Supplemental Material: Human–AI Research Process

Kyle Elliott Mathewson[1]

February 2026

[1]Faculty of Science, University of Alberta

Abstract

This supplemental document describes the human–AI collaboration process behind the companion paper "Extending Marginal Reputation to Persistent Markovian States." The main paper presents the mathematical results as a self-contained contribution; this supplement documents the research process, including the agent architecture, computational testing framework, timeline, and reflections on AI-assisted mathematical research.

Contents

1 Overview

The companion paper extends the main result of Luo & Wolitzky (2024) from i.i.d. states to persistent Markovian states, introducing the concepts of belief-robustness and the Markov commitment payoff. This supplement documents the research process that produced those results: an initial AI-assisted conjecture phase, expert critique, systematic computational verification, and iterative revision.

2 Phase 1: Initial Conjecture (Feb 16, 5:00–9:30 PM)

The initial extension was developed under a five-hour time constraint. Five AI agents—four instances of Claude Opus 4.6 and one instance of Claude Sonnet 4.5—worked under human coordination, with each agent assigned a specialized role.

Agent	**Role**	**Key Contribution**
Sonnet 4.5 Reader	Paper parsing	Multi-level summaries, equation extraction
Agent 840 (Opus)	First parse	Identified lifted-state approach, 5 interpretations
Agent 841 (Opus)	Proof coordinator	Directed 4 parallel subagents
Agent 852 (Opus)	Paper author	26-page LaTeX document
Agent 860 (Opus)	Peer reviewer	Identified continuation value subtlety

Within the time window, this architecture produced a 26-page paper, an interactive web demonstration, and a social media summary. The paper proposed that Theorem 1 extends to Markov states via the lifted state $\tilde{\theta}_t = (\theta_t, \theta_{t-1})$, correctly identifying several key mathematical tools—the process-independence of the KL chain rule, the well-defined stationary distribution on the lifted space, and the filter stability argument.

3 Phase 2: Expert Critique (Feb 16, 10:00–11:00 PM)

Within one hour of submission, Daniel Luo—co-author of Luo & Wolitzky (2024)—posted two threads of detailed technical feedback comprising 15 distinct points.

The most consequential observations identified a single core issue: the i.i.d. assumption disciplines short-run player information sets about the state. Under Markov dynamics with a state-revealing strategy, short-run player beliefs are given by the filtering distribution $F(\cdot|\theta_t)$ rather than the stationary distribution π, causing the Nash correspondence to become state-contingent.

The single most clarifying observation was:

> *"To make it clear: suppose s_1 just takes an action that reveals the state. In the iid case, this won't affect SR beliefs. But in the Markov case, this can cause beliefs to never settle into the stationary distribution."* — Daniel Luo

4 Phase 3: Computational Testing Framework

To determine precisely which elements of the proof extend and which require modification, we designed a systematic computational investigation organized into seven analysis areas (SA1–SA7).

4.1 Agent Architecture

The computational framework employed a hierarchical agent architecture:

- A reusable Python class (`Agent`) supporting task assignment, report generation, and hierarchical delegation.
- Seven subagents (SA1–SA7), each producing a synthesized report.
- Twenty-one sub-subagent scripts with detailed task specifications.

Each analysis script was designed around a *hypothesis* (what the paper claims), a *counter-hypothesis* (what the critique implies), and a *test* (what the simulation checks).

4.2 Execution

Four parallel subagents executed all 21 scripts simultaneously:

Batch	**Modules**	**Scripts**	**Runtime**
Batch 1	SA1 (Beliefs) + SA2 (State-revealing)	6	202s
Batch 2	SA3 (KL bound) + SA4 (Filter stability)	6	143s
Batch 3	SA5 (OT sensitivity) + SA6 (Nash dynamics)	6	19s
Batch 4	SA7 (Monotonicity) + orchestrator	3+2	47s

4.3 Results Summary

The findings divided sharply:

- **Surviving claims**: KL counting bound (SA3), filter stability (SA4), OT robustness (SA5), monotonicity characterization (SA7).
- **Failing claims**: SR belief convergence to π (SA1 refuted), static Nash correspondence (SA6 showed 37.2% disagreement), original payoff bound (SA6 showed 36.3% overestimation).

5 Phase 4: Manuscript Revision

The computational evidence guided a structured revision:

- The paper was decomposed into 12 modular LaTeX section files assembled by a master document.
- All quantitative claims were drawn from an auto-generated statistics file (`stats.tex`).
- An automated pipeline (`generate_paper.sh`) executes the full sequence: analysis scripts → statistics extraction → LaTeX compilation.

6 Timeline

Time	**Phase**	**Key Event**
Feb 16, 5:00–9:30 PM	Phase 1	5 AI agents produce initial draft
Feb 16, 10:00–11:00 PM	Phase 2	Expert posts 15-point technical feedback
Feb 17, 12:00–1:00 AM	Phase 3	Combined review, agent hierarchy designed
Feb 17, 1:00–2:00 AM	Phase 4	21 scripts, 40 figures, 28 reports
Feb 17, 2:00–3:00 AM	Phase 5	Corrected paper compiled

7 Reflections on AI-Assisted Mathematical Research

Several observations emerge from this process:

1. **AI + expert critique**: The combination of AI-assisted rapid exploration and human expert critique proved more productive than either alone. AI agents identified the lifted-state approach and the process-independent tools; the human expert identified the semantic gap between these tools and their game-theoretic interpretation.

2. **Computational triage**: Rather than attempting to determine *a priori* whether the critique invalidated the entire approach, the seven analysis modules produced quantitative evidence that cleanly separated the surviving claims from the failing ones.

3. **Scalable architecture**: The three-level hierarchical delegation (orchestrator, subagents, sub-subagents) with structured report aggregation scaled effectively.

4. **Tools vs. semantics**: AI systems can identify correct mathematical tools but may fail to interpret their role correctly within a larger argument. The KL bound *is* process-independent, but the proof's *use* of it depends on what it means for a period to be "non-distinguishing"—a semantic subtlety that required domain expertise.

Repository Structure

```
revisedTexPaper/
+-- main.tex              # Main paper
+-- supplemental_methodology.tex  # This document
+-- stats.tex             # Auto-generated statistics macros
+-- sections/             # Modular .tex files
+-- figures/              # Diagnostic figures
+-- scripts/              # 7 analysis + automation scripts
+-- response_letter_refineai.tex  # Response to reviewer feedback
```