THE UNIVERSITY OF CHICAGO

EFFICIENT BAYESIAN ANALYSIS OF HIGH DIMENSIONAL BIOLOGICAL SYSTEMS WITH MARKOV CHAIN MONTE CARLO

A THESIS SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PHYSICS

BY
CLAYTON W. SEITZ

CHICAGO, ILLINOIS
SPRING 20XX

TABLE OF CONTENTS

ABSTRACT

CHAPTER 1

PRIMER ON EXACT BAYESIAN METHODS AND VARIATIONAL INFERENCE

1.1 Markov Chain Monte Carlo

The invention of fast digital computers gives us the ability to simulate random processes and perform inference

1.1.1 Metropolis-Hastings and Gibbs sampling

1.1.2 Langevin Monte Carlo

The metropolis adjusted Langevin algorithm (MALA) aka Langevin Monte Carlo

1.1.3 Hamiltonian Monte Carlo

1.1.4 Stochastic Gradient Langevin Dynamics

1.2 Variational Inference

This won't be about deep learning, just introducing the concept so we can make a comparison with MCMC methods, in case that is relevant

1.2.1 The evidence lower bound

CHAPTER 2

A BAYESIAN APPROACH FOR INFERRING NEURONAL CONNECTIVITY FROM CA2+ IMAGING DATA AND MONTE CARLO SIMULATIONS

CHAPTER 3

BAYESIAN INFERENCE OF THE KINETIC PARAMETERS OF INTERFERON-GAMMA INDUCED TRANSCRIPTION

CHAPTER 4

STRATIFYING IMMUOGENIC TUMOR SUBSTRUCTURE

I am interested in the relationship between tumor substructure i.e. heterogenity and how it relates to the degree of T-cell inflammation. We can be confident that certain malignant clusters are more immunogenic than others, which, to a first approximation, can be understood by measuring T-cell quantity in the sample. This *could* then be supplemented by using T-cell signatures documented in the literature, differential expression, and clustering methods/mixture models to design fluorescent biomarkers. These will then be added to T-cell markers. In principle, this would give a more complete association of gene expression to the inflammaotry state, and would be a useful method in trying to understand changes during the evolution of cancer and following pharmacological treatment.

For the implementation, I will use Seurat for preprocessing and visualization. Malignant cells are isolated first using standard methods e.g, filtering, normalization, and UMAP clustering

This section will probably have to deviate significantly from the central theme of the dissertation, but it will be a major component of my work so it's necessary to include it