

# Scalability of ARS and Reinforcement Learning for Large-Scale Reasoning Suppression

Assignee Research

May 31, 2026

## Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: How does the scalability of ARS compare to reinforcement learning-based reasoning suppression methods when applied to larger models (e.g., >100B parameters) in terms of throughput and accuracy on the Large Reasoning Language Models (LRLMs or LRLMs) demonstrate remarkable capabilities in complex reasoning tasks, but suffer from significant computational inefficiencies due to overthinking phenomena. Existing efficient reasoning methods face the challenge of balancing reasoning. 8 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: ARS: Adaptive Reasoning Suppression for Efficient Large Reasoning Language Models. Research question: How does the scalability of ARS compare to reinforcement learning-based reasoning suppression methods when applied to larger models (e.g., >100B parameters) in terms of throughput and accuracy on the Big-Bench Hard (BBH) or ARC benchmarks?.

## 2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.5/10.

### 3 Results

11 papers retrieved. 8 claims extracted; 0 independently verified. Quality review score: 3.5/10.

### 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

### 5 Extracted Claims

Claim	Verified	Confidence
ARS achieves superior length reduction while maintaining competitive accuracy across all model scales.	×	0.13
ARS operates through three core components: Multi-checkpoint certainty estimation, Progressive threshold adaptation, and	×	0.14
ARS establishes multiple checkpoints at regular intervals during generation.	×	0.03
At each checkpoint, ARS estimates model certainty through tentative answer probing.	×	0.06
The heuristic difficulty estimation is used to schedule the mode of operation in ARS.	×	0.03
ARS uses different policies (CoDFastPolicy, ElasticModeratePolicy, DeepReflectPolicy) based on the difficulty of the que	×	0.02
ARS sets a maximum token limit of 1200 tokens per response.	×	0.03
ARS aims to minimize the expected output length while preserving reasoning accuracy.	×	0.06

### References

- <http://arxiv.org/abs/2508.04848v1>
- <http://arxiv.org/abs/2509.16679v1>
- <http://arxiv.org/abs/2510.00071v2>