Faculty of Science
Department of Mathematics

MA4198 PROJECT PROPOSAL (PROJECT CUM SEMINAR GROUP)

SUPERVISOR'S INFO

Name:	TONG Xin
Email:	xin.t.tong@nus.edu.sg
Tel number:	65162756
Office location:	S17-08-06

PROJECT ID: PS2610-04

TITLE

Sampling with Markov Chain Monte Carlo and Generative AI

BRIEF DESCRIPTION OF PROJECT

Sampling is the task of drawing realizations from a target probability distribution $p(x)$, which may be known only up to a normalizing constant or implicitly defined by a dataset. It arises throughout machine learning in Bayesian posterior inference, data augmentation, uncertainty quantification, and generative modeling. The core challenge is that p(x) is often high-dimensional and lacks a closed-form inverse CDF, making direct sampling intractable.

MCMC. Markov Chain Monte Carlo methods construct an ergodic Markov chain whose stationary distribution is p(x). At each step, a proposal is drawn and accepted with Metropolis–Hastings probability, ensuring detailed balance. After a burn-in period the chain mixes and successive iterates are approximate samples from p. Variants like Langevin dynamics and Hamiltonian Monte Carlo exploit score information to propose better moves, drastically reducing autocorrelation and improving efficiency in high dimensions.

Generative models. Rather than simulating a chain, generative models learn a parametric transformation from a simple base distribution (e.g., Gaussian) to approximate p(x). VAEs learn an encoder-decoder pair with a tractable ELBO objective; GANs train a generator adversarially against a discriminator; normalizing flows learn an invertible map so that the exact likelihood is tractable via the change-of-variables formula; diffusion models learn to reverse a gradual noising process via score matching.

EXPECTATION/S

The student should learn one (or two) sampling algorithm, its mathematical theory, and know how to implement it on some simple tasks.

PREREQUISITE/S (at level 3000 or below, with at most one course at level 3000)

MA2116/ST2131, MA3238/ST3236

READING REFERENCE/S

Sinho Chewi. *Log-Concave Sampling*
Kingma & Welling (2014). *Auto-Encoding Variational Bayes.* ICLR.
Goodfellow et al. (2014). *Generative Adversarial Nets.* NeurIPS.
Ho et al. (2020). *Denoising Diffusion Probabilistic Models.* NeurIPS