ICML
International Conference
On Machine Learning

Small Agent Group is the Future of Digital Health

Yuqiao Meng [1] Luoxi Tang [1] Dazheng Zhang [2] Rafael Brens [1] Elvys J. Romero [1] Nancy Guo [1] Safa Elkefi [1]
Zhaohan Xi [1]

BINGHAMTON UNIVERSITY
STATE UNIVERSITY OF NEW YORK

LEGES SINE MORIBUS VANAE

ICML
International Conference
On Machine Learning

Comparison of Clinical Reasoning

Single Giant LLM Agent

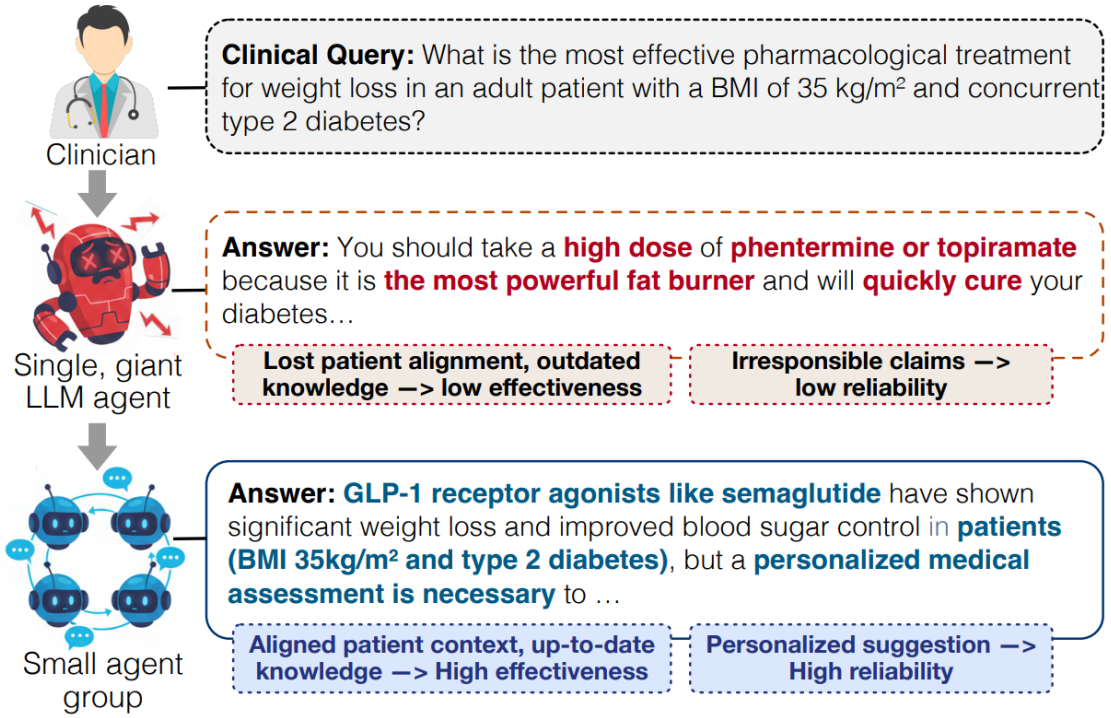
✕Failure Mode:

Outdated knowledge leads to irresponsible claims. Example: Recommending phentermine as a "cure" for diabetes in obese patients.

Small Agent Group (SAG)

✓Success Mode:

Up-to-date knowledge (GLP-1 agonists) with personalized suggestion of medical assessment. High alignment and reliability.



ICML
International Conference
On Machine Learning

Beyond the "Scaling-First" Paradigm

Scaling-First Monoliths

The assumption that clinical intelligence increases solely with model size. However, monolithic LLMs face effectiveness, reliability, and deployment barriers.

Small Agent Groups (SAG)

Distributing reasoning, evidence-based analysis, and audits through collaborative deliberation. Collective expertise vs. monolithic scale.

Synergistic reasoning can substitute for parameter growth in clinical settings.

ICML
International Conference
On Machine Learning

The Healthcare "Impossible Triangle"

Effectiveness

Clinical correctness across multidisciplinary expertise. Monoliths are bounded by training data; SAG uses parallel specialization.

Reliability

Self-critique and calibration. Agents audit each other to quantify uncertainty and reduce hallucinations in high-stakes contexts.

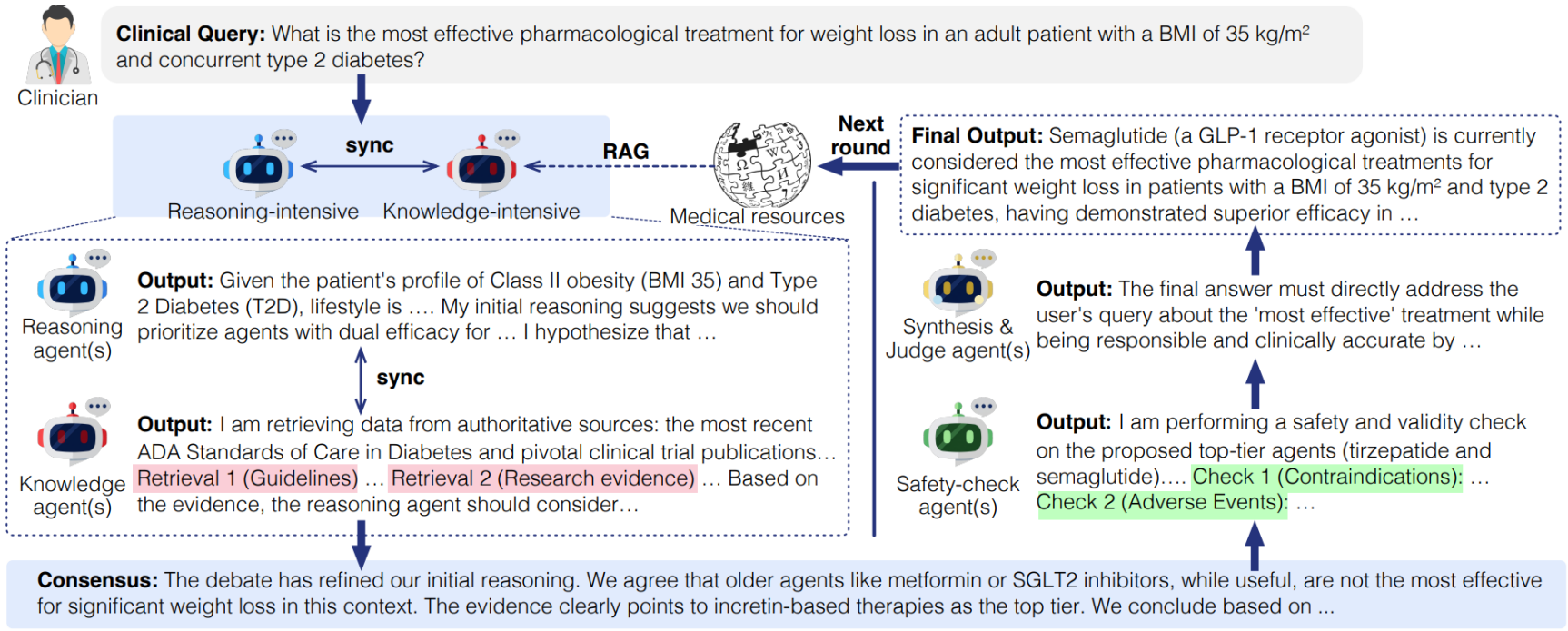
Deployability

Memory constraints for local institution hosting. Lightweight agents allow for manageable GPU footprints and IP protection.

ICML
International Conference
On Machine Learning

SAG Multi-Role Architecture

- **Reasoning**(A_R): Deductive hypothesis formation.
- **Knowledge-providing**(A_K): Evidence-based RAG from PubMed, CDC, FDA.
- **Safety-check**(A_S): Validity and consistency auditing.
- **Synthesis & Judge**(A_J): Final adjudication and consensus.



ICML
International Conference
On Machine Learning

Domain Optimization Strategies

Group Relative Policy Optimization(GRPO)
Optimizes agents based on relative trajectory quality compared to the group mean, stabilizing learning under noisy clinical feedback.

Centralized Training, Decentralized Execution (CTDE)
A centralized critic observes the full debate state during training to assign credit, while agents operate independently during deployment.

Evidence-Based RAG Grounding
Retrieval from Medline, CDC, and FDA ensures conclusions are grounded in latent-free authoritative evidence.

ICML
International Conference
On Machine Learning

Rigorous Evaluation Framework

Metric Dimension	Sub-Metrics	Benchmarks Used
Effectiveness	Correctness, Relevance, Fairness	MedQA, GPQA, PubMedQA, EquityMedQA
Reliability	Safety, Robustness, Consistency	MedSafetyBench, MMLU-Pro, NEJM-MedQA
Deployment Cost	Memory, FLOPs, Latency	Inference profiling (Peak Memory, Runtime)

ICML International Conference On Machine Learning

Superior Performance on Correctness

Method	Model	Llama Backbone (SAG: 3B each, Single: 70B)					Qwen Backbone (SAG: 4B each, Single: 72B)				
		M-QA	MCQA	NEJM	GPQA	Gap ↓	M-QA	MCQA	NEJM	GPQA	Gap ↓
Single, giant LLM	Pre-trained	59.8	51.2	42.4	43.6	17.4	72.0	63.0	55.7	41.1	30.9
	w/ PPO	73.4	70.1	54.6	61.3	18.8	77.2	72.5	57.9	72.0	19.3
	w/ DPO	81.5	74.2	50.8	62.0	30.7	82.4	75.1	72.5	67.3	15.1
Clinical specialist	Meditron	34.9	48.7	28.5	20.4	28.3	*Same results as the left*				
	Me-LLaMA	58.0	71.1	44.2	51.2	26.9	*(Same backbone for medical LLMs)*				
SAG	Pre-trained	84.6	77.8	64.9	70.1	19.7	86.0	79.6	68.1	73.3	17.9
	w/ GRPO	**90.3**	**85.2**	84.0	**88.6**	**6.3**	**91.4**	**86.1**	**85.6**	**92.6**	7.0
	w/ CTDE	89.6	84.7	**86.7**	79.4	10.2	91.0	85.8	**85.6**	87.3	**5.4**

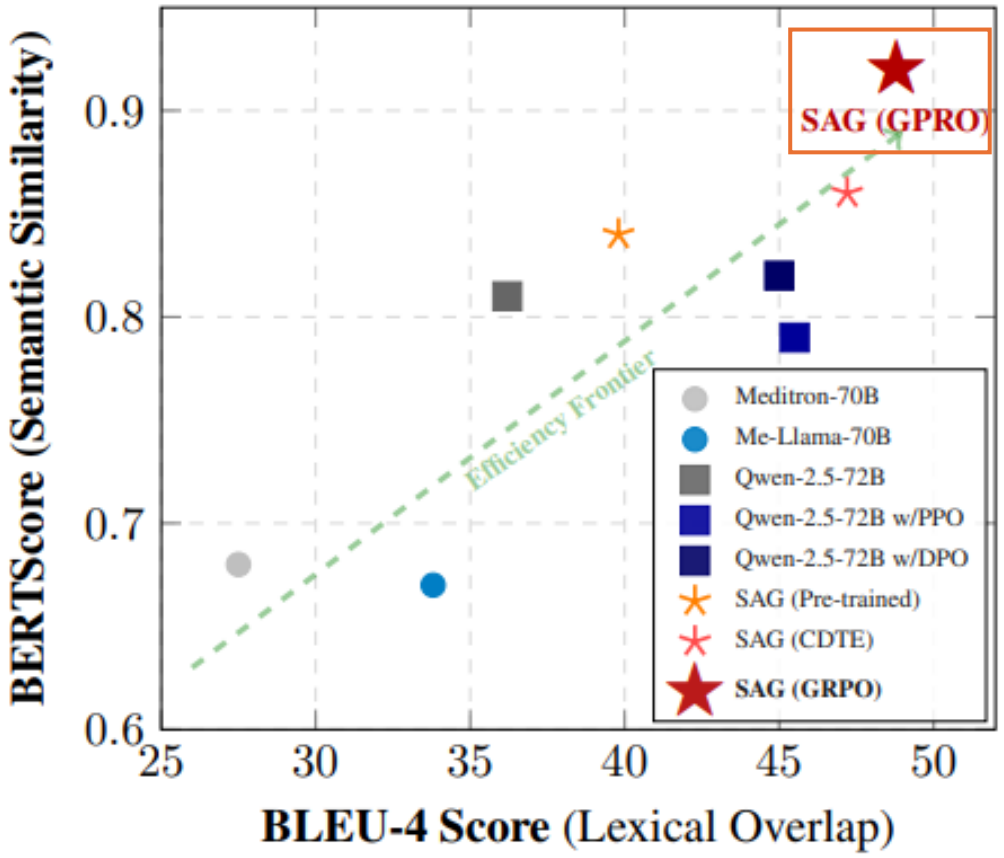
SAG consistently outperforms giant monoliths with significantly less parameters.

Clinical Relevance

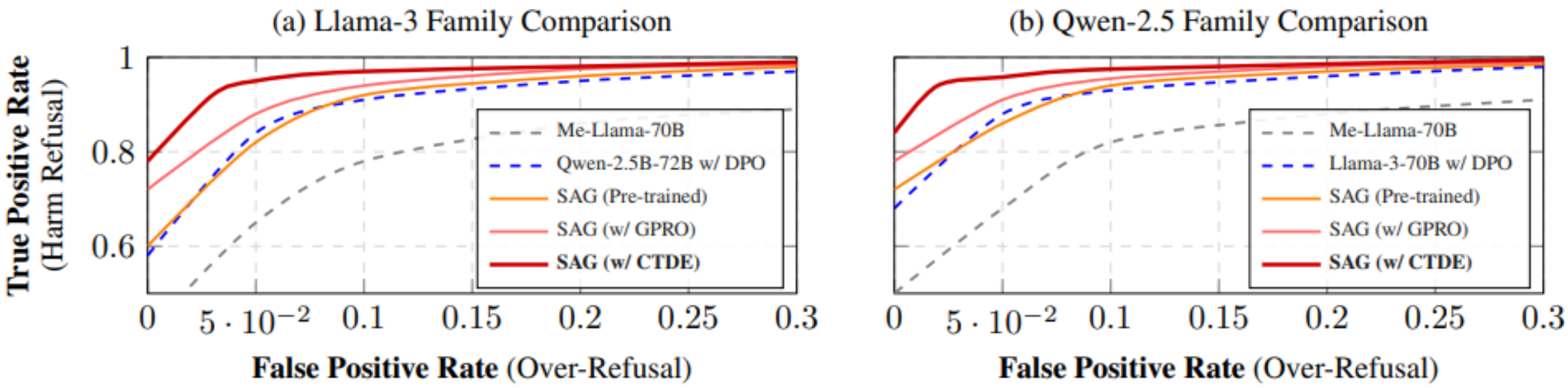
Semantic vs. Lexical Alignment

SAG dominates the efficiency frontier for evidence grounding.

- **BERTScore**: Higher semantic similarity to clinical truth.
- **BLEU-4**: Precise lexical overlap with medical guidelines.
- **Self-Correction**: Collaborative filtering reduces overconfident, hallucinated rationales.



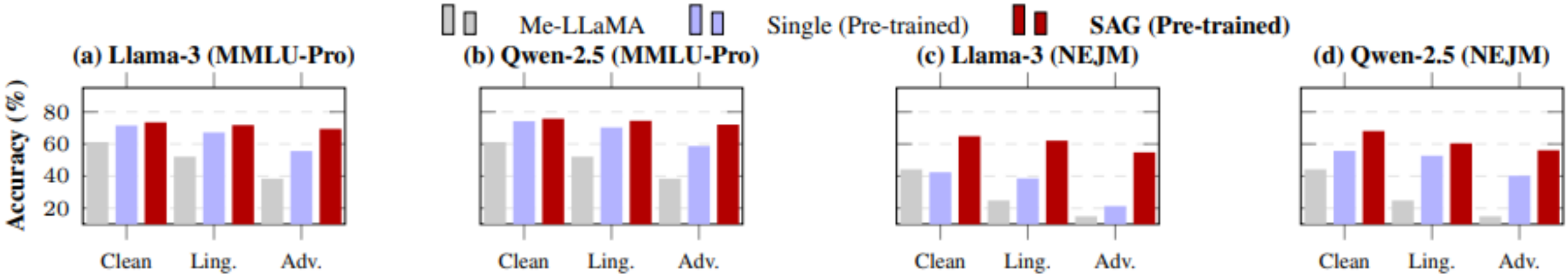
Expanding Safety Boundary



ROC Analysis on Safety

SAG provides a more precise separation between harmful and safe requests.

Robustness to Adversarial Noise



Resilience to Input Distractors

SAG behaves as an evidence-consistency verifier rather than a pattern matcher.

ICML
International Conference
On Machine Learning

Fairness Evaluation

Method	Race CDR ↓	Gender CDR ↓	Avg CDR ↓
Qwen-2.5-72B (Pre-trained)	2.6%	2.2%	2.4%
Qwen-2.5-72B (w/ PPO)	2.0%	1.7%	1.9%
Qwen-2.5-72B (w/ DPO)	1.7%	1.5%	1.6%
Meditron-70B	7.1%	3.0%	5.1%
Me-LLaMA-70B	5.4%	3.2%	4.3%
SAG (Pre-trained)	1.5%	1.3%	1.4%
SAG (w/ GRPO)	1.1%	1.0%	1.1%
SAG (w/ CTDE)	**0.8%**	**0.7%**	**0.8%**

ICML
International Conference
On Machine Learning

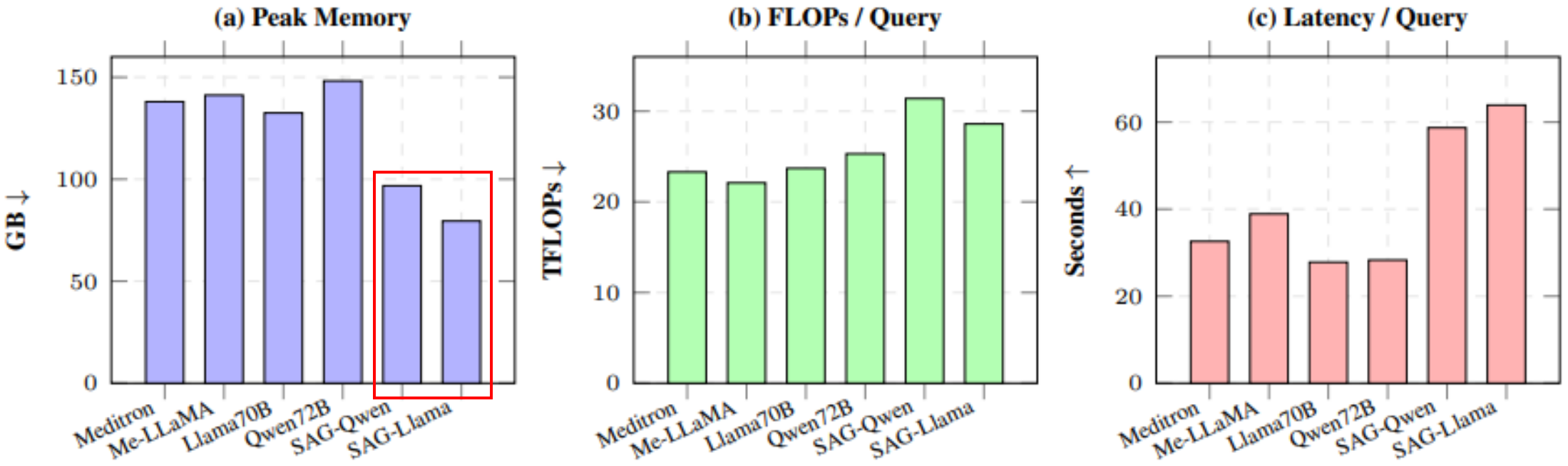
Ablation Study

Ablation	Llama Backbone					Qwen Backbone				
	M-QA	MCQA	NEJM	GPQA	Gap ↓	M-QA	MCQA	NEJM	GPQA	Gap ↓
w/o A_R (no reasoning agents)	70.2	63.1	41.8	34.7	35.5	72.4	66.0	45.7	38.9	33.5
w/o A_K (no RAG)	78.9	72.4	54.0	45.6	33.3	80.5	74.0	57.2	50.1	30.4
w/o A_S (no safety agents)	83.4	76.9	60.8	59.0	24.4	84.3	78.0	64.0	62.2	22.1
w/o A_J (no judgment agents)	76.5	69.0	58.3	47.2	29.3	78.0	71.6	61.5	49.0	29.0
Majority voting	74.8	67.5	52.1	44.0	30.8	76.9	69.8	55.1	47.5	29.4

Removing Reasoning Agent (A_R) causes the largest collapse in complex diagnosis.
Knowledge Agent (A_K) is critical for grounding
Synthesis Agent (A_J) is vital for arbitrator-led consensus.

ICML
International Conference
On Machine Learning

Deployment Efficiency



Thank you!

BINGHAMTON
UNIVERSITY
STATE UNIVERSITY OF NEW YORK