) = 1$ and $\alpha(D) = 0$ (if there is no such α, then $\mathbb{C} \vDash D \subseteq \overline{a} \vee D$). Such an α sets all $\overline{a}$ to true. Conversely, if $\mathbb{C} \vDash \overline{a} \vee D$ for all $a \in Lit(C)$ and α is such that $\alpha(\mathbb{C}) = \alpha(C) = 1$, it must hold that $\alpha(D) = 1$.

Part 2 follows from part 1 by induction. ☐

We introduce some space-saving notation. If $pred(r) = \{p, q\}$ we say that the *axioms for r* in $*Peb^d_G$ is the set $Ax^d(r) = \{\overline{p}_i \vee \overline{q}_j \vee \bigvee_{l=1}^{d} r_l \mid i, j \in [d]\}$. If r is a source, we define $Ax^d(r) = \{\bigvee_{i=1}^{d} r_i\}$. For V a set of vertices, let $Ax^d(V) = \{Ax^d(v) \mid v \in V\}$.

For v a vertex in T, we let $\mathbb{B}(v) = \bigvee_{i=1}^{d} v_i$. For $V \subseteq V(T)$, we define $\mathbb{B}(V) = \{\mathbb{B}(v) \mid v \in V\}$ and $A_V = \bigvee_{v \in V} \bigvee_{i \in [d]} v_i$. $\mathbb{B}(V)$ can be understood as "truth of all vertices in V" and A_V as "truth of some vertex in V".

This concludes the technical preliminaries. We next try to provide some intuition for how clause configurations are translated into pebble configurations.

Let us associate each vertex $v \in V(T)$ with the clauses $Ax^d(v)$. In a black-white pebbling of T, if at some time t there is an independent black pebble on v, an optimal pebbling will not pebble any vertex in T^v after time t. As an analogy of this, a clause configuration $\mathbb{C}_t$ should induce an independent black pebble on v only if no axioms from $Ax^d(T^v) = *Peb^d_{T^v}$ need be used to derive $\bigvee_{l=1}^{d} z_l$. This holds if and only if

$$\mathbb{C} \cup \big(*Peb^d_T \setminus *Peb^d_{T^v}\big) \vDash \bigvee_{l=1}^{d} z_l \tag{2}$$

by the implicational completeness of resolution. If (2) holds for v but not for $succ(v)$, we can interpret this by saying that the resolution derivation "has reached as far as v but not any farther" and indicate this fact by placing an independent black pebble on v.

It turns out that (2) is equivalent with the condition that $\mathbb{C}$ together with the truth of all vertices unrelated to v should imply truth of some vertex on the path from v to the root, or more concisely

$$\mathbb{C} \cup \mathbb{B}\big(T \setminus (T^v \cup P^v)\big) \models A_{P^v} \,, \tag{3}$$

and the condition (3) is more convenient to work with. In the next lemma, we prove the equivalence of (2) and (3). The lemma is intended only as a way to strengthen the intuition and motivate the formal definitions below. It will not be used in the following and is therefore optional reading.

Lemma 7.4. *Suppose that the clause configuration $\mathbb{C}$ is derived from $^*Peb^d_T$ for a complete binary tree T with root z and let r be an arbitrary vertex in $V(T)$. Then $\mathbb{C} \cup \mathbb{B}\big(T \setminus (T^r \cup P^r)\big) \models A_{P^r}$ if and only if $\mathbb{C} \cup \big({}^*Peb^d_T \setminus {}^*Peb^d_{T^r}\big) \models \bigvee_{l=1}^d z_l$.*

Proof. Note first that if $r = z$, the two implications are exactly the same. Assume therefore that r is not the root and that it has sibling s and successor u.

($\Rightarrow$) Suppose that $\mathbb{C} \cup \mathbb{B}\big(T \setminus (T^r \cup P^r)\big) \models A_{P^r}$. For all $v \in T \setminus \big(T^r \cup P^r\big)$ it holds that ${}^*Peb^d_T \setminus {}^*Peb^d_{T^r} \models \bigvee_{l=1}^d v_l$, since ${}^*Peb^d_{T^v} \subseteq {}^*Peb^d_T \setminus {}^*Peb^d_{T^r}$ and ${}^*Peb^d_{T^v} \models \bigvee_{l=1}^d v_l$, and the fact that resolution is implicationally complete means that these clauses are all derivable. Write $A_{P^r} = A_{P^r_*} \vee \bigvee_{i=1}^d r_i$. Resolve with $\overline{r}_i \vee \overline{s}_j \vee \bigvee_{l=1}^d u_l$ for all $i, j \in [d]$ to get $\big\{\overline{s}_j \vee A_{P^r_*} \mid j \in [d]\big\}$, derive $\bigvee_{j=1}^d s_j$ by implicational completeness and then resolve the clauses $\big\{\overline{s}_j \vee A_{P^r_*} \mid j \in [d]\big\}$ and $\bigvee_{j=1}^d s_j$ to get $A_{P^r_*}$. In the same way we can eliminate all vertices in $P^r \setminus \{z\}$ from $A_{P^r_*}$ and derive $\bigvee_{l=1}^d z_l$ using only axioms from ${}^*Peb^d_T \setminus {}^*Peb^d_{T^r}$. Since resolution is sound this implies that $\mathbb{C} \cup \big({}^*Peb^d_T \setminus {}^*Peb^d_{T^r}\big) \models \bigvee_{l=1}^d z_l$.

($\Leftarrow$) Rewrite the assumption as

$$\mathbb{C} \cup Ax^d\big(T \setminus (T^r \cup P^r)\big) \cup Ax^d\big(P^u_*\big) \cup \Big\{\overline{r}_i \vee \overline{s}_j \vee \textstyle\bigvee_{l=1}^d u_l \mid i, j \in [d]\Big\} \models \bigvee_{l=1}^d z_l.$$

Repeated use of Lemma 7.3 yields

$$\mathbb{C} \cup Ax^d\big(T \setminus (T^r \cup P^r)\big) \cup Ax^d(P^u_*) \models \textstyle\bigvee_{l=1}^d r_l \vee \bigvee_{l=1}^d z_l$$

and proceeding in the same way for all $w \in P^u_*$ we get

$$\mathbb{C} \cup Ax^d\big(T \setminus (T^r \cup P^r)\big) \models A_{P^r}.$$

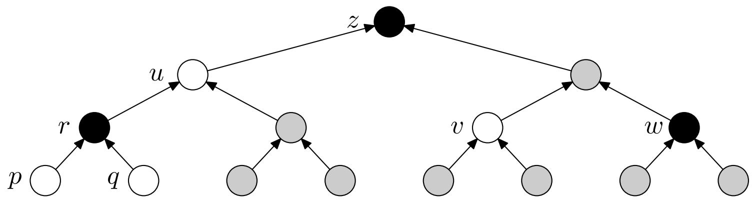
Any α satisfying $\mathbb{B}\big(T \setminus (T^r \cup P^r)\big)$ must satisfy $Ax^d\big(T \setminus (T^r \cup P^r)\big)$ and thus

$$\mathbb{C} \cup \mathbb{B}\big(T \setminus (T^r \cup P^r)\big) \models A_{P^r}.$$

□

Continuing our intuitive argument, the simplest case for a black pebble on a vertex v is when $\mathbb{C} \models \bigvee_{i=1}^d v_i$. Let us restrict our attention to this case and think of a black pebble on v as derived truth $\mathbb{B}(v) = \bigvee_{i=1}^d v_i$ of v. One way of looking at a dependent black pebble on v supported by white pebbles on W, or, in L-pebbling terminology, a subconfiguration $v\langle W\rangle$, is that given independent black pebbles on all $w \in W$ we can eliminate the white pebbles and get $v\langle\emptyset\rangle$. By analogy, a clause configuration $\mathbb{C}$ should induce a subconfiguration $v\langle W\rangle$ if we would get an induced independent black pebble on v by assuming the truth

$$\mathbb{C} = \left\{\overline{u}_i \vee \overline{v}_j \vee \bigvee_{l=1}^{d} z_l,\ \overline{p}_i \vee \overline{q}_j \vee \bigvee_{l=1}^{d} r_l,\ \bigvee_{l=1}^{d} w_l \mid 1 \le i,j \le d\right\}$$



$$\mathbb{L}(\mathbb{C}) = \{z\langle u, v\rangle,\ r\langle p, q\rangle,\ w\langle\emptyset\rangle\}$$

Figure 9: An example clause configuration $\mathbb{C}$ and induced L-configuration $\mathbb{L}(\mathbb{C})$.

of all $w \in W$, i.e., if $\mathbb{C} \cup \mathbb{B}\big(W\big) \models \bigvee_{i=1}^{d} v_i$. Figure 9 (which is Figure 3 but with renamed vertices) gives an example of this intuitive understanding of induced pebble configurations.

Our formal definitions follow the intuition presented above quite closely, modulo a few technical details.

Definition 7.5 (Support). Suppose for $\mathbb{C}$ a set of clauses, $v \in V\big(T\big)$ a vertex and $V \subseteq T \setminus P^v$ a set of vertices that $\mathbb{C} \cup \mathbb{B}\big(V\big) \models A_{P^v}$. Then V is a *support* for v with respect to $\mathbb{C}$, and if there is no $V' \subsetneqq V$ such that $\mathbb{C} \cup \mathbb{B}\big(V'\big) \models A_{P^v}$ the support is *minimal*. If V is a minimal support for v with respect to $\mathbb{C}$ such that $\mathbb{C} \cup \mathbb{B}\big(V\big) \not\models A_{P_*^v}$, we say that v is *maximal* with respect to $\mathbb{C}$ and V.

For V a support of v, we define the *supporting white pebbles* of v to be $swp(v, V) = \big\{w \in V \cap T_*^v \mid P_*^w \cap V = \emptyset\big\}$.

When it is clear from context, we sometimes omit which support or vertex is minimal or maximal with respect to what. Note that $swp(v, V)$ is a simple roof below v over $V \cap T_*^v$.

Definition 7.6 (Induced L-configuration). For $\mathbb{C}$ a set of clauses derived from $*Peb_T^d$, the *induced L-configuration* $\mathbb{L}(\mathbb{C})$ consists of all subconfigurations $v\langle V\rangle$ such that

1. there is a minimal support $V' \subseteq T \setminus P^v$ for v with respect to $\mathbb{C}$,
2. v is maximal with respect to $\mathbb{C}$ and V',
3. $V = swp(v, V')$.

Remark 7.7. The reason we use $V = swp(v, V')$ instead of $V' \cap T_*^v$ is that we need simple sets (Definition 5.3) to define our induced subconfigurations $v\langle V\rangle$, but the supporting sets V' are not necessarily simple. For instance, if we let

$$\mathbb{C}' = \left\{\overline{u}_i \vee \overline{v}_j \vee \overline{q}_l \vee \bigvee_{n=1}^{d} z_n,\ \overline{p}_i \vee \overline{q}_j \vee \bigvee_{n=1}^{d} r_n,\ \bigvee_{n=1}^{d} w_n \mid 1 \le i,j,l \le d\right\}$$

in Figure 9, the root z has the minimal supporting set $V' = \{u, v, q\}$. For technical reasons, it is simpler to ignore all but the topmost vertices in V', so by Definition 7.6 we get $\mathbb{L}(\mathbb{C}') = \mathbb{L}(\mathbb{C})$. Anyway, it seems very plausible that optimal resolution derivations should never result in clause configurations like $\mathbb{C}'$, so we probably do not lose anything by restricting the white pebbles to $V = swp(v, V')$ instead of $V' \cap T^v_*$.

Note also that a black pebble on v is defined in terms of A_{P^v}, not $\bigvee_{i=1}^{d} v_i$. This means that for

$$\mathbb{C}'' = \left\{ \overline{u}_i \vee \overline{v}_j \vee \bigvee_{n=1}^{d} z_n,\, \overline{p}_i \vee \overline{q}_j \vee \bigvee_{n=1}^{d} r_n,\, \bigvee_{n=1}^{d} w_n \vee \bigvee_{n=1}^{d} z_n \mid 1 \leq i, j \leq d \right\}$$

we again get an independent black pebble $w\langle\emptyset\rangle$ and $\mathbb{L}(\mathbb{C}'') = \mathbb{L}(\mathbb{C})$.

Recall that the goal of this section is to show that resolution derivations induce L-pebblings. Suppose that $\pi = \{\mathbb{C}_0, \ldots, \mathbb{C}_\tau\}$ is a resolution derivation of $\bigvee_{l=1}^{d} z_l$ from $*Peb^d_T$. For $\mathbb{C}_0 = \emptyset$ we obviously get $\mathbb{L}(\mathbb{C}_0) = \emptyset$, and it is not hard to see that at the end of the derivation $\mathbb{C}_\tau = \left\{ \bigvee_{n=1}^{d} z_n \right\}$ induces a single independent black pebble $\mathbb{L}(\mathbb{C}_\tau) = \{z\langle\emptyset\rangle\}$ on the root of T. Hence, we are done if we can prove that $\{\mathbb{L}(\mathbb{C}_0), \ldots \mathbb{L}(\mathbb{C}_\tau)\}$ forms the backbone of a legal L-pebbling $\mathcal{L}$, where the transitions $\mathbb{L}(\mathbb{C}_t) \rightsquigarrow \mathbb{L}(\mathbb{C}_{t+1})$ can be accomplished in accordance with the rules of the L-pebble game.

By the L-pebbling rules in Definition 5.7, any subconfiguration $v\langle V\rangle$ may be erased from $\mathcal{L}$ freely at any time. Consequently, we need not worry about subconfigurations $v\langle V\rangle \in \mathbb{L}(\mathbb{C}_t) \setminus \mathbb{L}(\mathbb{C}_{t+1})$ disappearing during the transition from $\mathbb{C}_t$ to $\mathbb{C}_{t+1}$. What we do need to check, though, is that no $v\langle V\rangle$ suddenly appears inexplicably in $\mathbb{L}(\mathbb{C}_{t+1})$ as a result of a resolution derivation step $\mathbb{C}_i \rightsquigarrow \mathbb{C}_{i+1}$, but that we can always derive any $v\langle V\rangle \in \mathbb{L}(\mathbb{C}_{t+1}) \setminus \mathbb{L}(\mathbb{C}_t)$ from $\mathbb{L}(\mathbb{C}_t)$ by the L-pebbling rules.

The rest of this section is devoted to proving this. We first make a pair of observations. The first observation relates subset containment of supporting sets and the order relation between corresponding subconfigurations.

Observation 7.8. *If $u \in P^v$ and $U', V' \subseteq T \setminus P^v$ are such that $U' \cap T^v_* \subseteq V' \cap T^v_*$, then $u\langle swp(u, U')\rangle \succeq v\langle swp(v, V')\rangle$.*

Proof. Using the characterization of $\preceq$ in Observation 5.6, it is sufficient to prove that $v \in T^u$ and $P^v \cap swp(u, U') = \emptyset$ and that $swp(v, V')$ is a simple roof below v over $swp(u, U') \cap T^v$.

The condition $v \in T^u$ is equivalent to $u \in P^v$, and since $U' \subseteq T \setminus P^v$ it clearly holds that $P^v \cap swp(u, U') \subseteq P^v \cap U' = \emptyset$. Both $swp(v, V')$ and $swp(u, U') \cap T^v$ are simple sets below v by assumption. The only nontrivial part is to establish that $swp(v, V')$ is a roof over $swp(u, U') \cap T^v$.

Suppose that $w \in swp(u, U') \cap T^v$. To prove that $swp(v, V')$ is a roof, we need to find a $w' \in P^w \cap swp(v, V')$. Since by assumption $swp(u, U') \cap T^v \subseteq U' \cap T^v_* \subseteq V' \cap T^v_*$, it holds that $w \in V' \cap T^v_*$. If $w \in swp(v, V')$ we are done, so suppose $w \notin swp(v, V')$. The reason that w is missing from $swp(v, V')$ must be that $P^w_* \cap V' \neq \emptyset$, but if we pick $w' \in P^w_* \cap V'$ of maximal height we get $P^{w'}_* \cap V' = \emptyset$ and $w' \in swp(v, V')$. This w' satisfies $w \in P^w \cap swp(v, V')$, which shows that $swp(v, V')$ is a roof over $swp(u, U') \cap T^v$. ☐

The second observation says that if a support V' is not minimal or a vertex v is not maximal with respect to a clause configuration $\mathbb{C}$, then this just means that $\mathbb{C}$ induces something stronger than $v\langle swp(v, V')\rangle$.

Observation 7.9. *If $\mathbb{C} \cup \mathbb{B}(V') \vDash A_{P^v}$ for $V' \subseteq T \setminus P^v$, then there is a subconfiguration $u\langle U\rangle \in \mathbb{L}(\mathbb{C})$ such that $v\langle swp(v, V')\rangle \preceq u\langle U\rangle$.*

Proof. Minimize $U' \subseteq V'$ and then maximize $u \in P^v$ so that $\mathbb{C} \cup \mathbb{B}(U') \models A_{P^u}$. Set $U = swp(u, U')$ and use Observation 7.8. ☐

With the help of these observations we can analyze how new subconfigurations $v\langle V\rangle$ can appear in $\mathbb{L}(\mathbb{C}_{t+1})$ after a resolution derivation step $\mathbb{C}_i \rightsquigarrow \mathbb{C}_{i+1}$.

Observation 7.10 (Inference). *If $\mathbb{C}_{t+1}$ is derived from $\mathbb{C}_t$ by inference, then $\mathbb{L}(\mathbb{C}_{t+1}) = \mathbb{L}(\mathbb{C}_t)$.*

Proof. $\mathbb{C}_t$ and $\mathbb{C}_{t+1}$ have the same logical consequences. ☐

Lemma 7.11 (Erasure). *Suppose that $\mathbb{C}_{t+1}$ is derived from $\mathbb{C}_t$ by erasure. Then for each $v\langle V\rangle \in \mathbb{L}(\mathbb{C}_{t+1})$ there is a $u\langle U\rangle \in \mathbb{L}(\mathbb{C}_t)$ such that $v\langle V\rangle \preceq u\langle U\rangle$.*

Proof. By assumption there is a $V' \subseteq T \setminus P^v$ such that $V = swp(v, V')$ and $\mathbb{C}_{t+1} \cup \mathbb{B}(V') \models A_{P^v}$. Certainly, the same implication holds for $\mathbb{C}_t \supseteq \mathbb{C}_{t+1}$. The lemma follows from Observation 7.9. ☐

In particular, all new subconfigurations resulting from an erasure $\mathbb{C}_t \rightsquigarrow \mathbb{C}_{t+1}$ can be obtained from $\mathbb{L}(\mathbb{C}_t)$ by reversal. One way of interpreting this is that no white pebbles can just disappear at an erasure step except if the black pebble that they support disappear as well. This is exactly the kind of "controlled removal" of white pebbles that the L-pebble game was designed to capture.

Lemma 7.12 (Axiom download). *If $\mathbb{C}_{t+1} = \mathbb{C}_t \cup \{C\}$ for an axiom clause $C \in Ax^d(r)$, then all subconfigurations $v\langle V\rangle \in \mathbb{L}(\mathbb{C}_{t+1}) \setminus \mathbb{L}(\mathbb{C}_t)$ can be obtained from $\mathbb{L}(\mathbb{C}_t) \cup r\langle pred(r)\rangle$ by reversals from $\mathbb{L}(\mathbb{C}_t)$ followed by mergers on $\{r\} \cup pred(r)$.*

Proof. By assumption, there is a minimal $V' \subseteq T \setminus P^v$ with $V = swp(v, V')$ such that $\mathbb{C}_t \cup \{C\} \cup \mathbb{B}(V') \models A_{P^v}$ for $C \in Ax^d(r)$. We will use repeatedly the fact that $\mathbb{B}(r) \models C$.

It is intuitively clear that axioms $C \in Ax^d(r)$ should not yield any interesting new subconfigurations $v\langle V\rangle$ if $r \in T \setminus T^v$, and for $r \in T^v$ we should be able to explain new subconfigurations with the help of $r\langle pred(r)\rangle$. We prove this by a case analysis over r.

$r \in T \setminus (T^v \cup P^v)$**:** We have $\mathbb{C}_t \cup \mathbb{B}(V' \cup \{r\}) \models A_{P^v}$ for $V' \cup \{r\} \subseteq T \setminus P^v$, so Observation 7.9 demonstrates that there is a $u\langle U\rangle \in \mathbb{L}(\mathbb{C}_t)$ such that $v\langle V\rangle = v\langle swp(v, V')\rangle = v\langle swp(v, V' \cup \{r\})\rangle \preceq u\langle U\rangle$.

$r \in P^v_*$**:** Write $C = \overline{p}_i \vee \overline{q}_j \vee \bigvee_{l=1}^{d} r_l$ for $\{p, q\} = pred(r) \neq \emptyset$ and assume without loss of generality that $P^v \cap pred(r) = \{p\}$. Using Lemma 7.3 to move p_i to the right of the implication sign yields $\mathbb{C}_t \cup \mathbb{B}(V') \models A_{P^v} \vee p_i = A_{P^v}$, and since V' is minimal it follows that $v\langle V\rangle \in \mathbb{L}(\mathbb{C}_t)$.

$r = v$**:** Note first that we are prepared to accept the introduction of $r\langle pred(r)\rangle$ without any explanation, so if $\mathbb{C}_t \cup \{C\} \cup \mathbb{B}(V') \models A_{P^r}$ for $pred(r) \subseteq V'$ no further analysis is needed for $r\langle swp(r, V')\rangle = r\langle pred(r)\rangle$. In particular, this is always the case if $pred(r) = \emptyset$, i.e., if r is a source.

Suppose that $v\langle V\rangle = r\langle swp(r, V')\rangle \in \mathbb{L}(\mathbb{C}_{t+1})$ for $V \neq pred(r) = \{p, q\}$, and write $C = \overline{p}_i \vee \overline{q}_j \vee \bigvee_{l=1}^{d} r_l$. We want to derive $r\langle V\rangle$ by the pebbling rules from $\mathbb{L}(\mathbb{C}_r) \cup r\langle pred(r)\rangle$. By symmetry, we get two subcases.

1. $p \in V, q \notin V$: By Definition 7.5, we have $p \in V'$ and $q \notin V'$. Observe that this implies that $V' \subseteq T \setminus P^q$. Also, we can use Lemma 7.3 to move q_j to the right-hand side of the implication sign and get $\mathbb{C}_t \cup \mathbb{B}(V') \models A_{P^r} \vee q_j \subseteq A_{P^r} \vee \bigvee_{j=1}^{d} q_j = A_{P^q}$. Plugging this into Observation 7.9 shows that there is a $w\langle W \rangle \in \mathbb{L}(\mathbb{C}_t)$ such that $q\langle V \setminus \{p\} \rangle = q\langle \mathit{swp}(q, V') \rangle \preceq w\langle W \rangle$. Thus we can derive $q\langle V \setminus \{p\} \rangle$ from $\mathbb{L}(\mathbb{C}_t)$ by reversal and then merge $r\langle \mathit{pred}(r) \rangle = r\langle p, q \rangle$ with $q\langle V \setminus \{p\} \rangle$ to obtain $r\langle (\{p, q\} \cup (V \setminus \{p\})) \setminus \{q\} \rangle = r\langle V \rangle$.
2. $p, q \notin V$: Again by Definition 7.5, we have $p, q \notin V'$. If we use Lemma 7.3 twice we get $\mathbb{C}_t \cup \mathbb{B}(V') \models A_{P^p} \wedge A_{P^q}$, and noting that $V' \subseteq T \setminus (P^p \cup P^p)$ we can apply Observation 7.9 to derive $p\langle V \cap T_*^p \rangle$ and $q\langle V \cap T_*^q \rangle$ from $\mathbb{L}(\mathbb{C}_t)$ by reversals. Merging these subconfigurations with $r\langle p, q \rangle$, we get $r\langle (V \cap T_*^p) \cup (V \cap T_*^q) \rangle = r\langle V \rangle$.

$r \in T_*^v$: By assumption, $\mathbb{C}_t \cup \{C\} \cup \mathbb{B}(V' \cup \{r\}) \models A_{P^v}$, and since $r \in T_*^v$ we have $\mathbb{C}_t \cup \mathbb{B}(V' \cup \{r\}) \models A_{P^v}$ for $V' \cup \{r\} \subseteq T \setminus P^v$. If $P^r \cap V' \neq \emptyset$, it holds that $\mathit{swp}(v, V' \cup \{r\}) = \mathit{swp}(v, V')$ and we can obtain $v\langle V \rangle$ from $\mathbb{L}(\mathbb{C}_t)$ by reversal according to Observation 7.9, so suppose $P^r \cap V' = \emptyset$.

Pick $U' \subseteq V' \cup \{r\}$ minimal and then $u \in P^v$ maximal with respect to U' such that $\mathbb{C}_t \cup \mathbb{B}(U') \models A_{P^u}$. By the minimality of V' we have $r \in U'$, and since $P_*^r \cap U' \subseteq P_*^r \cap V' = \emptyset$ it holds that $r \in \mathit{swp}(u, U')$. Consequently, we cannot use $u\langle U \rangle = u\langle \mathit{swp}(u, U') \rangle \in \mathbb{L}(\mathbb{C}_t)$ to derive $v\langle V \rangle \not\preceq u\langle U \rangle$ by reversal. However, since $U' \subseteq V' \cup \{r\}$, Observation 7.8 tells us that $v\langle (V \cup \{r\}) \setminus T_*^r \rangle = v\langle \mathit{swp}(v, V' \cup \{r\}) \rangle \preceq u\langle U \rangle$ can be derived by reversal from $\mathbb{L}(\mathbb{C}_t)$. If we could also derive $r\langle V \cap T_*^r \rangle$ from $\mathbb{L}(\mathbb{C}_t) \cup r\langle \mathit{pred}(r) \rangle$, a merger would produce the desired subconfiguration $v\langle (((V \cup \{r\}) \setminus T_*^r) \cup (V \cap T_*^r)) \setminus \{r\} \rangle = v\langle V \rangle$.

Hence, we are done if we can derive $r\langle V \cap T_*^r \rangle = r\langle \mathit{swp}(v, V') \cap T_*^r \rangle = r\langle \mathit{swp}(r, V') \rangle$ from $\mathbb{L}(\mathbb{C}_t) \cup r\langle \mathit{pred}(r) \rangle$. But $A_{P^r} \supseteq A_{P^v}$, so by assumption we have $\mathbb{C}_t \cup \{C\} \cup \mathbb{B}(V') \models A_{P^r}$ for $V' \subseteq T \setminus P^r$. This is almost exactly the case $r = v$ above, where we proved that $r\langle \mathit{swp}(r, V') \rangle$ is derivable from $\mathbb{L}(\mathbb{C}_t) \cup r\langle \mathit{pred}(r) \rangle$. The only difference is that now it is not necessarily true that V' is a minimal support and that r is maximal with respect to V'. But these assumptions were not used in the derivation of $r\langle \mathit{swp}(r, V') \rangle$ from $\mathbb{L}(\mathbb{C}_t) \cup r\langle \mathit{pred}(r) \rangle$ anyway, so we can reuse exactly the same proof here to get $r\langle \mathit{swp}(r, V') \rangle$. This concludes the analysis for $r \in T_*^v$.

Studying the pebbling moves in the case analysis above, we see that all subconfigurations $v\langle V \rangle \in \mathbb{L}(\mathbb{C}_{t+1}) \setminus \mathbb{L}(\mathbb{C}_t)$ can be obtained from $\mathbb{L}(\mathbb{C}_t) \cup r\langle \mathit{pred}(r) \rangle$ by a (possibly empty) sequence of reversals from $\mathbb{L}(\mathbb{C}_t)$, followed by a (possibly empty) sequence of mergers on $\{r\} \cup \mathit{pred}(r)$. ☐

Combining the results proven for axiom download, inference and erasure, we can show that a resolution derivation induces a legal L-pebbling.

Theorem 7.13. *Let $\pi = \{\mathbb{C}_0, \ldots, \mathbb{C}_\tau\}$ be a resolution derivation of $\bigvee_{l=1}^{d} z_l$ from $^*Peb_T^d$. Then $\{\mathbb{L}(\mathbb{C}_0), \ldots, \mathbb{L}(\mathbb{C}_\tau)\}$ is the backbone of a legal L-pebbling $\mathcal{L}$ of T such that $\max_{t \in [\tau]} \{\mathrm{cost}(\mathbb{L}(\mathbb{C}_t))\} = \Omega(\mathrm{cost}(\mathcal{L}))$.*

Proof. The fact that $\{\mathbb{L}(\mathbb{C}_0), \ldots, \mathbb{L}(\mathbb{C}_\tau)\}$ essentially is a legal L-pebbling was proven in Observation 7.10, Lemma 7.11 and Lemma 7.12, where it was explicitly indicated how the "holes" in $\mathbb{L}(\mathbb{C}_t) \rightsquigarrow \mathbb{L}(\mathbb{C}_{t+1})$ could be filled in by L-pebbling moves to get a legal pebbling $\mathcal{L}$.

The bound $\max_{t\in[\tau]}\{\mathrm{cost}(\mathbb{L}(\mathbb{C}_t))\} = \Omega(\mathrm{cost}(\mathcal{L}))$ does not follow immediately from this, however. The problem is that a single resolution derivation step $\mathbb{C}_t \rightsquigarrow \mathbb{C}_{t+1}$ may induce several L-pebbling moves to get from $\mathbb{L}(\mathbb{C}_t)$ to $\mathbb{L}(\mathbb{C}_{t+1})$ in $\mathcal{L}$. Therefore, we have to consider the possibility that the maximal pebbling cost in $\mathcal{L}$ is reached in some intermediate L-configuration $\mathbb{L}'$ in between $\mathbb{L}(\mathbb{C}_t)$ and $\mathbb{L}(\mathbb{C}_{t+1})$.

Since inference steps in π do not change the set of induced L-configurations, we get two cases.

1. $\mathbb{C}_t \rightsquigarrow \mathbb{C}_{t+1}$ is an erasure. The moves to get from $\mathbb{L}(\mathbb{C}_t)$ to $\mathbb{L}(\mathbb{C}_{t+1})$ are a series of reversals from $\mathbb{L}(\mathbb{C}_t)$ followed by a series of erasures from $\mathbb{L}(\mathbb{C}_t)$. In view of part 1 of Proposition 6.5, without loss of generality we can let $\mathbb{L}'$ be the L-configuration after all reversals but before all erasures. Then $\mathbb{L}' = \mathbb{L}(\mathbb{C}_t) \cup \mathbb{L}(\mathbb{C}_{t+1})$, and by part 2 of Proposition 6.5, we have $\mathrm{cost}(\mathbb{L}') \leq 2 \cdot (\mathrm{cost}(\mathbb{L}(\mathbb{C}_t)) + \mathrm{cost}(\mathbb{L}(\mathbb{C}_{t+1}))) \leq 4 \cdot \max_{i\in[t,t+1]}\{\mathrm{cost}(\mathbb{L}(\mathbb{C}_i))\}$.

2. $\mathbb{C}_t \rightsquigarrow \mathbb{C}_{t+1}$ is a download of $C \in Ax^d(v)$. In this case the moves to get from $\mathbb{L}(\mathbb{C}_t)$ to $\mathbb{L}(\mathbb{C}_{t+1})$ are a possible introduction of $v\langle pred(v)\rangle$ followed by a series of reversals from $\mathbb{L}(\mathbb{C}_t)$, then a series of mergers on $\{v\} \cup pred(v)$ and finally a series of erasures of subconfigurations not derived in the merger moves. Again by part 1 of Proposition 6.5, we may let $\mathbb{L}'$ be the L-configuration after all mergers but before the erasures.

   All pebbles in $Bl(\mathbb{L}') \cup Wh(\mathbb{L}')$ are present in either $\mathbb{L}(\mathbb{C}_t)$ or $\mathbb{L}(\mathbb{C}_{t+1})$, except possibly for the pebbles on $\{v\} \cup pred(v)$ which may have been introduced and then merged away. Since by construction all subconfigurations resulting from these mergers must be contained in $\mathbb{L}(\mathbb{C}_{t+1})$, the pebbles on $\{v\} \cup pred(v)$ are the only ones that can appear and then disappear during the intermediate pebbling steps. Arguing as in the proof of part 2 of Proposition 6.5, we see that if we remove $\{v\} \cup pred(v)$ from $Bl(\mathbb{L}') \cup Wh(\mathbb{L}')$ the pebbling cost cannot decrease by more than 5.

   Since all pebbles $Bl(\mathbb{L}') \setminus (\{v\} \cup pred(v))$ and $Wh(\mathbb{L}') \setminus (\{v\} \cup pred(v))$ are contained in $Bl(\mathbb{L}(\mathbb{C}_t)) \cup Bl(\mathbb{L}(\mathbb{C}_{t+1}))$ and $Wh(\mathbb{L}(\mathbb{C}_t)) \cup Wh(\mathbb{L}(\mathbb{C}_{t+1}))$, respectively, appealing to part 2 of Proposition 6.5 again we get that $\max_{i\in[t,t+1]}\{\mathrm{cost}(\mathbb{L}(\mathbb{C}_i))\} \geq \frac{1}{4}(\mathrm{cost}(\mathbb{L}') - 5)$.

This establishes that even if the maximal cost in the L-pebbling $\mathcal{L}$ induced by $\pi = \{\mathbb{C}_0, \ldots, \mathbb{C}_\tau\}$ is attained in some intermediate L-configuration $\mathbb{L}' \notin \{\mathbb{L}(\mathbb{C}_t) \mid t \in [\tau]\}$, it still holds that $\max_{t\in[\tau]}\{\mathrm{cost}(\mathbb{L}(\mathbb{C}_t))\} = \Omega(\mathrm{cost}(\mathcal{L}))$. The theorem follows. ☐

8 A Separation of Space from Width in Resolution

We have proven that $Sp(Peb^d_{T_h} \vdash 0) = Sp({*}Peb^d_{T_h} \vdash \bigvee_{l=1}^d z_l)$, and that each resolution derivation $\pi : {*}Peb^d_{T_h} \to \bigvee_{i=1}^d z_i$ induces a legal L-pebbling $\mathcal{L}$ of T_h such that $\max_{\mathbb{C}\in\pi}\{\mathrm{cost}(\mathbb{L}(\mathbb{C}))\} = \Omega(\mathrm{cost}(\mathcal{L}))$. From Sections 5 and 6 we know that $\mathrm{cost}(\mathcal{L}) = \Omega\bigl(BW\text{-}Peb(T)\bigr)$. The final component needed to piece together the proof of our lower bound on the refutation space of pebbling contradictions is to show that the number of pebbles in an induced L-configuration $\mathbb{L}(\mathbb{C})$ and the number of of clauses $\lvert\mathbb{C}\rvert$ are somehow connected.

We cannot expect a proof of this fact to work regardless of the pebbling degree d. The induced L-pebbling in Section 7 makes no assumptions about d, but we know that $Sp\bigl({*Peb}^1_G \vdash z_1\bigr) = Sp\bigl(Peb^1_G \vdash 0\bigr) = \mathrm{O}(1)$. If we look at the resolution refutation π of Peb^1_G in constant space sketched at the end of Section 4, we see that the induced L-pebbling starts by placing white pebbles on $pred(z)$ and a black pebble on z, i.e., introducing $z\langle pred(z)\rangle$, and then pushes the white pebbles downwards by introducing $v\langle pred(v)\rangle$ for all v in reverse topological order and merging until it reaches $z\langle S\rangle$ for S the source vertices of G. Finally, the white pebbles in S are eliminated one by one by introducing $s\langle\emptyset\rangle$ and merging. The reason that Peb^1_G can be refuted in constant space is that one single clause $\bigvee_{v\in V}\overline{v}_1 \vee z_1$ can induce an arbitrary number $|V|$ of white pebbles, or, phrasing it differently, that white pebbles are free for $d = 1$.

Below, we prove lower bounds $\lvert\mathbb{C}\rvert = \Omega\bigl(\sqrt{N}\bigr)$ for N induced independent black pebbles in Theorem 8.4 and for N induced white pebbles in Theorem 8.8. As we just observed, we will need $d \geq 2$ in the bound for the white pebbles, but for the black pebbles of Theorem 8.4 it turns out that the result holds for all $d \in \mathbb{N}^+$. We conclude the section by combining the two theorems to prove the lower bound on refutation clause space $Sp\bigl(Peb^d_{T_h} \vdash 0\bigr)$ in Theorem 1.1 and the separation of space and width in Corollary 1.2.

In the proofs, we will use the following definitions.

Definition 8.1. The vertex v is *represented positively* in the clause configuration $\mathbb{C}$ if $\bigl\{v_1,\ldots,v_d\bigr\} \cap Lit(\mathbb{C}) \neq \emptyset$ and *negatively* if $\bigl\{\overline{v}_1,\ldots,\overline{v}_d\bigr\} \cap Lit(\mathbb{C}) \neq \emptyset$. If v is represented positively but not negatively, we say that v is *purely positive* in $\mathbb{C}$, and v is *purely negative* if it is represented negatively but not positively.

For a clause C we use $V\bigl(C\bigr) = \bigl\{v \mid \{v_1,\ldots,v_d\} \cap Vars(C) \neq \emptyset\bigr\}$ to denote all vertices represented in C. Conversely, we let $Vars(V) = \bigl\{v_1,\ldots,v_d \mid v \in V\bigr\}$ be the set of all variables representing vertices in a vertex set V.

Definition 8.2. For v a vertex in T and α a truth value assignment, v is said to be *true under* α if $\alpha\bigl(\bigvee_{i=1}^d v_i\bigr) = 1$ and *false under* α if $\alpha\bigl(\bigvee_{i=1}^d v_i\bigr) = 0$. If $\alpha\bigl(v_i\bigr) = 1$ for all $i \in [d]$, we say that α makes v *supertrue*. We define

$$\alpha^{v=\nu}\bigl(u_i\bigr) = \begin{cases} \alpha\bigl(u_i\bigr) & \text{if } u \neq v, \\ \nu & \text{if } u = v \end{cases}$$

and say that $\alpha^{v=0}$ *flips* v *to false* and $\alpha^{v=1}$ *flips* v *to supertrue*.

Definition 8.3. A *restriction* ρ is a partial truth value assignment. We represent a restriction as the set of literals $\rho = \{a_1,\ldots,a_m\}$ set to true by ρ. For a clause C, the *ρ-restriction* of C is

$$C|_\rho = \begin{cases} 1 & \text{if } \rho \cap Lit(C) \neq \emptyset, \\ C \setminus \{\overline{a} \mid a \in \rho\} & \text{otherwise,} \end{cases}$$

where 1 denotes the trivially true clause, and the ρ-restriction $\mathbb{C}|_\rho$ of a set of clauses $\mathbb{C}$ is the union of the ρ-restrictions $C|_\rho \neq 1$ for $C \in \mathbb{C}$.

The proof of the lower bound for black pebbles is based on the observation that if $\mathbb{C}$ is sufficiently small and the set of induced pebbles W is sufficiently large, $\mathbb{C}$ contains too many positive literals from W to imply A_{P^w} for all $w \in W$.

Theorem 8.4. *Let $\mathbb{C}$ be a set of clauses and $W \subseteq V(T)$ a set of unrelated vertices. Suppose that $\mathbb{C}$ induces an independent black pebble $w\langle\emptyset\rangle$ on each $w \in W$. Then $|\mathbb{C}| \geq \sqrt{|W|/d}$.*

Proof. Let $S = |\mathbb{C}|$ and $N = |W|$. By Definition 7.6, if $w\langle\emptyset\rangle \in \mathbb{L}(\mathbb{C})$ there is a $V_w \subseteq T \setminus (T^w \cup P^w)$ such that $\mathbb{C} \cup \mathbb{B}(V_w) \models A_{P^w}$ but $\mathbb{C} \cup \mathbb{B}(V_w) \not\models A_{P^w_*}$. In other words, for each $w \in W$ there is a truth value assignment α_w such that $\alpha(\mathbb{C}) = \alpha_w(\mathbb{B}(V_w)) = 1$ but $\alpha_w(A_{P^w_*}) = 0$. Since $\mathbb{C} \cup \mathbb{B}(V_w) \models A_{P^w}$, this implies that $\alpha_w\left(\bigvee_{i=1}^{d} w_i\right) = 1$. Flipping w to false falsifies A_{P^w}, so it must hold that $\alpha_w^{w=0}(\mathbb{C} \cup \mathbb{B}(V_w)) = 0$, and $\mathit{Vars}(w) \cap \mathit{Vars}(V_w) = \emptyset$ shows that the falsified clause is in $\mathbb{C}$. It follows that all $w \in W$ are represented positively in $\mathbb{C}$.

Fix an arbitrary $w \in W$. We want the truth value assignment α_w for w described above to make as many vertices $w' \in W \setminus \{w\}$ as possible supertrue without falsifying $\mathbb{C}$. Obviously, $\mathbb{B}(V_w)$ is not falsified by flipping vertices to supertrue, and since the vertices in W are unrelated $A_{P^w_*}$ is not satisfied. Let $\alpha_0 = \alpha_w$. For all $w^1, \ldots, w^{N-1} \in W \setminus \{w\}$ in some arbitrary fixed order, if $\alpha_{j-1}^{w^j=1}(\mathbb{C}) = 1$, set $\alpha_j = \alpha_{j-1}^{w^j=1}$, otherwise set $\alpha_j = \alpha_{j-1}$. Every time we cannot flip a vertex w^j to supertrue there is some clause $C_j \in \mathbb{C}$ falsified by this flip, and since all variables w_i^j are left untouched in this case and fix all such C_j to true, these clauses cannot block the flipping of w^l to supertrue for $l \neq j$. Hence, α_{N-1} makes all except at most S vertices in $W \setminus \{w\}$ supertrue.

It follows that we can assume without loss of generality that for each $w \in W$, α_w is such that $\alpha(\mathbb{C}) = \alpha(\mathbb{B}(V_w)) = 1$, $\alpha(A_{P^w_*}) = 0$ and all but S vertices in $W \setminus \{w\}$ are supertrue under α_w.

Let us say that a clause $C \in \mathbb{C}$ is *positively W-sparse* if it holds that $\left|\{w_i \mid w \in W, i \in d\} \cap \mathit{Lit}(C)\right| \leq d(S+1)$, i.e., if there are at most $d(S+1)$ positive literals from vertices $w \in W$ in $\mathbb{C}$, and *positively W-dense* otherwise. Observe that there can be at most $dS(S+1)$ vertices in W represented positively in positively W-sparse clauses in $\mathbb{C}$.

Suppose that $N > dS(S+1)$. We proved above that all vertices in W are represented positively in $\mathbb{C}$, so there must exist a $w \in W$ that only occurs in positively W-dense clauses. Fix such a vertex w and consider the truth value assignment $\alpha_w^{w=0}$. By construction, it holds that $\alpha_w^{w=0}(A_{P^w}) = 0$ and $\alpha_w^{w=0}(\mathbb{B}(V_w)) = 1$, and for all $C \in \mathbb{C}$ where w is not purely positive we have $\alpha_w^{w=0}(C) = 1$. Consider $C \in \mathbb{C}$ such that w is purely positive in C. Then $\mathit{Lit}(C)$ contains at most d literals from w and at most dS positive literals from vertices $w' \in W \setminus \{w\}$ not supertrue under α_w. Since C is W-dense it must also contain a positive literal from some other vertex in W. By construction this vertex is supertrue, so such a literal fixes C to true. It follows that $\alpha_w^{w=0}(\mathbb{C}) = 1$. But this yields $\alpha_w^{w=0}(\mathbb{C} \cup \mathbb{B}(V_w)) = 1$ and $\alpha_w^{w=0}(A_{P^w}) = 0$, which contradicts the assumption that $\mathbb{C} \cup \mathbb{B}(V_w) \models A_{P^w}$. Thus, $N \leq dS(S+1) < d(S+1)^2$. $\square$

The idea behind the proof of the lower bound for white pebbles is similar, but here we want to flip the pebbled vertices $w \in \mathit{Wh}(v\langle V\rangle) = \mathit{Wh}(v\langle \mathit{swp}(v, V_w)\rangle)$ to false instead of supertrue. This is complicated by the fact that such flips may falsify the clauses $\mathbb{B}\big(V_w\big)$ of the support. To get around this problem, we need to bound the size $\big|V_w\big|$ of the supporting set in terms of $|\mathbb{C}|$.

The following result seems to be part of mathematical folklore.

Lemma 8.5. *A minimally unsatisfiable CNF formula F must have more clauses than variables.*

Proof. Study the bipartite graph on $F \times \mathit{Vars}(F)$. Since F is unsatisfiable there is no matching, so by Hall's theorem there is a $G \subseteq F$ such that $|G| > |N(G)|$. Pick G of maximal size and suppose $G \neq F$. Then G is satisfiable. Using Hall's theorem again, there is a matching between $F \setminus G$ and $\mathit{Vars}(F) \setminus N(G)$, so $F \setminus G$ and G are simultaneously satisfiable. Contradiction. $\square$

We can use Lemma 8.5 to get an upper bound on the size of a minimal supporting set, provided that the pebbling degree d is strictly greater than 1.

Lemma 8.6. *Suppose for a clause set $\mathbb{C}$ and a vertex v that there is a $V \subseteq T \setminus P^v$ such that $\mathbb{C} \cup \mathbb{B}\big(V\big) \vDash A_{P^v}$ but for all $V' \subsetneqq V$ it holds that $\mathbb{C} \cup \mathbb{B}\big(V'\big) \nvDash A_{P^v}$. Then $|\mathbb{C}| > (d-1)|V|$.*

Proof. For any restriction ρ, it clearly holds that $\mathbb{D}|_\rho \vDash D|_\rho$ if $\mathbb{D} \vDash D$. Define $\rho = \big\{\overline{u}_i \mid u \in P^v,\, i \in [d]\big\}$. Obviously, $A_{P^v}|_\rho = 0$, and since $V \subseteq T \setminus P^v$ we have $\mathbb{B}\big(V\big)|_\rho = \mathbb{B}\big(V\big)$, so $\mathbb{C}|_\rho \cup \mathbb{B}\big(V\big) \vDash 0$. By the minimality of V, for all $V' \subsetneqq V$ there is an α such that $\alpha\big(A_{P^v}\big) = 0$ but $\alpha\big(\mathbb{C} \cup \mathbb{B}\big(V'\big)\big) = 1$, and the fact that α and ρ coincide on $\mathit{Vars}(P^v)$ demonstrates that $\mathbb{C}|_\rho \cup \mathbb{B}\big(V'\big) \nvDash 0$. That is, $\mathbb{B}\big(V\big)$ must be contained in any minimally unsatisfiable subset of $\mathbb{C}|_\rho \cup \mathbb{B}\big(V\big)$. The set $\mathbb{B}\big(V\big)$ contains $|V|$ clauses and $d|V|$ variables, so by Lemma 8.5 we must have $\big|\mathbb{C}\big| \geq \big|\mathbb{C}|_\rho\big| > (d-1)|V|$. $\square$

For convenience, we also prove in a separate lemma that if $\mathbb{C}$ induces a white pebble on w, all literals $\overline{w}_i$, $i \in [d]$, are represented in $\mathit{Lit}(\mathbb{C})$.

Lemma 8.7. *Suppose for a clause set $\mathbb{C}$ and a vertex w that there is a $v \in P^w_*$ and a $V \subseteq T \setminus P^w_*$ such that $\mathbb{C} \cup \mathbb{B}\big(V\big) \vDash A_{P^v}$ but $\mathbb{C} \cup \mathbb{B}\big(V \setminus \{w\}\big) \nvDash A_{P^v}$. Then there is a subset $\big\{\overline{w}_i \vee C_i \mid i \in [d]\big\} \subseteq \mathbb{C}$ for which $\overline{w}_j \notin \mathit{Lit}(C_i)$ if $j \neq i$.*

Proof. Pick α such that $\alpha\big(\mathbb{C}\big) = \alpha\big(\mathbb{B}\big(V \setminus \{w\}\big)\big) = 1$ but $\alpha\big(A_{P^v}\big) = 0$. Then it must be the case that $\alpha\big(\bigvee_{i=1}^{d} w_i\big) = 0$. For all $i \in [d]$ we have $\alpha^{w_i=1}\big(\mathbb{B}\big(V\big)\big) = 1$ but $\alpha^{w_i=1}\big(A_{P^v}\big) = 0$, so flipping w_i while keeping w_j false for $j \neq i$ must falsify some clause in $\mathbb{C}$. This establishes that there are clauses $\overline{w}_i \vee C_i \in \mathbb{C}$ for all $i \in [d]$ such that $\overline{w}_j \notin \mathit{Lit}(C_i)$ for $j \neq i$. $\square$

With the help of Lemmas 8.6 and 8.7 we can establish a lower bound on $|\mathbb{C}|$ in terms of the number of induced white pebbles.

Theorem 8.8. *Suppose that $\mathbb{C}$ is a clause set and $W \subseteq V\big(T\big)$ an arbitrary set of vertices such that $\mathbb{C}$ induces white pebbles on all $w \in W$, i.e., $\mathit{Wh}\big(\mathbb{L}(\mathbb{C})\big) \supseteq W$. Then $\big|\mathbb{C}\big| \geq \sqrt{(d-1)|W|}$.*

Proof. Let $S = |\mathbb{C}|$ and $N = |W|$. If $\mathbb{C}$ induces a white pebble on w, by Definition 7.6 there is a $v_w\big\langle W_w\big\rangle = v_w\big\langle \mathit{swp}(v_w, V_w)\big\rangle \in \mathbb{L}(\mathbb{C})$ for which it holds that $v_w \in P^w_*$ and $w \in W_w \subseteq V_w$, where $V_w \subseteq T \setminus P^v$ is a supporting set such that $\mathbb{C} \cup \mathbb{B}(V_w) \models A_{P^{v_w}}$ but $\mathbb{C} \cup \mathbb{B}(V') \not\models A_{P^{v_w}}$ for all $V' \subsetneq V_w$. By Lemma 8.6 it holds that $|V_w| < \frac{S}{d-1}$.

Fix an arbitrary w, and let us write $D_w = A_{P^{v_w}}$ for brevity. By assumption, there exists an α_w such that $\alpha_w(\mathbb{C}) = \alpha_w(\mathbb{B}(V_w \setminus \{w\})) = 1$ but $\alpha_w(D_w) = 0$. Note that $\alpha_w(\bigvee_{i=1}^{d} w_i) = 0$. We want α_w to falsify as many as possible of the variables $w'_i \in Vars(W \setminus \{w\})$. For every $w' \in V_w$, the clauses $\mathbb{B}(V_w \setminus \{w\})$ force exactly one variable w'_i to true for $i \in [d]$ arbitrary, which gives a total of at most $\frac{S}{d-1} - 2$ variables, and trivially at most S more positive literals w'_i are needed in order to satisfy $\mathbb{C}$ (less if we can pick negative literals or variables from vertices $v \notin W$). All other variables in $Vars(W \setminus \{w\})$ can be flipped to false without falsifying $\mathbb{C} \cup \mathbb{B}(V_w)$ or satisfying D_w. Consequently, without loss of generality we may assume that α_w sets all but $\frac{dS}{d-1} - 2$ of the variables $w'_i \in Vars(W \setminus \{w\})$ to false.

Suppose that $N > \frac{S^2}{d-1}$. Let us say that a clause $C \in \mathbb{C}$ is *negatively W-sparse* if it holds that $\left|\{\overline{w}_i \mid w \in W, i \in [d]\} \cap Lit(C)\right| \leq \frac{d}{d-1}S$ and *negatively W-dense* otherwise. There are dN distinct negative literals $\overline{w}_i$ for $w \in W$, $i \in [d]$, and we know by Lemma 8.7 that all of them are present in $\mathbb{C}$. At most $\frac{d}{d-1}S^2 < dN$ literals can occur in negatively W-sparse clauses in $\mathbb{C}$, so there is some w with a literal $\overline{w}_i$ that occurs only in negatively W-dense clauses.

Fix such a literal w_i and consider the truth value assignment $\alpha_w^{w_i=1}$. We have $\alpha_w^{w_i=1}(D_w) = 0$ and $\alpha_w^{w_i=1}(\mathbb{B}(V_w)) = 1$, and the only clauses $C \in \mathbb{C}$ that can turn false are those where $\overline{w}_i \in Lit(C)$ and $\overline{w}_j \notin Lit(C)$ for $j \neq i$, since $\alpha_w^{w_i=1}(w_j) = 0$. Such a clause C is W-dense by assumption, and must therefore contain at least $\frac{dS}{d-1}$ other negative literals from vertices $w' \in W \setminus \{w\}$. At most $\frac{dS}{d-1} - 2$ of the the variables w'_j are true under $\alpha_w^{w_i=1}$, so we can find some satisfied literal $\overline{w}'_j \in Lit(C)$. It follows that C is true under $\alpha_w^{w_i=1}$, which implies that $\alpha_w^{w_i=1}(\mathbb{C}) = 1$. But this is impossible since $\mathbb{C} \cup \mathbb{B}(V_w) \models D_w$. Thus, $S \geq \sqrt{(d-1)N}$. □

At last, we can now prove the main theorem of this paper.

Theorem 1.1 (restated). *Let T_h denote the complete binary tree of height h and $Peb^d_{T_h}$ the pebbling contradiction of degree $d \geq 2$ defined on T_h. Then the space of refuting $Peb^d_{T_h}$ by resolution is bounded by $Sp(Peb^d_{T_h} \vdash 0) = \Omega(\sqrt{h})$.*

Proof. According to Observation 7.2, $Sp(Peb^d_G \vdash 0) = Sp({}^*Peb^d_G \vdash \bigvee_{i=1}^{d} z_i)$. Let $\pi = \{\mathbb{C}_0, \ldots, \mathbb{C}_\tau\}$ be a resolution derivation of $\bigvee_{i=1}^{d} z_i$ from ${}^*Peb^d_{T_h}$ in minimal clause space.

Combining Theorems 3.3, 5.12 and 7.13, we know that the derivation π induces a legal L-pebbling $\mathcal{L}$ of the tree T_h such that there is a clause configuration $\mathbb{C}_t \in \pi$ with $\mathrm{cost}(\mathbb{L}(\mathbb{C}_t)) = \Omega(\mathrm{cost}(\mathcal{L})) = \Omega(BW\text{-}Peb(T_h)) = \Omega(h)$. Fix such an induced L-configuration $\mathbb{L}_t = \mathbb{L}(\mathbb{C}_t)$ with $\mathrm{cost}(\mathbb{L}_t) = K = \Omega(h)$.

If $|Wh(\mathbb{L}_t)| \geq \frac{1}{3}K$, we have $|\mathbb{C}_t| = \Omega(\sqrt{K})$ by Theorem 8.8. Suppose that $|Wh(\mathbb{L}_t)| < \frac{1}{3}K$. Then there is an admissible choice of $B \subseteq Bl(\mathbb{L}_t)$ in the sense of Definition 5.8 such that $|B| > \frac{2}{3}K$.

Look at all closest related pairs of vertices $u, v \in B$, i.e., such that $u \in P^v_*$ but $(P^v_* \setminus P^u) \cap B = \emptyset$. For each such pair there must exist a white pebble $w_u \in Wh(\mathbb{L}_t) \cap (P^v \setminus P^u)$, since both u and v are admissible, and we can pick an arbitrary such w_u and associate it with u. Note that $w_u \neq w_{u'}$ if $u \neq u'$ since we consider closest pairs u, v. Remove all upper vertices u of such pairs from B to get $B' = \{v \in B \mid T_*^v \cap B = \emptyset\}$.

All vertices in B' are unrelated, and for each dependent black pebble $v \in B'$ there is a subconfiguration $v\langle V\rangle$ with $\emptyset \neq V \subseteq \mathit{Wh}(\mathbb{L}_t) \cap T_*^v$. Pick a white pebble $w_v \in V$ and associate it with v. For distinct dependent black pebbles $v_1, v_2 \in B'$, these white pebbles are clearly distinct since $T_*^{v_1} \cap T_*^{v_2} = \emptyset$, and they are also distinct from all w_u associated to $u \in B \setminus B'$ since $w_u \notin \bigcup_{v \in B'} T^v$. Remove all dependent black pebbles from B' to get $B'' = \{v \in B' \mid v\langle\emptyset\rangle \in \mathbb{L}_t\}$.

For each $u \in B \setminus B''$, we have identified a distinct associated $w_u \in \mathit{Wh}(\mathbb{L}_t)$. Consequently, we cannot have removed more than a total of $\frac{1}{3}K$ black pebbles, so $|B''| > \frac{1}{3}K$ and by appealing to Theorem 8.4 we get $|\mathbb{C}_t| = \Omega(\sqrt{K}) = \Omega(\sqrt{h})$.

It follows that $\mathit{Sp}(\mathit{Peb}_G^d \vdash 0) = \mathit{Sp}(\pi) \geq |\mathbb{C}_t| = \Omega(\sqrt{h})$. ☐

Since $W(\mathit{Peb}_G^d \vdash 0) = \mathrm{O}(1)$ for all pebbling contradictions by Theorem 4.2, this yields a separation of clause space from width. To get the corollary, let $F_n = \mathit{Peb}_{T_h}^2$ for $h = \lfloor \log n \rfloor$.

Corollary 1.2 (restated). *There is a family $\{F_n\}_{n=1}^{\infty}$ of k-CNF formulas of size $\mathrm{O}(n)$ such that $W(F_n \vdash 0) = \mathrm{O}(1)$ but $\mathit{Sp}(F_n \vdash 0) = \Omega(\sqrt{\log n})$.*

9 Conclusion and Open Problems

We have proven a non-constant lower bound on the refutation clause space of pebbling contradictions in resolution. Our result is the first lower bound on refutation space which is not the consequence of a lower bound on the refutation width for the same formulas, but instead separates the two measures.

This answers an open question in [7, 21, 23]. However, we believe that our answer can be strengthened in several ways.

Firstly, we conjecture that the lower bounds for $|\mathbb{C}|$ in terms of the number of induced pebbles N in Theorems 8.4 and 8.8 can be improved from $|\mathbb{C}| = \Omega(\sqrt{N})$ to $|\mathbb{C}| = \Omega(N)$. This would yield a tight linear bound measured in the tree height for the refutation clause space of pebbling contradictions over complete binary trees, which we strongly suspect to be the correct result.

Conjecture 9.1. *If $d \geq 2$, $\mathit{Sp}(\mathit{Peb}_{T_h}^d \vdash 0) = \Theta(\mathit{BW\text{-}Peb}(T_h)) = \Theta(h)$.*

Note that the proofs in Section 8 do not use in any way that $\mathbb{C}$ is derived from $^*\mathit{Peb}_T^d$ and that we therefore know quite a lot about the structure of the clauses in $\mathbb{C}$. For instance, it is not hard to show the following lemma, which we state without proof.

Lemma 9.2. *Suppose that D is a clause derived from $^*\mathit{Peb}_T^d$ by resolution. Then $D = D' \vee \bigvee_{i=1}^{d} s_i$ for some $s \in V(T)$ such that $V(D') \subseteq T_*^s$.*

Using this kind of information, we think that there should be room for improvement of the bounds in Section 8.

Secondly, we would like to generalize the lower bound on the refutation space of pebbling contradictions to the k-DNF resolution proof systems $\mathfrak{Res}(k)$ introduced in [25], where the clause configurations $\mathbb{C}$ consist of k-DNF formulas.

We believe that pebbling contradictions $Peb^{k+1}_{T_h}$ separate k-DNF resolution and $(k+1)$-DNF resolution with respect to space.

Conjecture 9.3. *For k-DNF resolution refutations of pebbling contradictions on complete binary trees, fixing k it holds that $W_{\mathfrak{Res}(k)}\big(Peb^{k+1}_{T_h} \vdash 0\big) = \mathrm{O}(1)$ and $Sp_{\mathfrak{Res}(k)}\big(Peb^{k+1}_{T_h} \vdash 0\big) = \Omega(h)$ but $Sp_{\mathfrak{Res}(k+1)}\big(Peb^{k+1}_{T_h} \vdash 0\big) = \mathrm{O}(1)$.*

Proving this conjecture would establish that the k-DNF resolution proof systems form a strict hierarchy with respect to clause space and improve the separation result in [21] for the restricted case of tree-like k-DNF resolution.

Thirdly, it would be nice to extend the bound on refutation space of pebbling contradictions to DAGs other than trees that have better size-pebbling price trade-off. For $L = L(F)$ the number of clauses in a formula F, we want to find a formula family which improves the the bound $Sp\big(F \vdash 0\big) = \Omega\big(\sqrt{\log L}\big)$ in Theorem 1.1 to $Sp\big(F \vdash 0\big) = \Omega\big(L\big)$ or at least $Sp\big(F \vdash 0\big) = \Omega\big(L^{\epsilon}\big)$ for some constant $\epsilon > 0$, but for which it still holds that $W\big(F \vdash 0\big) = \mathrm{O}(1)$.

Our guess is that the black-white pebbling price is a lower bound for pebbling contradictions over any DAG.

Conjecture 9.4. *For $d \geq 2$ and for G an arbitrary DAG with a unique target and with all vertices having indegree 0 or 2, $Sp\big(Peb^{d}_{G} \vdash 0\big) = \Omega\big(BW\text{-}Peb(G)\big)$.*

Since there are DAGs G_n of constant fan-in and size $\mathrm{O}(n)$ which have black-white pebbling price $BW\text{-}Peb(G_n) = \Theta\big(n/\log n\big)$ [26], a proof of Conjecture 9.4 would immediately yield the following corollary.

Corollary 9.5 (assuming Conjecture 9.4). *There is a family of unsatisfiable k-CNF formulas $\big\{F_n\big\}$ of size $\mathrm{O}(n)$ such that $W\big(F_n \vdash 0\big) = \mathrm{O}(1)$ but $Sp\big(F_n \vdash 0\big) = \Omega\big(n/\log n\big)$.*

A final question is whether refutation space can be separated from refutation length in the sense that there can be shown to exist a polynomial-size family of k-CNF formulas such that $Sp\big(F \vdash 0\big) = \omega\big(\sqrt{n \log L(F \vdash 0)}\big)$, where n is the number of variables in F. This would be an interesting contrast to the relation $W\big(F \vdash 0\big) = \mathrm{O}\big(\sqrt{n \log L(F \vdash 0)}\big)$ between length and width proven in [10]. We believe that such a formula family exists.

Conjecture 9.6. *There is a family of k-CNF formulas $\big\{F_n\big\}$ over n variables such that $Sp\big(F \vdash 0\big) = \omega\big(\sqrt{n \log L(F \vdash 0)}\big)$.*

Of course, if we could prove Conjecture 9.4, we would immediately get a positive answer to Conjecture 9.6 as well, using the same formula family as in Corollary 9.5.

Acknowledgements

I am grateful to my supervisor, Professor Johan Håstad, for giving feed-back to my ideas (sometimes supporting, sometimes criticizing them), and for providing a couple of ingenious ones of his own. Also, I would like to thank Douglas Wikström for interesting discussions about theoretical computer science in general and pebble games in particular. Finally, a special thanks to Joel Brynielsson, who introduced me to the mysteries of METAPOST.

References

[1] Michael Alekhnovich, Eli Ben-Sasson, Alexander A. Razborov, and Avi Wigderson. Space complexity in propositional calculus. *SIAM Journal on Computing*, 31(4):1184–1211, 2002.

[2] Michael Alekhnovich, Jan Johannsen, Toniann Pitassi, and Alasdair Urquhart. An exponential separation between regular and general resolution. In *Proceedings 34th Annual ACM Symposium on Theory of Computing (STOC '02)*, pages 448–456, May 2002.

[3] Albert Atserias and Victor Dalmau. A combinatorical characterization of resolution width. In *Proceedings 18th IEEE Annual Conference on Conference on Computational Complexity (CCC '03)*, pages 239–247, July 2003.

[4] Paul Beame. *Proof Complexity*. Lecture notes, Institute for Advanced Study, 2000.

[5] Paul Beame, Richard Karp, Toniann Pitassi, and Michael Saks. The efficiency of resolution and Davis-Putnam procedures. *SIAM Journal on Computing*, 31(4):1048–1075, 2002.

[6] Paul Beame and Toniann Pitassi. Propositional proof complexity: Past, present, and future. *Bulletin of the European Association for Theoretical Computer Science*, 65:66–89, June 1998.

[7] Eli Ben-Sasson. Size space tradeoffs for resolution. In *Proceedings 34th Annual ACM Symposium on Theory of Computing (STOC '02)*, pages 457–464, May 2002.

[8] Eli Ben-Sasson and Nicola Galesi. Space complexity of random formulae in resolution. *Random Structures and Algorithms*, 23(1):92–109, August 2003.

[9] Eli Ben-Sasson, Russell Impagliazzo, and Avi Wigderson. Near optimal separation of treelike and general resolution. *Combinatorica*, 24(4):585–603, September 2004.

[10] Eli Ben-Sasson and Avi Wigderson. Short proofs are narrow—resolution made simple. *Journal of the ACM*, 48(2):149–169, March 2001.

[11] A. Blake. *Canonical Expressions in Boolean Algebra*. PhD thesis, University of Chicago, 1937.

[12] Maria Luisa Bonet, Juan Luis Esteban, Nicola Galesi, and Jan Johannsen. On the relative complexity of resolution refinements and cutting planes proof systems. *SIAM Journal on Computing*, 30(5):1462–1484, 2000.

[13] Maria Luisa Bonet and Nicola Galesi. A study of proof search algorithms for resolution and polynomial calculus. In *Proceedings 40th Annual IEEE Symposium on Foundations of Computer Science (FOCS '99)*, October 1999.

[14] Josh Buresh-Oppenheim and Toniann Pitassi. The complexity of resolution refinements. In *Proceedings 18th IEEE Symposium on Logic in Computer Science (LICS 03)*, pages 138–147, June 2003.

[15] Samuel R. Buss, editor. *Handbook of Proof Theory.* Elsevier Science, Amsterdam, 1998.

[16] Vašek Chvátal and Endre Szemerédi. Many hard examples for resolution. *Journal of the ACM*, 35(4):759–768, October 1988.

[17] Stephen A. Cook and Robert Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44(1):36–50, March 1979.

[18] Stephen A. Cook and Ravi Sethi. Storage requirements for deterministic polynomial time recognizable languages. *Journal of Computer and System Sciences*, 13(1):25–37, 1976.

[19] Martin Davis, George Logemann, and Donald Loveland. A machine program for theorem proving. *Communications of the ACM*, 5(7):394–397, July 1962.

[20] Martin Davis and Hilary Putnam. A computing procedure for quantification theory. *Journal of the ACM*, 7(3):201–215, 1960.

[21] Juan Luis Esteban, Nicola Galesi, and Jochen Messner. On the complexity of resolution with bounded conjunctions. *Theoretical Computer Science*, 321(2-3):347–370, August 2004.

[22] Juan Luis Esteban and Jacobo Torán. Space bounds for resolution. *Information and Computation*, 171(1):84–97, 2001.

[23] Juan Luis Esteban and Jacobo Torán. A combinatorial characterization of treelike resolution space. *Information Processing Letters*, 87(6):295–300, 2003.

[24] Armin Haken. The intractability of resolution. *Theoretical Computer Science*, 39(2-3):297–308, August 1985.

[25] Jan Krajíček. On the weak pigeonhole principle. *Fundamenta Mathematicae*, 170(1-3):123–140, 2001.

[26] Thomas Lengauer and Robert Endre Tarjan. Upper and lower bounds on time-space tradeoffs. In *Proceedings 11th Annual ACM Symposium on Theory of Computing (STOC '79)*, pages 262–277, May 1979.

[27] Thomas Lengauer and Robert Endre Tarjan. The space complexity of pebble games on trees. *Information Processing Letters*, 10(4/5):184–188, July 1980.

[28] Nicholas Pippenger. *Pebbling.* Technical Report RC8258, IBM Watson Research Center, 1980. Appeared in Proceedings of the 5th IBM Symposium on Mathematical Foundations of Computer Science, Japan.

[29] Ran Raz. Resolution lower bounds for the weak pigeonhole principle. *Journal of the ACM*, 51(2):115–138, 2004.

[30] John Alan Robinson. A machine-oriented logic based on the resolution principle. *Journal of the ACM*, 12(1):23–41, January 1965.

[31] Jacobo Torán. Lower bounds for space in resolution. In *Proceedings 13th International Workshop Computer Science Logic (CSL '99)*, volume 1683 of *Lecture Notes in Computer Science*, pages 362–373. Springer-Verlag, 1999.

[32] Grigori Tseitin. On the complexity of derivation in propositional calculus. In A. O. Silenko, editor, *Structures in Constructive Mathematics and Mathematical Logic, Part II*, pages 115–125. Consultants Bureau, New York-London, 1968.

[33] Alasdair Urquhart. Hard examples for resolution. *Journal of the ACM*, 34(1):209–219, January 1987.

[34] Alasdair Urquhart. The complexity of propositional proofs. *Bulletin of Symbolic Logic*, 1(4):425–467, 1995.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'