MCP-Enabled LLM Agents for Closed-Loop Optimization in Real-Time Physical Experiments

Zekun Ren[1], Jiaen Yee[1,2], Hongzhao Tan[1], Qianxiao Li[3], and Kedar Hippalgaonkar[1,2]

[1]Berkeley Education Alliance for Research in Singapore (BEARS), Singapore, 138602
[2]School of Materials Science and Engineering, Nanyang Technological University (NTU), Singapore 639798
[3]Department of Mathematics, National University of Singapore (NUS), Singapore 117558

Abstract

Self-driving laboratories (SDLs) accelerate scientific discovery by automating the design–make–test–analyze (DMTA) cycle and closing the loop between decision-making and real experimental feedback [1, 2]. We present a modular multi-agent platform that enables large language models (LLMs) to execute physical experiments through the Model Context Protocol (MCP), a standard interface for connecting models to external tools and data [4, 5]. In our architecture, each laboratory instrument is encapsulated as an independent tool-agent (e.g., liquid handling, imaging, manipulation, weighing, and electrochemical measurement), and an orchestration layer composes these agents into end-to-end experimental workflows. This design separates device integration, protocol logic, and decision policies, enabling rapid reuse across tasks while preserving a consistent and auditable closed-loop interface between the LLM and the lab.

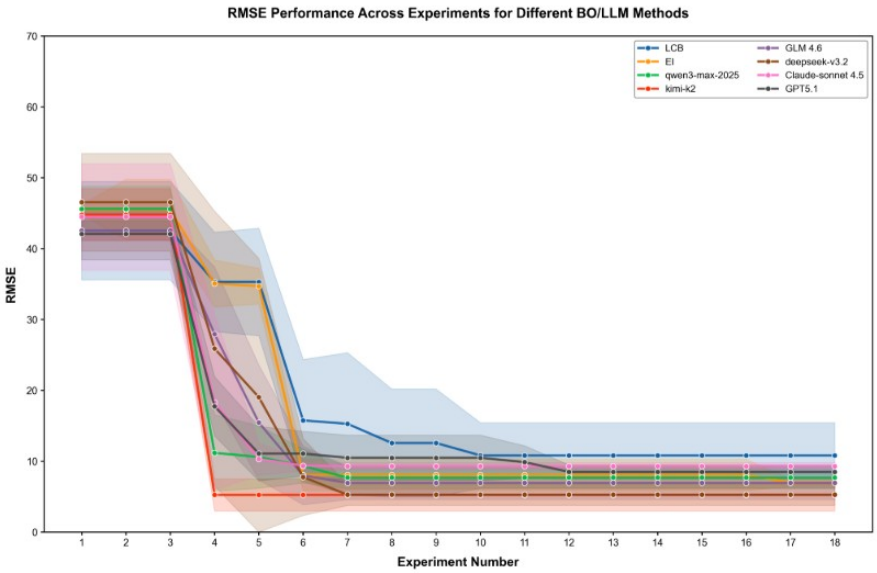

Figure 1: LLM-in-the-loop optimization benchmark on real-time physical experiments (LLM vs Bayesian Optimization).

We benchmark LLM-in-the-loop closed-loop optimization against a widely used SDL baseline, Bayesian optimization (BO), under identical physical constraints and sensing pipelines. Our first completed benchmark is a colour-mixing task with real-time camera feedback: the system iteratively proposes experimental conditions, executes them on hardware, measures outcomes from live sensor streams, computes an error signal, and updates its next action based on the measured result. In head-to-head physical runs,

LLM-driven policies consistently reach lower error in fewer iterations than BO, demonstrating improved sample efficiency when the optimizer can combine high-level reasoning with structured tool access (Figure 1). Beyond this benchmark, we apply the same MCP agent library and orchestration approach to viscous liquid handling and electrocatalyst measurements with an electrochemical station, forming a task suite that stresses robustness to non-ideal dynamics, sensor dropouts, and multi-step protocol compliance.

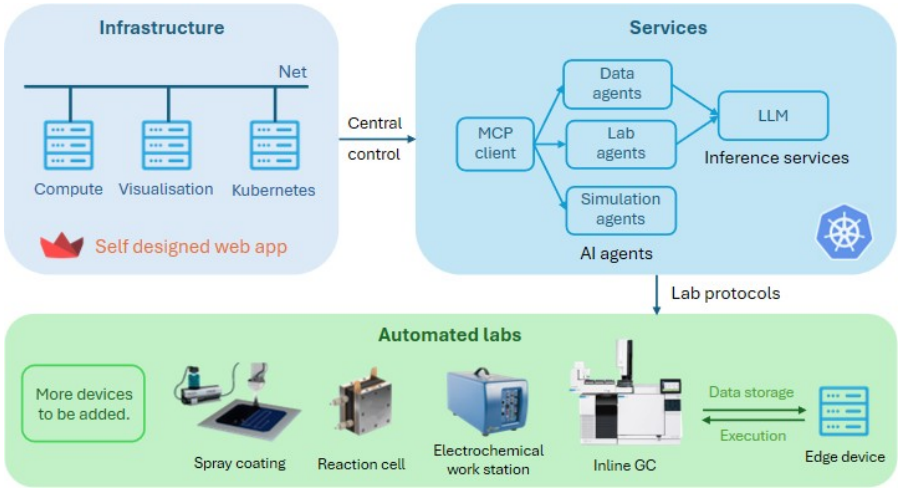

Figure 2: Proposed MCP-based multi-agent workflow for closed-loop electrochemistry measurement optimization.

Figure 2 outlines our proposed electrochemistry optimization workflow, where MCP tool-agents expose actuation and measurement primitives and an orchestrating controller closes the loop on experimental objectives. Overall, this work contributes an MCP-based blueprint for building instrument-agnostic LLM tool-agents and a benchmarking methodology that compares LLM policies with classical SDL optimizers on real-time physical experiments, complementing growing evidence that tool-augmented LLM systems can autonomously plan and execute experimental workflows [3].

References

[1] G. Tom *et al.* Self-Driving Laboratories for Chemistry and Materials Science. *Chemical Reviews*, 124(16):9633–9732, 2024. doi:10.1021/acs.chemrev.4c00055.

[2] B. P. MacLeod *et al.* Self-driving laboratory for accelerated discovery of thin-film materials. *Science Advances*, 6(20):eaaz8867, 2020. doi:10.1126/sciadv.aaz8867.

[3] D. A. Boiko, R. MacKnight, B. Kline, and G. Gomes. Autonomous chemical research with large language models. *Nature*, 624:570–578, 2023. doi:10.1038/s41586-023-06792-0.

[4] Anthropic. Introducing the Model Context Protocol. Nov 25, 2024. `https://www.anthropic.com/news/model-context-protocol` (accessed 2026-01-16).

[5] Model Context Protocol. MCP Specification (draft). `https://modelcontextprotocol.io/specification/draft` (accessed 2026-01-16).