Assignment 1: Sampler Synthesis

CDS DS 595 · Spring 2026 · Boston University

Out:	Wednesday, February 4	**Due:**	Wednesday, February 18
Weight:	15% of final grade	**Format:**	Individual

Overview

In this assignment, you will **design, implement, and critically analyze** a novel MCMC sampling method. You are encouraged to use AI assistants (e.g., Gemini, Claude, ChatGPT) freely throughout this process—learning to collaborate effectively with AI tools is itself a course objective.

The goal is not to produce a state-of-the-art method. Rather, the goal is to produce something new—even if simple—and to deeply understand what it's doing. A thoughtful analysis of a modest method is more valuable than a complex method that's fully vibe coded.

Learning Objectives. Practice algorithm design by proposing and implementing new ideas; learn to benchmark fairly using standard diagnostics and visual inspection; and develop effective strategies and a muscle for AI collaboration on technical problems.

Before starting, please read:

How AI assistance impacts the formation of coding skills
Anthropic, January 2026
https://www.anthropic.com/research/AI-assistance-coding-skills

Task

1. **Propose a method.** Working with an AI assistant, design a novel MCMC sampler. This could be:
   - A modification or extension of an existing method (e.g., adaptive step sizes, novel proposal distributions, hybrid approaches)
   - A combination of ideas from different methods
   - An entirely new approach suggested through your AI collaboration. You can prompt in creative ways—e.g., a "chemistry-inspired" sampler, or a proposal distribution you think might be interesting.

   Use the starter notebook provided with this assignment and elements of Lab 2 if useful as scaffolding for your implementation.

2. **Implement and test.** Implement your method in JAX (preferred) or NumPy. Test it on the following benchmark distributions, provided in the starter code notebook:

- **Rosenbrock (Banana)**: A curved, non-convex distribution that tests how well samplers handle strong correlations and curved degeneracies. Defined as $\log p(x, y) \propto -\frac{(1-x)^2}{20} - (y - x^2)^2$.
- **Neal's Funnel**: A hierarchical distribution where the scale varies dramatically across the space. The narrow neck is notoriously difficult for samplers with fixed step sizes. Defined as $v \sim \mathcal{N}(0, 9)$ and $x \sim \mathcal{N}(0, e^v)$.

3. **Analyze critically.** For each benchmark:
   - Visualize samples and compare to the target distribution
   - Check diagnostics: acceptance rates, effective sample size, autocorrelation
   - **Compare to baselines**: vanilla Metropolis-Hastings-Rosenbluth and HMC. You may use off-the-shelf implementations (e.g., BlackJAX) for the baselines, also provided in the starter code. Make sure to tune hyperparameters fairly for all methods (target a good acceptance ratio—see guidelines in starter code)!
   - Explore how your method's hyperparameters affect performance
   - Identify when your method works well and when it struggles
4. **Iterate.** Based on your analysis, refine your method. Document what you tried and what you learned from each iteration.

Deliverables

Submit the following via GitHub Classroom:

1. **Code/Notebook.** A well-documented Jupyter notebook or Python files containing:
   - Your method implementation with clear comments
   - Benchmark experiments with visualizations
   - Diagnostic plots (traces, corner plots, autocorrelation)
2. **Written Report.** A short report (maximum 3 pages, excluding references) in NeurIPS format describing:
   - Your method and its motivation
   - Experimental results and analysis
   - Discussion of strengths, weaknesses, and potential improvements
   - A section on **AI Collaboration** (see below)

   Use the NeurIPS 2025 LaTeX template. You can use Overleaf for editing. A template for preparing your report is provided in the `report` folder of the Assignment repo.

AI Collaboration

You are expected to use AI assistants throughout this assignment. In your report, include a section titled **"AI Collaboration"** (approximately half a page) that addresses:

- Which AI tools did you use and for what purposes?
- What prompting strategies were effective? What external resources (papers, code, documentation) did you provide, and what context did the AI have access to?
- Where did the AI provide useful ideas vs. where did you need to correct or guide it?
- What did you learn about working with AI on technical problems?

Suggested exercise. Try using the AI for **self-correction**: ask it to identify three potential ways your proposed sampler could fail (e.g., violating detailed balance, poor mixing in certain geometries, numerical instability). Then address at least one of those failure modes in your iteration step. This adversarial approach often leads to more robust designs than only asking the AI to generate ideas.

The goal is to reflect thoughtfully on the collaboration process.

Report Structure

Your written report should include the following sections. Don't feel pressured to write the page counts estimated below—a concise and clear exposition is all that's needed. **There are no page expectations as long as you stay under 3 pages.**

1. **Introduction** (∼0.5 page): What is your method and why did you design it this way? What problem or limitation of existing methods does it address? Is it similar to any existing methods?
2. **Method** (∼0.5 page): Technical description of your algorithm. Include pseudocode if helpful. **Briefly discuss whether your method satisfies detailed balance or ergodicity.** Even if you can't prove it rigorously, thinking through the conditions is important—"if it looks like a cloud of points, it must be a valid sample" is a common fallacy when working with AI-generated samplers.
3. **Experiments** (∼1 page): Results on both benchmarks. Include visualizations (samples, traces, corner plots) and quantitative diagnostics (ESS, acceptance rates). Compare to vanilla random walk MCMC and HMC baselines.
4. **Discussion** (∼0.5 page): Where does your method work well? Where does it struggle? What would you try next?
5. **AI Collaboration** (∼0.5 page): Reflect on your use of AI assistants (see above).

Grading

Total: 9 points (15% of final grade)

Category	Excellent (3)	Needs improvement (1)
Method & Implementation (3 pts)	Novel idea with clear motivation; code runs correctly and is well-documented	Trivial modification; no explanation; code has errors or is hard to follow
Analysis & Comparison (3 pts)	Thoughtful diagnostics; meaningful baseline comparison; identifies strengths and failure modes	Reports numbers without interpretation; no baseline comparison
AI Collaboration (3 pts)	Specific examples of what worked and didn't; insights about effective collaboration	Generic description ("I used ChatGPT"); no reflection on the process