On SGGS and Horn Clauses*

Maria Paola Bonacina[1,*,†], Sarah Winkler[2,†]

[1]Università degli Studi di Verona, Strada Le Grazie 15, 37134 Verona, Italy
[2]Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100 Bolzano, Italy

Abstract

SGGS (*Semantically-Guided Goal-Sensitive* reasoning) is a refutationally complete theorem-proving method that offers *first-order conflict-driven reasoning* and is *model complete in the limit*. This paper investigates the behavior of SGGS on *Horn clauses*, which are widely used in declarative programming, knowledge representation, and verification. We show that SGGS generates the *least Herbrand model* of a set of definite clauses, and that *SGGS terminates on Horn clauses if and only if hyperresolution does*, with the advantage that SGGS builds a model. We report on experiments applying the SGGS prototype prover Koala to Horn problems, with promising performances especially on satisfiable inputs.

Keywords

Theorem proving, Model building, SGGS, Horn theories

1. Introduction

SGGS (*Semantically-Guided Goal-Sensitive reasoning*) [1, 2, 3] is a first-order theorem-proving method with a rare combination of properties. SGGS is *semantically guided* by a fixed initial Herbrand interpretation I, that in this paper is either all-negative (I^-) or all-positive (I^+). SGGS is *model-based*, as it searches for a model of the input set of clauses by building candidate models, obtained by modifying I. Each candidate model is represented by a sequence of clauses, called a *trail*. SGGS is *conflict-driven*, as it applies inferences such as resolution only to explain conflicts between candidate model and clauses to be satisfied. Indeed, SGGS generalizes to first-order logic the CDCL (*Conflict Driven Clause Learning*) procedure for propositional satisfiability [4]. In addition to being *refutationally complete*, SGGS is *model complete in the limit*, which means that given a satisfiable input, the limit of a fair derivation represents a model.

Horn clauses are a standard language for declarative programming and knowledge representation, where definite clauses represent knowledge and negative clauses play the role of queries. Horn clauses are also an essential language for program verification [5]. A *Horn theory* is a theory presented by a set of definite clauses, and Horn theories are those such that

PAAR2022: 8th Workshop on Practical Aspects of Automated Reasoning, August 11–12, 2022, 9th FLoC, Haifa, Israel
* This research was funded in part by grant "Ricerca di base 2017" of the Università degli Studi di Verona.
* Corresponding author.
† These authors contributed equally.
mariapaola.bonacina@univr.it (M. P. Bonacina); sarwinkler@unibz.it (S. Winkler)
http://profs.sci.univr.it/~bonacina/ (M. P. Bonacina); http://cl-informatik.uibk.ac.at/users/swinkler/ (S. Winkler)
0000-0001-9104-2692 (M. P. Bonacina); 0000-0001-8114-3107 (S. Winkler)

the intersection of models is still a model [6, 7]. Thus, a set of definite clauses admits a *least Herbrand model*, the intersection of all its Herbrand models. The least Herbrand model is also the *least fixpoint* of a functional associated to a set of definite clauses [8, 9, 10].

In this paper we analyze the behavior of SGGS on Horn clauses and relate it to that of hyperresolution [11]. The interest in this parallel stems from the fact that SGGS with I^- or I^+ and hyperresolution have *sign-based semantic guidance* in common. Indeed, hyperresolution is the instance of semantic resolution [12] that adopts I^- or I^+ for semantic guidance. Positive hyperresolution employs I^- and negative hyperresolution employs I^+. In this paper we consider positive hyperresolution that we call hyperresolution for short. After an overview of SGGS (Sect. 2), we describe the behavior of SGGS with I^- or I^+ on Horn clauses, and show that SGGS with I^- *generates the least fixpoint model* (Sect. 3). It is well-known that hyperresolution has this property. We continue by proving that given a set of Horn clauses *SGGS halts if and only if hyperresolution halts*, but SGGS can *learn* from *Horn subproblems* useful information for termination on *first-order problems* (Sect. 4). As another well-known property of hyperresolution is that it behaves *linearly* if given a set of ground Horn clauses, we show that SGGS also has this property (Sect. 5). In the experiments we apply the SGGS prototype prover Koala [13] to Horn sets from the TPTP library [14], with promising results on satisfiable problems (Sect. 6).

2. Basic Definitions and Overview of SGGS

We assume the basic notions in theorem proving, and we use a, b for constants, $\mathsf{P}, \mathsf{Q}, \mathsf{R}, \mathsf{T}$ for predicates, f, g for functions, x, y, z for variables, t, u, v for terms, $\mathcal{V}ar(t)$ for the set of variables in t, and $top(t)$ for the top symbol of t. We also use L, M, P, Q for literals, $at(L)$ for L's atom, C, D, E for clauses, Greek lower case letters for substitutions, and I, J for Herbrand interpretations. The at notation extends to sets of literals and the $\mathcal{V}ar$ notation to literals. A clause is *positive* if all its literals are positive, *negative* if all its literals are negative, and *mixed* otherwise. A clause is *Horn* if it has at most one positive literal, and *definite* if it has exactly one positive literal. C^+ and C^- denote the disjunctions of the positive and negative literals in C.

We refer the readers to [1] for a simple exposition of SGGS, and to [2, 3] for a complete treatment. In SGGS, a clause C may have a constraint A, written $A \rhd C$. The atomic constraints have the form $true$, $false$, $top(t) = f$, or $t \equiv u$, where $\equiv$ is syntactic identity. The negation, conjunction, and disjunction of constraints is a constraint. Constraints may be omitted for brevity. The set $Gr(A \rhd C)$ of *constrained ground instances* (cgi's) of $A \rhd C$ is the set of the ground instances of C that satisfy A, and hence are the *effective ground instances*. Literals $A \rhd L$ and $B \rhd M$ *intersect* if $at(Gr(A \rhd L)) \cap at(Gr(B \rhd M)) \neq \emptyset$, and are *disjoint* otherwise.

SGGS is *semantically guided* by an *initial interpretation* I. Given a set S of clauses, if $I \not\models S$, SGGS seeks a model of S, by building candidate partial interpretations different from I, and using I as default to complete them. If the empty clause $\perp$ is generated, unsatisfiability is reported. If I is I^- (I^+) SGGS tries to discover which positive (negative) literals need to be true to satisfy S.

SGGS works with a *trail* of clauses $\Gamma = A_1 \rhd C_1[L_1], \ldots, A_n \rhd C_n[L_n]$, where $A_i \rhd C_i[L_i]$ occurs at index i and the brackets mean that literal L_i is *selected* in C_i. The length of Γ and its prefix of length j are denoted $|\Gamma|$ and $\Gamma|_j$. An SGGS trail Γ represents a partial interpretation

$I^p(\Gamma)$: if Γ is empty, denoted ε, $I^p(\Gamma){=}\emptyset$; otherwise, $I^p(\Gamma){=}I^p(\Gamma|_{n-1}) \cup pcgi(A_n \rhd L_n, \Gamma)$, where pcgi abbreviates *proper constrained ground instances*. A pcgi of $A_n \rhd C_n[L_n]$ is a cgi $C[L]$ that is not satisfied by $I^p(\Gamma|_{n-1})$ (i.e., $I^p(\Gamma|_{n-1}) \cap C[L] = \emptyset$) and can be satisfied by adding L as $\neg L \notin I^p(\Gamma|_{n-1})$. Thus, pcgi's are *productive* instances as they produce $I^p(\Gamma)$. For the selected literal, $pcgi(A_n \rhd L_n, \Gamma) = \{L\ :\ C[L] \in pcgi(A_n \rhd C_n[L_n], \Gamma)\}$. $I^p(\Gamma)$ is completed into an interpretation $I[\Gamma]$ by consulting I for the truth value of any literal undefined in $I^p(\Gamma)$.

A literal L is *uniformly false* in an interpretation J, if all $L' \in Gr(L)$ are false in J. Then, L is said to be I*-false* if it is uniformly false in I and I*-true* if it is true in I. A clause is I*-all-true* if all its literals are I-true and I*-all-false* if all its literals are I-false. I-false literals are preferred for selection to differentiate $I[\Gamma]$ from I. An I-true literal is selected only in an I-all-true clause.

A clause is a *conflict clause* if all its literals are uniformly false in $I[\Gamma]$. SGGS ensures that every I-all-true clause $C[L]$ in Γ is either a conflict clause or the *justification* of its selected literal L, meaning that all literals of $C[L]$ except L are uniformly false in $I[\Gamma]$, so that L must be true in $I[\Gamma]$ to satisfy $C[L]$. To this end, SGGS *assigns* I-true literals to clauses. Given trail $\Gamma = A_1 \rhd C_1[L_1], \ldots, A_n \rhd C_n[L_n]$, an I-true literal $M \in C_j[L_j]$ *is assigned* to $C_k[L_k]$ if $k{<}j$ and the selection of L_k makes M uniformly false in $I[\Gamma]$ (all literals in $Gr(A_j \rhd M)$ appear negated in $pcgi(A_k \rhd L_k, \Gamma)$ and hence in $I^p(\Gamma)$). SGGS requires that a non-selected I-true literal is assigned, while a selected I-true literal is assigned if possible. If all the literals of an I-all-true clause $C[L]$ in Γ are assigned, $C[L]$ is a conflict clause, and L is assigned to the largest index (or to the rightmost clause) among all literals in $C[L]$. If all the literals of $C[L]$ except L are assigned, $C[L]$ is the justification of L, and $C[L]$ is in the *disjoint prefix* of Γ. The disjoint prefix of Γ, denoted $dp(\Gamma)$, is the longest prefix such that $pcgi(A_j \rhd C_j[L_j], \Gamma) = Gr(A_j \rhd C_j[L_j])$ for all $A_j \rhd C_j[L_j]$ in $dp(\Gamma)$. Thus, all selected literals in $dp(\Gamma)$ are disjoint.

An *SGGS-derivation* is a series of trails $\Gamma_0 \vdash \Gamma_1 \vdash \ldots \Gamma_j \vdash \ldots$, where $\Gamma_0 = \varepsilon$, and $\forall j,\ j > 0$, an SGGS-inference generates Γ_j from Γ_{j-1} and S. If $\bot \notin \Gamma$ and $I[\Gamma] \not\models S$, SGGS has *two ways to make progress*. If $\Gamma = dp(\Gamma)$, the trail is in order, but since $I[\Gamma] \not\models S$ it means that $I[\Gamma] \not\models C'$ for a $C' \in Gr(C)$ and $C \in S$. Then, SGGS applies *SGGS-extension* to generate from C and Γ a clause $A \rhd E$, such that E is an instance of C and $C' {\in} Gr(A \rhd E)$. If $\Gamma \neq dp(\Gamma)$, the trail needs repair: either there are intersections between selected literals to be removed by *SGGS-splitting*, or there is a conflict. The *SGGS-extension rules* specialize the *SGGS-extension scheme* ([3, Def. 12]) that we instantiate for $I = I^+$ or $I = I^-$:

Definition 1. *Given input set S and trail Γ, if there is a clause $C{\in}S$ such that for all its I-true literals $L_1, \ldots, L_n$ ($n{\geq}0$) there are clauses $B_1 \rhd D_1[M_1], \ldots, B_n \rhd D_n[M_n]$ in $dp(\Gamma)$, such that literals $M_1, \ldots, M_n$ are I-false, and $\forall j,\ 1{\leqslant}j{\leqslant}n,\ L_j\alpha{=}\neg M_j\alpha$ with simultaneous most general unifier (mgu) α,* SGGS-extension *adds to Γ the* extension clause $A \rhd E = (\bigwedge_{j=1}^{n} B_j\alpha) \rhd C\alpha$, *assigning $L_1\alpha, \ldots, L_n\alpha$ to $D_1, \ldots, D_n$, respectively.*

Clause C is the *main premise* and $B_1 \rhd D_1[M_1], \ldots, B_n \rhd D_n[M_n]$ are the *side premises*. An SGGS-extension is *conflicting*, if $A \rhd E$ is a conflict clause, and *non-conflicting* otherwise, that is, if $A \rhd E$ has $pcgi$'s and therefore extends $I[\Gamma]$. If $A \rhd E$ is I-all-true then it is a conflict clause. A derivation starts with a non-conflicting SGGS-extension where C is I-all-false, so that α is empty, $n = 0$, and $E = C$. All such steps can be done as one and we assume they are.

Example 1. *Consider the following set S of definite clauses with I^- as initial interpretation:*

$$\begin{array}{llll} \mathsf{P}(\mathsf{f}(\mathsf{a},x)) & (i) & \mathsf{P}(\mathsf{g}(\mathsf{b},x)) & (ii) \\ \neg\mathsf{P}(\mathsf{f}(y,\mathsf{a})) \vee \mathsf{P}(\mathsf{g}(y,\mathsf{a})) & (iii) & \neg\mathsf{P}(\mathsf{g}(z,\mathsf{b})) \vee \mathsf{P}(\mathsf{f}(z,\mathsf{b})) & (iv). \end{array}$$

I^- satisfies clauses (iii) and (iv) because they have negative literals, but not the positive clauses (i) and (ii). SGGS extends the trail with (i) and (ii), so that $I[\Gamma_1]$ satisfies them. Thus, $I[\Gamma_1]$ satisfies neither the instance of (iii) with $\{y \leftarrow \mathsf{a}\}$ nor the instance of (iv) with $\{z \leftarrow \mathsf{b}\}$. SGGS extends the trail with these instances and halts, as $I[\Gamma_3]$ satisfies S:

$$\begin{array}{l} \Gamma_0 \colon \varepsilon \vdash \Gamma_1 \colon [\mathsf{P}(\mathsf{f}(\mathsf{a},x))],\ [\mathsf{P}(\mathsf{g}(\mathsf{b},x))] \\ \quad\vdash \Gamma_2 \colon [\mathsf{P}(\mathsf{f}(\mathsf{a},x))],\ [\mathsf{P}(\mathsf{g}(\mathsf{b},x))],\ \neg\mathsf{P}(\mathsf{f}(\mathsf{a},\mathsf{a})) \vee [\mathsf{P}(\mathsf{g}(\mathsf{a},\mathsf{a}))] \\ \quad\vdash \Gamma_3 \colon [\mathsf{P}(\mathsf{f}(\mathsf{a},x))],\ [\mathsf{P}(\mathsf{g}(\mathsf{b},x))],\ \neg\mathsf{P}(\mathsf{f}(\mathsf{a},\mathsf{a})) \vee [\mathsf{P}(\mathsf{g}(\mathsf{a},\mathsf{a}))],\ \neg\mathsf{P}(\mathsf{g}(\mathsf{b},\mathsf{b})) \vee [\mathsf{P}(\mathsf{f}(\mathsf{b},\mathsf{b}))]. \end{array}$$

Hyperresolution generates $\mathsf{P}(\mathsf{g}(\mathsf{a},\mathsf{a}))$ and $\mathsf{P}(\mathsf{f}(\mathsf{b},\mathsf{b}))$, and then halts. In the case of definite clauses the two methods appear close, as one selects positive literals on the trail and the other generates positive unit clauses. However, SGGS generates explicitly a model on the trail.

SGGS-deletion removes *all* disposable clauses, where $A_n \rhd C_n[L_n]$ is *disposable* if $I^p(\Gamma|_{n-1}) \models A_n \rhd C_n[L_n]$. *SGGS-splitting* aims at isolating intersections between selected literals, and it is the rule that introduces constraints. A *partition* of $A \rhd C[L]$ is a set $\{A_i \rhd C_i[L_i]\}_{i=1}^n$ such that $Gr(A \rhd C) = \bigcup_{i=1}^n \{Gr(A_i \rhd C_i)\}$, the L_i's are chosen consistently with L, and the $A_i \rhd L_i$'s are pairwise disjoint (cf. [3, Sect. 3.2]). Given trail clauses $A \rhd C[L]$ and $B \rhd D[M]$, a *splitting* of C by D, denoted $split(C, D)$, is a partition of $A \rhd C[L]$ such that $at(Gr(A_j \rhd L_j))$ for some j is the intersection of $A \rhd L$ and $B \rhd M$, and all other $A_i \rhd L_i$'s are disjoint from $B \rhd M$. *SGGS-splitting* replaces $A \rhd C[L]$ with $split(C, D)$. Clause $A \rhd C[L]$ is the *split clause*, and $A_j \rhd C_j[L_j]$ is the *representative* of $B \rhd D[M]$ in the splitting. *Splitting by similar literals* (*s-splitting*) and *splitting by dissimilar literals* (*d-splitting*) apply when $A \rhd C[L]$ occurs at a higher index than $B \rhd D[M]$, with similar and dissimilar referring to whether L and M have the same or opposite sign. After an s-splitting, D's representative in $split(C, D)$ is disposable (cf. [3, Lem. 3]) and hence removed by SGGS-deletion. After a d-splitting, the intersection between L and M is removed by *SGGS-resolution*.

If a clause $A \rhd C[L]$ added by SGGS-extension is in conflict with $I[\Gamma]$ and L is I-false, *SGGS-resolution explains the conflict* by resolving $A \rhd C[L]$ with the justification in $dp(\Gamma)$ whose selected literal makes L uniformly false in $I[\Gamma]$.

Definition 2 (SGGS-Resolution). *If Γ contains clauses $B \rhd D[M]$ and $A \rhd C[L]$ such that $B \rhd D[M]$ is I-all-true, is in $dp(\Gamma)$, and occurs at a smaller index than $A \rhd C[L]$, L is I-false, $L = \neg M\vartheta$ for some substitution ϑ, and $A \models B\vartheta$, then* SGGS-resolution *generates the* SGGS-resolvent *$A \rhd R$, where R is $(C \setminus \{L\}) \cup (D \setminus \{M\})\vartheta$, replaces C by $A \rhd R$, deletes all clauses with literals assigned to C, and assigns every I-true literal $P\vartheta$ in R to the same clause that P in C or D was assigned to (cf. [3, Def. 26]).*

In *conflict explanation* the resolvent is still a conflict clause. As SGGS-extension ensures that *all* I-false literals of a conflict clause can be resolved away (cf. [3, Def. 19]), conflict explanation generates either $\bot$ or an I-all-true conflict clause $H \rhd E[P]$. *Conflict solving* moves $H \rhd E[P]$ to the left of the clause $A \rhd C[L]$ in $dp(\Gamma)$ to which P is assigned (*SGGS-move*): the effect is to *flip* P from being uniformly false to being an implied literal. Then E resolves with C, so that the simplest *conflict explanation and solving sequence* has the form resolve* move resolve. Prior to the move, if the pcgi's of L contain more than the complements of the cgi's of P (i.e., $\neg Gr(H \rhd P) \subset pcgi(A \rhd L, \Gamma)$), *left splitting* replaces $A \rhd C[L]$ by $split(C, E)$ with P assigned to the representative of E in $split(C, E)$. Left splitting enables SGGS-move, which places E to the left of its representative in $split(C, E)$. Thus, the sequence has the form resolve* l-split? move resolve, where ? means at most once. If E contains another literal Q that is assigned to clause C and unifies with P with mgu ϑ, E cannot move because Q would have nowhere to be assigned. Then *SGGS-factoring* replaces E by $split(E, E\vartheta)$, where the *SGGS-factor* $E\vartheta$ is its own representative and literal assignments are inherited. As SGGS-factoring may repeat, the full form of the sequence is resolve* factor* l-split? move resolve.

Fairness of an SGGS-derivation involves several properties: an inference is applied whenever $\bot \notin \Gamma$ and $I[\Gamma] \not\models S$; no trivial splitting (e.g., splitting a ground clause) occurs; SGGS-deletion and other clause removals are applied eagerly; conflicts are solved before more SGGS-extensions; and inferences applying to shorter prefixes of the trail are not forever neglected in favor of others applying to longer prefixes (cf. [3, Defs. 32, 37, and 49]). The *limit* of a fair derivation $\Gamma_0 \vdash \Gamma_1 \vdash \ldots \Gamma_j \vdash \Gamma_{j+1} \vdash \ldots$ is the longest trail Γ_∞ such that $\forall i,\, i \leqslant |\Gamma_\infty|$, there exists an n_i such that $\forall j,\, j \geqslant n_i$, if $|\Gamma_j| \geq i$ then $\Gamma_j|_i$ is equivalent to $\Gamma_\infty|_i$ (cf. [3, Def. 50]). In words, all prefixes of the trail stabilize eventually, and $dp(\Gamma_\infty) = \Gamma_\infty$. Both derivation and Γ_∞ may be infinite, but Γ_∞ is Γ_k if the derivation halts at stage k. SGGS is *refutationally complete* and *model complete in the limit*: for all inputs S, initial interpretations I, and fair derivations, if S is satisfiable, $I[\Gamma_\infty] \models S$, and if S is unsatisfiable, $\bot \in \Gamma_k$ for some k (cf. [3, Thms. 9 and 11]).

3. On the Behavior of SGGS on Horn Clauses

Let S be a set of definite clauses and $\mathcal{A}$ its Herbrand base. The powerset $\mathcal{P}(\mathcal{A})$ of $\mathcal{A}$ is the set of all Herbrand interpretations viewed as sets of atoms. It is a complete lattice ordered by the subset relation $\subseteq$, with intersection $\bigcap$ as the greatest lower bound (glb) operator, union $\bigcup$ as the least upper bound (lub) operator, $\emptyset$ as bottom, and $\mathcal{A}$ as top element. A set of definite clauses admits a *least Herbrand model*, defined as the intersection of all Herbrand models of S, or, equivalently, as the fixpoint of the functional $T_S \colon \mathcal{P}(\mathcal{A}) \to \mathcal{P}(\mathcal{A})$ defined as follows. For all $J \in \mathcal{P}(\mathcal{A})$ and $L \in \mathcal{A}$, if there exist a clause $P \vee \neg Q_1 \vee \ldots \vee \neg Q_m$ $(m \geqslant 0)$ in S and a ground substitution σ, such that $L = P\sigma$ and $\{Q_1\sigma \ldots Q_m\sigma\} \subseteq J$, then $L \in T_S(J)$ and vice versa. As T_S is continuous (cf. [10, Prop. 6.3]) its least fixpoint $lfp(T_S)$ is given by the lub of the nondecreasing $\subseteq$-chain $\{T_S^k(\emptyset)\}_{k \geqslant 0}$ so that $lfp(T_S) = \bigcup_{k \geqslant 0} T_S^k(\emptyset)$ (cf. [10, Thm. 6.5]).

The initial interpretation I^- corresponds to the bottom $\emptyset$ of lattice $\mathcal{P}(\mathcal{A})$. SGGS with I^- reasons *forward* or *bottom-up*. The derivation starts with an SGGS-extension that puts on the trail all the positive unit clauses in S. If this addition of positive literals to $I^p(\Gamma)$ falsifies all the negative literals in instances of mixed clauses, SGGS-extensions with mixed clauses follow. Since I^--false (i.e., positive) literals are preferred for selection, and every clause has exactly one, *all selected literals are positive*, with *no choice of selected literal.* All extensions are non-conflicting extensions that add pcgi's to $I^p(\Gamma)$, which contains only positive literals and grows monotonically: $\forall j,\ j \geqslant 0,\ I^p(\Gamma_j) \subseteq I^p(\Gamma_{j+1})$. Since all selected literals are positive, the only splitting inferences are s-splitting steps, followed by deletions of representatives that remove intersections between selected literals, without affecting $I^p(\Gamma)$ and its monotonic growth. Since $I = I^-$, for an atom $L \in \mathcal{A}$, $I[\Gamma_\infty] \models L$ if and only if $L \in I^p(\Gamma_\infty)$. In other words, $I[\Gamma_\infty]$ seen as a set of atoms is $I^p(\Gamma_\infty)$. The next theorem shows that $I^p(\Gamma_\infty) = \mathit{lfp}(T_S)$.

Theorem 1. *Given an input set S of definite clauses, with Herbrand base $\mathcal{A}$ and functional $T_S \colon \mathcal{P}(\mathcal{A}) \rightarrow \mathcal{P}(\mathcal{A})$, for all fair SGGS-derivations with I^- as initial interpretation and limit Γ_∞, it is $I^p(\Gamma_\infty) = \mathit{lfp}(T_S)$.*

Proof: $\mathit{lfp}(T_S) \subseteq I^p(\Gamma_\infty)$: since SGGS is model complete in the limit, $I[\Gamma_\infty] \models S$, that is, $I^p(\Gamma_\infty) \models S$, writing Herbrand models as sets of atoms. Thus, $\mathit{lfp}(T_S) \subseteq I^p(\Gamma_\infty)$, because $\mathit{lfp}(T_S)$ is the least Herbrand model.
$I^p(\Gamma_\infty) \subseteq \mathit{lfp}(T_S)$: if $L \in I^p(\Gamma_\infty)$, it means that $L \in pcgi(A_i \rhd L_i, \Gamma_\infty)$ for some clause $A_i \rhd C_i[L_i]$ at index i in Γ_∞. This clause is placed at index i of the trail at some stage of the derivation. Since it is still there in Γ_∞, there exists a stage k such that for all stages j, $j \geqslant k$, clause $A_i \rhd C_i[L_i]$ is at index i in Γ_j and $L \in I^p(\Gamma_j)$. The proof is by induction on the smallest such k.
Base case: $k = 1$ and $L \in I^p(\Gamma_1)$. Trail Γ_1 is the result of the first SGGS-extension that adds all the I^--all-false input clauses. Since S is a set of definite clauses, the I^--all-false input clauses are the positive unit clauses of S. Thus, $L \in I^p(\Gamma_1)$ means that $L = P\sigma$ for a ground substitution σ and a positive unit clause $P \in S$ with $[P]$ in Γ_1. By definition of T_S, $L \in T_S(\emptyset)$ and $L \in \mathit{lfp}(T_S)$.
Induction hypothesis: for all k ($k \geqslant 1$), for all $n \leqslant k$, if n is the smallest stage of the SGGS-derivation such that for all j, $j \geqslant n$, $L \in I^p(\Gamma_j)$, then $L \in \mathit{lfp}(T_S)$.
Inductive case: take an $L \in \mathcal{A}$ for which $k+1$ is the smallest stage of the SGGS-derivation such that for all j, $j \geqslant k+1$, $L \in I^p(\Gamma_j)$. This means that $L \notin I^p(\Gamma_k)$. Thus, the step $\Gamma_k \vdash \Gamma_{k+1}$ must be a non-conflicting SGGS-extension with main premise $C = P \vee \neg Q_1 \vee \ldots \vee \neg Q_m$ in S ($m>0$), mgu α, and extension clause $C\alpha = [P\alpha] \vee \neg Q_1\alpha \vee \ldots \vee \neg Q_m\alpha$, such that $L \in pcgi(P\alpha, \Gamma_{k+1})$. Therefore, there exist a ground substitution σ' such that $L = P\alpha\sigma'$ and hence a substitution σ such that $C\sigma$ is ground and $L = P\sigma$. Furthermore, $\forall i,\ 1 \leqslant i \leqslant m,\ Q_i\sigma \in I^p(\Gamma_k)$, because otherwise $I[\Gamma_k] \models \neg Q_i\sigma$ (as $I^- \models \neg Q_i\sigma$), and hence $[P\sigma] \vee \neg Q_1\sigma \vee \ldots \vee \neg Q_m\sigma$ would not be in $pcgi([P\alpha] \vee \neg Q_1\alpha \vee \ldots \vee \neg Q_m\alpha, \Gamma_{k+1})$ and L would not be in $pcgi(P\alpha, \Gamma_{k+1})$. Then, for all i, $1 \leqslant i \leqslant m$, let h_i be the smallest stage such that $\forall j,\ j \geqslant h_i,\ Q_i\sigma \in I^p(\Gamma_j)$. Since $Q_i\sigma \in I^p(\Gamma_k)$, we have $h_i \leqslant k$ for all i, $1 \leqslant i \leqslant m$. Thus, by induction hypothesis, for all i, $1 \leqslant i \leqslant m$, $Q_i\sigma \in \mathit{lfp}(T_S)$. Since $\mathit{lfp}(T_S) \models S$, we have $\mathit{lfp}(T_S) \models P \vee \neg Q_1 \vee \ldots \vee \neg Q_m$ and hence $\mathit{lfp}(T_S) \models P\sigma \vee \neg Q_1\sigma \vee \ldots \vee \neg Q_m\sigma$. Therefore, $L = P\sigma \in \mathit{lfp}(T_S)$. □

For instance, in Example 1, $\mathit{lfp}(T_S)$ is $I^p(\Gamma_3)$. We describe next the behavior of SGGS when also negative clauses enter the picture, first with I^- and then with I^+.

Let S be a set of Horn clauses and $S_D \subset S$ the subset of the definite clauses. Unless a model of S is found, the consequence of adding positive ground literals to $I^p(\Gamma)$ towards building $\mathit{lfp}(T_{S_D})$ is that some instance of a negative clause in S becomes an I^--all-true conflict clause, that a conflicting SGGS-extension adds to the trail. Let C be the first conflict clause of this kind and Γ_k the first trail where C appears. The fact that C is in conflict with $I[\Gamma_k]$ means that C is in conflict with $I^p(\Gamma_k)$. By Theorem 1, this means that C is in conflict with a subset of $lfp(T_{S_D})$ and hence with $lfp(T_{S_D})$ itself. Since $lfp(T_{S_D})$ is the intersection of *all the Herbrand models of* S_D, it follows that C is in conflict with all of them. Therefore, the appearance of C on the trail reveals that S is unsatisfiable. By contrast, a set of first-order clauses does not admit a least Herbrand model, and the appearance of a negative conflict clause on a trail Γ_k in an SGGS-derivation with I^- simply means that $I[\Gamma_k] \not\models S$, so that SGGS solves the conflict and searches for another model.

Furthermore, because all the literals in C are assigned, it is possible to predict the number of SGGS-resolutions needed to go from C to $\bot$. This number turns out to be equal to the cardinality $|d_\Gamma(C)|$ of the *dependence set* $d_\Gamma(C)$ of C in Γ. This set is defined inductively as follows: if an I-true literal in C is assigned to clause D then $D \in d_\Gamma(C)$; if $D \in d_\Gamma(C)$ and an I-true literal in D is assigned to clause D' then $D' \in d_\Gamma(C)$; nothing else is in $d_\Gamma(C)$.

Lemma 1. *Given an input set S of Horn clauses and I^- as initial interpretation, for all fair SGGS-derivations, if an SGGS-extension adds to a trail Γ an I^--all-true conflict clause C, the derivation is a refutation, and the number of SGGS-resolution steps after the stage of Γ is $|d_\Gamma(C)|$.*

Proof: The proof is by induction on the cardinality $|d_\Gamma(C)|$.
Base case: $|d_\Gamma(C)| = 0$ and hence $d_\Gamma(C) = \emptyset$. Since C is an I^--all-true conflict clause, all its literals must be assigned. Thus, $d_\Gamma(C) = \emptyset$ implies that C is $\bot$ and both claims are fulfilled.
Induction hypothesis: if $|d_\Gamma(C)| = k$ for $k \geqslant 0$ then both claims hold.
Inductive case: let $|d_\Gamma(C)| = k + 1$, literal L be selected in C, and $D[M]$ be the clause which L is assigned to.
If SGGS-factoring applies to $C[L]$ (one or more times), let $C_f[L_f]$ and Γ^1 be the final SGGS-factor and trail, or $C_f[L_f] = C[L]$ and $\Gamma^1 = \Gamma$ otherwise. Either way, $d_{\Gamma^1}(C_f) = d_\Gamma(C)$, because SGGS-factoring unifies two literals of C that are both assigned to D and SGGS-factors inherit assignments, so that L_f is still assigned to D in Γ^1. Also, $d_{\Gamma^1}(D) = d_\Gamma(D)$ trivially holds.
If left-splitting applies to $C_f[L_f]$ and D, let $D'[M']$ and Γ^2 be the representative of C_f in $split(D, C_f)$ and the resulting trail, or $D'[M'] = D[M]$ and $\Gamma^2 = \Gamma^1$ otherwise. Either way, L_f is assigned to D' in Γ^2 and $d_{\Gamma^2}(D') = d_{\Gamma^1}(D) = d_\Gamma(D)$, because the I^--true literals in D' inherit their assignments from those in D. Thus, $d_{\Gamma^2}(C_f) = d_{\Gamma^1}(C_f) \setminus \{D\} \cup \{D'\} = d_\Gamma(C) \setminus \{D\} \cup \{D'\}$ as the only change is that L_f is assigned to D in Γ^1 and to D' in Γ^2 (†).
SGGS-move places $C_f[L_f]$ to the left of $D'[M']$ and then SGGS-resolution applies to $C_f[L_f]$ and $D'[M']$ generating the SGGS-resolvent $R = (C_f \setminus \{L_f\}) \cup (D' \setminus \{M'\})\vartheta$ that replaces D' in the resulting trail Γ^3. Every I^--true literal $P\vartheta$ in R is assigned to the same clause that P in C_f or in D' was assigned to (⋆). Since C_f is negative and D' is Horn, R is negative, it is another I^--all-true conflict clause, and *all* its literals are assigned.
We show that $d_{\Gamma^3}(R) = d_\Gamma(C) \setminus \{D\}$. For every clause $E \in d_{\Gamma^3}(R)$ due to a literal in $(C_f \setminus \{L_f\})\vartheta$, it is $E \in d_{\Gamma^2}(C_f)$ by (⋆), and $E \in d_{\Gamma^1}(C_f)$ by (†), and $E \in d_\Gamma(C)$ because $d_{\Gamma^1}(C_f) = d_\Gamma(C)$. For every clause $E \in d_{\Gamma^3}(R)$ due to a literal in $(D' \setminus \{M'\})\vartheta$, it is $E \in d_{\Gamma^2}(D')$ by (⋆), and $E \in d_\Gamma(D)$ as $d_{\Gamma^2}(D') = d_{\Gamma^1}(D) = d_\Gamma(D)$, and hence $E \in d_\Gamma(C)$ because the selected literal L of C is assigned to D in Γ. Up to here we have shown that

$d_{\Gamma^3}(R) \subseteq d_\Gamma(C)$. Next, $D \in d_\Gamma(C)$, but $D \notin d_{\Gamma^3}(R)$, because L_f was resolved away, and R replaced D' which either replaced D or is D itself. Thus, $d_{\Gamma^3}(R) \subset d_\Gamma(C)$. Furthermore, it is exactly $d_{\Gamma^3}(R) = d_\Gamma(C) \setminus \{D\}$, because by inspection of the above inferences nothing else has changed. Since $|d_\Gamma(C)| = k+1$, we have $|d_{\Gamma^3}(R)| = k$. By induction hypothesis the first claim holds, and the number of SGGS-resolutions after the stage of Γ^3 is k, so that the number of SGGS-resolutions after the stage of Γ is $k+1 = |d_\Gamma(C)|$. □

The initial interpretation I^+ is the top $\mathcal{A}$ of lattice $\mathcal{P}(\mathcal{A})$. Given a set S of definite clauses, SGGS with I^+ does not do anything as $I^+ \models S$. If S is a set of Horn clauses, where the negative clauses are the goal clauses, I^+ is *goal-sensitive* (it satisfies the assumptions, not the goal clauses), and hence SGGS with I^+ is goal-sensitive (all generated clauses are connected to goal clauses) [3]. Indeed, SGGS with I^+ reasons *backward* or *top-down*. The derivation starts with an SGGS-extension that puts on the trail all the negative clauses in S, each with a literal selected. If the addition of negative ground literals to $I^p(\Gamma)$ falsifies the positive literal in instances of a mixed clause in S, a non-conflicting SGGS-extension with mixed extension clause applies: a negative literal is selected in the extension clause and more negative ground literals are added to $I^p(\Gamma)$. Unless a model of S is found, the consequence of adding negative ground literals to $I^p(\Gamma)$ is that some instance of a positive unit clause in S becomes an I^+-all-true conflict clause, that a conflicting SGGS-extension adds to the trail.

Example 2. *Given* $S = \{\neg \mathsf{Q}(\mathsf{a}, x),\ \neg \mathsf{R}(x,y) \vee \neg \mathsf{R}(y,x) \vee \mathsf{Q}(x,y),\ \mathsf{R}(\mathsf{b},\mathsf{a}),\ \mathsf{R}(\mathsf{a},\mathsf{b})\}$*, the SGGS-derivation with* I^+ *starts as follows:*

$$\begin{array}{lr} \Gamma_0\colon \varepsilon \vdash \Gamma_1\colon [\neg \mathsf{Q}(\mathsf{a},x)] & \mathsf{extend} \\ \quad\vdash \Gamma_3\colon [\neg \mathsf{Q}(\mathsf{a},x)],\ [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y) & \mathsf{extend} \end{array}$$

and continues extending with I^+*-all-true conflict clause* $\mathsf{R}(\mathsf{a},\mathsf{b})$ *and solving the conflict:*

$$\begin{array}{lr} \vdash \Gamma_4\colon [\neg \mathsf{Q}(\mathsf{a},x)],\ [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y),\ [\mathsf{R}(\mathsf{a},\mathsf{b})] & \mathsf{extend} \\ \vdash \Gamma_5\colon [\neg \mathsf{Q}(\mathsf{a},x)],\ [\neg \mathsf{R}(\mathsf{a},\mathsf{b})] \vee \neg \mathsf{R}(\mathsf{b},\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},\mathsf{b}), & \\ \quad \mathit{top}(y) \neq \mathsf{b} \rhd [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y),\ [\mathsf{R}(\mathsf{a},\mathsf{b})] & \mathsf{l\text{-}split} \\ \vdash \Gamma_6\colon [\neg \mathsf{Q}(\mathsf{a},x)],\ [\mathsf{R}(\mathsf{a},\mathsf{b})],\ [\neg \mathsf{R}(\mathsf{a},\mathsf{b})] \vee \neg \mathsf{R}(\mathsf{b},\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},\mathsf{b}), & \\ \quad \mathit{top}(y) \neq \mathsf{b} \rhd [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y) & \mathsf{move} \\ \vdash \Gamma_7\colon [\neg \mathsf{Q}(\mathsf{a},x)],\ [\mathsf{R}(\mathsf{a},\mathsf{b})],\ [\neg \mathsf{R}(\mathsf{b},\mathsf{a})] \vee \mathsf{Q}(\mathsf{a},\mathsf{b}), & \\ \quad \mathit{top}(y) \neq \mathsf{b} \rhd [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y) & \mathsf{resolve} \end{array}$$

Then the derivation proceeds extending with I^+*-all-true conflict clause* $\mathsf{R}(\mathsf{b},\mathsf{a})$ *whose conflict-solving phase finds a refutation:*

$$\begin{array}{lr} \vdash \Gamma_8\colon [\neg \mathsf{Q}(\mathsf{a},x)],\ [\mathsf{R}(\mathsf{a},\mathsf{b})],\ [\neg \mathsf{R}(\mathsf{b},\mathsf{a})] \vee \mathsf{Q}(\mathsf{a},\mathsf{b}), & \\ \quad \mathit{top}(y) \neq \mathsf{b} \rhd [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y),\ [\mathsf{R}(\mathsf{b},\mathsf{a})] & \mathsf{extend} \\ \vdash \Gamma_9\colon [\neg \mathsf{Q}(\mathsf{a},x)],\ [\mathsf{R}(\mathsf{a},\mathsf{b})],\ [\mathsf{R}(\mathsf{b},\mathsf{a})],\ [\neg \mathsf{R}(\mathsf{b},\mathsf{a})] \vee \mathsf{Q}(\mathsf{a},\mathsf{b}), & \\ \quad \mathit{top}(y) \neq \mathsf{b} \rhd [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y) & \mathsf{move} \end{array}$$

$\vdash\ \Gamma_{10}\colon [\neg \mathsf{Q}(\mathsf{a},x)],\ [\mathsf{R}(\mathsf{a},\mathsf{b})],\ [\mathsf{R}(\mathsf{b},\mathsf{a})],\ [\mathsf{Q}(\mathsf{a},\mathsf{b})],$
$\quad top(y) \neq \mathsf{b} \rhd [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y)$ resolve

$\vdash\ \Gamma_{11}\colon [\neg \mathsf{Q}(\mathsf{a},\mathsf{b})],\ top(x) \neq b \rhd [\neg \mathsf{Q}(\mathsf{a},x)],\ [\mathsf{R}(\mathsf{a},\mathsf{b})],\ [\mathsf{R}(\mathsf{b},\mathsf{a})],$
$\quad [\mathsf{Q}(\mathsf{a},\mathsf{b})],\ top(y) \neq \mathsf{b} \rhd [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y)$ l-split

$\vdash\ \Gamma_{12}\colon [\mathsf{Q}(\mathsf{a},\mathsf{b})],\ [\neg \mathsf{Q}(\mathsf{a},\mathsf{b})],\ top(x) \neq b \rhd [\neg \mathsf{Q}(\mathsf{a},x)],\ [\mathsf{R}(\mathsf{a},\mathsf{b})],$
$\quad [\mathsf{R}(\mathsf{b},\mathsf{a})],\ top(y) \neq \mathsf{b} \rhd [\neg \mathsf{R}(\mathsf{a},y)] \vee \neg \mathsf{R}(y,\mathsf{a}) \vee \mathsf{Q}(\mathsf{a},y)$ move

$\vdash\ \Gamma_{13}\colon [\mathsf{Q}(\mathsf{a},\mathsf{b})],\ \bot,\ \ldots$ resolve

If the second literal is selected in the second clause of Γ_2*, the derivation is very similar, extending first with* $\mathsf{R}(\mathsf{b},\mathsf{a})$ *and then with* $\mathsf{R}(\mathsf{a},\mathsf{b})$.

4. On the Termination of SGGS and Hyperresolution

Both SGGS and hyperresolution halt on the clause set of Example 1. We show that this is a general result: given a Horn set, SGGS with I^- halts if and only if hyperresolution with subsumption, henceforth HR^+S, halts. A derivation by HR^+S has the form $S_0 \vdash S_1 \vdash \ldots S_j \vdash S_{j+1} \vdash \ldots$, where $S_0{=}S$; for all j, $j \geqslant 0$, S_{j+1} is obtained from S_j by either a hyperresolution or a subsumption step; and the limit S_∞ is the set $\bigcup_{i\geqslant 0}\bigcap_{j\geqslant i} S_j$ of the persistent clauses. In the Horn case, all hyperresolvents are unit clauses identified with their single literal. The next lemma establishes a correspondence between clauses in the respective limits of the two derivations.

Lemma 2. *Given a satisfiable set* S *of Horn clauses, for all fair derivations by SGGS with* I^- *and* HR^+S*, respectively, (i) for all clauses* $A \rhd C[L]$ *in* Γ_∞*, there exists a unit clause* $Q \in S_\infty$ *such that* $Gr(A \rhd L) \subseteq Gr(Q)$*, and (ii) for all unit clauses* $Q \in S_\infty$ *there exist clauses* $\{B_j \rhd D_j[M_j]\}_{j=1}^n$ *in* Γ_∞ *such that* $Gr(Q) \subseteq \bigcup_{j=1}^n Gr(B_j \rhd M_j)$.

Proof: As S is satisfiable and both methods are sound, neither derivation is a refutation. As described in Sect. 3 and by Lemma 1, the only inferences in the SGGS-derivation are non-conflicting SGGS-extensions, that add non-negative clauses where the single positive literal is selected, s-splitting, or SGGS-deletion steps.
Proof of Claim (i): by induction on the smallest stage k ($k \geqslant 1$) such that $A \rhd C[L]$ is in Γ_k.
Base case: $k = 1$ and $C[L]$ is an I^--all-false (i.e., positive, hence unit) input clause $L \in S$ added by the first SGGS-extension. If L is not subsumed in the HR^+S-derivation, $Q = L \in S_\infty$. Otherwise, there exists a $Q \in S_\infty$ that subsumes L, that is, $L = Q\sigma$ for some substitution σ, so that $Gr(L) \subseteq Gr(Q)$ holds.
Induction hypothesis: for all j, $0 \leqslant j \leqslant k$, for all clauses $A \rhd C[L]$ for which j is the smallest stage such that $A \rhd C[L]$ is in Γ_j, the claim holds.
Inductive case: let $A \rhd C[L]$ be a clause for which $k+1$ is the smallest stage such that $A \rhd C[L]$ is in Γ_{k+1}. This means that $A \rhd C[L]$ is placed on the trail by step $\Gamma_k \vdash \Gamma_{k+1}$. If $\Gamma_k \vdash \Gamma_{k+1}$ is an s-splitting step, $A \rhd C[L]$ is a clause in a partition of a split clause $B \rhd D[M]$ in Γ_k. By definition of partition, $Gr(A \rhd C[L]) \subseteq Gr(B \rhd D[M])$ and hence $Gr(A \rhd L) \subseteq Gr(B \rhd M)$. By induction hypothesis, there exists a $Q \in S_\infty$ such that $Gr(B \rhd M) \subseteq Gr(Q)$. Thus,

$Gr(A \rhd L) \subseteq Gr(B \rhd M) \subseteq Gr(Q)$ and the claim holds. Otherwise, $\Gamma_k \vdash \Gamma_{k+1}$ is an SGGS-extension. Let $C \in S$ be the main premise, $B_1 \rhd D_1[M_1], \ldots, B_n \rhd D_n[M_n]$ in $dp(\Gamma_k)$ the side premises, and $(\bigwedge_{i=1}^n B_i\alpha) \rhd C[L]\alpha$ the extension clause, where L is the only positive literal in C, and α is the simultaneous mgu of all the I^--true (i.e., negative) literals $\neg L_1, \ldots, \neg L_n$ of C with $M_1, \ldots, M_n$ (i.e., $L_i\alpha = M_i\alpha$ for all i, $1 \leqslant i \leqslant n$). By induction hypothesis, for all i, $1 \leqslant i \leqslant n$, there exists a $Q_i \in S_\infty$ such that $Gr(B_i \rhd M_i) \subseteq Gr(Q_i)$. This implies that for all i, $1 \leqslant i \leqslant n$, there exists a substitution σ_i (which can be a variable renaming or even empty) such that $M_i = Q_i\sigma_i$. The fact that the M_i's are instances of the Q_i's, together with the fact that the M_i's do not share variables ($\forall i\, \forall j,\ 1 \leqslant i \neq j \leqslant n,\ \mathcal{V}ar(M_i) \cap \mathcal{V}ar(M_j) = \emptyset$), because they come from distinct clauses, and the existence of the above simultaneous mgu α, imply the existence of a simultaneous mgu τ such that $\forall i,\ 1 \leqslant i \leqslant n,\ L_i\tau = Q_i\tau$. For all i, $1 \leqslant i \leqslant n$, let r_i be the stage such that $Q_i \in \bigcap_{j \geqslant r_i} S_j$ and let $r = max\{r_i\ :\ 1 \leqslant i \leqslant n\}$, so that $Q_1, \ldots, Q_n \in S_r$. Therefore, hyperresolution applies to nucleus $C \in S$ and satellites $Q_1, \ldots, Q_n \in S_r$, generating the unit hyperresolvent $L\tau$. Since τ is a unifier of more general literals, there exists a substitution σ such that $L\alpha = L\tau\sigma$. If $L\tau$ is persistent, $Q = L\tau \in S_\infty$, and from $L\alpha = L\tau\sigma$ we get $Gr((\bigwedge_{j=1}^n B_j\alpha) \rhd L\alpha) \subseteq Gr(L\alpha) \subseteq Gr(Q)$. If $L\tau$ is subsumed, there exist a $Q \in S_\infty$ and a substitution δ, such that $L\tau = Q\delta$. Thus, it is $Gr((\bigwedge_{j=1}^n B_j\alpha) \rhd L\alpha) \subseteq Gr(L\alpha) \subseteq Gr(L\tau) \subseteq Gr(Q)$.

Proof of Claim (ii): by induction on the smallest stage k ($k \geqslant 0$) such that $Q \in S_k$.

Base case: $k{=}0$ and $Q \in S_0 = S$ is a positive unit input clause. Thus, the first SGGS-extension that adds to the trail all I^--all-false clauses (i.e., all positive unit input clauses) places $[Q]$ on trail Γ_1 at some index i. If $[Q]$ is subject to neither SGGS-deletion nor s-splitting, $[Q]$ is in Γ_∞ and the claim holds. Since SGGS-move never applies, no clause is inserted to the left of $[Q]$ (i.e., at an index smaller than i) at any subsequent stage j ($j > 1$) of the SGGS-derivation. Thus, the only clauses that can s-split $[Q]$ or make it disposable are other positive unit input clauses that the first SGGS-extension places on the left of $[Q]$. By fairness, SGGS-deletion is applied before any other inference, and removes all disposable clauses in Γ_1 in one step. If a bunch of positive unit input clauses $[Q_1], \ldots, [Q_n]$ on the left of $[Q]$ in Γ_1 make it disposable, we have $Gr(Q) \subseteq \bigcup_{j=1}^n pcgi(Q_j, \Gamma_1) \subseteq \bigcup_{j=1}^n Gr(Q_j)$. If $[Q']$ on the left of $[Q]$ splits $[Q]$ into a partition $\{B_j \rhd [M_j]\}_{j=1}^n$, the representative of $[Q']$ is subsequently deleted by SGGS-deletion. Without loss of generality, let the representative of $[Q']$ be $B_n \rhd [M_n]$. Then $\{[Q']\} \cup \{B_j \rhd [M_j]\}_{j=1}^{n-1}$ is a partition of $[Q]$, that is, $Gr(Q) = Gr(Q') \cup \bigcup_{j=1}^{n-1} Gr(B_j \rhd M_j)$. Since the model-fixing phase between the first and the second SGGS-extension is finite, even if some of these clauses were subject in turn to s-splitting, the reasoning repeats and the process cannot go on forever, so that the claim is satisfied by clauses in Γ_∞.

Induction hypothesis: for all j, $0 \leqslant j \leqslant k$, for all unit clauses Q for which j is the smallest stage such that $Q \in S_j$, the claim holds.

Inductive case: let Q be a unit clause for which $k+1$ is the smallest stage such that $Q \in S_{k+1}$. This means that Q is a unit hyperresolvent generated by the step $S_k \vdash S_{k+1}$. Let $C = (P \vee \neg L_1 \vee \ldots \vee \neg L_m) \in S \subseteq S_k$ ($m > 0$) be the nucleus, $Q_1, \ldots, Q_m \in S_k$ the positive unit satellites, and τ the simultaneous mgu for this hyperresolution inference (i.e., $\forall i,\ 1 \leqslant i \leqslant m,\ L_i\tau = Q_i\tau$), so that $Q = P\tau$. By induction hypothesis, for all i, $1 \leqslant i \leqslant m$, there exist sets of clauses $\{B_{h_i} \rhd D_{h_i}[M_{h_i}]\}_{h_i=1}^{n_i}$ all in $\Gamma_\infty = dp(\Gamma_\infty)$, such that $Gr(Q_i) \subseteq$

$\bigcup_{h_i=1}^{n_i} Gr(B_{h_i} \rhd M_{h_i})$. Let r be the smallest stage of the SGGS-derivation where all these clauses are in $dp(\Gamma_r)$. By Lemma 3, at most $\Pi_{i=1}^{m} n_i$ SGGS-extensions, applying at stages equal to or greater than r, generate extension clauses $\{A_j \rhd C[P]\alpha_j\}_{j=1}^{h}$, with $h \leqslant \Pi_{i=1}^{m} n_i$, such that $Gr(Q) = Gr(P\tau) \subseteq \bigcup_{j=1}^{h} Gr(A_j \rhd P\alpha_j)$. By the Lifting Theorem for SGGS [3, Thm. 4], extension clauses are not disposable in the trail to which they are added. Since SGGS-move never applies, it cannot be that these clauses are made disposable by clauses that move to their left at subsequent stages. Thus, the clauses in $\{A_j \rhd ([P\alpha_j] \vee \neg L_1\alpha_j \vee \ldots \vee \neg L_m\alpha_j)\}_{j=1}^{h}$ are in Γ_∞, unless some of them is subject to s-splitting during the finite model-fixing phases that follow each SGGS-extension. As already shown in the base case, if this happens, each split clause get replaced by clauses in Γ_∞ with the same ground instances, so that the claim holds. □

The lemma invoked in the above proof is stated and proved below.

Lemma 3. *Given a set S of Horn clauses, assume that hyperresolution generates from nucleus $C = (P \vee \neg L_1 \vee \ldots \vee \neg L_m) \in S$ and satellites $Q_1, \ldots, Q_m$ the unit hyperresolvent $Q = P\tau$, where τ is the simultaneous mgu such that $\forall i,\ 1 \leqslant i \leqslant m,\ L_i\tau = Q_i\tau$. Suppose that for every satellite Q_i, SGGS with I^- has in $dp(\Gamma)$ a set of clauses $\{B_{h_i} \rhd D_{h_i}[M_{h_i}]\}_{h_i=1}^{n_i}$ with positive selected literals, such that $Gr(Q_i) \subseteq \bigcup_{h_i=1}^{n_i} Gr(B_{h_i} \rhd M_{h_i})$. Then at most $\Pi_{i=1}^{m} n_i$ SGGS-extensions with main premise C apply and generate extension clauses $\{A_j \rhd C[P]\alpha_j\}_{j=1}^{h}$ ($h \leqslant \Pi_{i=1}^{m} n_i$), such that $Gr(Q) = Gr(P\tau) \subseteq \bigcup_{j=1}^{h} Gr(A_j \rhd P\alpha_j)$.*

Proof: We show that $Gr(P\tau) \subseteq \bigcup_{j=1}^{h} Gr(A_j \rhd P\alpha_j)$. Let $L \in Gr(P\tau)$, that is, $L = P\tau\sigma$ for some ground instance $C\tau\sigma$ of C. By the first part of the hypothesis, for all i, $1 \leqslant i \leqslant m$, $L_i\tau = Q_i\tau$, and hence $L_i\tau\sigma = Q_i\tau\sigma$, so that $Q_i\tau\sigma$ is a ground instance of Q_i. By the second part of the hypothesis, for all i, $1 \leqslant i \leqslant m$, there exists an h_i ($1 \leqslant h_i \leqslant n_i$) such that $Q_i\tau\sigma \in Gr(B_{h_i} \rhd M_{h_i})$. Take as the side premises the $B_{h_i} \rhd M_{h_i}$, $1 \leqslant i \leqslant m$, such that $Q_i\tau\sigma \in Gr(B_{h_i} \rhd M_{h_i})$. Their positive selected literals $M_{h_1}, \ldots, M_{h_m}$ are simultaneously unifiable with the atoms $L_1, \ldots, L_m$, as witnessed by the unifier $\tau\sigma$. Let α be their most general unifier, and δ the substitution such that $\tau\sigma = \alpha\delta$. Then, SGGS-extension applies with main premise C, side premises $B_{h_i} \rhd M_{h_i}$, $1 \leqslant i \leqslant m$, and simultaneous mgu α, producing extension clause $(\bigwedge_{i=1}^{m} B_{h_i}\alpha) \rhd C[P]\alpha$, such that $L \in Gr((\bigwedge_{i=1}^{m} B_{h_i}\alpha) \rhd P\alpha)$, because $L = P\tau\sigma = P\alpha\delta$. Since an SGGS-extension with main premise C needs m side premises whose positive selected literals unify simultaneously with $L_1, \ldots, L_m$, and for each L_i, $1 \leqslant i \leqslant m$, there are n_i candidates, there are clearly at most $\Pi_{i=1}^{m} n_i$ such extensions. □

The main theorem follows.

Theorem 2. *Given a set of Horn clauses SGGS with I^- halts if and only if HR^+S halts.*

Proof: Since both SGGS and HR^+S are refutationally complete, it suffices to prove the claim for a satisfiable input set S of Horn clauses. Let $S_D \subseteq S$ be the subset of the definite clauses.
(⇒) By way of contradiction, assume that SGGS with I^- halts and HR^+S does not. SGGS halts at a stage k ($k > 0$) with finite limit $\Gamma_k = B_1 \rhd D_1[M_1], \ldots, B_m \rhd D_m[M_m]$ and $I^p(\Gamma_k) = \bigcup_{j=1}^{m} pcgi(B_j \rhd M_j, \Gamma_k) = \bigcup_{j=1}^{m} Gr(B_j \rhd M_j)$, as $\Gamma_k = dp(\Gamma_k)$. By Theorem 1, $I^p(\Gamma_k) = lfp(T_{S_D})$. The HR^+S-derivation is infinite with infinite limit S_∞ containing infinitely many persistent positive unit clauses $L_0, L_1, L_2, \ldots$. Also HR^+S generates

$\mathit{lfp}(T_{S_D})$, represented by $\bigcup_{i\geqslant 0} Gr(L_i)$. Thus, $\bigcup_{j=1}^{m} Gr(B_j \rhd M_j) = \bigcup_{i\geqslant 0} Gr(L_i)$. By Claim (i) in Lemma 2, $\forall j,\ 1\leqslant j\leqslant m$, there exists a $Q_j \in S_\infty$ such that $Gr(B_j \rhd M_j) \subseteq Gr(Q_j)$. Therefore, $\bigcup_{i\geqslant 0} Gr(L_i) = \bigcup_{j=1}^{m} Gr(B_j \rhd M_j) \subseteq \bigcup_{j=1}^{m} Gr(Q_j)$. Since the Q_j's are finitely many, this means that all but finitely many clauses in S_∞ can be subsumed, contradicting the assumption that S_∞ (the limit of a fair derivation) is infinite.
($\Leftarrow$) By way of contradiction, assume that HR^+S halts whereas SGGS with I^- does not. HR^+S halts at some stage k ($k\geqslant 0$) with finite limit S_k. The set $\{L_1,\ldots,L_m\}$ of the positive unit clauses in S_k represents $\mathit{lfp}(T_{S_D})$, in the sense that $\bigcup_{i=1}^{m} Gr(L_i) = \mathit{lfp}(T_{S_D})$. SGGS with I^- produces an infinite derivation with infinite limit $\Gamma_\infty = B_1 \rhd D_1[M_1],\ldots,B_j \rhd D_j[M_j], B_{j+1} \rhd D_{j+1}[M_{j+1}],\ldots$, where $\forall j,\ j\geqslant 0,\ M_i \in D_i^+$ and $I^p(\Gamma_\infty) = \bigcup_{j\geqslant 1} pcgi(B_j \rhd M_j, \Gamma_\infty) = \bigcup_{j\geqslant 1} Gr(B_j \rhd M_j)$ as $\Gamma_\infty = dp(\Gamma_\infty)$. By Theorem 1, $I^p(\Gamma_\infty) = \mathit{lfp}(T_{S_D})$. Thus, $\bigcup_{j\geqslant 1} Gr(B_j \rhd M_j) = \bigcup_{i=1}^{m} Gr(L_i)$. By Claim (ii) in Lemma 2, for all i, $1\leqslant i\leqslant m$, there exist clauses $\{B_{h_i} \rhd D_{h_i}[M_{h_i}]\}_{h_i=1}^{n_i}$ in Γ_∞ such that $Gr(L_i) \subseteq \bigcup_{h_i=1}^{n_i} Gr(B_{h_i} \rhd M_{h_i})$. Therefore, $\bigcup_{j\geqslant 1} Gr(B_j \rhd M_j) = \bigcup_{i=1}^{m} Gr(L_i) \subseteq \bigcup_{i=1}^{m}\bigcup_{h_i=1}^{n_i} Gr(B_{h_i} \rhd M_{h_i})$. Since the L_i's are finitely many, and each of them is covered by finitely many clauses in Γ_∞, this means that all but finitely many clauses in Γ_∞ are disposable, contradicting the assumption that Γ_∞ (the limit of a fair derivation) is infinite. □

Example 3. *Given the following set S of definite clauses*

$\mathsf{P}(x,\mathsf{f}(x))$	(i)	$\neg\mathsf{P}(x,\mathsf{f}(y)) \vee \mathsf{Q}(\mathsf{f}(y),x)$	(ii)
$\neg\mathsf{Q}(\mathsf{f}(x),y) \vee \mathsf{R}(x)$	(iii)	$\neg\mathsf{R}(\mathsf{f}(x)) \vee \mathsf{T}(x)$	(iv)
$\neg\mathsf{R}(x) \vee \mathsf{Q}(\mathsf{a},\mathsf{a})$	(v)		

SGGS with I^- halts after five extensions:

$$\begin{aligned}
\Gamma_0\colon \varepsilon \ &\vdash\ \Gamma_1\colon [\mathsf{P}(x,\mathsf{f}(x))]\\
&\vdash\ \Gamma_2\colon [\mathsf{P}(x,\mathsf{f}(x))],\ \neg\mathsf{P}(x,\mathsf{f}(x)) \vee [\mathsf{Q}(\mathsf{f}(x),x)]\\
&\vdash\ \Gamma_3\colon [\mathsf{P}(x,\mathsf{f}(x))],\ \neg\mathsf{P}(x,\mathsf{f}(x)) \vee [\mathsf{Q}(\mathsf{f}(x),x)],\ \neg\mathsf{Q}(\mathsf{f}(x),x) \vee [\mathsf{R}(x)]\\
&\vdash\ \Gamma_4\colon [\mathsf{P}(x,\mathsf{f}(x))],\ \neg\mathsf{P}(x,\mathsf{f}(x)) \vee [\mathsf{Q}(\mathsf{f}(x),x)],\ \neg\mathsf{Q}(\mathsf{f}(x),x) \vee [\mathsf{R}(x)],\\
&\qquad\quad \neg\mathsf{R}(\mathsf{f}(x)) \vee [\mathsf{T}(x)]\\
&\vdash\ \Gamma_5\colon [\mathsf{P}(x,\mathsf{f}(x))],\ \neg\mathsf{P}(x,\mathsf{f}(x)) \vee [\mathsf{Q}(\mathsf{f}(x),x)],\ \neg\mathsf{Q}(\mathsf{f}(x),x) \vee [\mathsf{R}(x)],\\
&\qquad\quad \neg\mathsf{R}(\mathsf{f}(x)) \vee [\mathsf{T}(x)],\ \neg\mathsf{R}(x) \vee [\mathsf{Q}(\mathsf{a},\mathsf{a})]
\end{aligned}$$

as $I[\Gamma_5] \models S$. Hyperresolution halts after generating the selected literals in Γ_5 as unit clauses.

Thanks to its model-based character and model representation via selected literals, SGGS can *learn* from a *Horn subproblem* of a non-Horn first-order problem a *literal selection strategy* that is useful for termination on the non-Horn problem.

Definition 3. *A clause set S_H is a* Horn subproblem *of a clause set S if for every $C\in S$ the set S_H contains a maximal subclause $C' \subseteq C$ that is Horn, and S_H contains no other clauses.*

Let S_H be a Horn subproblem of a non-Horn first-order problem S. Note that if $J \models S_H$ then $J \models S$ for a Herbrand interpretation J. A Horn subproblem S_H defines a *literal selection strategy*, that can be used when SGGS with I^- is applied to S, as follows. Whenever SGGS-extension adds an instance $C\alpha$ of a clause $C \in S$, such that $C\alpha$ has multiple positive literals that can be selected,[1] the strategy picks the positive literal $L\alpha$ such that L appears in the maximal Horn subclause $C' \subseteq C$ in S_H. While S may have more than one Horn subproblem S_H, the strategy is uniquely defined once an S_H is chosen. The next example portrays a situation where SGGS with I^- halts with the literal selection based on S_H and diverges otherwise.

Example 4. *Let* $S = \{(i)\ \mathsf{P}(x,\mathsf{a}),\ (ii)\ \neg\mathsf{P}(x,y) \vee \mathsf{R}(y) \vee \mathsf{P}(x,\mathsf{f}(y)),\ (iii)\ \neg\mathsf{R}(\mathsf{f}(x)) \vee \neg\mathsf{P}(x,\mathsf{f}(x))\}$ *and* $S_H = \{(i),\ (ii')\ \neg\mathsf{P}(x,y) \vee \mathsf{R}(y),\ (iii)\}$. *Given* S_H*, SGGS with* I^- *halts in two steps:*

$$\varepsilon \vdash [\mathsf{P}(x,\mathsf{a})] \vdash [\mathsf{P}(x,\mathsf{a})],\ \neg\mathsf{P}(x,\mathsf{a}) \vee [\mathsf{R}(\mathsf{a})].$$

If the literal selection strategy defined by S_H *is adopted, SGGS with* I^- *halts also on* S*:*

$$\varepsilon \vdash [\mathsf{P}(x,\mathsf{a})] \vdash [\mathsf{P}(x,\mathsf{a})],\ \neg\mathsf{P}(x,\mathsf{a}) \vee [\mathsf{R}(\mathsf{a})] \vee \mathsf{P}(x,\mathsf{f}(\mathsf{a})).$$

In both derivations $I^p(\Gamma_2)$ *is* $\{\mathsf{P}(\mathsf{f}^k(\mathsf{a}),\mathsf{a}) : k \geqslant 0\} \cup \{\mathsf{R}(\mathsf{a})\}$ *and* $I[\Gamma_2]$ *satisfies both* S_H *and* S*. On the other hand, if the last literal in* (ii)*'s instances is systematically selected, SGGS diverges:*

$$\begin{aligned}\varepsilon &\vdash [\mathsf{P}(x,\mathsf{a})] \vdash [\mathsf{P}(x,\mathsf{a})],\ \neg\mathsf{P}(x,\mathsf{a}) \vee \mathsf{R}(\mathsf{a}) \vee [\mathsf{P}(x,\mathsf{f}(\mathsf{a}))]\\ &\vdash [\mathsf{P}(x,\mathsf{a})],\ \neg\mathsf{P}(x,\mathsf{a}) \vee \mathsf{R}(\mathsf{a}) \vee [\mathsf{P}(x,\mathsf{f}(\mathsf{a}))],\ \neg\mathsf{P}(x,\mathsf{f}(\mathsf{a})) \vee \mathsf{R}(\mathsf{f}(\mathsf{a})) \vee [\mathsf{P}(x,\mathsf{f}^2(\mathsf{a}))]\\ &\vdash \ldots.\end{aligned}$$

Hyperresolution also halts on S_H*, but if given* S *it generates an infinite series of hyperresolvents where both depth and number of literals increase:* $\mathsf{R}(\mathsf{a}) \vee \mathsf{P}(x,\mathsf{f}(\mathsf{a}))$, $\mathsf{R}(\mathsf{a}) \vee \mathsf{R}(\mathsf{f}(\mathsf{a})) \vee \mathsf{P}(x,\mathsf{f}^2(\mathsf{a}))$, $\mathsf{R}(\mathsf{a}) \vee \mathsf{R}(\mathsf{f}(\mathsf{a})) \vee \mathsf{R}(\mathsf{f}^2(\mathsf{a})) \vee \mathsf{P}(x,\mathsf{f}^3(\mathsf{a})), \ldots$. *Furthermore, in this example, the nontermination of hyperresolution cannot be cured by an ordering* $\succ$ *on literals and the restriction to resolve only upon maximal literals in satellites. Indeed, the first satellite* (i) *is a unit clause, and in the generated hyperresolvents no ground literal can dominate the non-ground literal. For example,* $\mathsf{R}(\mathsf{a}) \not\succ \mathsf{P}(x,\mathsf{f}(\mathsf{a}))$ *as* $\mathsf{R}(\mathsf{a})$ *and* x *are incomparable.*

The following theorem generalizes the observation that termination of SGGS on a satisfiable Horn subproblem corresponds to termination on the original non-Horn problem.

Theorem 3. *Given a set* S *of clauses and* I^- *as initial interpretation, if* S *has a satisfiable Horn subproblem* S_H *for which there exists a fair, finite SGGS-derivation, then there exists a fair finite SGGS-derivation for* S *that selects the same literals and hence builds the same model.*

Proof: Let Θ and Θ' denote SGGS-derivations with initial interpretation I^- and input S and S_H respectively. We assume that Θ' is fair and finite: $\Gamma'_0 \vdash \Gamma'_1 \vdash \cdots \vdash \Gamma'_k$. We show by induction on the length k of Θ' that there exists a fair, finite derivation $\Theta\colon \Gamma_0 \vdash \Gamma_1 \vdash \cdots \vdash \Gamma_k$ of the

[1] The SGGS-extension is non-conflicting and $C\alpha$ has more than one positive literal with pcgi's, or $C\alpha$ is a non-I^--all-true conflict clause where any positive literal can be selected.

same length, such that for all stages j $(0 \leqslant j \leqslant k)$ and indices h $(1 \leqslant h \leqslant |\Gamma_j|)$ if Γ'_j has at index h clause $A_h \rhd C'_h[L_h]$ then Γ_j has at index h a clause $A_h \rhd C_h[L_h]$ with the same selected constrained literal. Since the induced partial interpretation is determined by the selected literals, it follows that for all stages j $(0 \leqslant j \leqslant k)$ it holds that $I^p(\Gamma_j) = I^p(\Gamma'_j)$ and hence $I[\Gamma_j] = I[\Gamma'_j]$.
Base case: $k = 0$: Γ'_0 is empty and so is Γ_0.
Induction hypothesis: the claim holds for length k.
Inductive case: let the length of Θ' be $k + 1$. As described in Sect. 3 and by Lemma 1, since S_H is a satisfiable Horn set, no conflicts can occur in Θ', so that $\Gamma'_k \vdash \Gamma'_{k+1}$ can only be a non-conflicting extension or an s-splitting or an SGGS-deletion step. If $\Gamma'_k \vdash \Gamma'_{k+1}$ is a non-conflicting extension with main premise $C' \in S_H$ and side premises $D'_1[M_1], \ldots, D'_n[M_n]$ in $dp(\Gamma'_k)$, it adds extension clause $C'[L]\alpha$, where α is the simultaneous mgu of all the I^--true (i.e., negative) literals $\neg L_1, \ldots, \neg L_n$ of C' with $M_1, \ldots, M_n$. Since S_H is a Horn subproblem of S, there must be a clause $C \in S$ such that $C' \subseteq C$. Furthermore, $(C')^- = C^-$ because C' is a maximal Horn subclause. By induction hypothesis, there exist clauses $D_1[M_1], \ldots D_n[M_n]$ in $dp(\Gamma_k)$ with the same selected literals. Therefore, Θ can be extended with a non-conflicting extension $\Gamma_k \vdash \Gamma_{k+1}$ with the same mgu and extension clause $C[L]\alpha$. If $\Gamma'_k \vdash \Gamma'_{k+1}$ is an s-splitting step, it replaces a clause $A_h \rhd C'_h[L_h]$ by $split(C'_h, C'_r)$ for some $A_r \rhd C'_r[L_r]$ $(r < h)$, where $split(C'_h, C'_r)$ is a partition $\{B_i \rhd D'_i[M_i]\}_{i=1}^n$ such that $Gr(A_h \rhd C'_h) = \bigcup_{i=1}^n \{Gr(B_i \rhd D'_i)\}$. By induction hypothesis, Γ_k has constrained clauses $A_h \rhd C_h[L_h]$ and $A_r \rhd C_r[L_r]$ at positions h and r respectively. Since the partition in SGGS-splitting is determined by the selected literals, Θ can be extended with an s-splitting step that replaces $A_h \rhd C_h[L_h]$ by $split(C_h, C_r)$, where $split(C_h, C_r)$ is a respective partition $\{B_i \rhd D_i[M_i]\}_{i=1}^n$ such that $Gr(A_h \rhd C_h) = \bigcup_{i=1}^n \{Gr(B_i \rhd D_i)\}$. If $\Gamma'_k \vdash \Gamma'_{k+1}$ deletes $A_h \rhd C'_h[L_h]$, it means that this clause is disposable, that is, $I^p(\Gamma'_k|_{h-1}) \models A_h \rhd C'_h[L_h]$. By induction hypothesis, Γ_k has a clause $A_h \rhd C_h[L_h]$ at index h and $I^p(\Gamma_k|_{h-1}) = I^p(\Gamma'_k|_{h-1})$. Thus, $A_h \rhd C_h[L_h]$ is disposable in Γ_k and Θ can be extended with a deletion step that removes it. Since Θ is made of inferences mirroring those in Θ' and Θ' is fair, also Θ is fair. □

Suppose that S is a first-order problem that is not in an SGGS-decidable fragment [13]. If S is unsatisfiable, SGGS halts by refutational completeness. Otherwise, one can look for a Horn subproblem S_H of S such that SGGS terminates on S_H, and apply SGGS to S with the literal selection dictated by S_H. If S_H is satisfiable, SGGS halts with a model of both S and S_H. If S_H is unsatisfiable, S may still be satisfiable, SGGS may halt or not, but the literal selection induced by S_H is not the cause of non-termination: if it does not lead to a model, conflicts arise, via conflict solving SGGS selects other literals, and searches elsewhere.

5. On the Length of SGGS-Derivations

Let S be a set of clauses and $\mathcal{A}$ its Herbrand base. A finite subset $\mathcal{B} \subseteq \mathcal{A}$ is a *finite basis.* An SGGS-derivation *is in* $\mathcal{B}$, if all cgi's of all clauses on the trail during the derivation are made of atoms in $\mathcal{B}$. A fair SGGS-derivation in a finite basis is finite [13, Thm. 1]. An SGGS-derivation is *ground*, if all clauses on the trail during the derivation are ground. The following example shows that a ground derivation may arise also if the input is not ground.

Example 5. *Let S_n be the following parametric set of Horn clauses ($n \geqslant 0$):*
$\{(i)\ \mathsf{P}(\mathsf{f}^n(\mathsf{a})),\ (ii)\ \neg\mathsf{P}(\mathsf{f}(x)) \vee \mathsf{P}(x),\ (iii)\ \neg\mathsf{P}(x) \vee \neg\mathsf{P}(\mathsf{f}(x)) \vee \ldots \vee \neg\mathsf{P}(\mathsf{f}^n(x))\}$. *These sets belong to an SGGS-decidable class named* restrained *where SGGS-derivations are ground and in a finite basis [13]. The finite basis for S_n is $\mathcal{B}_n = \{\mathsf{P}(\mathsf{f}^k(\mathsf{a})) : 0 \leqslant k \leqslant n\}$. The length of the SGGS-derivation with I^- is linear in n:*

$$
\begin{array}{llr}
\Gamma_0 \colon \varepsilon & \vdash \Gamma_1 \colon [\mathsf{P}(\mathsf{f}^n(\mathsf{a}))] & \text{extend } (i) \\
& \vdash \Gamma_2 \colon [\mathsf{P}(\mathsf{f}^n(\mathsf{a}))],\ \neg\mathsf{P}(\mathsf{f}^n(\mathsf{a})) \vee [\mathsf{P}(\mathsf{f}^{n-1}(\mathsf{a}))] & \text{extend } (ii) \\
& \vdash \Gamma_3 \colon [\mathsf{P}(\mathsf{f}^n(\mathsf{a}))],\ \neg\mathsf{P}(\mathsf{f}^n(\mathsf{a})) \vee [\mathsf{P}(\mathsf{f}^{n-1}(\mathsf{a}))], & \\
& \qquad \neg\mathsf{P}(\mathsf{f}^{n-1}(\mathsf{a})) \vee [\mathsf{P}(\mathsf{f}^{n-2}(\mathsf{a}))] & \text{extend } (ii) \\
& \vdash \ldots & \\
& \vdash \Gamma_{n+1} \colon \ \ldots,\ \neg\mathsf{P}(\mathsf{f}(\mathsf{a})) \vee [\mathsf{P}(\mathsf{a})] & \text{extend } (ii) \\
& \vdash \Gamma_{n+2} \colon \ \ldots, [\neg\mathsf{P}(\mathsf{a})] \vee \ldots \vee \neg\mathsf{P}(\mathsf{f}^n(\mathsf{a})) & \text{extend } (iii) \\
& \vdash \Gamma_{n+3} \colon \ \ldots, [\neg\mathsf{P}(\mathsf{a})] \vee \ldots \vee \neg\mathsf{P}(\mathsf{f}^n(\mathsf{a})), \neg\mathsf{P}(\mathsf{f}(\mathsf{a})) \vee [\mathsf{P}(\mathsf{a})] & \text{move} \\
& \vdash \Gamma_{n+4} \colon \ \ldots, [\neg\mathsf{P}(\mathsf{f}(\mathsf{a}))] \vee \ldots \vee \neg\mathsf{P}(\mathsf{f}^n(\mathsf{a})) & \text{resolve} \\
& \vdash \ldots & \\
& \vdash \Gamma_{3n+4} \colon \ \bot, \ldots & \text{resolve}
\end{array}
$$

*where the derivation length is $3n+4 = n+4+2n$ as it takes 2 steps (*move *and* resolve*) to eliminate each of the n literals in the last clause in Γ_{n+4}. Hyperresolution generates $\mathsf{P}(\mathsf{f}^{n-1}(\mathsf{a})), \ldots, \mathsf{P}(\mathsf{a})$ and then $\bot$. Unrestricted resolution does not stay in $\mathcal{B}$ (e.g., resolving (i) upon $\neg\mathsf{P}(\mathsf{f}^{n-1}(x))$ in (iii) gets a clause including $\neg\mathsf{P}(\mathsf{f}^{n+1}(\mathsf{a}))$), and may generate exponentially many clauses in the worst case (e.g., the 2^{n+1} subclauses of the instance of (iii) where $x \leftarrow \mathsf{a}$).*

Theorem 4. *Given a set S of Horn clauses, for all fair SGGS-derivations with I^- as initial interpretation, if the derivation is ground and in a finite basis $\mathcal{B}$, its length is linear in $|\mathcal{B}|$.*

Proof: By [13, Lem. 1], for all stages j ($j \geqslant 0$) of a fair derivation in a finite basis $\mathcal{B}$, $|\Gamma_j| \leqslant |\mathcal{B}|+1$ and $|\Gamma_j| \leqslant |\mathcal{B}|$ if $dp(\Gamma_j) = \Gamma_j$. Let Θ denote a fair ground derivation in $\mathcal{B}$. Since Θ is ground, $dp(\Gamma_j) = \Gamma_j$ holds at all stages, because ground literals have no intersection. Also, Θ is finite by [13, Thm. 1]. Recall the behavior of SGGS with I^- from Sect. 3. Since Θ is ground, no SGGS-splitting (and hence no SGGS-deletion) applies, and the model-constructing phase of Θ is made only of SGGS-extensions. If S is satisfiable, no conflict ever arises, Θ itself is made only of SGGS-extensions, and $|\Gamma_j| \leqslant |\mathcal{B}|$ at its final stage j means that the number of SGGS-extensions and hence $|\Theta|$ is in $\mathcal{O}(|\mathcal{B}|)$. If S is unsatisfiable, a conflict with an I^--all-true conflict clause must arise. Let $\Gamma_i \vdash \Gamma_{i+1}$ be the conflicting extension that adds to the trail such a conflict clause. By Lemma 1, Θ is a refutation, and the number of SGGS-resolutions after stage $i+1$ is equal to the cardinality of the dependence set of the conflict clause, which is bounded by $|\Gamma_i|$, hence by $|\mathcal{B}|$. Since Θ is ground, SGGS-factoring and left splitting do not apply during the conflict-solving phase. As there is at most one application of SGGS-move for every resolution step, the length of the conflict-solving phase is in $\mathcal{O}(|\mathcal{B}|)$. Since the length of both model-constructing and conflict-solving phases is in $\mathcal{O}(|\mathcal{B}|)$, the claim follows. □

problem class	# sets	Koala (I^-) SAT	Koala (I^-) UNS	Koala (I^+) SAT	Koala (I^+) UNS	E 2.4 SAT	E 2.4 UNS	Vampire 4.4 SAT	Vampire 4.4 UNS	iProver 3.5 SAT	iProver 3.5 UNS
Horn	1,220	131	581	66	467	43	889	79	969	106	970

Table 1
Results of Koala, E, Vampire, and iProver on Horn problems within 300 sec of wall-clock time.

	all	SAT	UNS	EPR	Stratified	Restrained	Sort-restrained
average	148	260	123	227	191	21	175
median	37	120	31	121	58	7	40

Table 2
Average and median length of the derivations by Koala on Horn problems.

As a corollary, if S is *ground*, no inference introduces new atoms, and $\mathcal{B}$ is given by the set of the ground atoms occurring in S. While also hyperresolution behaves linearly, given a set of ground Horn clauses, other resolution-based strategies (e.g., positive or negative resolution, ordered resolution) can still generate an exponential number of clauses [15, Ch. 1].

6. Experiments

We focus on Horn problems as Horn logic is the subject of this paper. We wrote a script that considers all 5,000 problems without equality in TPTP 7.4.0, transforms problems into clausal form if needed and detects Horn problems. This resulted in 1,220 Horn problems that we submitted to the SGGS prototype prover Koala,[2] with I^- or I^+ as initial interpretation, E, Vampire, and iProver. All experiments were run single-threaded on a 12-core Intel i7-5930K 3.50GHz machine with 32GB of main memory.

Table 1 reports how many problems were found satisfiable (SAT) or unsatisfiable (UNS) by each tool. Since Koala with I^- succeeded on 712 of the 1,220 Horn sets (58% success rate), whereas Koala with I^+ succeeded on 633 of them (51% success rate), the results suggest that I^- is more effective than I^+ on Horn problems. In comparison with the other provers, Koala is ahead on the satisfiable instances but remains behind on the unsatisfiable ones. Indeed, SGGS works by building models, whereas resolution-based strategies are notoriously very good at discovering the unsatisfiability of Horn sets.

Table 2 displays data about the length of SGGS-derivations produced by Koala, distinguishing between satisfiable and unsatisfiable Horn problems, and among four classes that are SGGS-decidable because they admit finite bases [13].

[2]Koala is available at https://github.com/bytekid/koala, the experimental data at http://cl-informatik.uibk.ac.at/users/swinkler/koala/horn.html or http://profs.sci.univr.it/~bonacina/sggs.html.

7. Discussion

We studied what happens when SGGS with sign-based semantic guidance is applied to Horn clauses. If the input is Horn, SGGS reasons *forward* (a.k.a. *bottom-up*) or *backward* (a.k.a. *top-down*), depending on whether the guiding interpretation is *all-negative* (I^-) or *all-positive* (I^+). SGGS with I^- generates the least fixpoint model of a set of definite clauses, and if the input is unsatisfiable, the first conflict with negative conflict clause announces a refutation: while hyperresolution may get $\perp$ in one step, SGGS may get it with a multi-step conflict explanation and solving phase, because SGGS builds a model and has to fix it when a conflict arises.

SGGS with I^- and hyperresolution with subsumption have the same termination behavior on Horn (note that SGGS-deletion and subsumption are not equivalent, because SGGS-deletion depends on the order of appearance of clauses on the trail, e.g. [3, Ex. 2 and Lem. 1]). Furthermore, SGGS can *learn* from a Horn subproblem of a non-Horn problem a literal selection strategy useful for termination on the non-Horn problem. The model-based character of SGGS pays off on satisfiable inputs, as seen in the experiments comparing the Koala prototype with mature theorem provers such as Vampire [16], E [17], and iProver [18].

Horn problems are not the playground that SGGS was designed for, and hence it is plausible that they do not exercise all its features, such as its conflict-driven nature and its similarity to CDCL. This is not surprising, because the approaches to generalize features of DPLL-CDCL from propositional to first-order logic (e.g., first-order splitting [19], the model evolution calculus [20], SGGS [3]) were designed having non-Horn problems in mind, as resolution-based strategies are efficient on Horn problems. Thus, the results of this paper show that SGGS behaves well also in a realm that was not its motivating target.

A main direction for future work is to endow SGGS with equality reasoning by building the equality axioms in both inference system and model representation. The model-based and conflict-driven style of SGGS also makes it an attractive candidate for integration in frameworks for satisfiability modulo theories such as CDSAT [21, 22]. The Koala prototype may be developed in many ways, including preprocessing the input w.r.t. unit resolution, unit-resulting resolution, and subsumption, and devising caching techniques to avoid recomputations.

Acknowledgements This work was done while the first author was visiting the Computer Science Laboratory of SRI International, whose support is greatly appreciated.

References

[1] M. P. Bonacina, D. A. Plaisted, SGGS theorem proving: an exposition, in: S. Schulz, L. D. Moura, B. Konev (Eds.), Proc. PAAR-4 (2014), volume 31 of *EPiC Series in Computing*, EasyChair, 2015, pp. 25–38. doi:10.29007/m2vf.

[2] M. P. Bonacina, D. A. Plaisted, Semantically-guided goal-sensitive reasoning: model representation, J. Autom. Reason. 56 (2016) 113–141. doi:10.1007/s10817-015-9334-4.

[3] M. P. Bonacina, D. A. Plaisted, Semantically-guided goal-sensitive reasoning: inference system and completeness, J. Autom. Reason. 59 (2017) 165–218. doi:10.1007/s10817-016-9384-2.

[4] J. Marques Silva, I. Lynce, S. Malik, Conflict-driven clause learning SAT solvers, in: A. Biere, M. Heule, H. Van Maaren, T. Walsh (Eds.), Handbook of Satisfiability, volume 185 of *Frontiers in AI and Appl.*, IOS Press, 2009, pp. 131–153. doi:10.3233/978-1-58603-929-5-131.

[5] N. Bjørner, A. Gurfinkel, K. McMillan, A. Rybalchenko, Horn clause solvers for program verification, in: L. D. Beklemishev, A. Blass, N. Dershowitz, B. Finkbeiner, W. Schulte (Eds.), Fields of Logic and Computation II – Essays Dedicated to Yuri Gurevich, volume 9300, Springer, 2015, pp. 24–51. doi:10.1007/978-3-319-23534-9_2.

[6] J. C. C. McKinsey, The decision problem for some classes of sentences without quantifiers, J. Symb. Log. 8 (1943) 61–76. doi:10.2307/2268172.

[7] A. Horn, On sentences which are true of direct unions of algebras, J. Symb. Log. 16 (1951) 14–21. doi:10.2307/2268661.

[8] R. A. Kowalski, M. H. Van Emden, The semantics of predicate logic as a programming language, J. ACM 23 (1976) 733–742. doi:10.1145/321978.321991.

[9] K. R. Apt, M. H. Van Emden, Contributions to the theory of logic programming, J. ACM 29 (1982) 841–862. doi:10.1145/322326.322339.

[10] J. W. Lloyd, Foundations of Logic Programming, Symbolic Computation-Artificial Intelligence, Second Extended ed., Springer, Berlin, 1987. doi:10.1007/978-3-642-83189-8.

[11] J. A. Robinson, Automatic deduction with hyper-resolution, Int. J. Comput. Math. 1 (1965) 227–234. doi:10.2307/2272384.

[12] J. R. Slagle, Automatic theorem proving with renamable and semantic resolution, J. ACM 14 (1967) 687–697. doi:10.1007/978-3-642-81955-1_4.

[13] M. P. Bonacina, S. Winkler, SGGS decision procedures, in: N. Peltier, V. Sofronie-Stokkermans (Eds.), Proc. IJCAR-10, volume 12166 of *LNAI*, Springer, 2020, pp. 356–374. doi:10.1007/978-3-030-51074-9_20, Long version entitled *Semantically-guided goal-sensitive reasoning: decision procedures and the Koala prover* submitted for publication and available at http://profs.sci.univr.it/~bonacina/sggs.html.

[14] G. Sutcliffe, The TPTP problem library and associated infrastructure. From CNF to TH0, TPTP v6.4.0, J. Autom. Reason. 59 (2017) 483–502. doi:10.1007/s10817-009-9143-8.

[15] D. A. Plaisted, Y. Zhu, The Efficiency of Theorem Proving Strategies, Friedr. Vieweg & Sohn, 1997. doi:10.1007/978-3-663-07847-0.

[16] L. Kovàcs, A. Voronkov, First order theorem proving and Vampire, in: N. Sharygina, H. Veith (Eds.), Proc. CAV-25, volume 8044 of *LNCS*, Springer, 2013, pp. 1–35. doi:10.1007/978-3-642-39799-8_1.

[17] S. Schulz, S. Cruanes, P. Vukmirović, Faster, higher, stronger: E 2.3, in: P. Fontaine (Ed.), Proc. CADE-27, volume 11716 of *LNAI*, Springer, 2019, pp. 495–507. doi:10.1007/978-3-030-29436-6_29.

[18] A. Duarte, K. Korovin, Implementing superposition in iProver, in: N. Peltier, V. Sofronie-Stokkermans (Eds.), Proc. IJCAR-10, volume 12167 of *LNAI*, Springer, 2020, pp. 388–397. doi:10.1007/978-3-030-51054-1_24.

[19] C. Weidenbach, Combining superposition, sorts and splitting, in: J. A. Robinson, A. Voronkov (Eds.), Handbook of Automated Reasoning, volume 2, Elsevier, 2001, pp. 1965–2013. doi:10.1016/b978-044450813-3/50029-1.

[20] P. Baumgartner, B. Pelzer, C. Tinelli, Model evolution with equality - revised and imple-

mented, J. Symb. Comput. 47 (2012) 1011–1045. doi:10.1016/j.jsc.2011.12.031.

[21] M. P. Bonacina, S. Graham-Lengrand, N. Shankar, Conflict-driven satisfiability for theory combination: transition system and completeness, J. of Autom. Reason. 64 (2020) 579–609. doi:10.1007/s10817-018-09510-y.

[22] M. P. Bonacina, S. Graham-Lengrand, N. Shankar, Conflict-driven satisfiability for theory combination: lemmas, modules, and proofs, J. of Autom. Reason. 66 (2022) 43–91. doi:10.1007/s10817-021-09606-y.