

LatentChem: From Textual CoT to Latent Thinking in Chemical Reasoning

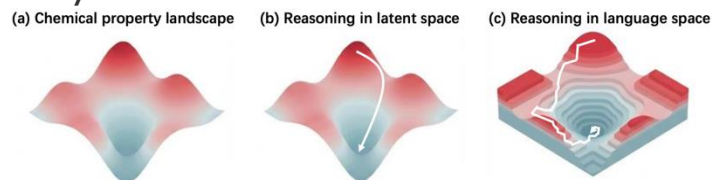
LIT @ ICLR 2026 | Xinwu Ye*, Yicheng Mao*, Jia Zhang, Yimeng Liu, Li Hao, Fang Wu, Zhiwei Li, Yuxuan Liao, Zehong Wang, Zhiyuan Liu, Zhenfei Yin, LiYuan, Philip Torr, Huan Sun, Xiangxiang Zeng, Mengdi Wang, Le Cong, Shenghua Gao, Xiangru Tang



MOTIVATION & PROBLEM

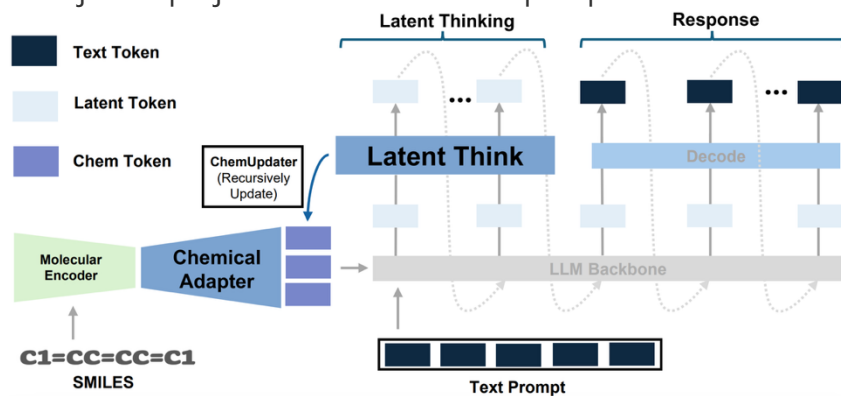
Is CoT always optimal for chemistry?

• Forcing continuous chemical reasoning to be verbalized creates a "continuity-discretization gap."



METHOD OVERVIEW: How LatentChem Works

1. ChemAdapter: turns molecules into ChemTokens.
2. Latent thinking (ChemUpdater & Latent Projector): reasons in latent states.
3. ChemUpdater: updates molecule features each step.
4. Latent Projector: projects latents to LLM's input space.

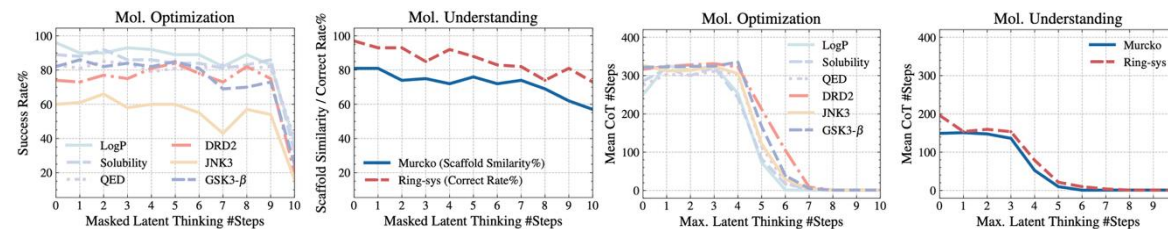


🌟 **Core Takeaway:** LatentChem does not "remove reasoning"; instead, it internalizes chemical reasoning from visible textual steps into continuous latent-space computation, substantially reducing reasoning cost while preserving or even improving accuracy.

CORE FINDINGS

The Model "Spontaneously Stops Writing CoT"

- Outcome-driven GRPO shifts reasoning into latent space.
- Noise on early latent steps steadily hurts performance.
- Low latent budget brings explicit CoT back.



KEY RESULTS & OBSERVATIONS

• Largest gains appear on **open-ended** tasks, while closed-ended performance remains competitive (e.g., 63.73% vs. 50.00% on ChemCoTBench).

• LatentChem cuts reasoning step overhead by 5.4x–29.9x on average tasks, with a 10.84x mean reduction.

Table 2 Main results on chemical benchmarks. We report the non-tie win rate (R_{win}^{\uparrow}) of each method compared to the reference CoT baseline (Stage 1+2+4). The "All" column denotes the aggregated win rate on each benchmark. Bold and underlined values indicate the best and second-best performance, respectively. * denotes statistical significance with $p < 0.01$ according to the sign test.

Category	Method	ChemCoTBench			Mol-Instructions			ChemLLMBench			ChEBI-20
		Open	Closed	All	Open	Closed	All	Open	Closed	All	Open
Text-only	Qwen-3-8B	13.98*	2.80*	8.61*	58.65*	21.87*	23.97*	66.67*	19.58*	22.78*	74.86*
	Qwen-3-8B (SFT)	28.08*	20.63*	25.75*	73.13*	35.92*	38.39*	83.14*	37.51*	41.11*	80.85*
Chem. LLMs	Stage 1	43.84*	36.76*	41.73*	63.00*	50.49	51.25*	68.47*	50.56	51.57	64.80*
	Stage 1+2	38.35*	28.44*	35.50*	50.36	43.69*	44.10*	59.77	45.12*	46.21*	54.60*
Latent Chem. LLMs	Coconut-Chem	34.71*	26.32*	32.11*	57.94*	39.45*	40.57*	75.00*	42.46*	44.82*	65.58*
	LatentChem (Ours)	63.73*	50.00	59.88*	72.00*	48.36*	49.88	87.37*	52.87	55.58*	85.26*

