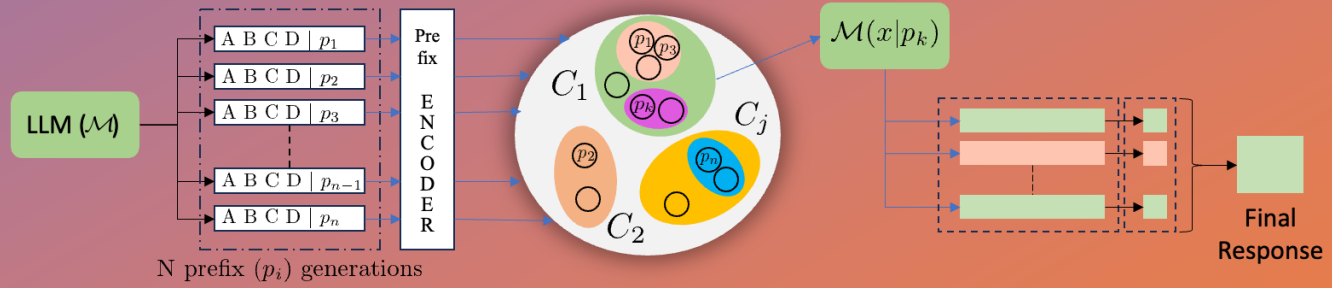
# The Path of Least Resistance: Guiding LLM Reasoning Trajectories With Prefix Consensus

**Ishan Jindal**[a*], Sai Prashanth Akuthota[b], Jayant Taneja[b], Sachin Dev Sharma[b]

[a]Fujitsu Research India

[b]Samsung R&D Institute India

[*]Work conducted while at Samsung R&D Institute India



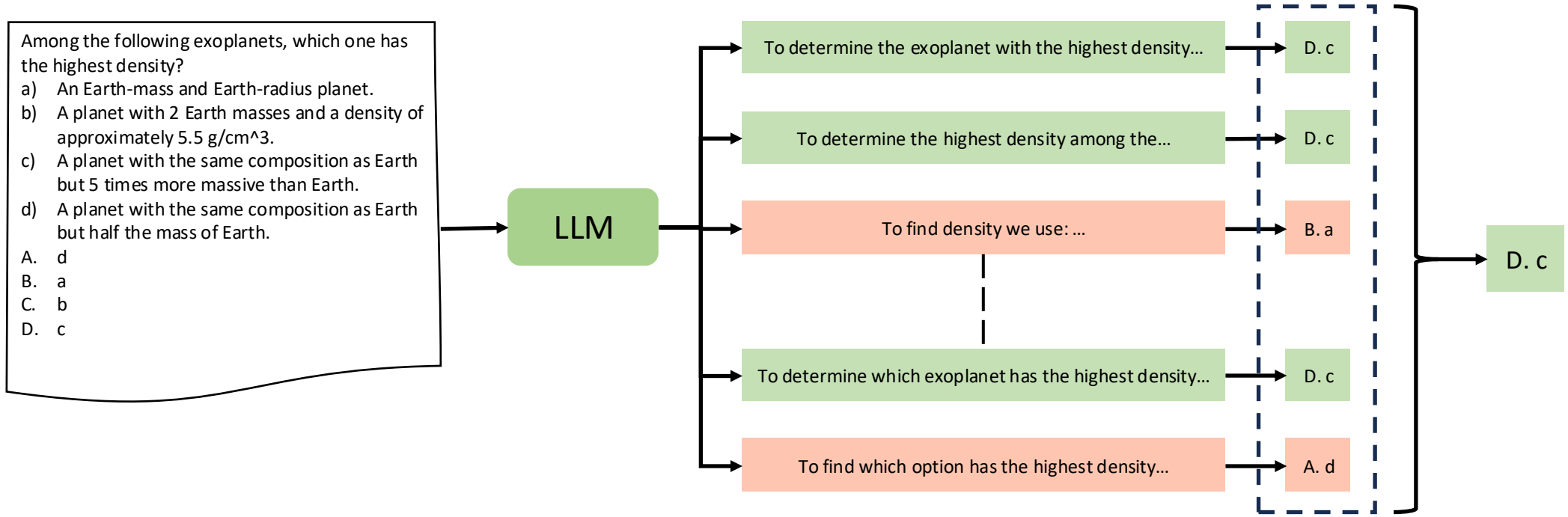

ICLR

# Introduction

PoLR

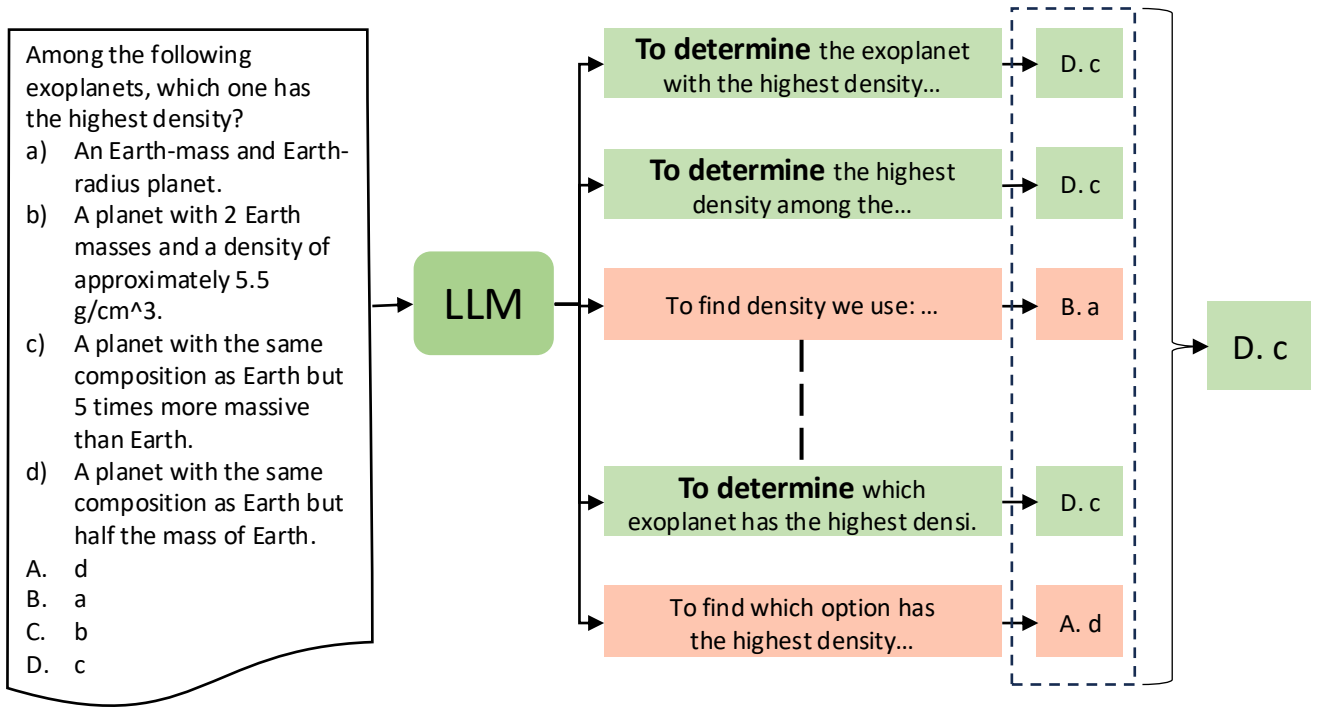
# The Cost of Good Reasoning

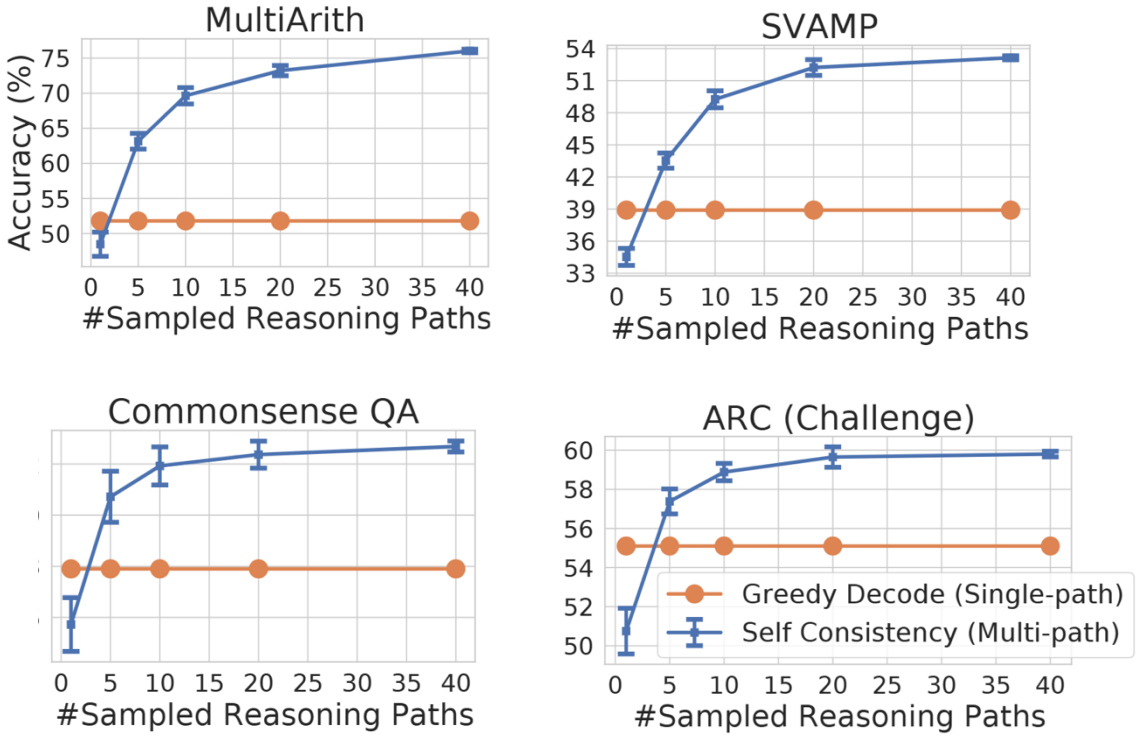
## Background: Self Consistency (SC) Decoding



# The Cost of Good Reasoning

## Background: Self Consistency (SC) Decoding




Wang et al., 2023
(arXiv:2203.11171v4)

# Prior Art

## Exploiting Answer Consistency

### Adaptive Consistency by Aggarwal et al. (2023)

Stop sampling when answer consensus forms

→ Full reasoning trace generation needed

### Early Stopping Self-Consistency by Li et al. (2024)

Halts once a confident majority detected

→ Full reasoning trace generation needed

### Reasoning-Aware Self-Consistency by Wan et al. (2025)

Evaluate reasoning paths, aggregate answers using weighted majority voting

→ Full reasoning trace generation needed

# An Observation

**Question**: Melinda has three empty boxes and 12 textbooks, three of which are mathematics textbooks. One box will hold any three of her textbooks, one will hold any four of her textbooks, and one will hold any five of her textbooks. If Melinda packs her textbooks into these boxes in random order, the probability that all three mathematics textbooks end up in the same box can be written as m n , where m and n are relatively prime positive integers. Find m+n.

The first 32 words (prefix substring) of all answers:

A1: To determine the probability that all three mathematics textbooks end up in the same box, we need to consider the total number of ways to distribute the 12 textbooks into the three
A2: To determine the probability that all three mathematics textbooks end up in the same box, we need to follow these steps: 1. **Calculate the total number of ways to pack the textbooks
A3: To determine the probability that all three mathematics textbooks end up in the same box, we start by calculating the total number of ways to distribute the 12 textbooks into the three
A4: To determine the probability that all three mathematics textbooks end up in the same box, we need to consider the total number of ways to distribute the 12 textbooks into the three
A5: To determine the probability that all three mathematics textbooks end up in the same box, we start by calculating the total number of ways to distribute the 12 textbooks into the three
…

ICLR

# Prior Art

## Exploiting Answer Consistency

### Adaptive Consistency by Aggarwal et al. (2023)

Stop sampling when answer consensus forms

→ Full reasoning trace generation needed

### Early Stopping Self-Consistency by Li et al. (2024)

Halts once a confident majority detected

→ Full reasoning trace generation needed

### Reasoning-Aware Self-Consistency by Wan et al. (2025)

Evaluate reasoning paths, aggregate answers using weighted majority voting

→ Full reasoning trace generation needed

## Exploiting Early Prefixes

### Path Consistency by Zhu et al. (2024)

Estimate the confidence of partial reasoning paths and guide subsequent generations toward promising branches

→Relies on external confidence estimators

### UPFT by Ji et al. (2025)

Shows that prefixes contain rich signals and uses them at *training time* for supervision

→Train-time approach

# Do Prefixes Encode Early Consensus?

DeepSeek-Distill-Qwen-7B

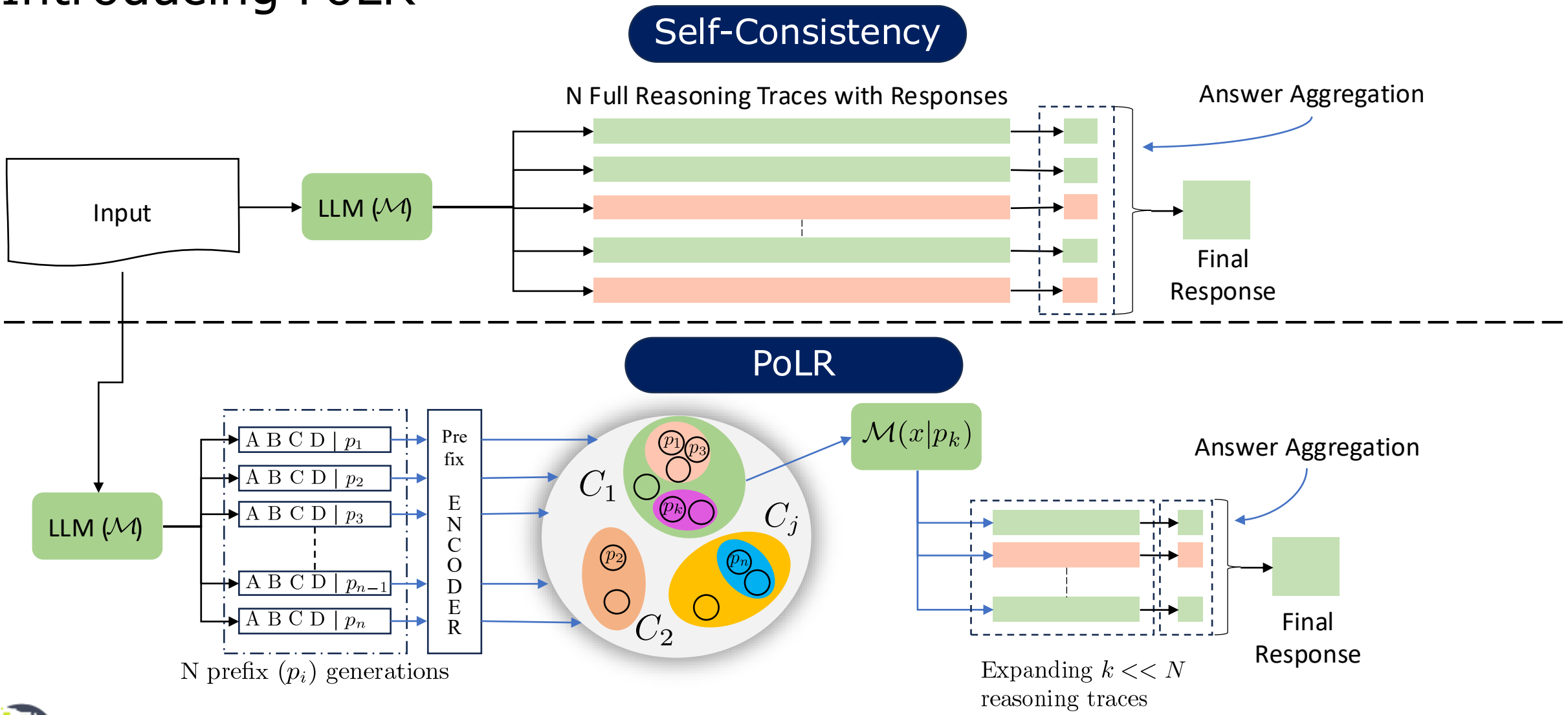
| | $L_p$ | Expansion rate | Accuracy | EPM |
|---|---|---|---|---|
| MATH500 | SC | 1.00 | 89.8 | – |
| | 32 | 0.64 | 89.8 | 125 |
| | 64 | 0.58 | 89.6 | 63 |
| | 128 | 0.48 | 89.2 | 5 |
| | 256 | 0.45 | 89.2 | 0 |

LLMs encode structural agreement well before generating complete answers

# Method

PoLR

# Introducing PoLR



# Experiments

PoLR

# Setup

| Models | Benchmarks | Metrics |
|---|---|---|
| DeepSeek-R1-Distill-Qwen-1.5B | GSM8K | Exact Match (EM) -- GSM8K |
| DeepSeek-R1-Distill-Qwen-7B | MATH500 | Pass@1 -- MATH500, AIME |
| QWQ32B | AIME24 | Accuracy – GPQA-DIAMOND |
| MiMo-7B-RL-0530 | AIME25 | Token Efficiency $\eta = 1 - \frac{T_{PoLR}}{T_{SC}}$ |
| Phi-4 | GPQA-DIAMOND | Path Expansion (PExp) |
| Qwen2.5-Math-7B | STRATEGYQA | PoLR Overhead ($\kappa_t$) |

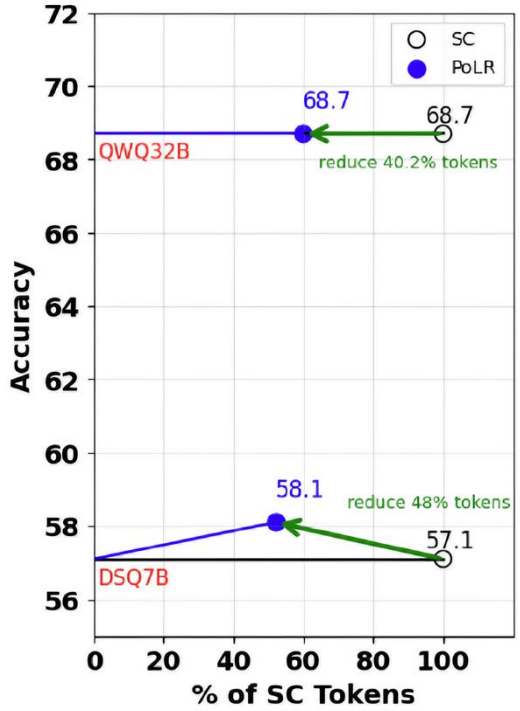
**Defaults**
* Unless stated otherwise

$L_p = 256$ | Agglomerative Clustering | $N \in \{11, 31, 51\}$ | 10 random seeds per experiment

**Baselines**

Self-Consistency (SC) | Adaptive-Consistency (AC) | Single-Sample CoT | Early-Stopping Self-Consistency (ESC)

# Token Savings Without Accuracy Loss




| Dataset | Model | N | SC Acc | PoLR Acc | $\eta$ | Overhead $\kappa_t$ |
|---|---|---|---|---|---|---|
| GSM8K | QwQ32B | 51 | 90.8% | 90.5% | 47.6% | 11 ms |
| Math500 | QwQ32B | 51 | 91.8% | 92.0% | 51.8% | 11 ms |
| GPQA-Diamond | QwQ32B | 51 | 68.7% | 70.2% | 53.8% | 17 ms |

40-60% Token Reduction • <20 ms Overhead • Accuracy Preserved or Improved

ICLR

# Generalizability

| | DSQ1.5B | | DSQ7B | | QWQ32B | | Phi-4-15B | | MiMo-RL-7B | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\Delta_{acc}$ | $\eta$ | $\Delta_{acc}$ | $\eta$ | $\Delta_{acc}$ | $\eta$ | $\Delta_{acc}$ | $\eta$ | $\Delta_{acc}$ | $\eta$ |
| GSM8K | +0.0 | 40.1% | +0.2 | 26.5% | -0.3 | 47.6% | | | | |
| MATH500 | -0.8 | 52.4% | -0.4 | 48.7% | +0.2 | 51.8% | | | | |
| AIME24 | | | -6.7 | 50.9% | +0.0 | 59.7% | +3.3 | 49.5% | | |
| AIME25 | | | +0.0 | 48.8% | -10.0 | 56.8% | +0.0 | 54.7% | | |
| GPQA Diamond | | | -1.5 | 57.1% | +1.5 | 53.8% | | | -0.5 | 51.4% |

* AIME25/QWQ32B −10% = 3 problems out of 30 → all cases where SC itself barely succeeds (consensus <63%)

Consistent across architectures (Qwen, DeepSeek, Phi, MiMo), scales (1.5B–32B), and training paradigms (distillation, RL, SFT)

# Further Insights

PoLR

# Compounding Gains with Adaptive Methods

**N Short Prefixes Generated**
$p_1, p_2, ...p_N$ (each truncated at $L_p$ tokens)

↓

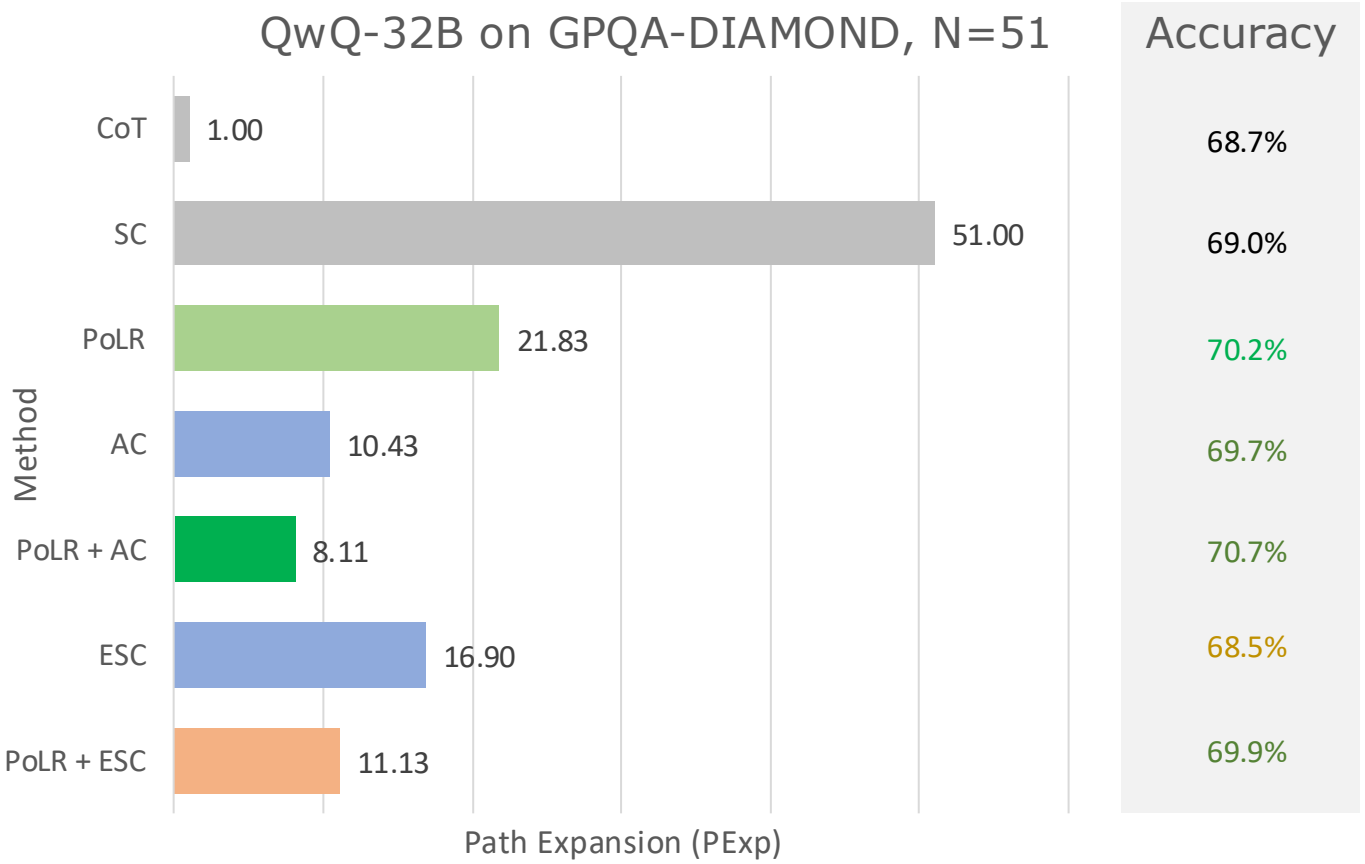
**PoLR: embed, cluster, select C***
Expand all $K$ traces in C* (parallely)

↓

**AC/ESC: Adaptive stopping within C***
Halt confident majority detection(sequential checks)

↓

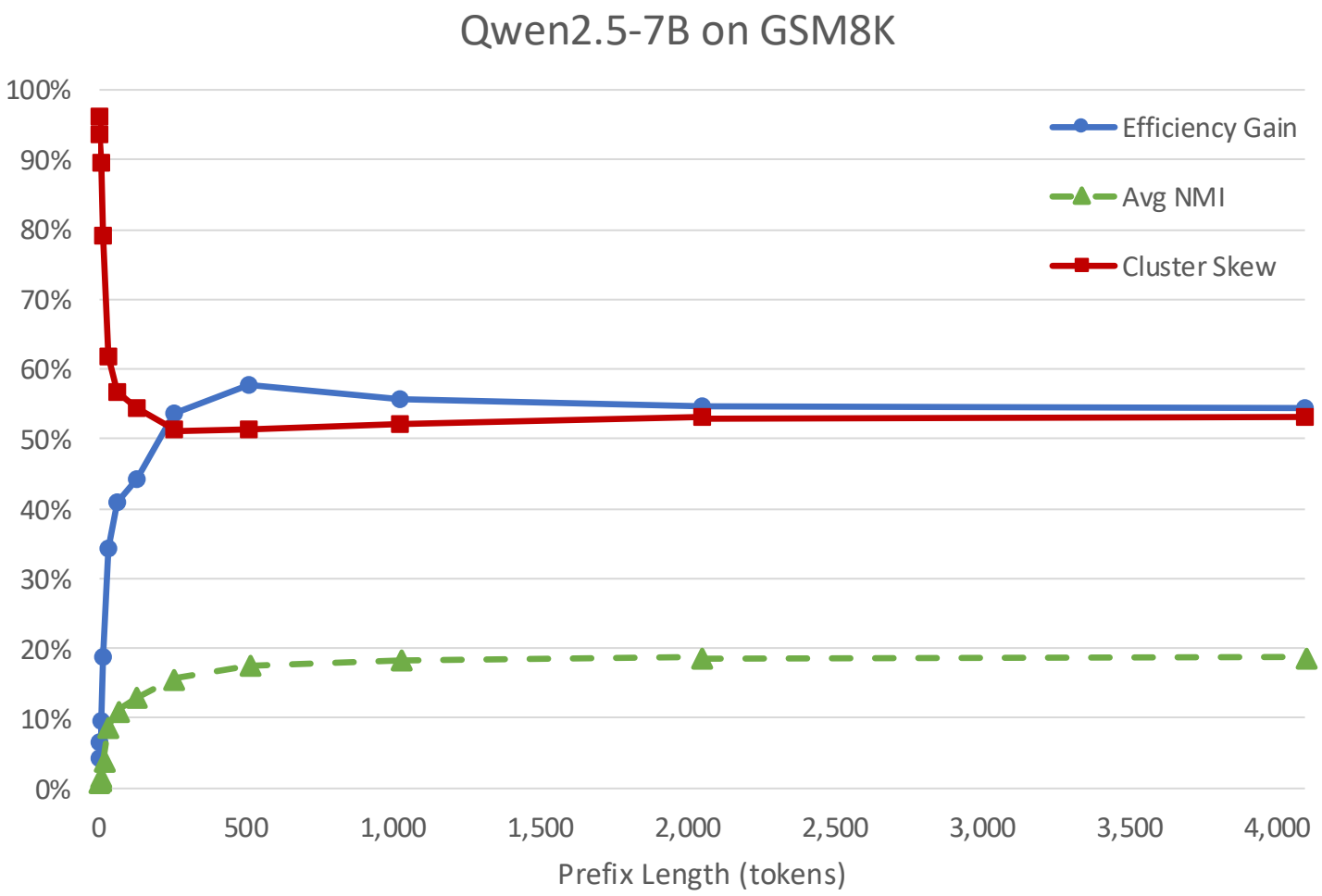
**Majority Vote → Final Answer**




PoLR + adaptive methods achieve 75% token savings on average (compared to SC)

ICLR

# Effect Of Prefix Length




## Key Insights

### Efficiency Plateau

Efficiency plateaus at ~512 tokens (~58% Savings over SC)
→ Longer prefixes don't help further

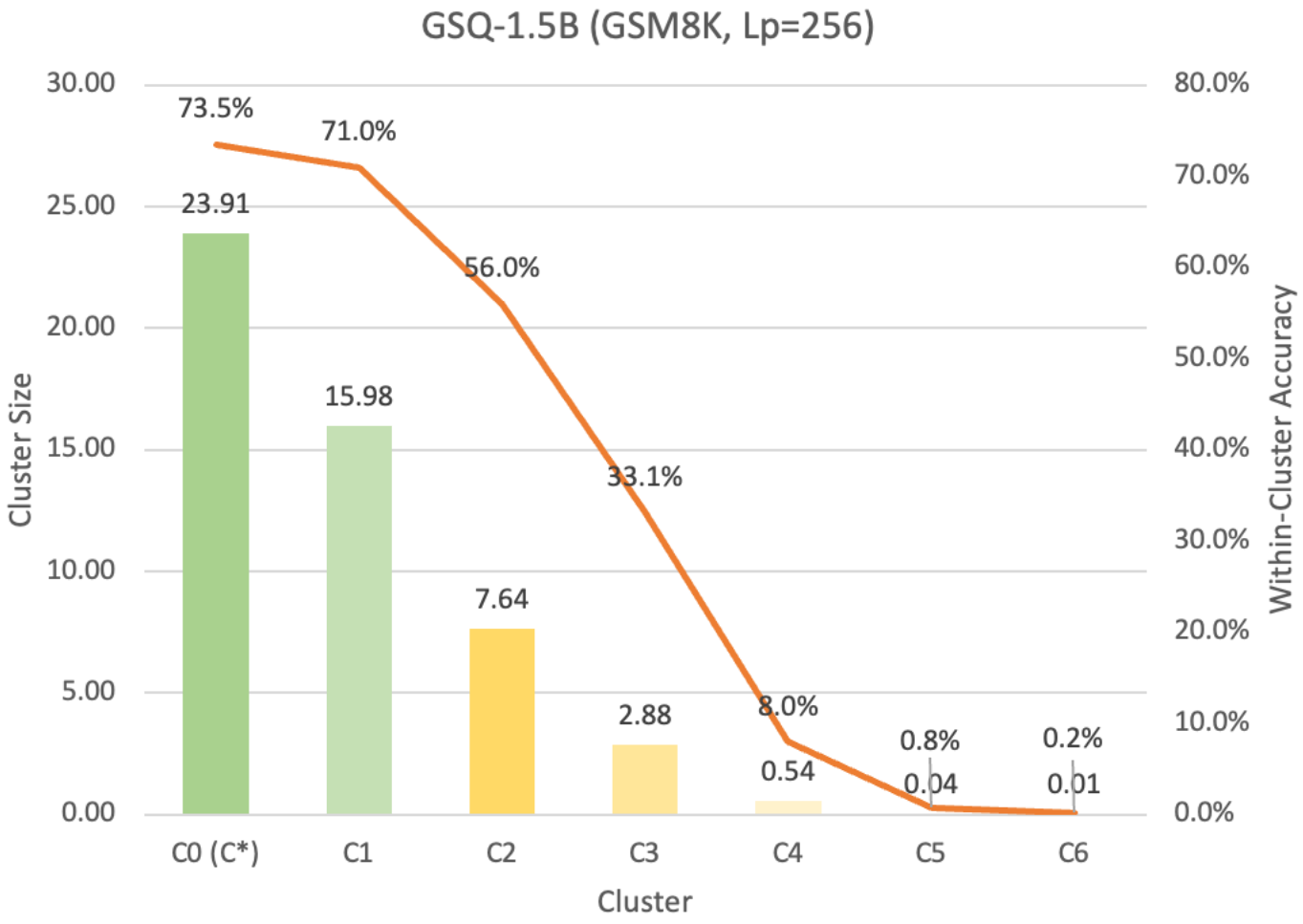
### Homogeneity-Informativeness Tradeoff

Cluster skew drops from ~0.96 (Lp=2) to ~0.52 (Lp=256), then stabilizes.
→ Short prefixes are highly homogeneous but less informative

### Skew Drives Efficiency

Avg NMI remains relatively low (~0.18 at Lp=4096)
→ Correctness alignment is sufficient even at lower prefix lengths

# Cluster Size Predicts Accuracy

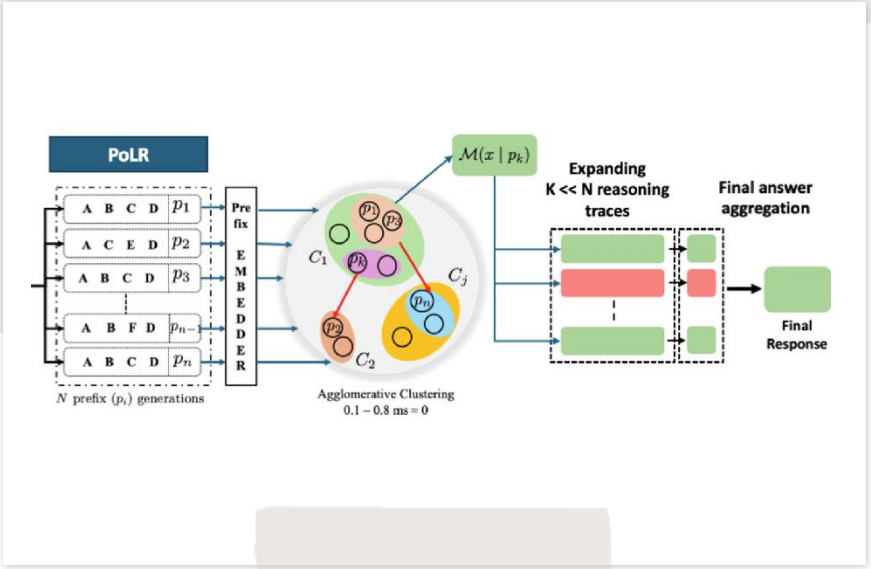



Cluster Sizes and Accuracy Correlation

| | DSQ-7B | QwQ-32B |
|---|---|---|
| GSM8K | 0.90 | 0.90 |
| MATH500 | 0.89 | 0.88 |
| AIME24 | 0.75 | 0.87 |
| AIME25 | 0.76 | 0.78 |
| GPQA | 0.90 | 0.78 |

ICLR

# Summary




**Drop-in SC Replacement:**

- No fine-tuning
- No extra models
- Works at inference

**40-60% Token Savings, up-to 50% latency reduction**

Consistent across tasks, model families and model sizes

**Universal Pre-Filter:**

Composable with AC and ESC, with even higher token savings

**Robust to:**

- Clustering methods
- Prefix lengths
- Cluster Selection strategies

Sai Prashanth Akuthota
a.prashanth@samsung.com

Ishan Jindal
ishan.jindal@fujitsu.com

ICLR

Sachin Dev Sharma
sachin.dev@samsung.com

Jayant Taneja
j.taneja@samsung.com

# Authors

The Path of Least Resistance: Guiding LLM Reasoning Trajectories With Prefix Consensus