We would like to thank the reviewers for their constructive feedback. In the following sections, we discuss how we have addressed your comments and where in the revised paper we have introduced changes:

Reviewer 1

Comment	Changes
While individual works are clearly described, the comparisons remain largely qualitative. Incorporating quantitative analyses such as performance metrics and scalability benchmarks would significantly strengthen the survey. A consolidated table summarizing dataset size, learning regime, and reported accuracy could help readers quickly grasp the empirical landscape of neurosymbolic ASP frameworks	We added a new section with an in-depth comparative analysis on pages 28-31 in lines 914-1034. It includes two new tables summarising the accuracies of every framework on all major datasets. We analyse the performances with reference to the structures of frameworks and the difficulty of the tasks. Our comprehensive comparison discusses the strengths and weaknesses of different methods and establishes areas that need improvement.
in the Scalability Issues section, the discussion is primarily qualitative. As a result, it is unclear to what extent these systems actually face scalability bottlenecks. Including quantitative evidence such as runtime statistics, solver timeouts, or memory usage would help ground the discussion in measurable performance rather than remaining at a high conceptual level.	We have added a runtime analysis in the scalability section on pages 34-36 in lines 1131-1170. A new table compiles the training times of all papers that have reported this information. Again, we analyse why timeouts occur, what structures are more scalable and what open challenges exist.

Reviewer 2: Robin Manhaeve

Comment	Changes
Restricting attention solely to ASP excludes related and relevant neurosymbolic systems. Without a stronger justification for this focus, the contribution risks being too narrow.	We have added a thorough discussion of other neurosymbolic frameworks on pages 4-6 in lines 98-168 and 186-194. It includes systems using propositional and first-order logic, as well prominent papers like logic tensor networks or probabilistic circuits. We put an extended focus on logic programming methods

(i) broaden or better justify its exclusive focus on ASP	including DeepProbLog, Scallop, ABL and MetaAbd. We now justify our focus on ASP in the paper. We reference its high expressivity, allowing practitioners to model commonsense reasoning for example. The main reason we focus on ASP, however, is that it gives us the ability to cover every framework in detail. Widening the focus to all of neurosymbolic logic programming would at least double the number of frameworks to discuss. It would require new graphical representations for all these frameworks and significantly increase the complexity of our comparative analysis, as there is a wider range of different capabilities to consider.
A stronger conclusion section could articulate open challenges and research directions to inspire and guide the community. (iii) expand the conclusion to identify challenges and future directions	We have added open challenges and promising research directions to the conclusion section on page 36. We argue that the biggest open challenge is scalability and that both better implementations and new methods are required. Other challenges include reducing reliance on hard-coding and creating more realistic datasets. We end with a more inspiring call to action geared towards the ability of neurosymbolic AI to discover new knowledge, rather than just summarising existing information.
The paper often emphasizes the superior expressivity of ASP over other formalisms, but it is not clear how frequently this added expressivity is practically relevant in NeSy applications. For many tasks, simpler languages such as Datalog are sufficient, and acknowledging this would provide a more balanced perspective.	We provide a more balanced perspective on the strengths and drawbacks of ASP compared to other formalisms on pages 5-6 in lines 169-185. Rather than just mentioning the increased expressivity of ASP, we also acknowledge superior scalability of languages like Datalog, the merits of guided search in abduction, the readability advantages of query-driven languages like Prolog and unique features such as probabilistic facts in ProbLog.
Much of the discussion is descriptive rather than analytical. The reader is left without a clear comparative synthesis of why certain methods succeed or fail, what the main trade-offs are, and where the field should head next. The section on performance analysis is relatively weak. Claims are not	We added a new section with an in-depth comparative analysis on pages 28-31 in lines 914-1034. It includes two new tables summarising the accuracies of every framework on all major datasets. We analyse the performances with reference to the structures of frameworks and the difficulty of the tasks. Our comprehensive comparison discusses the strengths and weaknesses

well supported with systematic comparisons, and the survey misses an opportunity to synthesize insights from experimental results across different frameworks. (ii) move beyond description to provide deeper comparative and critical analysis (iv) strengthen the performance and related work sections with references and positioning.	of different methods and establishes areas that need improvement. We have also added a runtime analysis on pages 34-36 in lines 1131-1170. A new table compiles the training times of all papers that have reported this information. Again, we analyse why timeouts occur, what structures are more scalable and what open challenges exist.
Likewise, the related work section does not situate this survey with respect to existing, broader NeSy surveys. Readers would benefit from a short discussion of how this work complements or differs from prior surveys in the area (iv) strengthen the performance and related work sections with references and positioning.	On pages 3-4 in lines 57-97, we position our survey in the landscape of existing overview papers. We include over a dozen existing surveys and describe their focus and relation to our work. We conclude that, among the numerous surveys about neurosymbolic AI in general, there are none with an in-depth analysis of ASP frameworks.
The formal definition of semantics (Section 2) uses sets of ground literals, whereas answer sets are typically defined as sets of atoms.	We have adapted the definitions to use atoms rather than literals.