

To the Graduate Council:

I am submitting herewith a dissertation written by Kelechi Akwataghbe entitled “Paradigms of Scientific Machine Learning: From Model-Based Inference in NMR Spectroscopy to Predictive Automation in Agricultural Robotics.” I have examined the final paper copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Applied Mathematics.

Ioannis Sgouralis, Major Professor

We have read this dissertation
and recommend its acceptance:

Christopher Strickland

Xinyue Zhao

Carlos Steren

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

To the Graduate Council:

I am submitting herewith a dissertation written by Kelechi Akwataghbe entitled “Paradigms of Scientific Machine Learning: From Model-Based Inference in NMR Spectroscopy to Predictive Automation in Agricultural Robotics.” I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Applied Mathematics.

Ioannis Sgouralis, Major Professor

We have read this dissertation
and recommend its acceptance:

Christopher Strickland

Xinyue Zhao

Carlos Steren

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Paradigms of Scientific Machine Learning: From Model-Based Inference in NMR Spectroscopy to Predictive Automation in Agricultural Robotics

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Kelechi Akwataghibe

August 2025

© by Kelechi Akwataghibe, 2025
All Rights Reserved.

Dedicated to my parents, Emmanuel and Josephine Akwataghbe who made great sacrifices for myself and my siblings to thrive.

Acknowledgements

I would like to thank my Lord and Savior, Jesus Christ, for the fortitude to get to this point and for giving me a natural bent for seeing the big picture.

I am grateful for my advisor, Dr. Ioannis Sgouralis, who has been an inspiration to me—a model of what a scientist should strive to be. I am particularly grateful for his patience in bringing me from the cradle of research to a point where I can stand and, at least, take steps on my own.

I am also grateful to everyone who made my journey at UTK a delight: to Pam, for her invaluable guidance and advice as I journeyed through the program; to Ben, who was always willing to provide technical support whenever I needed it; to Dr. Vaselios Maroulas for providing career guidance to me when I first joined the department of Mathematics and to every other person whose presence added color to my stay.

Lastly, to my lovely wife, Deborah, our parents and siblings, my mentors, friends, and well-wishers, thank you for your support, prayers, and understanding throughout this journey.

All models are wrong, but some are useful.

George Box

Abstract

Data-driven modeling is crucial for extracting insights from complex scientific systems. Modern machine learning provides two powerful but philosophically distinct paradigms: model-based inference, which aims to understand physical mechanisms, and predictive modeling, which optimizes for forecasting accuracy. This dissertation explores and advances these two paradigms through their application in two different domains: Nuclear Magnetic Resonance spectroscopy (NMR) for biochemical analysis and computer vision for agricultural robotics.

The first section presents two distinct Bayesian inference models I developed to analyze 1D NMR spectra. The first is a nonparametric model which enables robust spectral deconvolution and automatic peak detection of general 1D NMR spectra which we apply to flourine-19 (^{19}F), while the second is a parametric model used to quantify the kinetics of chemical exchange. Both models provide not only estimates of spectral parameters but also a rigorous quantification of their associated uncertainty.

The second section focus on a deep learning approach I developed to address the challenge of automated feed monitoring in precision agriculture. I describe a Convolutional Neural Network (CNN) which was trained to accurately predict leftover feed quantities from 3D images; a crucial step in the development of individualized feeding schedules.

Ultimately, these distinct lines of inquiry converge on a central theme: the choice between inferential and predictive modeling is dictated by the scientific

objective, whether that objective is achieving deep physical insight or enabling robust automation.

Table of Contents

1	Introduction	1
1.1	Data-driven approaches for modeling physical systems	2
1.1.1	Bayesian models for NMR spectroscopy	3
1.1.2	Deep Learning for predictive modeling in agricultural robotics	4
1.2	A unified methodological perspective	4
1	Bayesian Inference for Parameter Estimation in NMR Spectroscopy	6
2	Introduction to NMR spectroscopy and Bayesian modeling	7
2.1	Fundamental principles of NMR spectroscopy	7
2.1.1	Nuclear spins	8
2.1.2	^{19}F NMR experiments	9
2.1.3	Existing analysis approaches to 1D NMR	11
2.2	A Bayesian framework for parameter inference	14
2.2.1	Computational inference via Markov Chain Monte Carlo (MCMC) Sampling	15
3	Bayesian nonparametric statistical learning for resonance identification and characterization in ^{19}F-NMR	17
3.1	Disclosure	18
3.2	Introduction	18
3.3	Methods	21

3.3.1	Model overview	21
3.3.2	Model description	21
3.3.3	Model inference	23
3.4	Results	24
3.4.1	Application to raw data from hTRPM8-VSLD experiments . .	25
3.4.2	Validation with synthetic data	28
3.4.3	Benchmarking with synthetic data	31
3.4.4	Comparison with existing NMR analysis methods	34
3.5	Discussion	39
4	A statistical learning framework for probing chemical exchange	43
4.1	Disclosure	44
4.2	Introduction	44
4.3	Methods	47
4.3.1	Data generation in NMR experiments	47
4.3.2	The Bloch-McConnell equations for K-Site exchange	48
4.3.3	Forward Model: Simulating the FID signal	49
4.3.4	Forward model: Incorporating temperature-dependent Chemical Exchange rates	53
4.4	Results	56
4.4.1	Validation with synthetic data	56
4.4.2	Benchmarking with synthetic data	59
4.4.3	Discussion	60
2	Deep Learning for predictive modeling in precision agriculture	62
5	Development of precision feed AI systems	63
5.1	Disclosure	64

5.2	Introduction	64
5.3	Methods	66
5.3.1	Animals and materials	66
5.3.2	Data organization and analysis	67
5.3.3	Computational methods	70
5.4	Results	73
5.4.1	System Validation	73
5.4.2	Herd feeding pattern	76
5.4.3	Individual animal feeding patterns	76
5.5	Discussion	81
5.6	Limitations and future directions	81
5.6.1	Acknowledgments	82
6	Conclusions	83
	Bibliography	86
A	Bayesian nonparametric statistical learning for resonance identification and characterization in ^{19}F-NMR	99
B	Probing statistical learning with chemical exchange	112
C	Development of precision feed AI systems	116
	Vita	119

List of Tables

2.1	Discrete Nuclear Spin Orientations Based on Spin Quantum Number (I)	8
5.1	Comprehensive summary of visit events. The reported values indicate mean $\pm$ one standard deviation.	77
5.2	Summary of consumption related events during the course of this study. The reported values indicate mean $\pm$ one standard deviation.	77
C.1	Comprehensive summary of visit events and feed consumption for all animals during the time course of this study. The reported values indicate mean $\pm$ one standard deviation.	117
C.2	Comprehensive summary of daily events and feed consumption for all animals during the time course of this study. The reported values indicate mean $\pm$ one standard deviation.	118

List of Figures

3.1	Spectra representations of deconvolved resonances. Plots 0A-4A: Individual resonances deduced from posterior estimates of the FID data from 1024 scans, of ^{19}F hTRPM8-VSLD. Plots 0B-4B: Frequency domain interpretation of plots 0A-4A	27
3.2	Histograms showing SNR across all resonances for each dataset. Plots 1A-4A, 1B-4B, 1C-3C show SNR for individual resonances of the dataset containing 1024 scans, 256 scans, 64 scans respectively. . . .	28
3.3	Effect of the number of scans. 0A-0E: Frequency domain spectrum showing the deconvolutions of resonances deduced from posterior estimates of the FID data from 1024, 256, 64, 16, and 4 scans of ^{19}F hTRPM8-VSLD. Plots 1A-4A, 1B-4B, 1C-4C, 1D-4D, and 1E-3E are histograms showing the distribution of chemical shifts for individual resonances of the FID data from 1024, 256, 64, 16, and 4 scans respectively	29
3.4	Validation with synthetic data. Histograms of relevant parameters resulting from the posterior distribution in the analysis of the FID data simulated with identical resonance parameters of the dataset containing 1024 scans of ^{19}F hTRPM8-VSLD. Plot 0A shows the deconvolution of all the resonances. Plots 1A-4A, 1B-4B and 1C-4C respectively correspond to the chemical shifts, amplitudes and relaxation rates of each individual resonance. The vertical lines are the ground truth values used in the simulation of the data.	30

3.5	FID with synthetic data showing effect of decreasing SNR. A: Baseline SNR, B: SNR with 2x noise variance, C: SNR with 4x noise variance, D: SNR with 8x noise variance, E: SNR with 16x noise variance, F: SNR with 32x noise variance.	32
3.6	Frequency domain spectra showing decreasing distance of separation between two identical peaks d' . A: $d' = 78.72$, B: $d' = 39.36$, C: $d' = 19.68$, D: $d' = 9.84$, E: $d' = 4.92$ (Failure), ...	35
3.7	Frequency domain spectra of synthetic data showing decreasing amplitude of one shoulder peak. 1A-4A: The left shoulder peak is constructed with 1x,0.5x,0.25x, and 0.125x the baseline amplitude. 1B-4B: Corresponding histograms showing the number of resonances identified by the method (ground truth = 2)	36
3.8	Frequency domain spectra of synthetic data showing increasing complexity. 1A-4A: Spectra constructed with 5,9,17, and 33 resonances respectively, 1B-4B: Corresponding histograms showing the number of resonances identified by the method. Spectrum in 4A is unusually complex and, as a result, exposes a failure mode in our method. . . .	37
3.9	Data analysis of cleanest dataset with MestReNova. Panel A shows the frequency domain spectrum of raw data. Panel B shows the frequency domain spectrum of data after preprocessing steps (phasing and apodization). Panel C shows resonances detected using qGSD, MestReNova's peak picking algorithm	38

3.10	Data analysis of cleanest dataset with Deep Picker. Panel A compares the MAP spectrum generated by my method with the spectrum produced by Deep Picker at a noise-level cutoff of 5.5, where Deep Picker identified 27 resonances. Panel B compares the MAP spectrum generated by my method with the spectrum produced by Deep Picker at a noise-level cutoff of 10, where Deep Picker identified 9 resonances. Panel C compares the MAP spectrum generated by my method with the spectrum produced by Deep Picker at a noise level cutoff of 18, where Deep Picker identified 5 resonances. Panel D compares the MAP spectrum generated by my method with the spectrum produced by Deep Picker at a noise level cutoff of 19, where Deep Picker identified 4 resonances.	40
4.1	Synthetic data against the model signal generated from the Maximum A Posteriori (MAP) parameters at five temperatures, increasing from 311 K (bottom) to 371 K (top). Panels 1A-4A show the frequency-domain FID, and panels 1B-4B depict the corresponding exchange-rate posterior distributions (histograms) with the true value marked by a vertical black line. As temperature increases, the two distinct peaks at low temperature broaden and merge, illustrating accelerated chemical exchange. This is reflected in the broadening of the posterior histograms, which indicates increased uncertainty.	57
4.2	Panels 1A-5A and 1B-5B represent synthetic data and MAP reconstructions for the lowest-temperature spectrum of each noise level. These are followed by normalised posterior histograms for the relaxation rate (r) (panels 1C-5C), enthalpic ($\Delta H^\ddagger$) and entropic ($\Delta S^\ddagger$) activation parameters(panels 1D-5D and 1E-5E respectively). Rows are ordered by decreasing SNR from top to bottom.	58

4.3	Root-mean-square error (RMSE) of the phase parameter (p) as a function of the signal-to-noise ratio (SNR). Each coloured marker corresponds to one synthetic dataset and uses the same row colours employed in fig. 4.2. The SNR axis is reversed (high SNR on the left) to maintain consistency with the top-to-bottom ordering in the spectra panel	59
5.1	Data pipeline summarizing procedures and the successive stages of data transformations applied in this work.	67
5.2	Example images of the feed bin: (A) empty bin (0kg), (B) full bin (3.9kg), (C) image shown in panel B after registration with respect to the image of panel A, and (D) difference image between full (panel C) and empty bins (panel A), enhancing contrast for processing through the CNN. In all panels, the highlighted region indicates the crop region. Although the area of the crop region represents only a small portion of the full field of view, due to the operation of 3D cameras, the actual imaged volume nearly covers the volume of the entire bin.	69
5.3	Configuration of CNN that converts 3D images to feed amount estimates.	71
5.4	Portion of the timeline of this study with marked visit events and acquired images. As can be seen, visit events last for ≈ 7 min each and are separated by intervals of ≈ 1.5 min. Each event is aligned with ≈ 22 images acquired successively at intervals of 0.5 sec. The time and duration of the visit events depend on animal behavior and show random patterns. Occasional gaps in the train of acquired images are due to images failing prescreening or image registration which are excluded from our analysis.	74

5.5	Histograms of the metrics in table 5.1, the vertical line identifies the mean values. For clarity, in Panel B, times between successive visit events above 600s are binned at 600s.	77
5.6	Feed consumption estimates. Panel A displays feed consumed (NOT feed leftover) vs feed delivered over all events of this study and highlights 4 specific instances, representing two pairs of successive visits (1a/1b and 2a/2b). Panel B shows the timeline of these visit events while the bin images in panel C show the leftover feed at the end of each visit event.	78
5.7	Histograms of the metrics in table 5.1, the vertical line identifies the mean values. For clarity, in Panel B, times between successive visit events above 600s are binned at 600s.	80
5.8	Comparison of average daily feed delivered and consumed per animal to herd daily averages. (Left) Bar chart showing individual animals with their respective average daily delivered (blue) and consumed (red) feed. (Right) Aggregated comparison of the overall average delivered and consumed feed across all animals.	80

Chapter 1

Introduction

Modern science increasingly relies on computational tools to extract information from complex, and often high-dimensional data [LeCun et al. \[1998\]](#), [Hinton and Salakhutdinov \[2006\]](#), [McInnes et al. \[2018\]](#), [Rabiner \[1989\]](#). In this realm of data-driven modeling, an interesting tension exists between two distinct paradigms: The first of which is model-based inference, which seeks to understand the underlying mechanisms of a system, while the other is predictive modeling, which aims to accurately forecast future behavior. The first prioritizes the interpretability of physical parameters, and the robust quantification of uncertainty, while the second prioritizes predictive power which it usually does by leveraging complex, non-linear function approximators.

This dissertation presents a systematic study of these two paradigms via their application in two disparate scientific domains: the analysis of Nuclear Magnetic Resonance (NMR) spectra for biochemical understanding and the prediction of feed consumption for agricultural robotics. While both domains have differing scientific objectives and data modalities, they serve as powerful case studies. The NMR project illustrates the inferential approach, by using Bayesian models to estimate the parameters of a physical process. On the other hand, the agricultural project

showcases the predictive approach, by using deep learning to build a forecasting tool for automating an agricultural process. By juxtaposing these applications, I aim to show by this work the respective strengths, limitations and complementary nature of inferential and predictive modeling in solving real-world problems.

1.1 Data-driven approaches for modeling physical systems

In recent years, advances in machine learning have revolutionized the ability to model time-dependent processes [Vaswani et al. \[2017\]](#), [Hochreiter and Schmidhuber \[1997\]](#). While often grouped together, the underlying methodologies can be broadly categorized based on their primary objective: inference or prediction.

1. **Bayesian inference for scientific understanding:** Bayesian inference, which belongs to the class of inferential methods, is centered on using data to update our knowledge about the parameters of a generative (often physical) model [Gelman et al. \[2013\]](#), [Ghahramani \[2015\]](#). Bayesian methods provide a rigorous mathematical framework for incorporating such prior knowledge in order to yield a full *posterior distribution* for each model parameter. This *posterior distribution* provides a shape of uncertainty for each model parameter given the data provided. This "white box" approach to mathematical modelling is quite effective for parameter estimation in noisy or sparse data environments where understanding the physical meaning as well as uncertainty of the model components is paramount [Breiman \[2001\]](#), [Gelman et al. \[2013\]](#).
2. **Neural networks for predictive modeling:** Methods in this paradigm excel at learning complex, non-linear relationships directly from high dimensional data, often without an explicit underlying physical model [Goodfellow et al. \[2016\]](#), [Li et al. \[2021\]](#). These methods have become the state-of-the art for tasks where predictive accuracy is the primary metric of success (such as computer

vision). These "black box" models are powerful function approximators, and consequently, are ideal for building automated systems that must respond to complex patterns in input data.

These distinct yet complementary approaches form the methodological framework of this dissertation, and in subsequent chapters, I explore applications to spectroscopy and agricultural robotics.

1.1.1 Bayesian models for NMR spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy uses magnetic resonance to probe molecular dynamics and chemical exchange processes in chemical and biological compounds [Jacobsen \[2007\]](#), [Bain \[2003\]](#). However, accurately quantifying exchange rates and spectral parameters from NMR data still remains a challenging task due to spectral overlap, noise, and the complexity of underlying kinetics [Rule and Hitchens \[2006b\]](#), [Bain et al. \[2001\]](#). In this dissertation, I introduce two distinct Bayesian modeling approaches to address different aspects of this challenge, using fluorine-19 (F-19) data as a case study.

- The first is a non-parametric [Sgouralis et al. \[2018\]](#), [Jazani et al. \[2019b\]](#), [Tavakoli et al. \[2020b,a\]](#) Bayesian model that estimates not only the spectral parameters, but also the number of resonances present in a given 1D NMR spectra. Although this approach does not explicitly model the kinetics associated with chemical exchange, it provides a highly flexible method for spectral decomposition in complex biological compound.
- The second is a parametric [Ghahramani \[2015\]](#) Bayesian model in which the user supplies the number of resonances a priori. This model incorporates the effects of chemical exchange between spin populations in the given compound and estimates the relevant parameters associated with each resonance.

The particular Bayesian framework used here allows for flexible modeling of spectral components and robust, automatic uncertainty quantification of estimated parameter values.

1.1.2 Deep Learning for predictive modeling in agricultural robotics

In an entirely different domain, precision agriculture has adopted machine learning for optimizing livestock management and feed distribution [Liakos et al. \[2018\]](#), [Porto et al. \[2013\]](#), [Zheng et al. \[2018\]](#). One such utility is predicting feed leftover in dairy robotic feeding systems using image data of leftover feed for the purpose of developing custom feeding regimes for livestock. This dissertation describes a neural network model [Goodfellow et al. \[2016\]](#) developed to predict feed consumption patterns in dairy. Ultimately, this is a first step towards a unified precision system for dairy cows. The model enhances automation in dairy farming by providing accurate, real-time predictions, thereby reducing food waste.

1.2 A unified methodological perspective

Although there appears to be a domain disconnectedness between NMR spectroscopy and agricultural robotics, they provide a powerful lens through which we may compare the roles of inference and prediction in data driven science. The NMR studies display the power of Bayesian inference to deconstruct a complex signal into its physically meaningful components and to provide a robust measure of certainty in those findings. On the other hand, the feed prediction projects highlights the predictive power of deep learning to result in automation, even when the underlying patterns in the data are too complex to be captured by a simple mechanistic model. In the end, this dissertation argues that the choice of a data-driven model doesn't just depend on the technicality of the problem. There appears to be a critical philosophical detail

dictated by the scientific question at hand: Is the goal to understand or to predict? In the pages that follow, I detail my contributions and contend that both choices are valid and powerful, and the modern computational scientist will be tasked in knowing which tool to deploy and how to interpret its results in the context of the scientific goal. Ultimately, the insights are not only in the data, but in the questions we are ask of it and the methods we choose to answer them.

On a fundamental level, both models I describe in this dissertation capture a time evolution of a physical object (in the case of feed level in a feeding bin for cows) and state (in the case of FID signal for NMR). They both utilize data to derive insights, although they go about it in different manners. This is a crucial point with data-driven methods because for this particular reason, the model insights are heavily impacted by data.

PART 1

Bayesian Inference for Parameter Estimation in NMR Spectroscopy

Chapter 2

Introduction to NMR spectroscopy and Bayesian modeling

2.1 Fundamental principles of NMR spectroscopy

The idea behind NMR spectroscopy is very similar to other forms of spectroscopy [Jacobsen \[2007\]](#), [Rule and Hitchens \[2006a\]](#). In spectroscopy generally, a photon of electromagnetic radiation induces a transition between two discrete energy states: from a lower-energy ground state to a higher energy excited state. For instance, in visible light spectroscopy, electrons within atoms or molecules absorb photon energy, which enables them to move to higher-energy states. These energy states are specific, fixed energy levels that, in the case of visible light, electrons in atoms or molecules can occupy.

NMR spectroscopy, operates in an analogous manner except that this transition occurs in the nuclei of the atoms on quantum objects called nuclear spins [Jacobsen \[2007\]](#), [Kaplan \[2012\]](#). In addition, the spins transition to higher energy levels when their containing nucleus experience a certain external magnetic field.

Unlike other forms of spectroscopy, the generation of the ground and excited NMR states requires the existence of an external magnetic field for probing *nuclear spin* states [Jacobsen \[2007\]](#), [Kaplan \[2012\]](#).

2.1.1 Nuclear spins

Atomic nuclei consists of nucleons which refer to a composition of protons and neutrons [Rutherford and Mott \[1932\]](#). These nucleons give rise to a quantum mechanical property called the nuclear spin. The spins of nuclei behave like tiny magnets in the sense that when placed in a strong external magnetic field, they can take distinct discrete orientations to the field ranging from parallel to antiparallel. These orientations are called spin states, such that the parallel orientation is the state with the lowest energy while the antiparallel orientation, the state with the highest energy.

The *spin quantum number*, I , quantifies the intrinsic spin angular momentum of atomic nuclei. For atomic nuclei with an even number of protons and neutrons (such as ^{12}C , ^{16}O), the spin quantum number is zero, in which case, there is no net spin and the nuclei do not exhibit NMR signals. Similarly, nuclei with an odd total number of protons and neutrons (such as ^1H , ^{13}C), I is half-integer. Nuclei with even numbers of protons and neutrons tend to have integer spins. Examples of such nuclei are ^{14}F which will be a primary focus of one of the projects described in this dissertation.

According to quantum theory, a nucleus with spin I can adopt $2I + 1$ possible spin orientations in an external magnetic field [Atkins and De Paula \[2006\]](#). Table 2.1 below shows examples of spin orientations for common nuclei.

Table 2.1: Discrete Nuclear Spin Orientations Based on Spin Quantum Number (I)

Spin quantum number (I)	Possible orientations (m_I)	Description
$\frac{1}{2}$	$+\frac{1}{2}, -\frac{1}{2}$	Parallel and antiparallel orientations
1	$+1, 0, -1$	Parallel, perpendicular, and antiparallel orientations
$\frac{3}{2}$	$+\frac{3}{2}, +\frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}$	Four discrete orientations

2.1.2 ^{19}F NMR experiments

In a typical NMR experiment, a biological sample placed in an NMR spectrometer is irradiated with electromagnetic radiation of a specified frequency [Jacobsen \[2007\]](#), [Kaplan \[2012\]](#). Specific nuclei, called *NMR active nuclei* experience a transition from a lower energy state to a higher energy state. When the energy of this radiation matches the energy difference between the spin states, absorption occurs and the spins at the lower energy state transition to a higher energy state.

The energy difference (ΔE) between nuclear spin states is the foundation of the NMR phenomenon. However, an NMR spectrometer measures frequencies, not energy. We find a connection between frequencies and energy established by the Planck-Einstein relation [Atkins and De Paula \[2006\]](#), which states that the energy of a photon (E) is directly proportional to its frequency (ν):

$$E = h\nu \tag{2.1}$$

where h is the Planck's constant.

Each nucleus type (for instance ^{13}C , ^1H) has a gyromagnetic ratio (γ) which describes how magnetic that particular nucleus is relative to its spin. The energy difference between spin states is directly related to the product of the gyromagnetic ratio and the strength of the applied magnetic field (B_0) according to the following equation:

$$\Delta E = \gamma \hbar B_0 = \frac{\gamma \hbar B_0}{2\pi} \tag{2.2}$$

where $\hbar = \frac{h}{2\pi}$. Solving for the frequency, ν , we arrive at a cornerstone equation in NMR, the *Larmor frequency* (ν_L):

$$\nu_L = \frac{\gamma B_0}{2\pi} \tag{2.3}$$

The Larmor frequency is the specific resonant frequency at which a nucleus with a gyromagnetic ratio γ will absorb energy when placed in a magnetic field of strength B_0 . However, in a real molecule, the Larmor frequency is not a single, fixed value for all nuclei of the same type (e.g. all ^{19}F atoms). The local cloud of electrons surrounding each nucleus generates its own small magnetic field that opposes the main field B_0 . This phenomenon, known as *chemical shielding* (σ), causes each nucleus to experience a slightly different effective magnetic field, B_{eff} :

$$B_{eff} = (1 - \sigma)B_0 \quad (2.4)$$

This means that nuclei in different chemical environments will have a unique Larmor frequency. Consequently, to get a complete picture of the sample, an experiment does not excite just one frequency but a continuous range of them simultaneously.

Manipulating Nuclear Spins

At thermal equilibrium, the population of nuclei is not evenly distributed between the spin states. A slight excess of nuclei will usually occupy the lower-energy (parallel) state compared to the higher-energy (antiparallel) state [Tolman \[1979\]](#), [Bloch \[1946\]](#). This small population difference results in a net bulk magnetization (M_0) that is aligned with the main magnetic field (B_0). This alignment is referred to as the *longitudinal magnetization*, as it lies along the primary axis of the magnetic field (conventionally the z-axis). In this state of equilibrium, no signal can be detected.

To generate a measurable NMR signal, this equilibrium must be disturbed. This disturbance is accomplished through a *pulse experiment*. In a pulse experiment, a short and intense pulse of radiofrequency (RF) energy is directed at the sample. To address the challenge of exciting the entire range of Larmor frequencies caused by chemical shift, this pulse is not monochromatic. Instead, it is designed to contain a broad band of frequencies.

When this hard pulse is applied, its associated magnetic field (B_1) exerts a torque on the net magnetization vector (M_0), tipping it away from its stable alignment along the z-axis. A precisely calibrated 90° pulse, which is commonly used in NMR experiments, has the exact duration and power to rotate the magnetization by 90 degrees, moving it entirely into the xy-plane.

The component of magnetization now rotating in the xy-plane is called the transverse magnetization. This rotating magnetic field is the source of the NMR signal [Rule and Hitchens \[2006b\]](#), [Jacobsen \[2007\]](#). As it precesses around the main B_0 field at the various Larmor frequencies of the excited nuclei, it induces a small, oscillating electrical current in a detector coil placed around the sample.

Immediately after the RF pulse is turned off, the signal is at its strongest. At this moment, the net magnetization vector has been tipped into the xy-plane, and just as importantly, the individual nuclear spins that constitute this vector are precessing in a coherent phase. Once the pulse ends, the system immediately begins to relax back to equilibrium. The individual nuclear spins, initially precessing in a coherent phase, now precess freely, and lose their synchrony in a process called *dephasing*. This loss of coherence causes the net transverse magnetization to decrease. This decaying, oscillating signal is called the *Free Induction Decay* (FID), which is the time-domain data the methods I describe later aim to analyze.

2.1.3 Existing analysis approaches to 1D NMR

The analysis of 1D NMR spectra, which involves the extraction of chemically meaningful information from the raw spectral data has progressed from the painstaking manual interpretation of relatively simple spectra in the early days of NMR to the highly sophisticated, often automated, computational methodologies employed today which are able to analyze highly complex data [Gutowsky et al. \[1951\]](#), [Ernst and Anderson \[1966\]](#), [Li et al. \[2021\]](#), [Cheng et al. \[2014\]](#). In other to place the methods I describe later within an appropriate context, in this section, I present key historical

milestones in the analysis of 1D NMR data. The story begins with the introduction of Fourier Transform (FT) NMR by Richard R. Ernst and Weston A. Anderson [Ernst and Anderson \[1966\]](#) which fundamentally revolutionized the manner in which NMR data were acquired and processed, catapulting the technique into a new era of sensitivity and versatility. Unlike Continuous wave-NMR [Rule and Hitchens \[2006a\]](#) which performed a sequence-by-sequence frequency scanning of nuclei, FT-NMR (as described above) employs a short, intense pulse of radiofrequency energy that simultaneously excites all nuclei of a given type (e.g., all protons) within a defined frequency range [Rule and Hitchens \[2006a\]](#).

The advent of FT-NMR, with its ability to generate vast quantities of digital data rapidly, brought about a parallel evolution in computational methods for data processing and spectral interpretation [Ernst and Anderson \[1966\]](#), [Lindon and Ferrige \[1980\]](#). Most methods that analyze NMR data perform specific preprocessing steps before the analysis: apodization, where a weighting function is applied to the time-domain data to enhance either Signal-to-Noise ratio (S/N) or resolution at the expense of the other; zero-filling, where the NMR spectrum is augmented with zeros prior to Fourier transformation to improve the digital resolution of the resulting spectrum; the Fourier transformation itself; phase correction, a critical step to ensure that all spectral lines have a pure absorptive lineshape, which is essential for accurate integration and interpretation; and baseline correction, to remove any broad, underlying distortions from the spectrum [Li et al. \[2022\]](#), [Kaplan \[2012\]](#), [Rule and Hitchens \[2006b\]](#), [Chitode \[2008\]](#)

The fundamental analysis task is to extract the key parameters that define each resonance. These are its frequency (chemical shift), intensity, and lineshape. [Bain et al. \[2001\]](#), [Rule and Hitchens \[2006b\]](#). The process often begins with peak picking which entails the accurate identification of the frequencies (chemical shifts) and intensities of the NMR signals. Early approaches involved manual visual inspection and annotation of peaks. However, as spectra became more complex, the need for

automation grew, and this led to the development of various automated peak picking algorithms. These earlier algorithms often relied on identifying local maxima in the spectrum, using derivative methods to locate inflection points, or employing more sophisticated signal processing techniques [Kleywegt et al. \[1990\]](#), [Liu et al. \[2012\]](#).

The limitations of these local maxima methods, especially in crowded spectral regions, prompted a paradigm shift away from simple peak detection towards more holistic approaches known as spectral deconvolution or lineshape fitting [Koradi et al. \[1998\]](#), [Schmid et al. \[2023\]](#), [Alipanahi et al. \[2009\]](#). This new paradigm models a region of the spectrum as a summation of theoretical lineshapes (typically Lorentzian or Gaussian functions) [Bloch \[1946\]](#), [Rule and Hitchens \[2006a\]](#) to extract the parameters of the underlying signals. Modern NMR software packages, like Mestrelab Research S.L., frequently incorporate advanced peak picking routines, such as those based on Global Spectral Deconvolution (GSD), which can effectively identify and parameterize even highly overlapped peaks [Mestrelab Research S.L., Liu et al. \[2023\]](#). More recently, Bayesian statistical methods and machine learning algorithms are being increasingly applied to enhance the accuracy of peak picking, particularly in distinguishing genuine peaks from noise, baseline artifacts, or instrumental glitches. For instance, the UnidecNMR approach utilizes a Bayesian deconvolution algorithm for robust peak detection [Buchanan et al. \[2025\]](#), while other methods model the spectrum as a mixture of Gaussian and/or Lorentzian density functions to identify underlying signals [Li et al. \[2021\]](#), [Cheng et al. \[2014\]](#). Deep learning models, such as NMRNet and DEEP Picker, trained on large datasets of experimental or synthetic spectra, are also demonstrating considerable promise in achieving human-level accuracy in peak picking [Li et al. \[2021, 2022\]](#). Despite these advances, accurate and reliable peak picking remains a significant challenge, especially in spectra characterized by extensive peak overlap (crowded regions), low signal-to-noise ratios, or the presence of complex artifacts [Li et al. \[2021\]](#). For instance, most of the methods described above are parametric, and so, they require the user to pre-specify the number of peaks in

a given region which is an extremely difficult task in spectra with extensive peak overlap. Furthermore, while machine learning models show promise, they can be "black boxes" and typically struggle with novel signal types not present in their training data. Crucially, many existing methods provide only point estimates of parameters, failing to robustly quantify the uncertainty associated with these values, which is especially critical in spectra with low signal-to-noise ratios. The persistent need for expert-guided intervention and the lack of rigorous uncertainty quantification for complex spectra motivate the development of a new class of analysis tools. The Bayesian non-parametric method presented in this dissertation is designed to directly address these limitations.

2.2 A Bayesian framework for parameter inference

Scientific inquiry fundamentally involves creating mathematical models of the physical world. These models are parameterized by physical quantities and have been shown to affect the underlying system being studied. A foundation model for NMR data is provided by the Bloch's equations, which are extended by the Bloch-McConnell's equations [Bloch \[1946\]](#), [McConnell \[1958\]](#) to describe systems with chemical exchange. Both equations are parameterized by physical quantities such as reaction rates and relaxation times and the goal of data analysis is to infer the values of these parameters from experimental data. Bayesian inference provides a rigorous probabilistic framework for doing this, thereby allowing us to reason about parameter uncertainty in a principled manner [Gelman et al. \[2013\]](#).

The foundation of this framework is *Bayes' Rule*, which describes how our knowledge about a set of parameters, θ , is updated in light of observed data, D :

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \tag{2.5}$$

The terms in this equation describe a complete cycle of learning:

- *Prior Probability, $P(\theta)$* : This distribution captures our prior knowledge about the model parameters before observing the data. At this level, we can enforce physical constraints, such as requiring the rate to be positive, or the phase to be within a certain range (say $-\frac{\pi}{2}$ to $\frac{\pi}{2}$). The stronger the restrictions placed on the prior distributions, the *more informative* it is, and the greater influence it will have on our inference.
- *Likelihood, $P(D|\theta)$* : This is the forward model that connects the parameters to the data. It answers the question: "Assuming a particular set of parameter values θ , what is the probability of observing our experimental data D ?"
- *Posterior Probability, $P(\theta|D)$* : This is the result of the inference. The posterior distribution represents our updated knowledge about the parameters after taking the data into account. It rigorously combines the prior knowledge with information contained in the data to provide a complete picture of which parameter values are most plausible and what their associated uncertainties are. In the next pages, I describe two different posteriors: one formulates its likelihood from Bloch's equations and the other from the Bloch-McConnell's equations. Both of them predict the set of parameters that describe a given NMR signal.

2.2.1 Computational inference via Markov Chain Monte Carlo (MCMC) Sampling

Despite its elegance, for practical uses, Bayes' rule rarely gives a posterior distribution, $P(\theta|D)$, that can be described by a simple analytical formula. This is because the posterior is a product of the likelihood and prior distributions, and any non-trivial model typically results in a complex, high-dimensional shape that is not a known, named distribution [Gelman et al. \[2013\]](#). Hence, we are not able to plot the posterior from an equation and make deductions.

To solve this problem, we use computational methods to represent the posterior distribution [Gelman et al. \[2013\]](#), [Metropolis et al. \[1953\]](#). This is done by *sampling*. In sampling, instead of describing the posterior distribution with an equation, we draw a large amount of samples from it. The density of these collected samples in any region of the parameter space will be proportion to the posterior probability in that region. Markov Chain Monte Carlo (MCMC) is a class of algorithms designed to perform this robust sampling [Metropolis et al. \[1953\]](#), [Hastings \[1970\]](#), [Gelfand and Smith \[1990\]](#), [Gelman et al. \[2013\]](#). An MCMC sampler constructs a *random walk* through the parameter space that is guaranteed to be *ergodic*. A random walk is said to be ergodic if with infinite time, it can reach all parts of the distribution with non-zero probability. The collection of samples visited during this walk forms a chain that, in the long-run, will converge to the true posterior distribution.

A significant challenge in MCMC sampling is the potential existence of multiple, well-separated regions of high probability in the posterior distribution [Robert et al. \[1999\]](#). A simple sampler, like the Metropolis algorithm [Chib and Greenberg \[1995\]](#) which is based on a basic random walk proposal will usually have difficulty traversing the low-probability regions between these maxima. This is because when it enters one of the local maxima, it tends to stay there for a long time; making the exploration of the parameter space inefficient.

To overcome this challenge and ensure a more robust exploration of the posterior landscape, we employed more sophisticated approaches. The two main strategies used to sample the parameters for our models were *slice sampling*, developed in [Neal \[2003\]](#), which automatically adapts its step size to the local shape of the distribution, and a multiplicative random walk, where proposals were generated using an elliptical slice sampling scheme [Murray et al. \[2010\]](#) and further adapted to dynamical systems in [McCoy et al. \[2025\]](#).

Chapter 3

Bayesian nonparametric statistical learning for resonance identification and characterization in ^{19}F -NMR

3.1 Disclosure

This chapter is adapted from a manuscript currently being prepared for journal submission has been revised to fit coherently with the broader narrative flow.

3.2 Introduction

Fluorine is a significant element in chemistry, biology, and medicine because it can distinctively modify the properties and functions of molecules but most importantly because it is magnetically active. Fluorine NMR is a type of nuclear magnetic resonance spectroscopy that uses the ^{19}F nucleus to detect and identify fluorine-containing compounds [Chekmenev et al. \[2008\]](#), [Kitevski-LeBlanc and Prosser \[2012\]](#), [Zientek et al. \[2014\]](#). ^{19}F -NMR can provide information about the structure, dynamics, interactions, and mechanisms of fluorinated molecules or molecular systems [Rule and Hitchens \[2006a\]](#).

Important advantages of ^{19}F -NMR over the more common ^1H - or ^{13}C -NMR include: high sensitivity as ^{19}F is the third most receptive NMR nucleus after ^3H and ^1H , and it is 100% naturally abundant; large chemical shift range, which allows fine separation of different ^{19}F sites; and development of characteristic coupling patterns with other nuclei, such as ^1H , ^{13}C , ^{31}P , and ^{15}N , which can reveal the connectivity and stereochemistry of the molecules of interest. Due to these advantages, ^{19}F -NMR is nowadays used to study complex biological samples, such as enzymes, proteins, nucleic acids, and membranes, [Chekmenev et al. \[2008\]](#), [Kitevski-LeBlanc and Prosser \[2012\]](#), [Zientek et al. \[2014\]](#).

However, widespread adoption of ^{19}F -NMR is currently limited due to poor peak separation and inaccurate integration caused by the very wide spectral range of ^{19}F and substantial background signals from fluorinated impurities or solvents. Current research seeks to improve peak separation and integration during data acquisition,

where experiments are conducted with high-purity reagents and deuterated solvents that are free of fluorine contamination. Similarly, current research aims to suppress background signals during data acquisition in experiments using specialized pulse sequences or filters. Nevertheless, both approaches require demanding protocols and non-standard supplies.

Alternatively, the limitations of ^{19}F -NMR could also be avoided or at least improved by appropriate data analysis methods adapted to the ^{19}F signals. To this end, current approaches use artificial intelligence (AI) and advanced machine learning to train specialized neural networks to recognize NMR signals [Li et al. \[2021, 2022\]](#), [Jahangiri et al. \[2023\]](#), [Schmid et al. \[2023\]](#). Nevertheless, the apparent lack of existing training data specifically covering ^{19}F -NMR conditions or reliable pre-trained AI models targeting ^{19}F do not allow fully automated analysis workflows.

Conventionally, NMR data, including ^{19}F data, are summarized in spectra, which represent the intensity of the NMR signal as a function of frequency. Such spectra are made up of a series of resonances, each associated with a unique nuclear spin possessed by a particular atom type or atomic group in the molecule of interest. In a spectrum, each resonance peak is uniquely determined by its center frequency, often given in chemical shifts, its shape, and its amplitude and phase [Bruce et al. \[2000\]](#), [Ramani et al. \[1976\]](#).

A complete analysis of an NMR dataset in the frequency domain involves identifying the complete set of resonance peaks in its spectrum and assigning each peak to its relevant atomic group [Rubtsov and Griffin \[2007\]](#), [Cheng et al. \[2014\]](#). To date, this analysis has been done largely with the intervention of a human expert. This is due to the fact that the spectra of biological NMR data often contain overlapping peaks that make the characterization process difficult without human assistance [Teng \[2012\]](#). Moreover, in certain NMR experiments, including ^{19}F -NMR, noise has a strong interference with the actual signal, resulting in substantial ambiguity during characterization.

Over the years, different methods have been proposed for identifying the resonance peaks as well as for the subsequent parameter fitting. Early methods used matrix factorization approaches as well as singular value decomposition to identify resonances [Korzhnev et al. \[2001\]](#), [Tikole et al. \[2014\]](#), [Alipanahi et al. \[2009\]](#). These methods required engineered spectral features to be extracted after preprocessing, and using a heuristic process, each data point was classified as belonging to a resonance or not. However, the reproducibility of these methods is limited, as heuristics or non-principled feature engineering are primarily based on human experience. Alternative parametric approaches, such as those employing Bayesian statistics [Rubtsov and Griffin \[2007\]](#), [Matviychuk et al. \[2017\]](#), typically model spectra as mixtures of multivariate Gaussian densities and attempt to filter them in accordance with relevant peak parameters. However, due to their parametric nature, these methods face the model selection problem and require predetermination of the total number of peaks present before characterization can occur [Rubtsov and Griffin \[2007\]](#). More recent methods have utilized machine learning via neural networks with specialized architectures to predict chemical shifts, and also identify peaks [Li et al. \[2021\]](#). As is typical with neural network approaches, the robustness of these methods depends on the quality and size of the training dataset which are typically unavailable for complex biomolecules probed under realistic experimental conditions.

In this study, I introduce a new technique that takes advantage of Bayesian nonparametrics, a field of statistics that has only become widely used in physical applications in recent years [Hines \[2015\]](#), [Tavakoli et al. \[2019\]](#), [Sgouralis et al. \[2018\]](#), [Jazani et al. \[2019a\]](#), [Tavakoli et al. \[2020c,b,a\]](#). A key innovation of my method is that it can account for an arbitrary number of resonances in a given NMR dataset, including ^{19}F -NMR, thus entirely eliminating the need for human input. Using the proposed method, I can detect the resonance peaks and accurately estimate their characteristics.

In the sections that follow, I provide a detailed description of my method and then demonstrate the application of this method to raw experimental data. I cap the results section off by providing detailed validation with synthetic data.

3.3 Methods

3.3.1 Model overview

My analysis begins with an experimentally obtained time series of FID data and derives estimates of all the quantities that characterize the resonances present in the data. These quantities include: (i) the number of resonances present, (ii) resonance frequencies, decay rates, amplitudes, and phases. In the following, I explain each one in detail. Computational details of the entire method are available in the Appendix.

3.3.2 Model description

The starting point of my analysis is the raw data, namely the FID measurements $z_{1:N}$ where each datapoint $z_n = (x_n, y_n)$ denotes the free induction decay detected at time t_n by the spectrometer in quadrature detection. At the timescale of interest, individual detections of the FID signal occur stochastically and independently from each other. I assume that the signal recorded by the spectrometer consists of the true FID measurements contaminated with additive noise, see Appendix [appendix A](#). Accordingly, each data point follows a Gaussian distribution

$$\begin{aligned} x_n &\sim \text{Normal} \left(P + \sum_{k=1}^K X_n^k, \frac{1}{\tau} \right) \\ y_n &\sim \text{Normal} \left(Q + \sum_{k=1}^K Y_n^k, \frac{1}{\tau} \right). \end{aligned} \quad n = 1, \dots, N \quad (3.1)$$

Here, P and Q are the spectrometer offsets in the real and imaginary channels respectively, while X_n^k and Y_n^k are representative models of a resonance's contribution.

According to Bloch's equations [Bloch \[1946\]](#), they are given as follows:

$$X_n^k = p^k e^{-r^k T_n} \cos(w^k T_n) - q^k e^{-r^k T_n} \sin(w^k T_n), \quad (3.2)$$

$$Y_n^k = q^k e^{-r^k T_n} \cos(w^k T_n) + p^k e^{-r^k T_n} \sin(w^k T_n), \quad (3.3)$$

where $T_n = t_n - t_{\min}$ is the time elapsed since the onset of the detection phase $t_{\min}$; while, p^k and q^k are the real and imaginary parts of the k^{th} resonance's amplitude, r^k represents its relaxation rate, and w^k its frequency. In my formulation, resonances are labeled with superscripts $k = 1, \dots, K$ and time-points are labeled with subscripts $n = 1, \dots, N$. The value of K is *unknown* as, typically, the number of resonances is unidentified. Finally, the noise variance, which is equal to $1/\tau$ in my formulation, represents the standard deviation of the FID measurements in both real and imaginary channels.

Although my analysis method remains exclusively in the time domain, to help with the interpretation, I can readily convert my results to the frequency domain. Namely, the k^{th} resonance's contribution, modeled by eqs. (3.2) and (3.3), in the frequency domain corresponds to a spectrum of the form

$$B^k(\nu) = (p^k + iq^k) \frac{e^{-(r^k + i(2\pi\nu - w^k))t_{\min}}}{r^k + i(2\pi\nu - w^k)}. \quad (3.4)$$

Moreover, the entire FID corresponds to the spectrum

$$B(\nu) = \sum_{k=1}^K B^k(\nu). \quad (3.5)$$

For details see Appendix [appendix A](#). I note that, since the derivation of these spectra does not rely on discrete Fourier transforms [Winograd \[1978\]](#), my results are *unaffected* by aliasing, filtering, or windowing choices.

3.3.3 Model inference

The quantities I wish to estimate, such as the resonance frequencies w^k , relaxation rates r^k , and amplitudes (p^k, q^k) , are introduced as model variables in the preceding formulation. I follow the Bayesian paradigm to estimate the values of these variables [Brooks et al. \[2011\]](#). Each of these variables, being model parameters, requires priors. I make clear choices for these priors for interpretation and computational convenience, adopting the distributions described in [Appendix appendix A](#).

To allow the complexity of my model to fluctuate based on the number of resonances present in the data, I turn to the nonparametric formulation. Namely, I recast [eq. \(3.2\)](#) and [eq. \(3.3\)](#) as

$$X_n^k = b^k \left(p^k e^{-r^k T_n} \cos(w^k T_n) - q^k e^{-r^k T_n} \sin(w^k T_n) \right), \quad (3.6)$$

$$Y_n^k = b^k \left(q^k e^{-r^k T_n} \cos(w^k T_n) + p^k e^{-r^k T_n} \sin(w^k T_n) \right). \quad (3.7)$$

I restrict the newly introduced variables b^k , one for each model resonances, and in my formulation, they may take only values 1 or 0. Specifically, the possibility that $b^k = 0$, coinciding with the case where the k^{th} resonance does not contribute to the data, allows me to introduce a technically infinite number of model resonances. With the introduction of b^k , I can estimate the number of resonances that contribute to the data (termed “active” to distinguish them from those that do not contribute, termed “inactive”) simultaneously with the rest of the parameters simply by treating each b^k as a separate parameter and estimating its value (of 1 for active resonances and 0 for inactive ones) [Jazani et al. \[2019b\]](#), [Paisley and Carin \[2009\]](#).

To estimate the value of each b^k , I apply a prior

$$b^k | q^k \sim \text{Bernoulli}(q^k), \quad k = 1, \dots, K$$

and subsequently a hyperprior on q^k in order to learn precisely how many model resonances are active in my data. For the latter, I choose $q^k \sim \text{Beta}(A, B)$ with specifically chosen hyperparameters A and B . Both steps can be combined by invoking the Beta-Bernoulli process [Paisley and Carin \[2009\]](#), which is described in detail in the Appendix appendix [A](#).

Once the choices for the priors above are made, I form a joint posterior probability distribution, given by

$$p\left(\tau, P, Q, p^{1:K}, q^{1:K}, r^{1:K}, w^{1:K}, b^{1:K} \mid z_{1:N}\right), \quad (3.8)$$

encompassing all unknown variables which I may wish to determine [van de Schoot et al. \[2021\]](#), [Von Toussaint \[2011\]](#), [Sivia and Skilling \[2006\]](#), [Gelman et al. \[2013\]](#), [Ghahramani \[2015\]](#). The nonparametric prior on b^k excludes analytic forms for my posterior. For this reason, as I detail in the Appendix appendix [A](#), I develop a computational scheme that takes advantage of the Monte Carlo Markov chain that can be used to generate pseudorandom samples from this posterior [Brooks et al. \[2011\]](#).

Synthetic data

I obtain the synthetic data presented in the Results section through standard pseudorandom computer simulations that mimic the Bloch’s equation formulation of the FID.

3.4 Results

The proposed method can be used for the analysis of 1D NMR data and becomes particularly powerful in the analysis of novel signals such as those of ^{19}F -NMR that remain poorly understood. As I explain in section [3.3](#), I built upon the Bayesian paradigm of data analysis [Dou and Hodgson \[1995\]](#). In particular, the

method develops and computes a full posterior probability distribution over an unknown number of resonances and all unknown parameters of such resonances. The parameters of interest include (i) frequencies, (ii) relaxation rates, and (iii) amplitudes and phases. However, instead of working with the Fourier transformed spectrum, which most spectroscopists are familiar with, I analyze the raw free-induction decay (FID) forming the NMR signal as recorded by the spectrometer directly in the time domain. Using raw FID measurements directly, I avoid preprocessing steps, such as filtering or windowing of the FID, thus avoiding artifacts that may be introduced [Chitode \[2008\]](#). Such aspects are particularly important in ^{19}F -NMR, whose chemical-shift range far exceeds that of conventional ^1H experiments, and subsequently demands wider spectral widths and careful calibration.

In this section, I begin by showing how this novel method performs on raw NMR data from real-life experiments. Subsequently, I demonstrate and validate the method and testing against ground truth using synthetic (simulated) data that mimic the properties of real ones.

I provide the first two results in time domain format since model’s data are FID measurements in time domain. Subsequent spectral results are provided in the frequency domain as these are the easiest to interpret.

3.4.1 Application to raw data from hTRPM8-VSLD experiments

With selective labeling

First, I apply the method to the analysis of ^{19}F -data of hTRPM8-VSLD, a protein that is responsible for sensing thermal stimuli [Yin et al. \[2018\]](#). My input data consists of the FID measurements obtained in quadrature detection as a real time series and an imaginary time series that evolve over time following a single pulse excitation [Rule and Hitchens \[2006a\]](#). Figure [fig. 3.1](#) illustrates the FID measurements in both the

time domain and the frequency domain, highlighting the dual representation nature of the dataset. Subsequent figures mainly employ the more commonly used frequency-domain representation.

I assessed my method’s robustness in precisely identifying and characterizing resonances. This was done by analyzing NMR signals from the protein, which underwent fluorine labeling at four specific sites.

Full characterization

Next, I obtained and analyzed NMR signals of the same protein under different signal-to-noise ratio (SNR) conditions. For clarity, I calculate the SNR of the k^{th} resonance with the formula:

$$\text{SNR}^k = \sqrt{\tau((p^k)^2 + (q^k)^2)}. \quad (3.9)$$

where the parameters p^k , q^k , τ are described in section 3.3. These quantify amplitude and noise. From fig. 3.2, the signals observed from datasets with more repeats appear to have higher signal-to-noise ratio as expected.

The lowest resolution signal consisted of 4 scans, the highest resolution 1024 scans, and each intermediate signal was obtained by quadrupling the number of scans in the previous one. In this context, a scan is a repetition of an NMR pulsed experiment. For each scan, FID measurements are added, yielding an increase in signal intensity. During this signal-averaging process, random noise is suppressed proportionally to the square root of the number of scans [Jacobsen \[2007\]](#).

As a consequence of the increase in SNR, I obtain sharper posterior estimates for each of the parameters, reflecting greater certainty on the results. This behavior is accurately reflected by the method as shown in fig. 3.1 and fig. 3.3.

For example, in fig. 3.1, I show a characterization of FID data for the cleanest dataset by the proposed method. It strongly identifies 4 resonances whose chemical shifts are

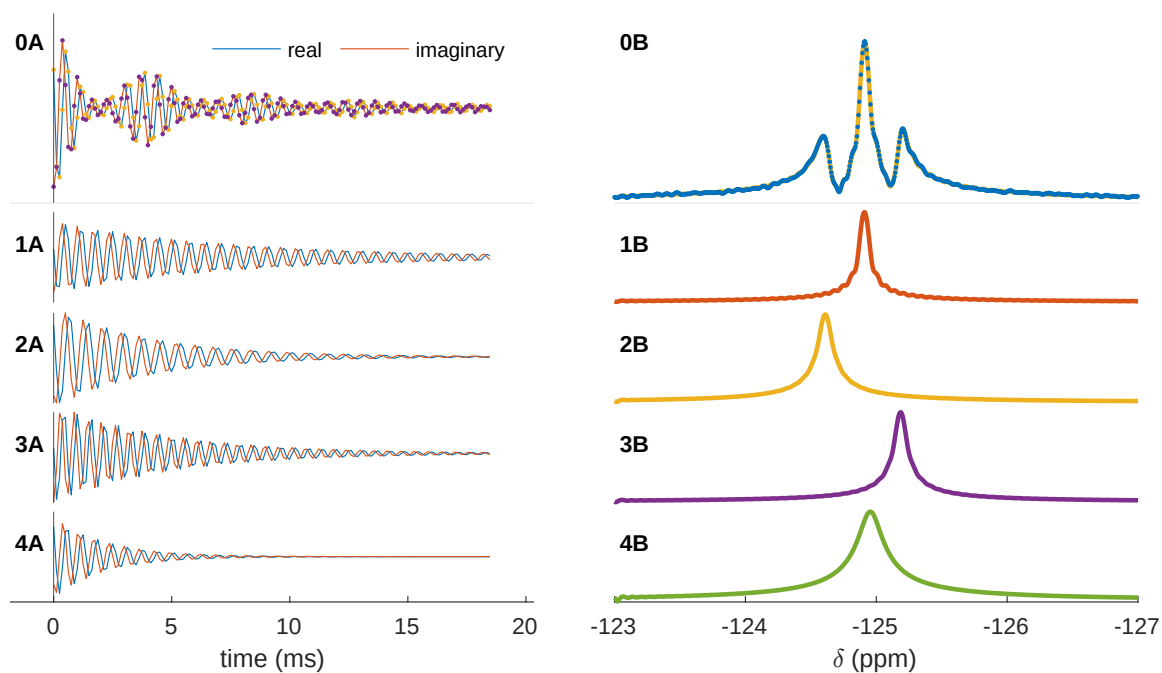


Figure 3.1: Spectra representations of deconvolved resonances. Plots 0A-4A: Individual resonances deduced from posterior estimates of the FID data from 1024 scans, of ^{19}F hTRPM8-VSLD. Plots 0B-4B: Frequency domain interpretation of plots 0A-4A

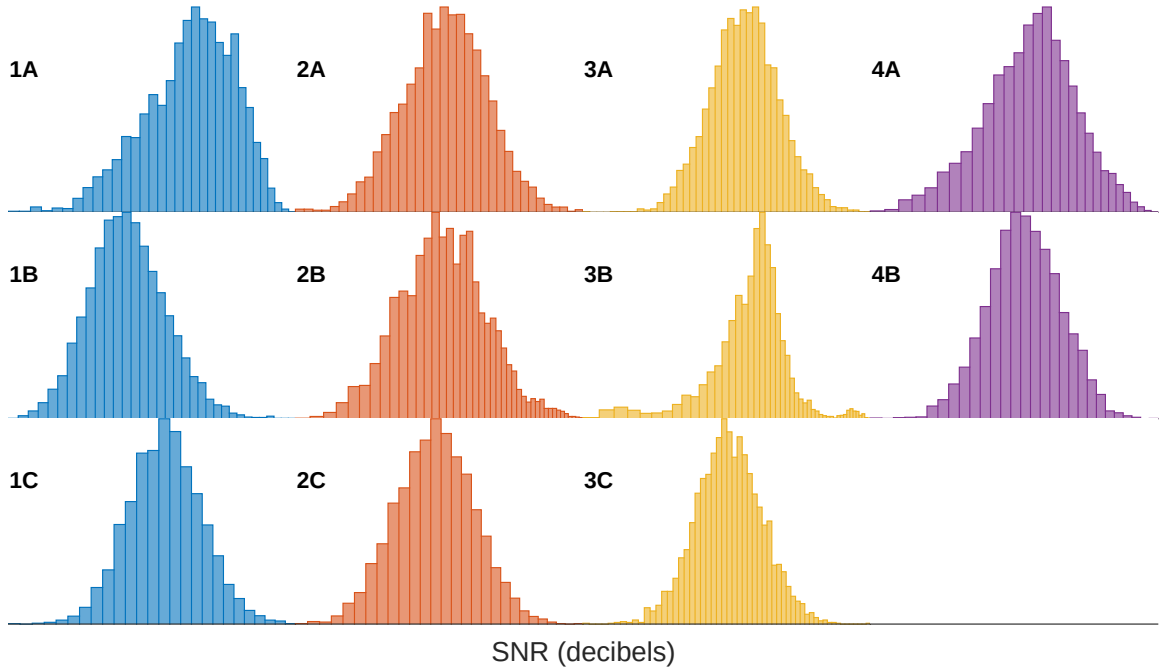


Figure 3.2: Histograms showing SNR across all resonances for each dataset. Plots 1A-4A, 1B-4B, 1C-3C show SNR for individual resonances of the dataset containing 1024 scans, 256 scans, 64 scans respectively.

described in fig. 3.3. As the SNR decreases in the lower panels, the method cannot identify the peak that has the smallest relaxation rate, figs. 3.1 and 3.3.

3.4.2 Validation with synthetic data

To demonstrate the validity of this method, I simulate NMR data using as ground truth the estimates of the posterior distribution obtained earlier, this way I obtain ideal test cases that match real-life experiments. Using synthetic data, I can easily compare my results with ground truth, and by fixing ground truth in the earlier estimates, I ensure similar characteristics to a real-life experiment having the characteristics of ^{19}F -NMR. The corresponding analyses of this data by the method are shown in fig. 3.4. As can be seen, an excellent agreement is obtained between the posterior estimates and the ground truth. This indicates that the method is robust and accurately recovers the underlying parameters for realistic ^{19}F -NMR data.

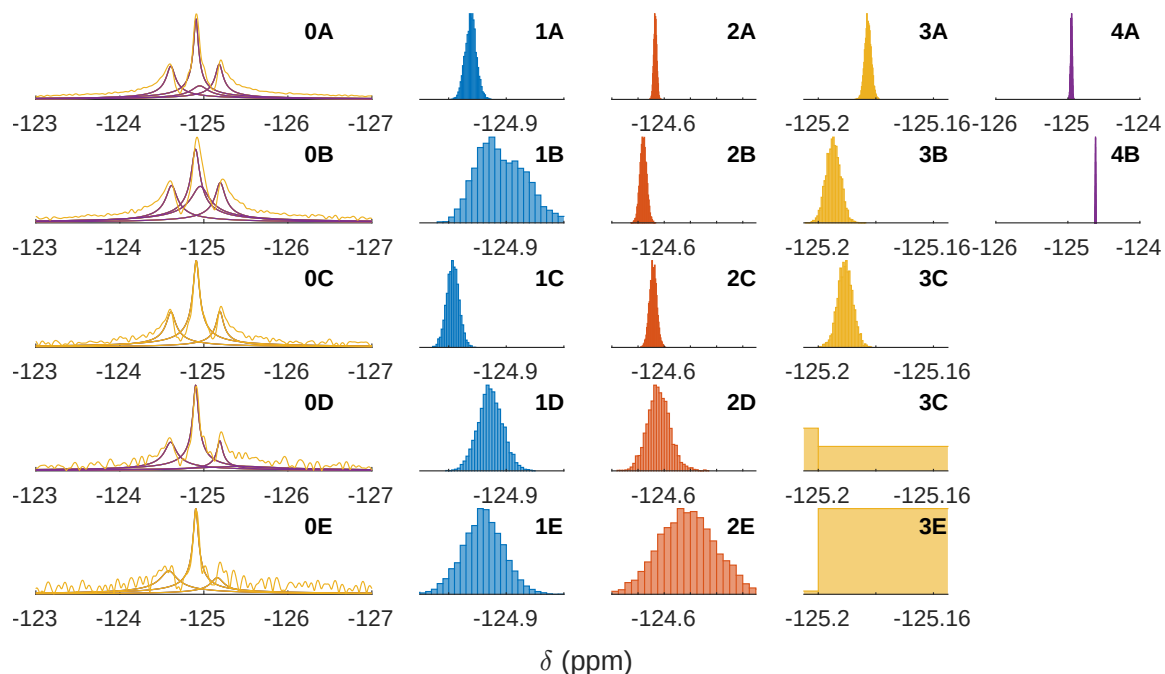


Figure 3.3: Effect of the number of scans. 0A-0E: Frequency domain spectrum showing the deconvolutions of resonances deduced from posterior estimates of the FID data from 1024, 256, 64, 16, and 4 scans of ^{19}F hTRPM8-VSLD. Plots 1A-4A, 1B-4B, 1C-4C, 1D-4D, and 1E-3E are histograms showing the distribution of chemical shifts for individual resonances of the FID data from 1024, 256, 64, 16, and 4 scans respectively

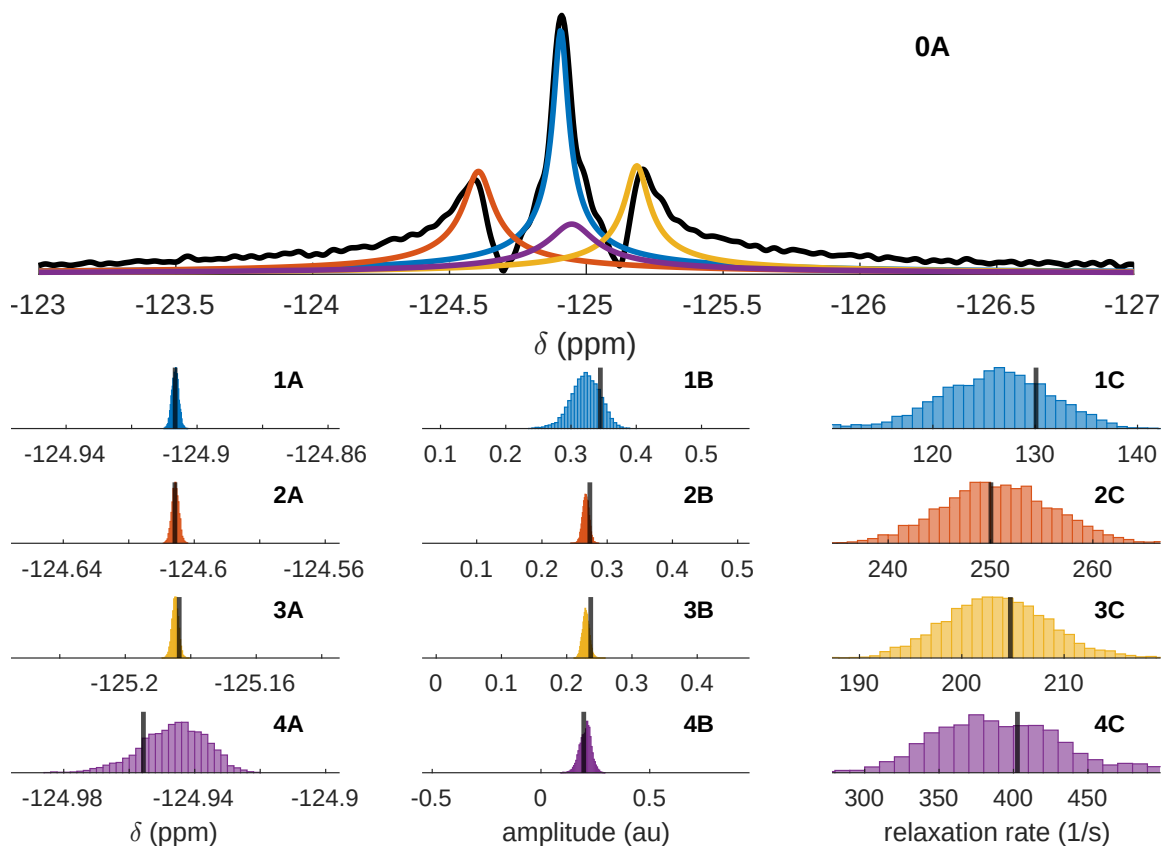


Figure 3.4: Validation with synthetic data. Histograms of relevant parameters resulting from the posterior distribution in the analysis of the FID data simulated with identical resonance parameters of the dataset containing 1024 scans of ^{19}F hTRPM8-VSLD. Plot 0A shows the deconvolution of all the resonances. Plots 1A-4A, 1B-4B and 1C-4C respectively correspond to the chemical shifts, amplitudes and relaxation rates of each individual resonance. The vertical lines are the ground truth values used in the simulation of the data.

3.4.3 Benchmarking with synthetic data

To further demonstrate the robustness of my proposed approach method, I simulate ^{19}F -NMR spectra in different ranges of (i) noise, (ii) peak separation, and (iii) spectrum complexity. I describe each case next.

Noise

In this study, I generated simulated data with a noise profile similar to that of the experimental data analyzed in Section section 3.4.1. Next, I generate additional datasets with the same parameters, but increased noise levels. For comparison purposes, the generated noise variances equal 1x, 2x, 4x, 8x, 16x, and 32x their baseline. These correspond to SNR that are $\sqrt{1}\text{x}$, $\sqrt{2}\text{x}$, $\sqrt{4}\text{x}$, $\sqrt{8}\text{x}$, $\sqrt{16}\text{x}$ and $\sqrt{32}\text{x}$ respectively less than the baseline.

The results of my analysis for all datasets are presented in Figure fig. 3.5. As can be seen, the method successfully identifies all resonance peaks in the cleanest dataset (1x) in excellent agreement with the ground truth. However, as S/N decreases (lower panels), the method fails to identify one of the peaks.

Peak separation

In this case, I analyze datasets of two identical resonance peaks with varying degrees of peak separation. Identical peaks are the hardest cases to identify, as they are the most difficult to pick out when they coincide. As shown in fig. 3.6, the method identifies the two distinct peaks, up to a small threshold distance. I define distance in this context as the ratio of the distance of separation between the peaks' maxima and the full width at half-maximum for each peak. The full width at half maximum for each peak is proportional to the relaxation rate, and without loss of generality, I

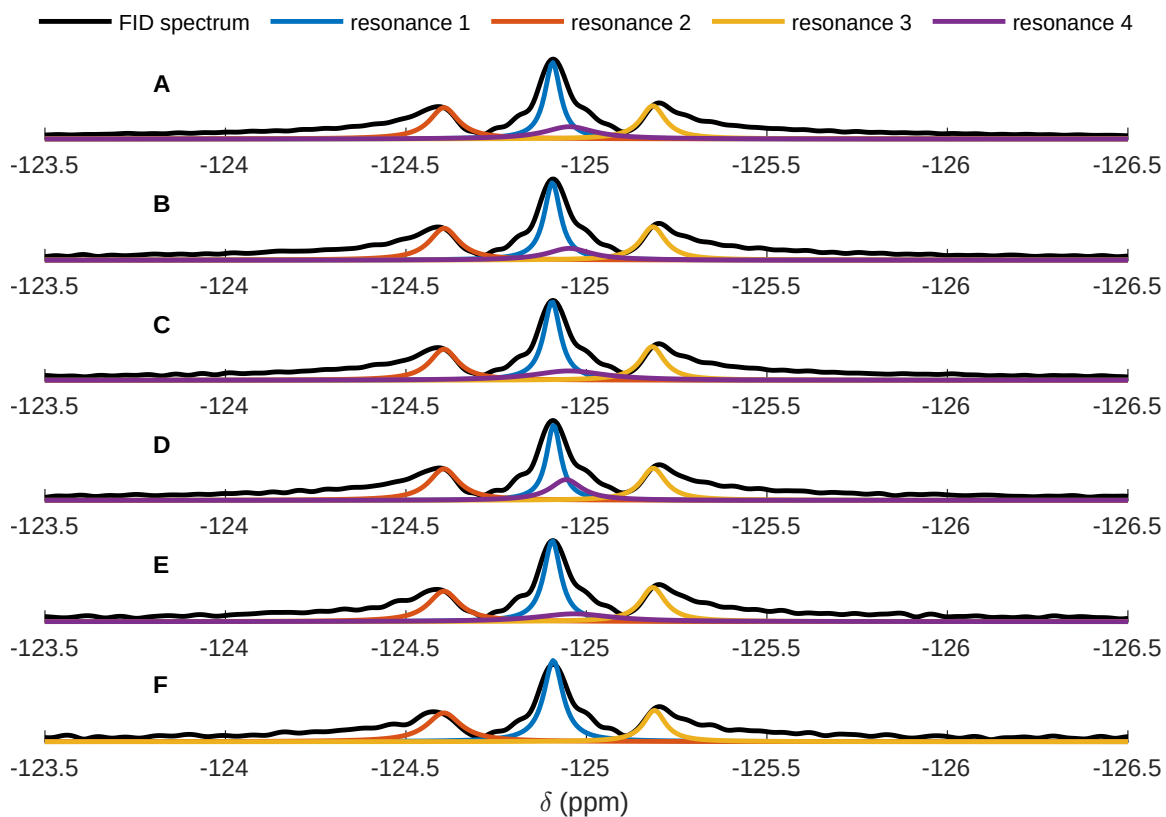


Figure 3.5: FID with synthetic data showing effect of decreasing SNR. A: Baseline SNR, B: SNR with 2x noise variance, C: SNR with 4x noise variance, D: SNR with 8x noise variance, E: SNR with 16x noise variance, F: SNR with 32x noise variance.

can let the constant of proportionality be 1. So that this distance is given as

$$d' = \frac{f_2}{r_2} - \frac{f_1}{r_1} \quad (3.10)$$

For the particular case examined, the relaxation rate is identical for both peaks. Hence

$$d' = \frac{\Delta\text{frequency}}{\text{relaxation rate}(r)} \quad (3.11)$$

As the distance of separation reaches a threshold value ($d' = 4.92$), the FID spectrum visually shows only one distinct resonance peak. Despite encountering a specific instance of failure, it is important to acknowledge that the method consistently exceeds human performance in most of the cases tested.

Decreasing amplitudes

In this experiment, I examined datasets that contain two overlapping peaks. In all cases analyzed, I reduced the amplitude of a specific peak by half while keeping all other parameters constant.

As illustrated in fig. 3.7, the proposed method successfully distinguished the two distinct peaks until the amplitude approached 0 up to a threshold of 0.0074 adu. At that point, the method could only identify a single peak.

Spectrum complexity

In this final benchmarking, I test the performance of the method against spectra of increasing complexity-designed specifically to mimic the challenges of complex NMR data. The first of such spectra is generated with 2 identical resonance peaks at a distance that ensures clear separation. Subsequent spectra are generated by introducing identical peaks at equidistant positions from the currently existing ones.

With this construction, the number of resonance peaks present in a spectrum is one less than twice the number of resonance peaks present in its previous spectrum.

Figure 3.8 below shows how the method performs as the complexity increases. As can be seen, the estimated number of resonance peaks is in excellent agreement with the ground truth.

3.4.4 Comparison with existing NMR analysis methods

Comparison to MestReNova’s peak picking algorithm

In order to illustrate the advantages of this fully automated method, I show how I repeat the analysis with the popular NMR Analysis software, MestReNova [Mestrelab Research S.L.](#). MestReNova’s auto peak-picking is based on quantitative Global Spectrum Deconvolution (qGSD) which uses a generalized Lorentzian function to model the experimental lineshapes. Like most NMR data analysis methods, qGSD analyzes frequency-domain NMR spectra and therefore requires a fair amount of preprocessing to transform the raw data into a method-friendly format.

In fig. 3.9, I show the steps required to analyze the cleanest dataset using MNova. All transformation steps were performed in MNova, and the data was displayed with MATLAB. The qGSD algorithm gives a total of three resonances, without a priori information. However, with structural information on the molecule of interest, I can inject an additional peak and require the method to help localize the four peaks.

Finally, I want to point out that the analysis of my method is completely automated and does not require filterings or manual adjustments.

Comparison to Deep Picker

To further illustrate the effectiveness of this approach, in fig. 3.10, I compare it with Deep Picker, a state-of-the-art deep learning-based peak-picking method recently developed by Li et al. [Li et al. \[2021\]](#). Deep Picker utilizes a neural network

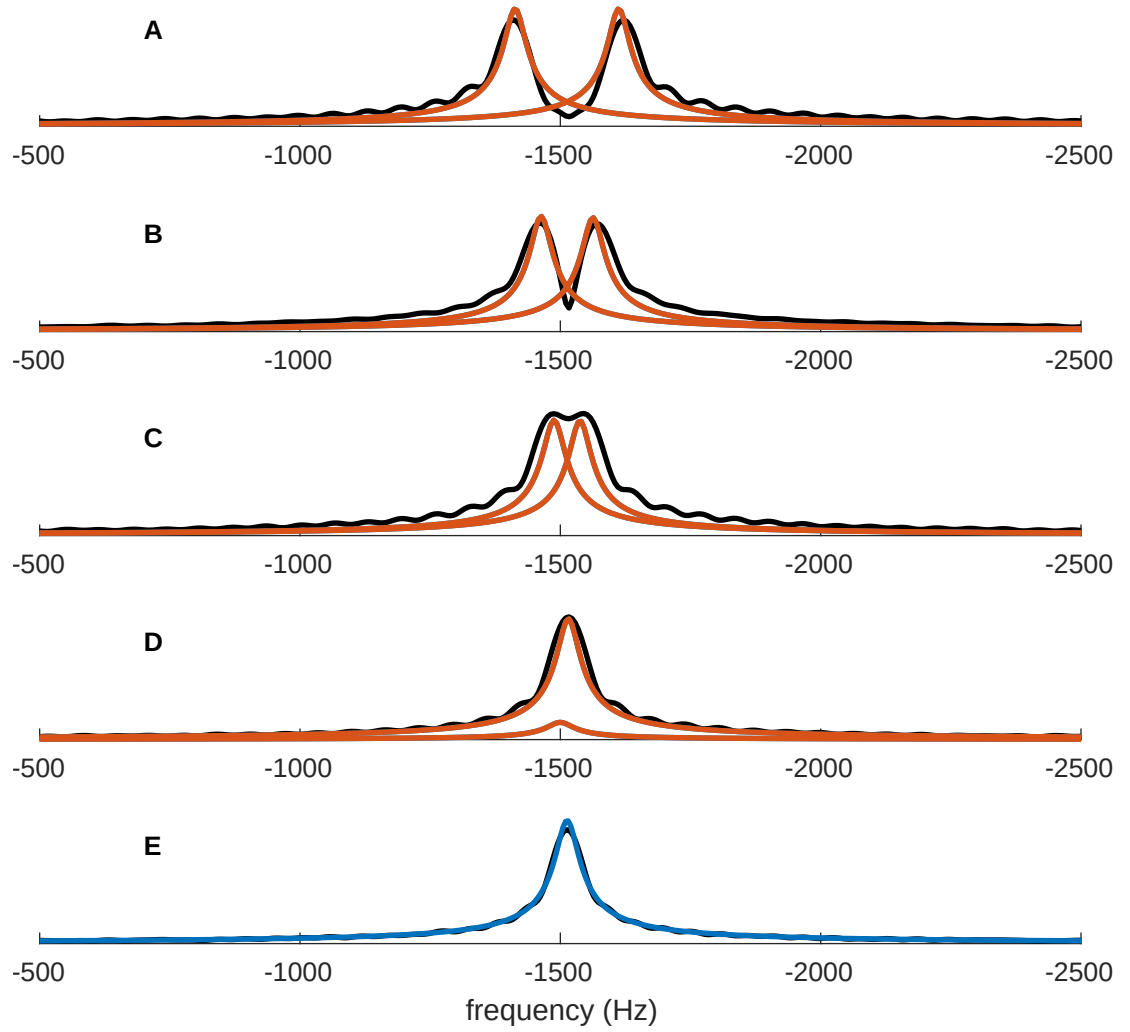


Figure 3.6: Frequency domain spectra showing decreasing distance of separation between two identical peaks d' . A: $d' = 78.72$, B: $d' = 39.36$, C: $d' = 19.68$, D: $d' = 9.84$, E: $d' = 4.92$ (Failure),

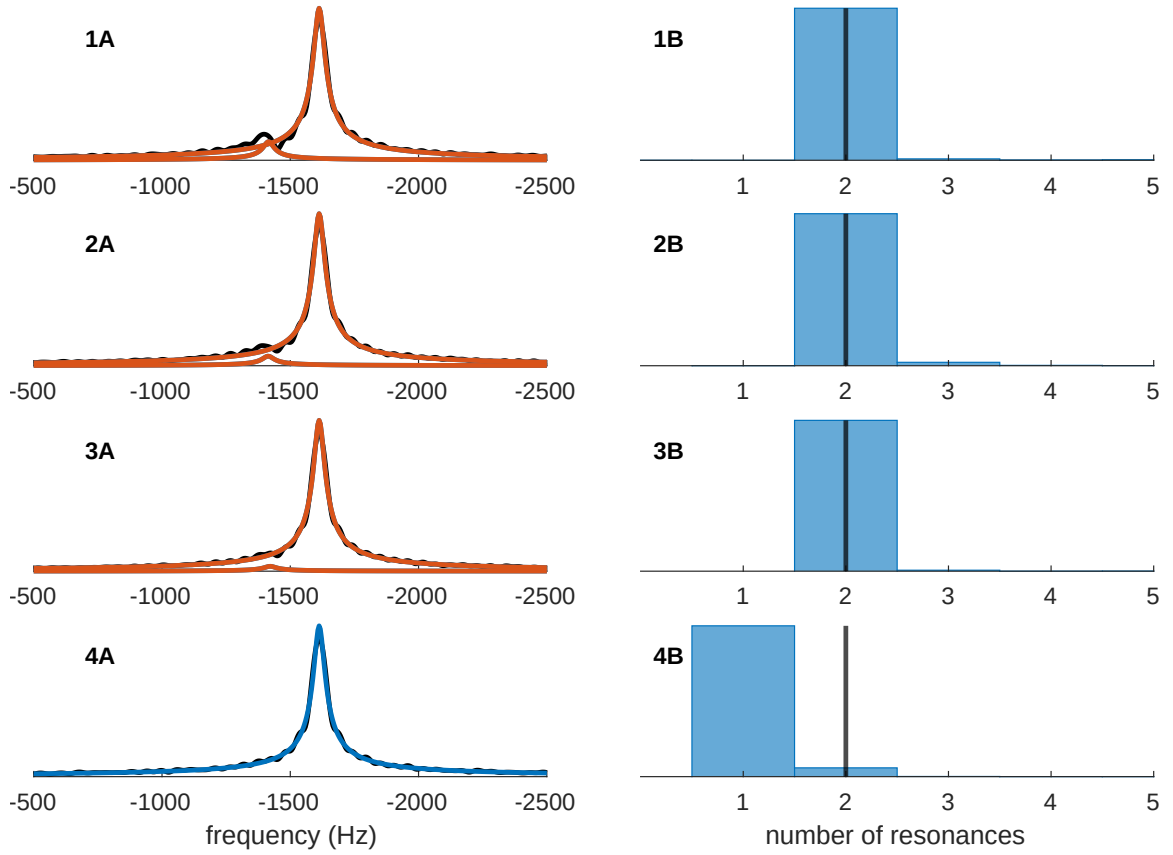


Figure 3.7: Frequency domain spectra of synthetic data showing decreasing amplitude of one shoulder peak. 1A-4A: The left shoulder peak is constructed with $1x, 0.5x, 0.25x$, and $0.125x$ the baseline amplitude. 1B-4B: Corresponding histograms showing the number of resonances identified by the method (ground truth = 2)

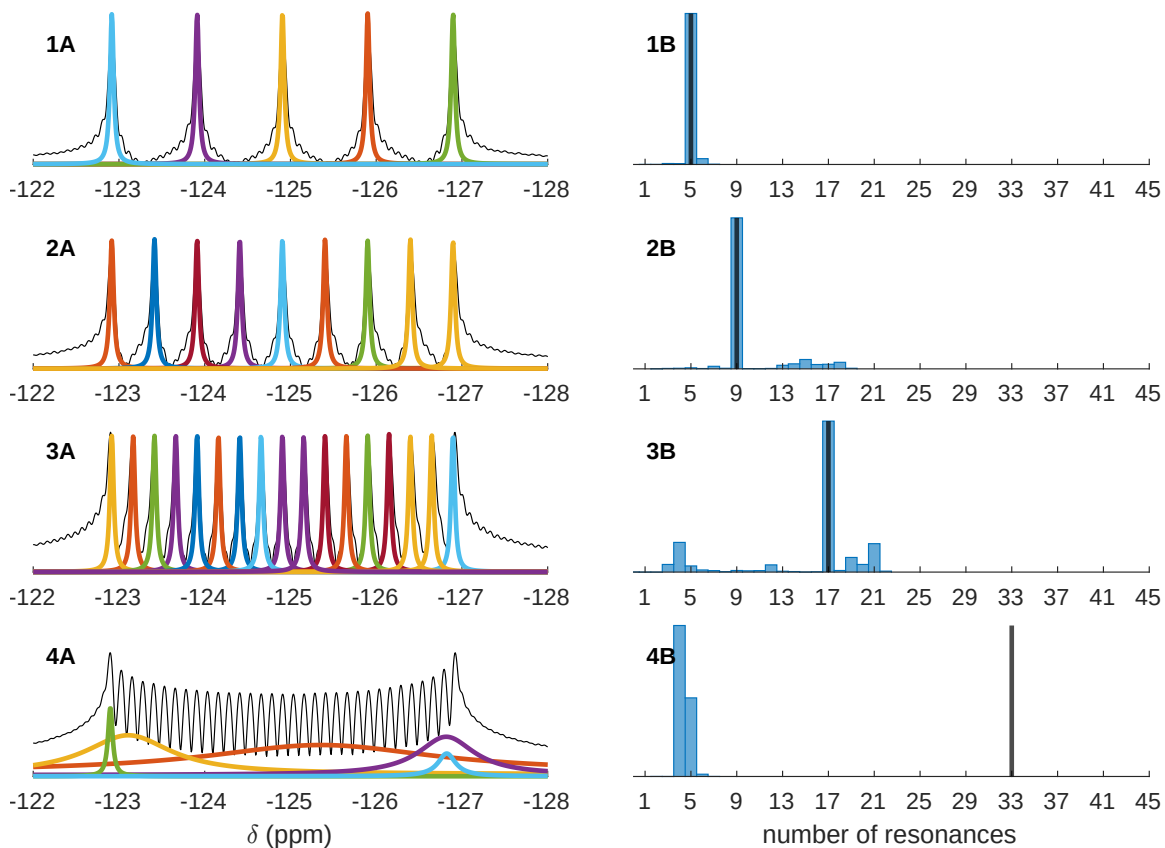


Figure 3.8: Frequency domain spectra of synthetic data showing increasing complexity. 1A-4A: Spectra constructed with 5,9,17, and 33 resonances respectively, 1B-4B: Corresponding histograms showing the number of resonances identified by the method. Spectrum in 4A is unusually complex and, as a result, exposes a failure mode in our method.

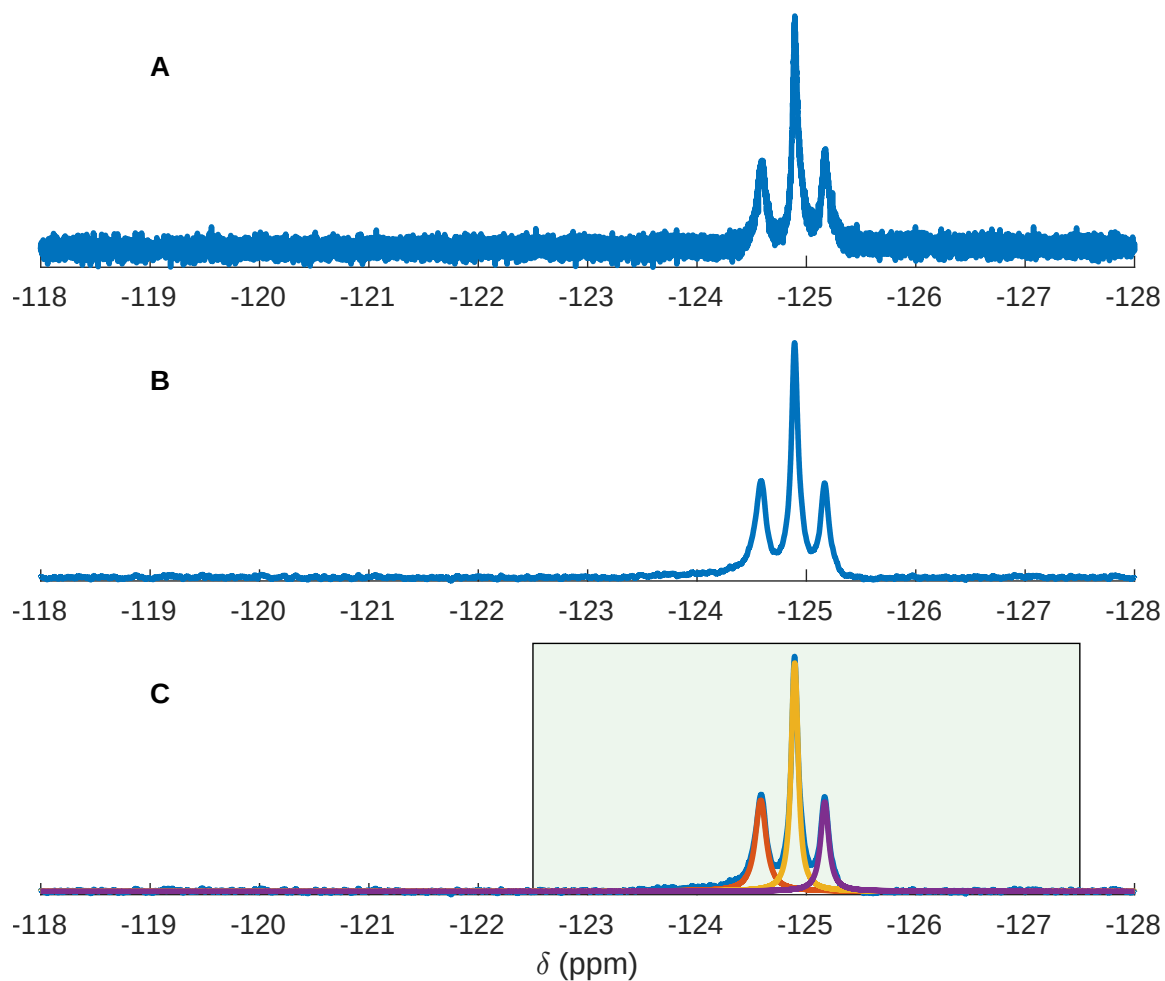


Figure 3.9: Data analysis of cleanest dataset with MestReNova. Panel A shows the frequency domain spectrum of raw data. Panel B shows the frequency domain spectrum of data after preprocessing steps (phasing and apodization). Panel C shows resonances detected using qGSD, MestReNova's peak picking algorithm

architecture trained on preprocessed NMR data to estimate peak locations and characteristics. To do this comparison, I carefully preprocess the NMR data in accordance with the specifications provided by the authors, and then input these preprocessed data into the Deep Picker system. Using the default settings, Deep Picker overestimated the number of peaks, suggesting an inherent tendency to overfit noise under these conditions. For this reason, I incrementally increased the noise threshold in three other separate trials. As I did this, the method began to stabilize, but the results remained suboptimal: the number of identified peaks never reached the correct count of 4 and instead plateaued at 3. Although these were three out of the four peaks, the estimated spectrum Deep Picker generated deviated noticeably from the original data, showing an inability of the deep picker to reverse the preprocessing done.

My proposed method, on the other hand, accurately identified all 4 peaks without any need for preprocessing or manual parameter tuning. The estimated spectrum produced by my approach showed excellent overlap with the original, demonstrating its ability to faithfully capture the underlying resonances while rejecting noise. This difference is particularly significant because it highlights the ability of the method to adaptively determine the appropriate model complexity directly from the data, a feature that inherently avoids the pitfalls of both overfitting and underfitting.

3.5 Discussion

In this study, I have presented a novel method to identify and characterize resonances in raw NMR data. This method is particularly relevant for ^{19}F -NMR, where the existing understanding and user experience to guide manual choices is limited. The method employs Bayesian nonparametrics to generate posterior probability distributions over all possible combinations of resonances and their characteristic

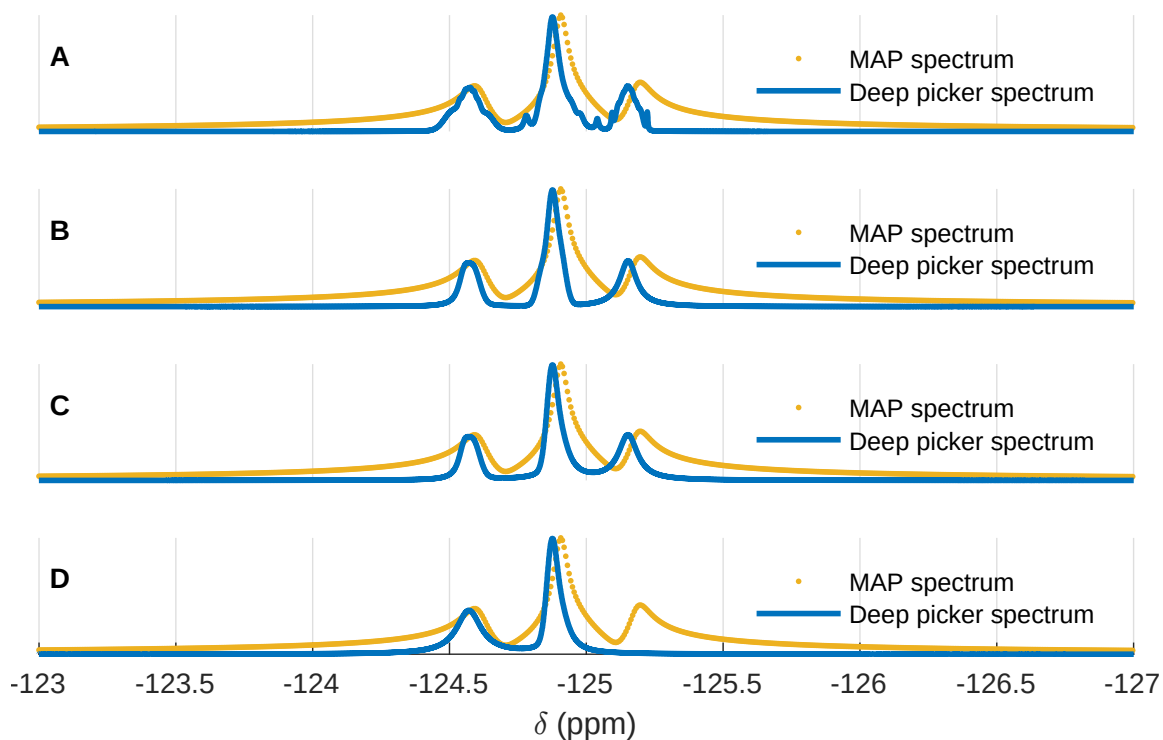


Figure 3.10: Data analysis of cleanest dataset with Deep Picker. Panel A compares the MAP spectrum generated by my method with the spectrum produced by Deep Picker at a noise-level cutoff of 5.5, where Deep Picker identified 27 resonances. Panel B compares the MAP spectrum generated by my method with the spectrum produced by Deep Picker at a noise-level cutoff of 10, where Deep Picker identified 9 resonances. Panel C compares the MAP spectrum generated by my method with the spectrum produced by Deep Picker at a noise level cutoff of 18, where Deep Picker identified 5 resonances. Panel D compares the MAP spectrum generated by my method with the spectrum produced by Deep Picker at a noise level cutoff of 19, where Deep Picker identified 4 resonances.

parameters. Such posterior distributions combine our initial assumptions about the resonance characterization with the likelihood of obtaining the observed FID data.

Unlike conventional Bayesian methods that rely on probability distributions over a fixed number of parameters [van de Schoot et al. \[2021\]](#), Bayesian nonparametrics makes the method more robust because it addresses the model selection problem [Ghahramani \[2015\]](#), [Gelman et al. \[2013\]](#). The model selection problem refers to the challenge of choosing the best model for our data among a set of candidate models [Ghahramani \[2015\]](#). Specifically in NMR data analysis, each candidate model is distinguished by the number of resonance peaks it can fit. Instead of arbitrarily deciding on the candidate model beforehand and then attempting to fit the parameters to the model, my method allows for the simultaneous choice of the model and fitting of its parameters. This approach eliminates the need for human intervention and takes a significant step towards full automation of NMR data analysis.

Another advantage of my method over other techniques is that it exclusively uses data in the time domain. This approach completely avoids errors that could be introduced during the Fourier transformation process, which is typically used to convert FID data from the time domain to the frequency domain. Working directly with the data in the time domain, the method offers improved precision and reliability in resonance identification and characterization regardless of windowing and filtering choices [Jacobsen \[2007\]](#), [Rubtsov and Griffin \[2007\]](#). In addition, the time domain formulation also allows the method to provide interpretable estimation of errors (i.e., posterior variance) determined exclusively from the FID supplied (i.e., length and noise) as opposed to ad hoc metrics like residuals and chi-square errors.

In recent years, novel deep learning approaches, such as Deep Picker [Li et al. \[2021\]](#), have also been applied to address similar problems. Although Deep Picker is a non-parametric method, its neural network architecture requires high-quality training data for effective model training. However, it is challenging to obtain training data that cover all potential use cases. In contrast, my approach does not require model

training, thus eliminating the need for training data altogether. Furthermore, as I showed in a previous section, deep picker still requires some parameter tuning in order to work well.

Beyond deep learning, my approach also offers distinct advantages over traditional signal processing techniques, which often rely on user-defined thresholds for peak identification. As shown in a previous section, commercial software packages, such as Mnova, typically utilize such methods. A key benefit of my model is that it entirely bypasses the need for a threshold. This removes a source of potential error, as choosing an appropriate threshold can be challenging, especially in spectra with low signal-to-noise ratios or overlapping resonances.

My method employs a promising Bayesian model formulation that holds the potential for extension to multidimensional NMR. The primary challenge is the computational cost; for the Flourine dataset discussed in this paper, with only 4 resonances, our method required about 5 hours to generate enough samples to converge to the posterior distribution. Generally, the analysis of a spectrum can range from a few hours to a few days depending on spectral complexity. Therefore, ongoing work focuses on achieving a computationally efficient representation of the model to enable practical applications, especially those involving specialized pulse sequences and non-Gaussian profiles. This pursuit opens up exciting possibilities for advancing the field of NMR research.

Chapter 4

A statistical learning framework for probing chemical exchange

4.1 Disclosure

This chapter is adapted from a manuscript currently being prepared for journal submission has been revised to fit coherently with the broader narrative flow.

4.2 Introduction

Biological function emerges from the dynamic ability of molecules to alter their conformation and engage with binding partners. From the conformational changes that activate a protein enzyme to the binding and release of a drug from its target, these dynamic processes occur on a wide range of timescales [Steigel and Spiess \[2012\]](#), [Bain \[2003\]](#), [Solomon \[1955\]](#). For several decades, Nuclear Magnetic Resonance (NMR) spectroscopy has been extremely powerful in characterizing these dynamics at atomic resolution. It achieves this by observing a phenomenon known as chemical exchange, where nuclei move between distinct chemical environments [Bain \[2003\]](#), [Kaplan \[2012\]](#), [Keeler \[2011\]](#). Accurate probing of this chemical exchange process provides insights into the kinetic and thermodynamic landscapes that govern molecular function [Bain \[2003\]](#), [Kaplan \[2012\]](#), [Keeler \[2011\]](#).

The information about chemical exchange is encoded in the shape of NMR resonance peaks, otherwise called the *lineshape* [Bain et al. \[2001\]](#), [Higinbotham and Marshall \[2001\]](#), [Keeler \[2011\]](#). In a system undergoing chemical exchange, the appearance of the spectrum is determined by the relationship between the relevant exchange rate constants and the frequency separation of the interconnected states [Bain et al. \[2001\]](#).

In the slow exchange regime, where these exchange rates constants are significantly smaller than the relevant frequency separation, the spectrum outlines distinct peaks for each molecular state [Bain et al. \[2001\]](#), [Keeler \[2011\]](#), [Kaplan \[2012\]](#). As the exchange rates increase and become comparable to the frequency separation, the system enters the intermediate exchange regime, where the resonance peaks

broaden significantly and begin to coalesce [Bain et al. \[2001\]](#), [Keeler \[2011\]](#), [Kaplan \[2012\]](#). Finally, in the fast exchange regime, the individual resonances merge into a single, sharp peak at a population weighted average frequency [Bain et al. \[2001\]](#), [Keeler \[2011\]](#), [Kaplan \[2012\]](#). To extract the underlying thermodynamic and kinetic parameters of the system, scientists primarily analyze the transformation of these lineshapes, often as a function of temperature [Bain \[2003\]](#), [Gultekin and Gore \[2005\]](#), [Carr and Purcell \[1954\]](#).

Within NMR studies on chemically exchanging systems, the task of spectral analysis is multifaceted. It typically begins with the identification of a full set of resonance peaks, each representing distinct states or conformations inherent to the system [Rule and Hitchens \[2006b\]](#). However, recognizing these peaks is only the preliminary step. Embedded within these spectral features are the unique kinetic and thermodynamic parameters that demand careful extraction [Bain et al. \[2001\]](#), [Steigel and Spiess \[2012\]](#), [Allerhand et al. \[1966\]](#).

Across the years, chemical exchanging systems have been studied primarily through the meticulous analysis of NMR lineshape functions. Foundational work by Gutowsky and Saika [Gutowsky and Saika \[1953\]](#) and McConnell [McConnell \[1958\]](#) laid the theoretical groundwork for understanding how chemical exchange rates affect NMR lineshapes. Their analyses showed that as the rate of exchange between distinct chemical environments changes, the corresponding 1D NMR signals undergo predictable transformations: from well-resolved, separate peaks at slow exchange rates, through broadened spectrum at an intermediate state, to a single, population-weighted average peak under fast exchange conditions [McConnell \[1958\]](#). By fitting these temperature-dependent lineshapes, the lineshape analysis of 1D NMR spectra became the gold standard for scientists seeking to extract relevant kinetic parameters. A significant advancement, especially in the domain of slower exchange, included the development of saturation transfer (ST) spectroscopy by Forsen and Hoffman [Forsén and Hoffman \[1963\]](#). This technique has been employed to measure rate

constants by observing the intensity changes when one exchanging site is irradiated. Further refinements include quantitative lineshape fitting routines which usually use Non-Linear Least Squares (NLLS) algorithms, such as the Levenberg-Marquadt method, to iteratively adjust model parameters and minimize the sum of squared differences between an experimental spectrum and a spectrum simulated using Bloch McConnell’s equations [McConnell \[1958\]](#).

Despite being powerful, these NLLS fitting methods are vulnerable to the complex, high-dimensional parameter space of an NMR experiment [Sima and Van Huffel \[2007\]](#). These methods often struggle because there are many possible solutions, and the likelihood of getting to the result that best represents the chemical process heavily depends on initial guesses provided by researchers [Osborne \[1975\]](#), [Dennis Jr and Welsch \[1978\]](#), [Johnson \[1994\]](#), [Bard \[1974\]](#). This sensitivity introduces hidden uncertainties, making it hard to reproduce or trust final results.

To overcome these fundamental challenges, I introduce a new paradigm for chemical exchange analysis: a Bayesian parametric model of the time-domain FID signal [Pressé and Sgouralis \[2023\]](#). Instead of seeking a single point of best fit, my Bayesian framework explores the entire valid parameter space to learn the full posterior probability distribution for every model parameter [Bishop and Nasrabadi \[2006\]](#). The resulting posterior distributions provide not just the most probable values for each parameter but also a rigorous and comprehensive quantification of uncertainties which eliminates the sensitivity to initial guesses that plagues NLLS methods. A key innovation of my model is its ability to perform a global analysis of multiple datasets simultaneously. Here, an entire temperature series of spectra can be fit at once to directly infer the global thermodynamic parameters that govern the exchange process, ensuring all data are leveraged in a cohesive and statistically robust manner.

I demonstrate the effectiveness and robustness of the method using carefully designed synthetic datasets that closely mimic realistic experimental conditions. I first validate my approach with a synthetic dataset simulating two peaks undergoing

chemical exchange across several temperature points, transitioning across exchange regimes. Afterwards, I evaluate the model’s performance with noisy synthetic data to demonstrate its robustness. Finally, I compare my model against simpler, commonly-used alternatives, showing superior accuracy and reliability.

4.3 Methods

4.3.1 Data generation in NMR experiments

In a typical 1D NMR experiment, the free induction decay (FID) signal is acquired during the detection phase, where the net transverse magnetization of nuclear spins precesses and relaxes over time [Rule and Hitchens \[2006a\]](#), [Kaplan \[2012\]](#). The signal is sampled at increasing time points $t_1 < t_2 < \dots < t_N$ within the time window $[t_{\min}, t_{\max}]$.

The acquired FID signal at each time t_n is a complex value $z_n = x_n + iy_n$, where x_n and y_n represent the real and imaginary components of the transverse magnetization, respectively [Rule and Hitchens \[2006a\]](#). These are influenced by physical factors such as precession frequency, transverse relaxation, and the presence of chemical exchange between spin pools.

The measured signal is modeled as:

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} P \\ Q \end{pmatrix} + \lambda \begin{pmatrix} M'_x(t_n) \\ M'_y(t_n) \end{pmatrix} + \zeta_n \quad (4.1)$$

where P, Q are the baseline offsets, λ is the gain, $M'_x(t), M'_y(t)$ are the true transverse magnetization components predicted by Bloch-type equations, and $\zeta_n \sim \mathcal{N}_2(0, S)$ represents additive Gaussian noise.

4.3.2 The Bloch-McConnell equations for K-Site exchange

The time evolution of the macroscopic magnetization vector

$$\vec{M}^k(t) = [M_x^k(t), M_y^k(t), M_z^k(t)]^\top$$

for the k -th spin pool (where $k = 1, \dots, K$) in a system undergoing chemical exchange is described by the Bloch-McConnell equations [McConnell \[1958\]](#):

$$\begin{aligned} \frac{d}{dt} \vec{M}^k(t) = & \gamma(1 - \sigma^k) \left(\vec{M}^k(t) \times \vec{B}^\#(t) \right) + \begin{bmatrix} -R_2^k & 0 & 0 \\ 0 & -R_2^k & 0 \\ 0 & 0 & -R_1^k \end{bmatrix} \left(\vec{M}^k(t) - M_0^k \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \\ & + \sum_{l=1}^K C^{lk} \left(\frac{\vec{M}^l(t)}{M_0^l} - \frac{\vec{M}^k(t)}{M_0^k} \right) \end{aligned}$$

Here:

- γ is the gyromagnetic ratio.
- σ^k is the chemical shielding constant for pool k .
- $\vec{B}^\#(t)$ is the effective magnetic field vector, including the static field B_0 and any radiofrequency (RF) pulses.
- R_1^k and R_2^k are the longitudinal and transverse relaxation rates for pool k , respectively. These are positive rate constants. Note the sign convention used here places the negative sign explicitly in the relaxation matrix, consistent with decay processes.
- M_0^k is the equilibrium magnetization of pool k , proportional to its population.
- C^{lk} represents the equilibrium flux between sites l and k (i.e., $\text{Rate}_{l \rightarrow k}^{\text{eq}} = \text{Rate}_{k \rightarrow l}^{\text{eq}} = C^{lk}$). I assume that $C^{lk} = C^{kl}$ due to the chemical exchange process.

These rates are linked to the underlying chemical kinetics and equilibrium populations.

During the FID acquisition period following an initial excitation pulse, there are no RF pulses applied, and the effective magnetic field is typically assumed constant along the z-axis: $\vec{B}^\#(t) = [0, 0, B_0]^\top$. I am primarily interested in the evolution of the transverse magnetization, which is conveniently represented by the complex quantity $M^k(t) = M_x^k(t) + iM_y^k(t)$. Under these conditions, the Bloch-McConnell equations for the transverse components simplify.

The precession term becomes $\gamma(1 - \sigma^k)(M_x^k B_0 \hat{j} - M_y^k B_0 \hat{i}) = -i\omega_k M^k(t)$, where $\omega_k = \gamma(1 - \sigma^k)B_0$ is the Larmor frequency of pool k . The transverse relaxation term becomes $-R_2^k M_x^k - iR_2^k M_y^k = -R_2^k M^k(t)$. The exchange term for $M^k(t)$ is $\sum_{l=1}^K C^{lk} \left(\frac{M^l(t)}{M_0^l} - \frac{M^k(t)}{M_0^k} \right)$.

Combining these terms yields a system of K coupled linear ordinary differential equations (ODEs) for the complex transverse magnetizations:

$$\frac{d}{dt} M^k(t) = (-R_2^k - i\omega_k) M^k(t) + \sum_{l=1}^K C^{lk} \left(\frac{M^l(t)}{M_0^l} - \frac{M^k(t)}{M_0^k} \right).$$

This system describes the evolution of the transverse magnetizations due to Larmor precession, transverse relaxation, and chemical exchange during the FID period [Bain et al. \[2001\]](#), [Rule and Hitchens \[2006a\]](#).

4.3.3 Forward Model: Simulating the FID signal

The system of ODEs governing the transverse magnetizations can be expressed in matrix form, facilitating its solution. Let $\vec{M}(t)$ be the column vector $[M^1(t), \dots, M^K(t)]^\top$. The system can be written as:

$$\frac{d}{dt} \vec{M}(t) = A \vec{M}(t)$$

where A is a $K \times K$ complex matrix whose elements capture the combined effects of relaxation, precession, and exchange. Based on the ODE derived above, the elements of A are:

$$A = \begin{bmatrix} -R_2^1 - i\omega_1 - \sum_{l \neq 1} \frac{C^{l1}}{M_0^1} & \frac{C^{12}}{M_0^2} & \frac{C^{13}}{M_0^3} & \cdots & \frac{C^{1K}}{M_0^K} \\ \frac{C^{21}}{M_0^1} & -R_2^2 - i\omega_2 - \sum_{l \neq 2} \frac{C^{l2}}{M_0^2} & \frac{C^{23}}{M_0^3} & \cdots & \frac{C^{2K}}{M_0^K} \\ \frac{C^{31}}{M_0^1} & \frac{C^{32}}{M_0^2} & -R_2^3 - i\omega_3 - \sum_{l \neq 3} \frac{C^{l3}}{M_0^3} & \cdots & \frac{C^{3K}}{M_0^K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{C^{K1}}{M_0^1} & \frac{C^{K2}}{M_0^2} & \frac{C^{K3}}{M_0^3} & \cdots & -R_2^K - i\omega_K - \sum_{l \neq K} \frac{C^{lK}}{M_0^K} \end{bmatrix}$$

Let k_{lk} be the first-order rate constant $l \rightarrow k$ and k_{kl} be the first-order rate constant $k \rightarrow l$. Assuming chemical equilibrium, detailed balance requires that $k_{lk}M_0^l = k_{kl}M_0^k$. If $C^{lk} = C^{kl}$ is the flux, then $k_{lk}M_0^l = k_{kl}M_0^k$ and $k_{lk} = C^{kl}/M_0^l, k_{kl} = C^{kl}/M_0^k$. Using these definitions, the ODE can be rewritten as:

$$\frac{d}{dt}M^k(t) = (-R_2^k - i\omega_k)M^k(t) + \sum_{l=1}^K (k_{lk}M^l(t) - k_{kl}M^k(t))$$

The matrix elements then become:

$$A = \begin{bmatrix} -R_2^1 - i\omega_1 - \sum_{l \neq 1} k_{l1} & k_{12} & k_{13} & \cdots & k_{1K} \\ k_{21} & -R_2^2 - i\omega_2 - \sum_{l \neq 2} k_{l2} & k_{23} & \cdots & k_{2K} \\ k_{31} & k_{32} & -R_2^3 - i\omega_3 - \sum_{l \neq 3} k_{l3} & \cdots & k_{3K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ k_{K1} & k_{K2} & k_{K3} & \cdots & -R_2^K - i\omega_K - \sum_{l \neq K} k_{lK} \end{bmatrix}$$

Let the system matrix be constructed as:

$$A = \Lambda_{\text{relax/precess}} + G \quad (4.2)$$

where $\Lambda_{\text{relax/precess}}$ is a diagonal matrix containing the relaxation and precession terms, and G is the generator matrix for the chemical exchange process.

$$\Lambda_{\text{relax/precess}} = \begin{bmatrix} -R_2^1 - i\omega_1 & 0 & 0 & \dots & 0 \\ 0 & -R_2^2 - i\omega_2 & 0 & \dots & 0 \\ 0 & 0 & -R_2^3 - i\omega_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -R_2^K - i\omega_K \end{bmatrix}$$

The exchange generator matrix G is constructed to satisfy detailed balance [Norris \[1998\]](#), [Yang et al. \[2006\]](#), [Joshi \[2013\]](#) with respect to the equilibrium populations $\boldsymbol{\pi} = [\pi_1, \dots, \pi_K]^\top$, where π_k is the equilibrium fraction of pool k ($\sum_k \pi_k = 1$). This is achieved by defining G based on a symmetric matrix S (where $S_{kl} = S_{lk} \geq 0$ for $k \neq l$, and $S_{kk} = 0$) which contains the fundamental exchange rate information. The $\frac{K(K-1)}{2}$ unique off-diagonal elements of S constitute the parameters defining the exchange rates. The elements of G are given by:

$$G = \begin{bmatrix} -\sum_{l \neq 1} \frac{S_{1l}}{\pi_1} & \frac{S_{12}}{\pi_1} & \frac{S_{13}}{\pi_1} & \dots & \frac{S_{1K}}{\pi_1} \\ \frac{S_{21}}{\pi_2} & -\sum_{l \neq 2} \frac{S_{2l}}{\pi_2} & \frac{S_{23}}{\pi_2} & \dots & \frac{S_{2K}}{\pi_2} \\ \frac{S_{31}}{\pi_3} & \frac{S_{32}}{\pi_3} & -\sum_{l \neq 3} \frac{S_{3l}}{\pi_3} & \dots & \frac{S_{3K}}{\pi_3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{S_{K1}}{\pi_K} & \frac{S_{K2}}{\pi_K} & \frac{S_{K3}}{\pi_K} & \dots & -\sum_{l \neq K} \frac{S_{Kl}}{\pi_K} \end{bmatrix}$$

This construction defines the entries of the generator matrix G to satisfy two critical properties. First, by ensuring that $\sum_k G_{kl} = 0$ for all l (i.e., the columns sum to zero), the model adheres to a fundamental requirement of continuous-time Markov chains that conserves total probability. Secondly, the construction is designed to satisfy detailed balance, a stricter condition that makes the process time-reversible [Yang et al. \[2006\]](#). The formal proof that this construction meets the detailed balance

condition is provided in Section B. Substituting these into Eq. (4.2), the elements of the full system matrix A become:

$$A = \begin{bmatrix} -R_2^1 - i\omega_1 - \sum_{l \neq 1} \frac{S_{1l}}{\pi_1} & \frac{S_{12}}{\pi_1} & \frac{S_{13}}{\pi_1} & \cdots & \frac{S_{1K}}{\pi_1} \\ \frac{S_{21}}{\pi_2} & -R_2^2 - i\omega_2 - \sum_{l \neq 2} \frac{S_{2l}}{\pi_2} & \frac{S_{23}}{\pi_2} & \cdots & \frac{S_{2K}}{\pi_2} \\ \frac{S_{31}}{\pi_3} & \frac{S_{32}}{\pi_3} & -R_2^3 - i\omega_3 - \sum_{l \neq 3} \frac{S_{3l}}{\pi_3} & \cdots & \frac{S_{3K}}{\pi_3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{S_{K1}}{\pi_K} & \frac{S_{K2}}{\pi_K} & \frac{S_{K3}}{\pi_K} & \cdots & -R_2^K - i\omega_K - \sum_{l \neq K} \frac{S_{Kl}}{\pi_K} \end{bmatrix}$$

The solution to the linear ODE system $\frac{d}{dt}\vec{M}(t) = A\vec{M}(t)$ with an initial condition $\vec{M}(0) = \vec{M}_0$ is given by the matrix exponential:

$$\vec{M}(t) = e^{At}\vec{M}_0$$

For computational efficiency, the matrix exponential is calculated using eigendecomposition of A Hallak and Dalal [2023]. Let $\mathbf{\Lambda} = [\lambda_1, \dots, \lambda_K]^\top$ be the vector of eigenvalues and E be the matrix whose columns are the corresponding right eigenvectors:

$$AE = E \text{diag}(\mathbf{\Lambda}t) \implies At = E \text{diag}(\mathbf{\Lambda}t)E^{-1}.$$

Generate FID

The free induction decay (FID) can be represented as $f(t) = [1, \dots, 1] \cdot \vec{M}(t)$ Using the eigendecomposition of A , I can write the FID as:

$$f(t) = \mathbf{1}^\top (E \text{diag}([e^{\lambda_1 t}, \dots, e^{\lambda_K t}])E^{-1}) \vec{M}_0$$

where $\vec{M}_0$ is a vector containing the equilibrium magnetization of all pools $k = 1, \dots, K$

The core likelihood model assumed that the observed FID data points $z_{1...N}$ are generated from the model f in the following manner;

$$z_n \sim \text{Normal}(f_n, \frac{1}{h})$$

where h is the precision. Assuming independence across N complex FID points, the full likelihood is represented as:

$$z|f, h \sim \left(\frac{h}{2\pi}\right)^N \exp\left(-\frac{h}{2} \sum_{n=1}^N |z_n - f_n|^2\right) = \left(\frac{h}{2\pi}\right)^N \exp\left(-\frac{h}{2} \|z - f\|^2\right).$$

4.3.4 Forward model: Incorporating temperature-dependent Chemical Exchange rates

I now extend the forward model to incorporate data collected at various temperatures. Each FID signal, acquired at temperature T , is characterized by temperature-dependent chemical exchange rates and viscosities [Bloembergen et al. \[1948\]](#). I then integrate the temperature dependence into the forward model via the *Eyring* equation and a viscosity-based scaling for the chemical exchange rates.

Specifically, for each FID at temperature T , the chemical exchange rate constants are recalculated using an effective rate scaling factor r , which accounts for viscosity variations:

$$r(T) = r_{ref} \left(\frac{T_{ref}}{T}\right) \exp\left(\frac{E_{a,visc}}{R} \left(\frac{1}{T} - \frac{1}{T_{ref}}\right)\right) \quad (4.3)$$

Here:

- r_{ref} is an arbitrarily set reference rate constant at temperature T_{ref} ,
- $E_{a,visc}$ is the activation energy associated with viscosity changes,

- R is the universal gas constant.

The exchange rate matrix S used to construct the generator matrix G is defined element-wise using the Eyring equation. This helps incorporate thermodynamic parameters specific to each chemical exchange pathway:

$$S_{kl}(T) = \frac{k_B T}{h} \exp\left(\frac{\Delta S_{kl}^\ddagger}{R}\right) \exp\left(-\frac{\Delta H_{kl}^\ddagger}{RT}\right) \quad (4.4)$$

where:

- k_B is Boltzmann's constant,
- h is Planck's constant,
- $\Delta S_{kl}^\ddagger$ and $\Delta H_{kl}^\ddagger$ represent the entropy and enthalpy of activation for the exchange between pools k and l , respectively.

Therefore, for each experimental FID at temperature T , the full system matrix $A(T)$ becomes:

$$A(T) = \Lambda_{\text{relax/precess}} + r(T)G(T) \quad (4.5)$$

The evolution of the magnetization vector at temperature T is then solved via the eigendecomposition approach described previously:

$$\vec{M}(t, T) = e^{A(T)t} \vec{M}_0 \quad (4.6)$$

The simulated FID at each temperature T is given by:

$$f(t, T) = \mathbf{1}^\top (E(T) \text{diag}([e^{\lambda_1(T)t}, \dots, e^{\lambda_K(T)t}]) E^{-1}(T)) \vec{M}_0 \quad (4.7)$$

The likelihood model assumes that the observed FID data points $z_{1,\dots,N}$ at temperature T are generated from the model $f(t, T)$ as follows:

$$z_n \sim \text{Normal}(f_n, \frac{1}{h})$$

where h is the precision parameter. Assuming independence across N complex FID points, the full likelihood at temperature T is represented as:

$$z|f, h \sim \left(\frac{h}{2\pi}\right)^N \exp\left(-\frac{h}{2} \sum_{i=1}^N |z_i - f_i|^2\right) = \left(\frac{h}{2\pi}\right)^N \exp\left(-\frac{h}{2} \|z - f\|^2\right).$$

The overall likelihood function across multiple temperatures is the product of individual likelihoods:

$$\mathcal{L}(\text{data}|\text{parameters}, h) = \prod_{T \in \text{temperatures}} \left(\frac{h}{2\pi}\right)^N \exp\left(-\frac{h}{2} \|z(T) - f(T)\|^2\right)$$

Bayesian Model

Figure ?? below provides a graphical model showing the interdependence of parameters in the model.

To further develop this Bayesian model, I place priors on the unknown parameters:

$$\omega_k \sim \text{Beta}(E, F, \omega_{\min}, \omega_{\max})$$

$$r_{ref} \sim \text{Gamma}(\alpha_r, \beta_r)$$

$$q_k \sim \text{Gamma}(\alpha_q, \beta_q)$$

$$p \sim \text{Beta}(A, B, p_{\min}, p_{\max})$$

$$h \sim \text{Gamma}(\alpha, \beta)$$

$$dS_{kl} \sim \text{Gamma}(\alpha_{dS}, \beta_{dS})$$

$$dH_{kl} \sim \text{Gamma}(\alpha_{dH}, \beta_{dH})$$

where the new parameters are given by:

$$\begin{aligned}\omega_k &= \gamma(1 - \sigma^k) \\ r_k &= R_2^k \\ q_k &= \pi_k \\ dS_{kl} &= \Delta S^\ddagger \\ dH_{kl} &= \Delta H^\ddagger\end{aligned}$$

p is the global phase offset.

I use MCMC sampling to draw from the posterior distribution over model parameters. The implementation is described in [appendix B](#)

4.4 Results

4.4.1 Validation with synthetic data

First, I simulate NMR data to serve as a proof of concept for my analysis pipeline. The data modeled a two-site system with populations fixed in a 60:40 ratio, with data generated at five temperatures that spanned the slow and intermediate regimes.

The corresponding analyses of this data by the method are shown in [fig. 4.1](#) below.

[fig. 4.1](#) shows an excellent agreement between the posterior estimates obtained from the method and the ground truth. In particular, the posteriors remain extremely narrow and centered on the truths, showing that my Bayesian model computes each parameter with high precision even in regions with spectral overlap.

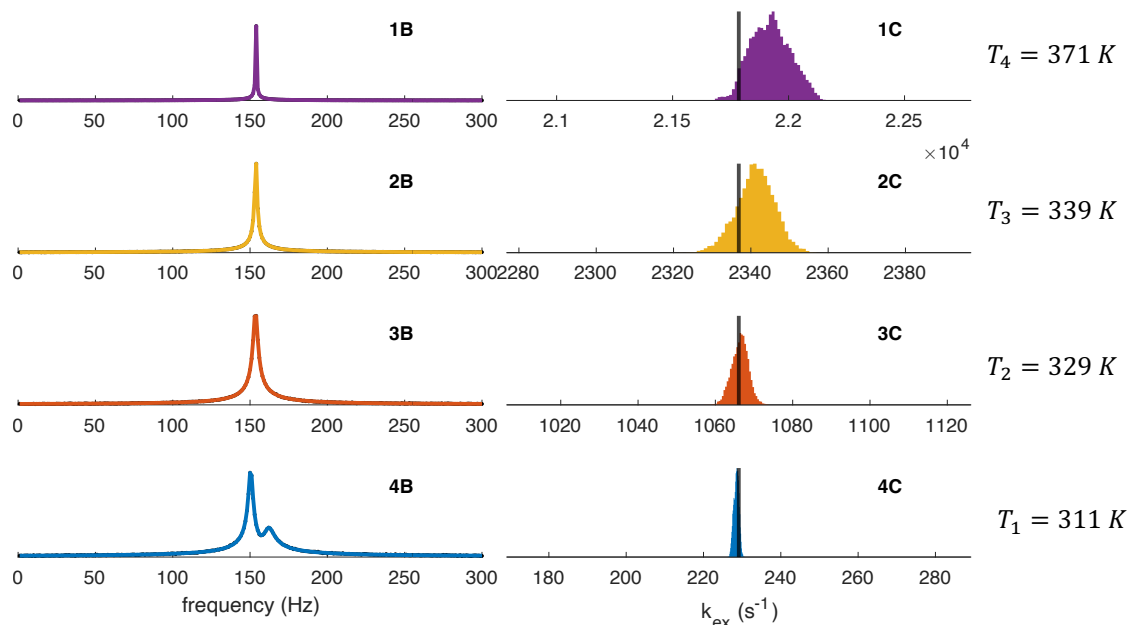


Figure 4.1: Synthetic data against the model signal generated from the Maximum A Posteriori (MAP) parameters at five temperatures, increasing from 311 K (bottom) to 371 K (top). Panels 1A-4A show the frequency-domain FID, and panels 1B-4B depict the corresponding exchange-rate posterior distributions (histograms) with the true value marked by a vertical black line. As temperature increases, the two distinct peaks at low temperature broaden and merge, illustrating accelerated chemical exchange. This is reflected in the broadening of the posterior histograms, which indicates increased uncertainty.

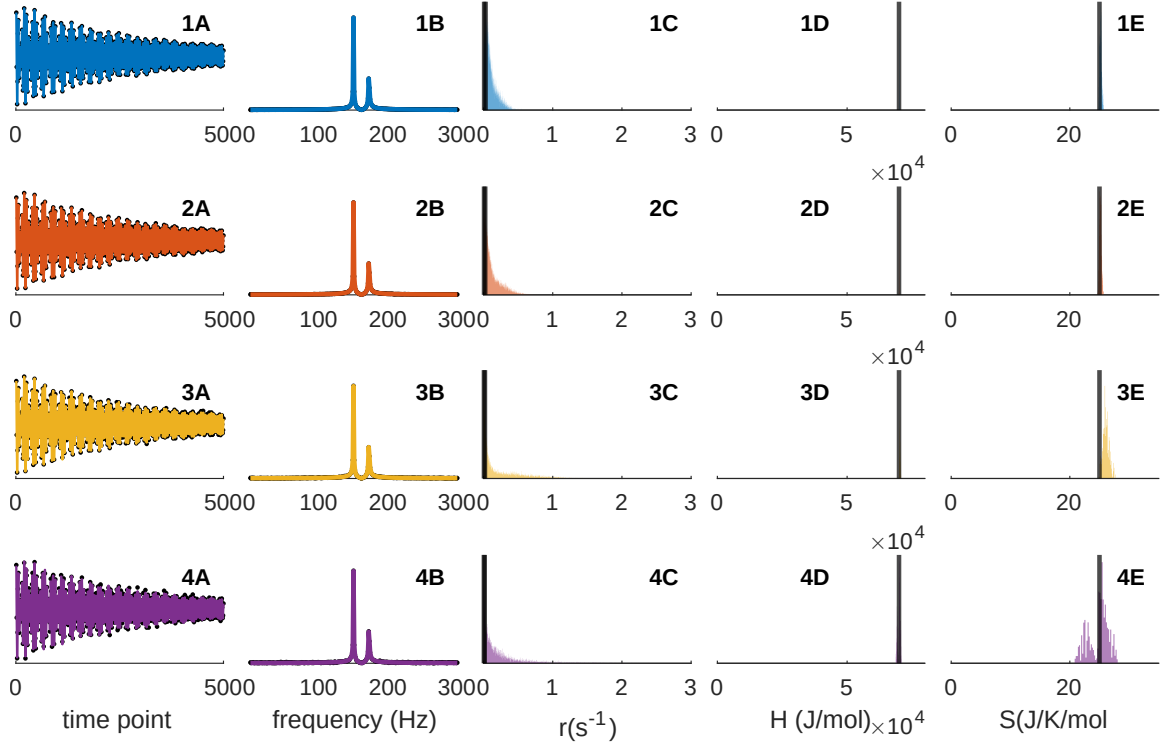


Figure 4.2: Panels 1A-5A and 1B-5B represent synthetic data and MAP reconstructions for the lowest-temperature spectrum of each noise level. These are followed by normalised posterior histograms for the relaxation rate (r) (panels 1C-5C), enthalpic ($\Delta H^\ddagger$) and entropic ($\Delta S^\ddagger$) activation parameters (panels 1D-5D and 1E-5E respectively). Rows are ordered by decreasing SNR from top to bottom.

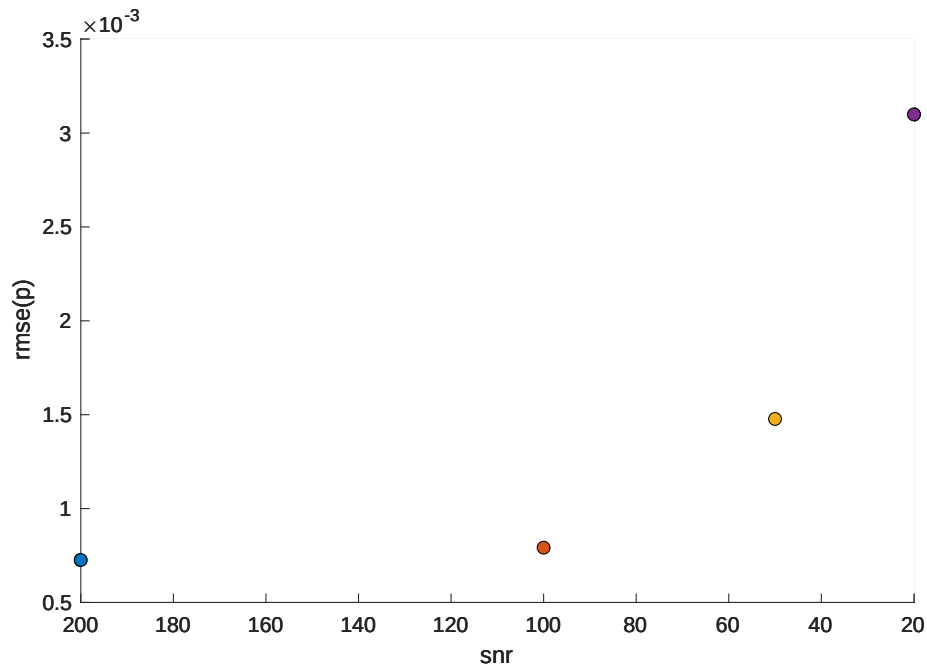


Figure 4.3: Root-mean-square error (RMSE) of the phase parameter (p) as a function of the signal-to-noise ratio (SNR). Each coloured marker corresponds to one synthetic dataset and uses the same row colours employed in fig. 4.2. The SNR axis is reversed (high SNR on the left) to maintain consistency with the top-to-bottom ordering in the spectra panel

4.4.2 Benchmarking with synthetic data

In certain physical regimes, NMR exchange experiments may generate data from which it is inherently difficult or impossible to disentangle specific parameters Binsch [1968]. The following experiments aim to assess how my model behaves in such known challenging regimes. More precisely, I test whether the model correctly identifies lack of information by producing appropriately uncertain posterior distributions (e.g. wide or correlated distributions) rather than converging to potentially incorrect but precise estimates. I simulate spectra under different: (i) noise profiles, (ii) temperature-sampling strategies (uniform grid, sparse grid, and a cluster centered on the coalescence region).

Noise

In this study, I generated simulated data with a noise profile similar to that of the data analyzed in Section section 4.4.1. Next, I generate additional datasets with the same parameters, but increased noise levels. These new datasets correspond to SNR that are 1x, 2x, 4x, respectively less than the baseline. The SNR was defined as :

$$\text{SNR} = \frac{\max_t |f_{\text{clean}}(t)|}{\sigma_{\text{noise}}},$$

where $f_{\text{clean}}(t)$ is the noise-free FID at $T = 290$ K. I then drew posterior samples of the phase parameter $\{p^{(n)}\}_{n=1}^N$ and the root mean-square-error (RMSE) was evaluated as

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N (p^{(n)} - p_{\text{true}})^2}.$$

I present the results of the analysis in fig. 4.2 and fig. 4.3.

In fig. 4.2, I observe that as the the SNR diminishes, the posterior distributions systematically broaden, illustrating how the method accurately depicts uncertainty while remaining centered on the true parameters. Similarly, the resulting curve in fig. 4.3 shows a smooth, monotonic increase in RMSE as the SNR is lowered. This trend mirrors the widening posteriors in the histogram panel and demonstrates that the robustness of my method to inflate uncertainty in step with declining data quality while retaining quantitative accuracy over a broad range of noise conditions.

4.4.3 Discussion

The results from the synthetic data benchmarks demonstrate that the proposed Bayesian framework is not only capable of accurately inferring physical parameters from complex data, but also provides a robust quantification of uncertainty, especially in low-information NMR regimes.

The robustness of my framework in the presence of noise highlights a key advantage than traditional point-estimation methods [Reeves and Shaw \[1970\]](#), [Abergel and Palmer \[2004\]](#). As demonstrated in [fig. 4.2](#) and [fig. 4.3](#), the model’s response to decreasing data quality is ideal: the posterior distributions for the estimated parameters widen around the ground truth. This behavior is crucial for experimental NMR studies, as it signifies that the model does not become misleadingly overconfident when the data is poor. The smooth, monotonic increase in the RMSE further substantiates this point—that the precision of the estimate degrades in a predictable manner commensurate with the signal-to-noise Ratio. This ensures that any single parameter estimate derived from the posterior (such as the mean or median) remains a reliable indicator of the true value, while the posterior width provides an accurate and trustworthy measure of the confidence one should have in that estimate. Ultimately, these findings portray my method not just as a data-fitting tool, but as a framework for evaluating the quality and information content of the experiment itself as it provides not only parameter estimates but a clear picture of the certainty of these estimates, which is intrinsically linked to the quality of the experiment.

PART 2

Deep Learning for predictive modeling in precision agriculture

Chapter 5

Development of precision feed AI systems

5.1 Disclosure

This chapter is adapted from a manuscript currently being prepared for journal submission has been revised to fit coherently with the broader narrative flow.

5.2 Introduction

Nutritionists have evolved over time from balancing dietary crude protein (CP) with metabolized protein (MP), to providing individual metabolized amino acids (AA). More recently, research has aimed to achieve greater efficiency gains through more precise matching of MP and metabolizable AA supplies with animal requirements [Moraes et al. \[2018\]](#). To this end, the NASEM [National Academies of Sciences, Engineering, and Medicine \[2021\]](#) is a deterministic model that estimates the supply and requirement of AA of individual animals. Today, the emerging field of precision feeding can use new technologies to refine and adjust metabolizable AA supply to meet individual animal needs [Gaillard et al. \[2020\]](#). Precision dairy cattle feeding depends on identification of the nutrient requirements for each cow in the herd and utilizes technological advancements to adjust the animal’s diet and, subsequently, the nutrient supply to match its requirements feed supply [Pomar et al. \[2021\]](#).

The development of precision dairy cow feeding methods (PDCF) requires automatic collection and high throughput processing of data, and system monitoring in real time to estimate and deliver a diet in the correct amount and composition adjusted to match individual animal nutrient requirements. New feeding technologies (e.g. sensors, feeders, automatic milking systems (AMS) and small wearable devices) and data handling solutions (cloud storage, real-time communication, and automated processing) allow group-housed animals to be fed according to their individual requirements, improving herd feeding efficiency. Currently in the dairy industry, however, predictive models of nutrient requirements and excretion (e.g. NASEM

2021 Council et al. [2001]) are critically required to select the best feeding strategies for the herd. For the dairy industry, PDCF is a promising solution to optimize metabolizable AA use efficiency and reduce waste by adjusting the supply of crude protein (CP) to individual cows at different points in the productive cycle of each cow Prestegard-Wilson et al. [2021], Kaufman et al. [2017], Cabrera et al. [2020]. Precision feeding of dairy cows has the potential to minimize metabolizable AA overfeeding and underfeeding by predicting the individual cow requirements for AA in real time.

The use of AI solutions and machine learning (ML) can be implemented to analyze animal and environmental data in real time to accurately predict AA requirements for each animal in the milking herd. Automatic milking systems could be used to acquire data during each milking event. These data could be processed to do analytics over milk production responses to feeding levels of metabolizable AA for each cow in the herd. Machine learning could perform AMS animal and non-animal data analytics to refine nutrient supply, improve feeding management, resulting in increased biological and economic efficiency of dairy farms. In AMS, cows are fed a combination of a partial mixed ration (PMR) in the feed bunk and with feed in the milking robot. The feed delivered in the robot varies based on animal factors e.g. productivity, however, individual cow intake of the feed delivered by the robot is not measured. This limitation must be addressed to balance individual cow inputs and outputs of nutrients.

The goal of this work is to determine the amount and variability of feed offered by the robot, feed leftover in the robot by individual cows at the end of each milking event, and feed intake at each milking. Furthermore, I aim to characterize the individual cow variability and to characterize the size of the variability and its dynamic across time and animal traits (e.g. milk production, parity, and days in lactation).

5.3 Methods

5.3.1 Animals and materials

All procedures were approved by the Institutional Animal Care and Use Committee of the University of Tennessee (protocol 2973-0423). The animal trial was carried out at the East Tennessee AgResearch and Education Center - Little River Animal and Environmental Unit (ETREC – LRAEU; Walland, TN).

The farm facilities included a Lely A5 Astronaut automatic milking system (AMS) equipped with automated grain feeders fed by 4 individual grain bins.

A total of 54 lactating Holstein cows were housed in a free-stall barn with ad libitum access to water and free access to the AMS. The cows were grouped into a single pen equipped with 50 free stalls. Cows were identified in the AMS using radio frequency identification ear tags, and chewing activity was monitored using rumination collars (Lely, Harmelen, The Netherlands). The cows produced 45.4 ± 11.7 kg milk/d; weighed 729 ± 82 kg BW; and were 166 ± 64 DIM at the beginning of the period.

The animals were fed a common PMR that was offered 2x/day (8 AM and 3 PM) in adequate amounts to achieve a target of 2% refusals. The PMR contained corn silage, wheat straw, and a grain mixture. Water was added to the PMR to maintain the 40% dry matter content (DM).

The feed offered on the AMS was tailored for each cow according to individual predictions of DMI, animal size, animal production, and experimental objectives using the NASEM (2021) model.

A time of flight 3-D camera (ifm O3D303) was installed inside the robot feed bin to collect images before and after each milking event to estimate feed delivered, leftover, and consumed by cows at each milking event (see next section). The camera

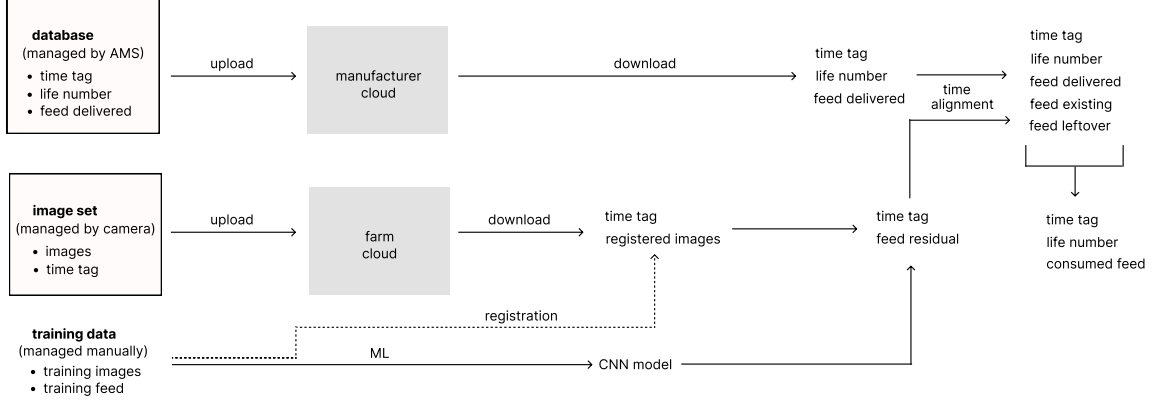


Figure 5.1: Data pipeline summarizing procedures and the successive stages of data transformations applied in this work.

was connected to a computer (Lenovo XX) to store temporarily images of the feed bin acquired continuously and the computer had internet connection to online data storage (Microsoft OneDrive) for permanent storage.

5.3.2 Data organization and analysis

Our data processing pipeline is summarized in fig. 5.1 and is detailed below.

Raw data acquisition

During the course of this study, individual animals entered the AMS voluntarily in a sequential manner, producing visit events. Information on visits (such as time and duration of the event, animal triggering the event, amounts of feed provided, and milk produced) was provided by two separate data sources. The first is the database managed internally by the AMS. The second is an image set acquired by the purpose-installed 3D camera monitoring the AMS’s bin.

The AMS’s database is automatically managed by the AMS and provides individualized physiological and behavioral information on the animal triggering each event. This information is organized in visit records that indicates the time of the visit event (time tag and duration), the animal that triggers the event (life number), the amount

of milk collected (milk yield) for the events that resulted in milking and the amount of feed delivered (delivered feed) for the events that resulted in feeding. Visit records are available for download from the AMS manufacturer at 24 hr intervals.

The 3D camera is managed by custom scripts that run on a local computer at the farm site separately from the AMS. The camera automatically acquires images at intervals of 0.5 sec. Custom scripts are used to pre-screen the acquired images in real-time. Since the AMS employs a moving bin, pre-screening allows for selection and retention of only the images that image the entire bin as opposed to occasionally parts of it. The pre-screened images are time-tagged and uploaded in real time to on-line cloud storage. The stored images are available for download on demand.

Feed estimation

To obtain the actual feed consumed during a milking event, I combined data from the database (feed delivered) with data from 3D images (feed residual). The latter are obtained by processing the stored images with a custom convolutional neural network (CNN) that is developed and employed as described below. The CNN uses the supplied images to provide an estimate of the feed remaining in the AMS' bin at the time of the image capture. This time coincides with the time between successive visit events. For this reason, the residual feed estimates are used twice: once as feed existing in the bin before the following milking event and once as feed leftover after the previous milking event.

The residual feeds are associated with the visit records captured by the AMS based on time following temporal alignment. Temporal alignment accounts for matching the acquisition time of each image (recorded by the camera) with the time span of a visit event (recorded by the AMS). This way, temporal alignment ensures that the images, which are obtained and managed independently of the AMS, are associated with the correct AMS data and so residual feeds are correctly designated as existing and leftover with respect to the visit events.

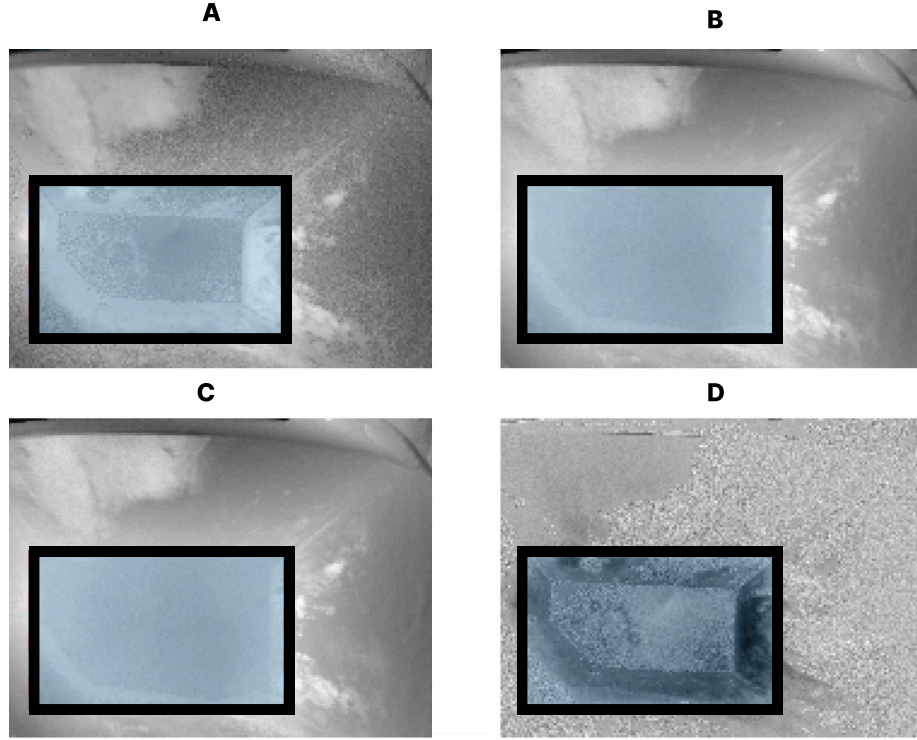


Figure 5.2: Example images of the feed bin: (A) empty bin (0kg), (B) full bin (3.9kg), (C) image shown in panel B after registration with respect to the image of panel A, and (D) difference image between full (panel C) and empty bins (panel A), enhancing contrast for processing through the CNN. In all panels, the highlighted region indicates the crop region. Although the area of the crop region represents only a small portion of the full field of view, due to the operation of 3D cameras, the actual imaged volume nearly covers the volume of the entire bin.

During this study, visit events lasted for several minutes each, resulting in multiple images stored per event. Since each visit event is associated with multiple stored images and so multiple residual feed amounts, the existing and leftover feed estimates are obtained by computing the median prediction across all images associated with an event.

After temporal alignment, the estimated feed consumed at a given visit event resulted from

$$\text{feed consumed} = \text{feed delivered} + \text{feed existing} - \text{feed leftover} \quad (5.1)$$

where the existing feed indicates feed remaining in the AMS’s bin at the end of the previous visit event and feed leftover indicates feed remaining in the AMS’s bin at the end of the current event. That is, our calculations correct the feed delivered by the AMS according to the feed adjustment

$$\text{feed adjustment} = \text{feed existing} - \text{feed leftover} \quad (5.2)$$

which is the net change in the residual feed within the bin across the visit.

5.3.3 Computational methods

Model deployment

I used a convolutional neural network (CNN) to estimate the amount of feed remaining in the AMS’s bin between successive visit events. The network receives 3D image data, captured between events when the AMS’s bin is retracted, and outputs the amount of feed that remains in the bin during imaging.

The repeated opening and closing of the feeding bin during feeding events caused small shifts in the camera’s position over time. Although these shifts did not drastically change the field of view, they introduced slight misalignments that prevented a pixel-wise comparison between acquired images and training images. To correct for this

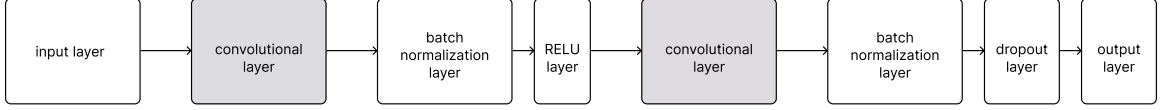


Figure 5.3: Configuration of CNN that converts 3D images to feed amount estimates.

misalignment, I cropped the images from an original size of 137x176 to 66x106 to emphasize the most relevant viewing area (see fig. 5.2) and registered [Muthukumaran and Sivakumar \[2017\]](#) all acquired images with respect to a baseline reference image from the training data set, cropped identically. The viewing area approximates the smallest rectangle that bounds the largest leftover feed in the training data (about 3.9 kg). This ensured that the trained model reliably interpreted changes in the feed weight without being affected by positional variations in the camera.

A small percentage of images with artifacts ($\approx 0.1\%$), which escaped prescreening, yielded negative feed values. These images were excluded from further processing.

Model development

The relationship between images and feed weight is too complicated to model from first principles. For this reason, I leverage the power of modern AI tools and apply a purpose-built data-driven model; an image model, which takes the form of a CNN [Goodfellow et al. \[2016\]](#) as described below, receives images and computes the feed remaining in the bin between successive visit events.

At input, the network receives registered 3D images of the feed bin, which have been cropped to a size of 66 x 106 (as mentioned above) with the baseline image of the empty bin subtracted (see fig. 5.2, panel D) to enhance contrast.

The CNN architecture presented here was selected after an exploratory phase where we tested multiple designs for optimal performance. The final model, illustrated in fig. 5.3, is designed to process 3D image data stored in grayscale format (i.e. 1 scalar value per pixel). The architecture consists of standard convolutional layers, ReLU

activation functions, normalization layers, and a fully connected regression output 2,3. The first convolutional layer applies a set of 2x2 kernels, which are small matrices of learnable weights. This operation allows the model to capture local spatial patterns in the feed distribution, and a bias term is added to the output of each convolution. Subsequently, a batch normalization layer is introduced to stabilize activations and accelerate convergence. The ReLU activation function introduces nonlinearity, which enables the network to learn complex feature representations such as bin and feed height. A second convolutional layer, identical in kernel size to the first, is introduced to learn more abstract spatial dependencies such as bin wall and feed texture. This is followed by an additional batch normalization layer and a ReLU activation layer. To prevent overfitting, I include a dropout layer to deactivate a subset of neurons during training. The network then passes the extracted features through a fully connected layer, which maps the learned representations to a single scalar output corresponding to the estimated feed.

Training data is obtained by manually filling the bin with a pre-determined amount of feed. Specifically, I used a ladder of 52 feed steps that ranged from an empty bin to a maximum amount of 3.9 kg. For the first step of the ladder, i.e. empty bin, I acquired 200 images. For the subsequent 51 steps, I acquired 50 images per step. This procedure resulted in 2750 image/feed pairs that I used to train the CNN. For model training, I use the root-mean-square propagation learning algorithm which improves upon standard gradient descent by using an adaptive learning rate schedule. Unlike plain gradient descent which applies a single, fixed learning rate to all parameters, RMSprop adjusts the rate for each parameter individually based on a moving average of recent gradient magnitudes. This allows for more stable and efficient adjustments as the training progresses. The trained CNN demonstrated high predictive accuracy, achieving an R^2 value of 0.9881 on an unseen dataset. This result signifies a near-perfect fit between the model's predictions and the true data.

5.4 Results

5.4.1 System Validation

As detailed in section 5.3, on a visit basis, feed consumed is determined by the sum of feed delivered and a feed adjustment. The former is reported by the AMS and the latter is derived by the difference between existing and leftover feeds which, in turn, stem from the residual feeds that are calculated based on the 3D images. From these, “feed delivered” denotes the amount automatically determined by the AMS based on pre-set criteria, “feed existing” term represents the residual feed from the preceding visit event and “feed leftover” denotes the residual feed remaining in the bin at the end of the current event.

A characteristic timeline in fig. 5.4 below shows a sample overlay of images taken by the 3D camera alongside visit events recorded in the AMS system. Here, I see image bursts associated with each visit event. This confirms that the system is robust enough to pair the visit events recorded in the AMS (fig. 5.1) with the images taken from the 3D camera.

Analysis of all visit events provides an overview of how individual cows interact with the AMS and the image acquisition system provides an accurate estimation of their feeding characteristics. A summary is shown in table 5.1. There were 5160 total visit events during the study period, indicating that each animal visited the AMS an average of 3.2 times per day. This number is in agreement with general behavioral traits.

Our results also showed that the average visit duration was 6.95 ± 1.91 min while the time elapsed between successive visit events averaged 1.63 ± 7.23 min. The histogram in fig. 5.5 corroborates this fact, as the duration for visit events peaked around 6.5 minutes and was skewed right towards higher values with a few outliers occurring after 15 mins. The camera captured an average of 27.54 ± 19.32 images per event.

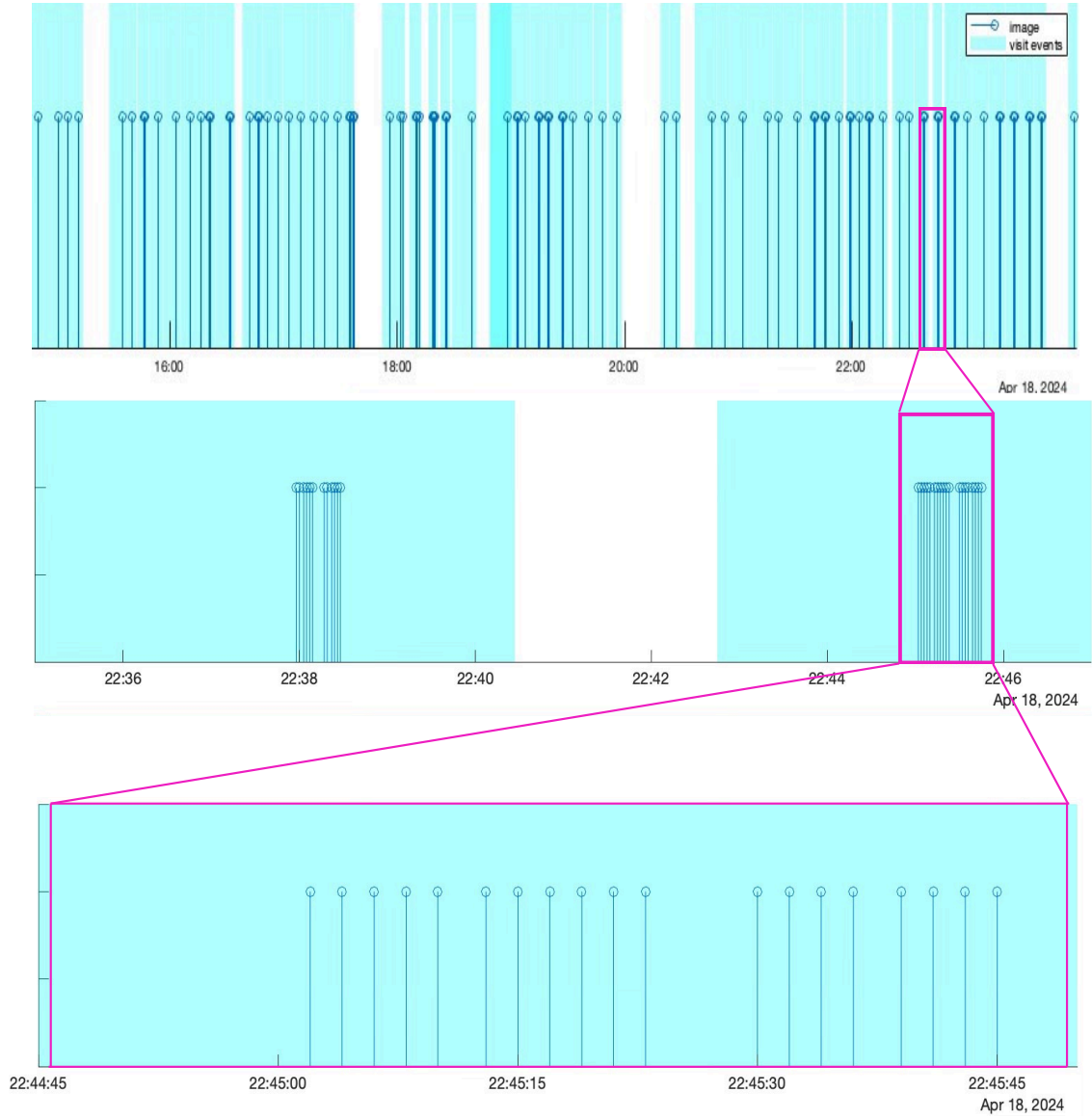


Figure 5.4: Portion of the timeline of this study with marked visit events and acquired images. As can be seen, visit events last for ≈ 7 min each and are separated by intervals of ≈ 1.5 min. Each event is aligned with ≈ 22 images acquired successively at intervals of 0.5 sec. The time and duration of the visit events depend on animal behavior and show random patterns. Occasional gaps in the train of acquired images are due to images failing prescreening or image registration which are excluded from our analysis.

A plot showing how feed consumed correlated with feed delivered over the visit events of this study is shown in fig. 5.6 (panel A). It is important to note that the plot shows the amount of feed consumed, not the amount of feed leftover. The maximum value for feed consumed in our dataset is 3.9kg. In this plot, points correspond to individual visits. Placement along the horizontal axis (delivered feed) is determined by AMS's data and placement along the vertical axis (consumed feed) is determined by a combination of AMS's and camera's data. Points falling near the diagonal indicate visits that delivered and consumed feed are of the same amount. Similarly, points above the diagonal indicate visits that consumed feed exceeds delivered feed (possibly due to excessive existing feed) and points below the diagonal indicate visits that consumed feed lacks delivered feed (possibly due to excessive leftover feed). Overall, points above and below the diagonal are roughly balanced, indicating that events with feed overconsumption are balanced out by events with feed underconsumption.

Approximate balancing of the feed consumption across successive events is exemplified by the highlighted points 1a,1b and 2a,2b in fig. 5.6. Specifically, a large quantity of leftover feed following one event (e.g., panel A, point 1a) corresponds to a higher consumption in the subsequent event (e.g., panel A, point 1b), as expected since the animal in the second visit can consume both delivered feed in own visit and a portion of the leftover of the previous visit. Conversely, minimal leftover feed from one event (e.g., panel A, point 2a) leads to a lower calculated consumption in the next (e.g., panel A, point 2b). Panel B provides a timeline positioning these consecutive event pairs (1a,1b and 2a,2b) within the sequence of visit events; while, panel C, includes images captured after each visit event, showing the residual feed remaining in the bin.

5.4.2 Herd feeding pattern

Histograms illustrating the distribution of per-event metrics are presented in fig. 5.7. The distribution of feed consumed per event was almost symmetric, peaking at ≈ 1800 g, indicating that overfeeding events are roughly balanced by underfeeding events over the entire herd. In contrast, feed delivered per event was skewed right, showing a peak at ≈ 1500 g and an outlier at ≈ 5000 g. This is consistent with the operation of the AMS which senses the animal triggering each event and responds by adjusting the feed to be delivered based on a predetermined daily target.

More quantitative results are shown in table 5.2. On average, animals in the herd received 1832.11 ± 577.66 g of feed per visit. From this, they consumed an average of 1831.12 ± 658.84 g, resulting in an average of 818.28 ± 485.29 g of residual feed remaining in the bin at the end of each visit. Our analysis equally showed an average feed adjustment of 0.15 ± 409.45 g, signaling once more that over the entire herd feed overconsumption is balanced out by feed underconsumption. As explained above, this is not surprising since residual feed is accounted twice: once as leftover feed from one visit that contributes to underconsumption, and also as existing feed for the next visit that contributes to overconsumption.

5.4.3 Individual animal feeding patterns

While overall herd averages provide a general understanding of the animals' interaction with the AMS, substantial variation in feeding behavior exists among individual animals as seen in table C.1 and table C.2 (see appendix C).

Considering per-visit statistics (table C.1 supplement), average feed consumed per visit across the herd ranged from approximately 1300.4 g to 3920.2 g. However, I discovered considerable variation in the consistency of consumption for individual animals, as indicated by standard deviations ranging from 329.9 g to 1077.3 g. This showed that while some animals had a seemingly consistent feeding pattern, some

Table 5.1: Comprehensive summary of visit events. The reported values indicate mean $\pm$ one standard deviation.

Metric	Value
First event in this study	04-17-2024 00:09:52
Last event in this study	05-17-2024 23:56:46
Total visit events	5160
Duration for visit event (min)	6.95 ± 1.91
Time between successive visit events (min)	1.63 ± 7.23
Number of images per visit event	27.54 ± 19.32

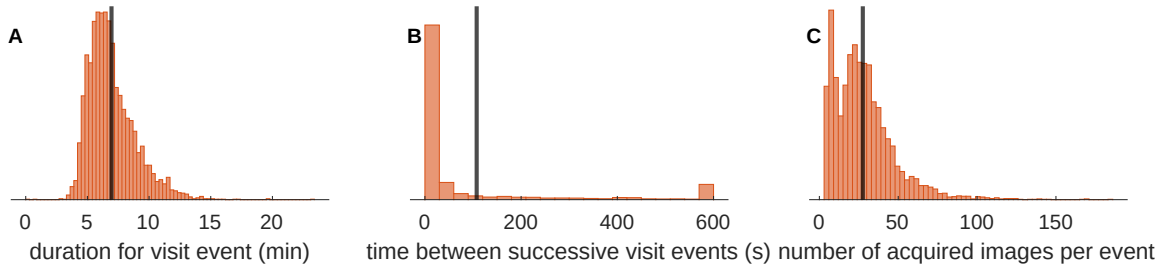


Figure 5.5: Histograms of the metrics in table 5.1, the vertical line identifies the mean values. For clarity, in Panel B, times between successive visit events above 600s are binned at 600s.

Table 5.2: Summary of consumption related events during the course of this study. The reported values indicate mean $\pm$ one standard deviation.

Metric	Value
Feed consumed per event (g)	1831.12 ± 658.84
Feed delivered per event (g)	1832.11 ± 577.66
Leftover feed per event (g)	818.28 ± 485.29
Adjustment in feed per event (g)	0.15 ± 409.45

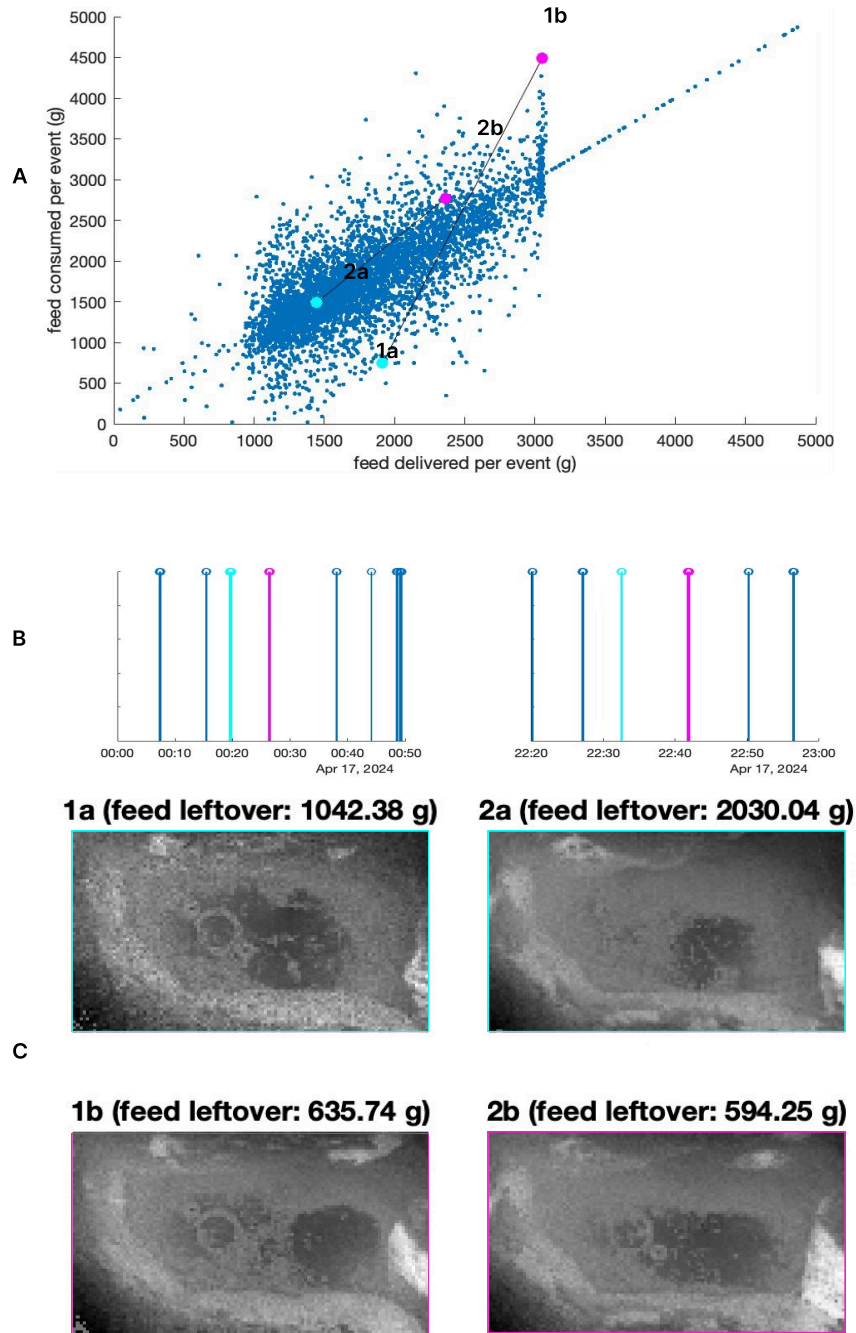


Figure 5.6: Feed consumption estimates. Panel A displays feed consumed (NOT feed leftover) vs feed delivered over all events of this study and highlights 4 specific instances, representing two pairs of successive visits (1a/1b and 2a/2b). Panel B shows the timeline of these visit events while the bin images in panel C show the leftover feed at the end of each visit event.

others might have been more inconsistent. I observed similar variations for delivered feed (means ranging from 1288.6 g to 3431.3 g and standard deviations from 177.1 g to 1035.8 g) and residual feed (means ranging from 517.0 g to 1202.5 g and standard deviations from 278.0 g to 656.4 g).

Aggregating AMS and image data to daily feed totals provides insight into daily intake patterns (table C.2 supplement). The mean daily feed consumed per animal ranged from 3007.2 g to 8744.4 g, while mean daily feed delivered ranged from 3511.9 g to 8744.4 g. The variability in daily totals for individual cows is reflected in the standard deviations of these daily sums (consumed standard deviations: 819.3 g to 2031.4 g; delivered standard deviations: 595.8 g to 1666.3 g). Mean daily leftover feed ranged from 1242.4 g to 3888.8 g (standard deviations: 533.3 g to 2265.4 g), and mean daily adjusted feed ranged from -1245.7 g to 916.2 g (standard deviations: 434.2 g to 1176.1 g).

Figure 5.7 shows variations in daily feed consumption and delivery averages for individual animals across the observed period. It also demonstrates how, collectively for all cows, the amount of feed consumed per visit closely matches the amount delivered by the AMS. While the AMS balances feed delivery with feed consumption at the herd level, as illustrated by the right panel, it fails to achieve the same balance at the individual animal level, as illustrated in the left panel. Specifically, on average, 44.4% of the animals of this study consumed more feed than delivered, while 55.6% of the animals received more feed than they consumed.

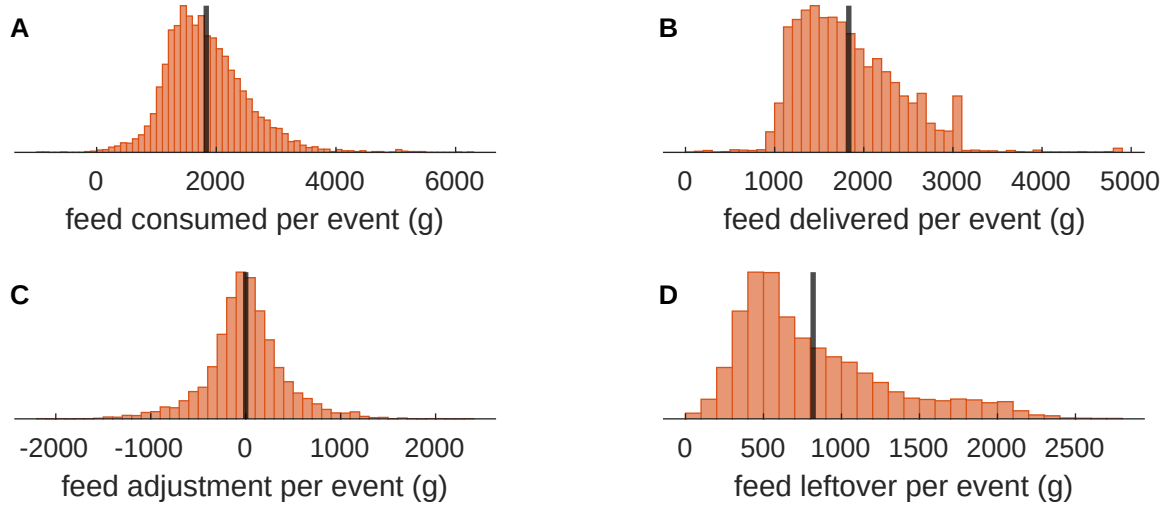


Figure 5.7: Histograms of the metrics in table 5.1, the vertical line identifies the mean values. For clarity, in Panel B, times between successive visit events above 600s are binned at 600s.

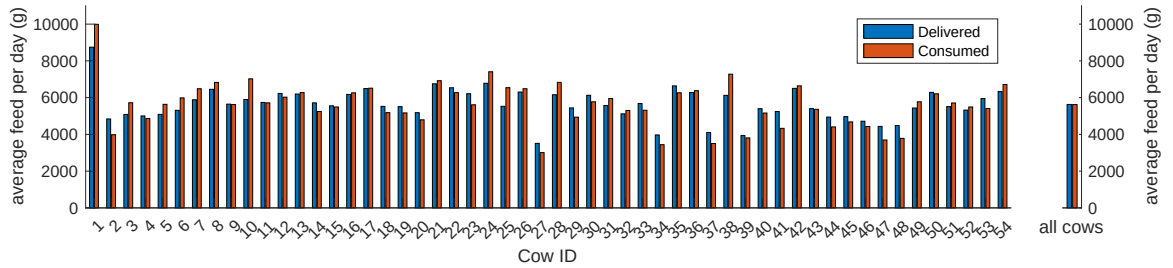


Figure 5.8: Comparison of average daily feed delivered and consumed per animal to herd daily averages. (Left) Bar chart showing individual animals with their respective average daily delivered (blue) and consumed (red) feed. (Right) Aggregated comparison of the overall average delivered and consumed feed across all animals.

5.5 Discussion

This study was aimed at quantifying the amount and variability of feed delivered, and consumed by individual dairy cows within an automated milking system (AMS). Our findings show a notable discrepancy between the feed amounts programmed in the AMS and the actual feed consumption observed at the cow level. This discussion will analyze these discrepancies in the context of individual cow behavior, and established industry benchmarks.

A significant discovery of this work was the consistent presence of feed leftovers, with an average of $818.28 \pm 485.29\text{g}$ remaining in the feed bin after a milking event. This observation refutes the common assumption that cows will consume all feed offered during feeding, and this again highlights the importance and need for the measuring leftovers to determine the true amount of feed consumed.

The primary factor limiting full consumption of feed appears to be the time available for eating. The average visit duration in our study was 7 mins. As noted by Kertz et al. [Kertz et al. \[1981\]](#), industry standards suggest that cows can consume feed at a rate of 250 to 400 g/min. Based on our observed milking duration, this implies a maximum potential consumption of 1.75 to 2.8 kg per visit. So, when the feed delivered to the cow exceeded this capacity, leftovers were inevitable.

Furthermore, I observed high-intercow variability in the amount of feed in the amount of feed leftover. This suggests that factors other than time, such as stage of lactation, individual biological composition play a crucial role, as has been discussed elsewhere.

5.6 Limitations and future directions

I conducted this study on a single farm with one type of AMS and a concentrate composite of 4 different feed types. The observed delivery accuracy and intake behaviors may differ from other system configurations or feed types. Future work

will need to expand this methodology to multiple herds and system configurations to validate these findings. Furthermore, integrating wearable sensor data could provide deeper insights into the specific cow behaviors that lead to feed refusal and thus, allow for more targeted management strategies.

5.6.1 Acknowledgments

This research has been fully supported by a seed grant from the AI Tennessee Initiative at the University of Tennessee Knoxville.

Chapter 6

Conclusions

I began with the premise that AI and modern computational science navigate a crucial and productive tension between two distinct paradigms: model-based inference, which seeks mechanistic understanding, and predictive modeling, which aims to forecast outcomes. Having now presented case studies in both biochemical analysis and precision agriculture, I return to the central theme to synthesize the findings and reflect on the respective strengths, limitations, and ultimately, the complementary nature of these two powerful approaches.

In demonstrating the inferential paradigm, this work delved into the complex analytical landscape of Nuclear Magnetic Resonance (NMR) spectroscopy. Through the development of two distinct Bayesian frameworks, I addressed fundamental challenges in extracting physical meaning from spectral data. My Bayesian non-parametric model provided a robust solution to the critical model selection problem, delivering an automated, principled method for identifying an unknown number of resonances. Complementing this, my parametric model for chemical exchange replaced the unreliable methods of the past with a thermodynamically consistent global analysis. In both instances, the primary achievement was not merely a result or estimation of parameters, but the generation of interpretable physical

parameters—each accompanied by a rigorous quantification of its uncertainty. This work demonstrated the power of the inferential approach to deliver deep, trustworthy, and mechanistic insight into complex physical processes.

I also explored the predictive paradigm through the challenge of detecting feed consumption in agricultural robotics. Acknowledging that accurate prediction relies on high-quality data, I first conducted a foundational study that refuted the flawed “zero-leftover” assumption in automated feeding systems and quantified the true nature of individual cow feed intake. This essential groundwork enabled the subsequent development of the dissertation’s predictive centerpiece: a deep learning model engineered to accurately forecast the feed consumption of individual cows. In alignment with the goals of predictive modeling, this tool prioritized forecasting power, leveraging a complex, non-linear function approximator to create a practical tool aimed at optimizing and automating a real-world agricultural process.

By placing these case studies side-by-side, this dissertation offers a clear perspective on the two paradigms. The inferential NMR models provide in-depth scientific understanding but demand significant domain expertise to construct. The predictive agricultural model offers high forecasting accuracy, essential for automation, but with less inherent physical interpretability. Yet, the most significant insight from this comparative exploration is not in their differences, but the way they complement each other.

The most important finding of this work therefore, is that these class of approaches should not necessarily be forced into a single, unified framework. Rather, their value is in their distinction. The path for future work should focus on sharpening these specialized tools and understanding the landscape where each is most appropriately deployed.

This might take the form of developing rigorous frameworks for model selection and diagnosis that helps data analysts choose between these paradigms. These

methodologies should be able to robustly determine when a system is sufficiently understood to warrant an inferential approach versus when its complexity and problem type demands for the application of a predictive approach. This is not a trivial undertaking as it fundamentally seeks to proffer wisdom not merely models.

In the end, I argue in this dissertation a more broader point that the field of AI is a mature discipline equipped with diverse paradigms for approaching problems. It seems to me that true progress is achieved not by searching for a single master algorithm, but by embracing the diversity and thoughtfully applying the right method to the right problem whether that problem lies in the spins of a biological sample, or in the feeding of a dairy cow.

Bibliography

- Daniel Abergel and Arthur G Palmer. Approximate solutions of the bloch–mcconnell equations for two-site chemical exchange. *ChemPhysChem*, 5(6):787–793, 2004. [61](#)
- Babak Alipanahi, Xin Gao, Emre Karakoc, Logan Donaldson, and Ming Li. Picky: a novel svd-based nmr spectra peak picking method. *Bioinformatics*, 25(12):i268–i275, 2009. [13](#), [20](#)
- Adam Allerhand, HS Gutowsky, J Jonas, and RA Meinzer. Nuclear magnetic resonance methods for determining chemical-exchange rates1. *Journal of the American Chemical Society*, 88(14):3185–3194, 1966. [45](#)
- Peter Atkins and Julio De Paula. *Physical chemistry*, volume 1. Macmillan, 2006. [8](#), [9](#)
- Alex D Bain. Chemical exchange in nmr. *Progress in nuclear magnetic resonance spectroscopy*, 43(3-4):63–103, 2003. [3](#), [44](#), [45](#)
- Alex D Bain, Darrell M Rex, and Rashida N Smith. Fitting dynamic nmr lineshapes. *Magnetic Resonance in Chemistry*, 39(3):122–126, 2001. [3](#), [12](#), [44](#), [45](#), [49](#)
- Yonathan Bard. *Nonlinear parameter estimation*, volume 1209. Academic press New York, 1974. [46](#)
- Gerhard Binsch. The study of intramolecular rate processes by dynamic nuclear magnetic resonance. *Topics in Stereochemistry*, pages 97–192, 1968. [59](#)

- Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006. [46](#)
- Felix Bloch. Nuclear induction. *Physical review*, 70(7-8):460, 1946. [10](#), [13](#), [14](#), [22](#), [100](#)
- Nicolaas Bloembergen, Edward Mills Purcell, and Robert V Pound. Relaxation effects in nuclear magnetic resonance absorption. *Physical review*, 73(7):679, 1948. [53](#)
- Leo Breiman. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, 16(3):199–231, 2001. [2](#)
- Tamara Broderick, Michael I Jordan, and Jim Pitman. Beta processes, stick-breaking and power laws. 2012. [110](#)
- Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of markov chain monte carlo*. CRC press, 2011. [23](#), [24](#), [111](#)
- Stephen D Bruce, John Higinbotham, Ian Marshall, and Paul H Beswick. An analytical derivation of a popular approximation of the voigt function for quantification of nmr spectra. *Journal of Magnetic Resonance*, 142(1):57–63, 2000. [19](#)
- Charles Buchanan, Gogulan Karunanithy, Olga Tkachenko, Michael Barber, Michael T Marty, Timothy J Nott, Christina Redfield, and Andrew J Baldwin. Unidecnmr: automatic peak detection for nmr spectra in 1-4 dimensions. *Nature Communications*, 16(1):449, 2025. [13](#)
- Victor E Cabrera, Jorge A Barrientos-Blanco, Hector Delgado, and Liliana Fadul-Pacheco. Symposium review: Real-time continuous decision making using big data on dairy farms. *Journal of dairy science*, 103(4):3856–3866, 2020. [65](#)
- Herman Y Carr and Edward M Purcell. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical review*, 94(3):630, 1954. [45](#)

- Eduard Y Chekmenev, Siu-Kei Chow, Daniel Tofan, Daniel P Weitekamp, Brian D Ross, and Pratip Bhattacharya. Fluorine-19 nmr chemical shift probes molecular binding to lipid membranes. *The Journal of Physical Chemistry B*, 112(20):6285–6287, 2008. [18](#)
- Yichen Cheng, Xin Gao, and Faming Liang. Bayesian peak picking for nmr spectra. *Genomics, proteomics & bioinformatics*, 12(1):39–47, 2014. [11](#), [13](#), [19](#)
- Siddhartha Chib and Edward Greenberg. Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335, 1995. [16](#)
- JS Chitode. *Digital signal processing*. Technical publications, 2008. [12](#), [25](#)
- National Research Council, Committee on Animal Nutrition, and Subcommittee on Dairy Cattle Nutrition. *Nutrient requirements of dairy cattle: 2001*. National Academies Press, 2001. [65](#)
- John E Dennis Jr and Roy E Welsch. Techniques for nonlinear least squares and robust regression. *Communications in Statistics-simulation and Computation*, 7(4):345–359, 1978. [46](#)
- Lixin Dou and RJW Hodgson. Bayesian inference and gibbs sampling in spectral analysis and parameter estimation. i. *Inverse problems*, 11(5):1069, 1995. [24](#)
- Richard R Ernst and Weston A Anderson. Application of fourier transform spectroscopy to magnetic resonance. *Review of Scientific Instruments*, 37(1):93–102, 1966. [11](#), [12](#)
- Sture Forsén and Ragnar A Hoffman. Study of moderately rapid chemical exchange reactions by means of nuclear magnetic double resonance. *The Journal of Chemical Physics*, 39(11):2892–2901, 1963. [45](#)
- Charlotte Gaillard, Ludovic Brossard, and Jean-Yves Dourmad. Improvement of feed and nutrient efficiency in pig production through precision feeding. *Animal Feed Science and Technology*, 268:114611, 2020. [64](#)

- Alan E Gelfand and Adrian FM Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409, 1990. [16](#)
- Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*. CRC press, 2013. [2](#), [14](#), [15](#), [16](#), [24](#), [41](#), [110](#), [115](#)
- Zoubin Ghahramani. Probabilistic machine learning and artificial intelligence. *Nature*, 521(7553):452–459, 2015. [2](#), [3](#), [24](#), [41](#)
- Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016. [2](#), [4](#), [71](#)
- Thomas L Griffiths and Zoubin Ghahramani. The indian buffet process: An introduction and review. *Journal of Machine Learning Research*, 12(4), 2011. [110](#)
- David H Gultekin and John C Gore. Temperature dependence of nuclear magnetization and relaxation. *Journal of magnetic resonance*, 172(1):133–141, 2005. [45](#)
- H Si Gutowsky and A Saika. Dissociation, chemical exchange, and the proton magnetic resonance in some aqueous electrolytes. *The Journal of Chemical Physics*, 21(10):1688–1694, 1953. [45](#)
- HS Gutowsky, DW McCall, and CP Slichter. Coupling among nuclear magnetic dipoles in molecules. *Physical Review*, 84(3):589, 1951. [11](#)
- Assaf Hallak and Gal Dalal. On the products of stochastic and diagonal matrices. *arXiv preprint arXiv:2304.11634*, 2023. [52](#)
- W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. 1970. [16](#)

- John Higinbotham and Ian Marshall. Nmr lineshapes and lineshape fitting procedures. 2001. [44](#)
- Keegan E Hines. A primer on bayesian inference for biophysical systems. *Biophysical journal*, 108(9):2103–2113, 2015. [20](#)
- Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006. [1](#)
- Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997. [2](#)
- Neil E Jacobsen. *NMR spectroscopy explained: simplified theory, applications and examples for organic chemistry and structural biology*. John Wiley & Sons, 2007. [3](#), [7](#), [8](#), [9](#), [11](#), [26](#), [41](#)
- Amir Jahangiri, Xiao Han, Dmitry Lesovoy, Tatiana Agback, Peter Agback, Adnane Achour, and Vladislav Orekhov. Nmr spectrum reconstruction as a pattern recognition problem. *Journal of Magnetic Resonance*, 346:107342, 2023. [19](#)
- Sina Jazani, Ioannis Sgouralis, and Steve Pressé. A method for single molecule tracking using a conventional single-focus confocal setup. *The Journal of chemical physics*, 150(11), 2019a. [20](#)
- Sina Jazani, Ioannis Sgouralis, Omer M Shafriz, Sanjeevi Sivasankar, and Steve Pressé. An alternative framework for fluorescence correlation spectroscopy. *Biophysical Journal*, 116(3):282a, 2019b. [3](#), [23](#)
- Michael L Johnson. [1] use of least-squares techniques in biochemistry. In *Methods in Enzymology*, volume 240, pages 1–22. Elsevier, 1994. [46](#)
- Badal Joshi. A detailed balanced reaction network is sufficient but not necessary for its markov chain to be detailed balanced. *arXiv preprint arXiv:1312.4196*, 2013. [51](#)

- Jerome Kaplan. *NMR of chemically exchanging systems*. Elsevier, 2012. [7](#), [8](#), [9](#), [12](#), [44](#), [45](#), [47](#)
- JD Kaufman, KR Kassube, and AG R  s. Lowering rumen-degradable protein maintained energy-corrected milk yield and improved nitrogen-use efficiency in multiparous lactating dairy cows exposed to heat stress. *Journal of Dairy Science*, 100(10):8132–8145, 2017. [65](#)
- James Keeler. *Understanding NMR spectroscopy*. John Wiley & Sons, 2011. [44](#), [45](#)
- AF Kertz, BK Darcy, and LR Prewitt. Eating rate of lactating cows fed four physical forms of the same grain ration. *Journal of Dairy Science*, 64(12):2388–2391, 1981. [81](#)
- Julianne L Kitevski-LeBlanc and R Scott Prosser. Current applications of 19f nmr to studies of protein structure and dynamics. *Progress in nuclear magnetic resonance spectroscopy*, 62:1–33, 2012. [18](#)
- Gerard J Kleywegt, Rolf Boelens, and Robert Kaptein. A versatile approach toward the partially automatic recognition of cross peaks in 2d 1h nmr spectra. *Journal of magnetic Resonance*, 88(3):601–608, 1990. [13](#)
- Reto Koradi, Martin Billeter, Max Engeli, Peter G  ntert, and Kurt W  thrich. Automated peak picking and peak integration in macromolecular nmr spectra using autopsy. *Journal of Magnetic Resonance*, 135(2):288–297, 1998. [13](#)
- Dmitry M Korzhnev, Ilghiz V Ibraghimov, Martin Billeter, and Vladislav Yu Orekhov. Munin: application of three-way decomposition to the analysis of heteronuclear nmr relaxation data. *Journal of biomolecular NMR*, 21:263–268, 2001. [20](#)
- Yann LeCun, L  on Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. [1](#)

- Da-Wei Li, Alexandar L Hansen, Chunhua Yuan, Lei Bruschweiler-Li, and Rafael Brüschweiler. Deep picker is a deep neural network for accurate deconvolution of complex two-dimensional nmr spectra. *Nature communications*, 12(1):5229, 2021. [2](#), [11](#), [13](#), [19](#), [20](#), [34](#), [41](#)
- Da-Wei Li, Lei Bruschweiler-Li, Alexandar L Hansen, and Rafael Brüschweiler. Deep picker1d and voigt fitter1d: A versatile tool set for the automated quantitative spectral deconvolution of complex 1d nmr spectra. *Magnetic Resonance Discussions*, 2022:1–20, 2022. [12](#), [13](#), [19](#)
- Konstantinos G Liakos, Patrizia Busato, Dimitrios Moshou, Simon Pearson, and Dionysis Bochtis. Machine learning in agriculture: A review. *Sensors*, 18(8):2674, 2018. [4](#)
- John C Lindon and AG Ferrige. Digitisation and data processing in fourier transform nmr. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 14(1):27–66, 1980. [12](#)
- Qian-Bao Liu, Jiazheng Liu, Jing-Guang Lu, Ming-Rong Yang, Wei Zhang, Wen-Jia Li, Zheng-Ming Qian, Zhi-Hong Jiang, and Li-Ping Bai. Quantitative 1h nmr with global spectral deconvolution approach for quality assessment of natural and cultured cordyceps sinensis. *Journal of Pharmaceutical and Biomedical Analysis*, 235:115603, 2023. [13](#)
- Zhi Liu, Ahmed Abbas, Bing-Yi Jing, and Xin Gao. Wavpeak: picking nmr peaks through wavelet-based smoothing and volume-based filtering. *Bioinformatics*, 28(7):914–920, 2012. [13](#)
- Yevgen Matviychuk, Erik von Harbou, and Daniel J Holland. An experimental validation of a bayesian model for quantification in nmr spectroscopy. *Journal of Magnetic Resonance*, 285:86–100, 2017. [20](#)
- Harden M McConnell. Reaction rates by nuclear magnetic resonance. *The Journal of chemical physics*, 28(3):430–431, 1958. [14](#), [45](#), [46](#), [48](#)

- Stephen McCoy, Daniel McBride, D. Katie McCullough, Benjamin C. Calfee, Erik Zinser, David Talmy, and Ioannis Sgouralis. Quantitative assessment of biological dynamics with aggregate data, 2025. URL <https://arxiv.org/abs/2504.02581>. 16, 115
- Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018. 1
- Mestrelab Research S.L. Mestrelab: Nmr analysis software. <https://mestrelab.com/main-product/nmr>. Accessed: 29 January 2025. 13, 34
- Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953. 16
- LE Moraes, E Kebreab, JL Firkins, RR White, R Martineau, and H Lapierre. Predicting milk protein responses and the requirement of metabolizable protein by lactating dairy cows. *Journal of dairy science*, 101(1):310–327, 2018. 64
- Iain Murray, Ryan Adams, and David MacKay. Elliptical slice sampling. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 541–548. JMLR Workshop and Conference Proceedings, 2010. 16, 111, 115
- D Muthukumaran and M Sivakumar. Medical image registration: a matlab based approach. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol*, 2(1):29–34, 2017. 71
- National Academies of Sciences, Engineering, and Medicine. *Nutrient Requirements of Dairy Cattle: Eighth Revised Edition*. The National Academies Press, Washington, DC, 2021. ISBN 978-0-309-71449-5. doi: 10.17226/25806. URL <https://nap.nationalacademies.org/catalog/25806/nutrient-requirements-of-dairy-cattle-eighth-revised-edition>. 64

- Radford M Neal. Slice sampling. *The annals of statistics*, 31(3):705–767, 2003. [16](#), [111](#), [115](#)
- James R Norris. *Markov chains*. Number 2. Cambridge university press, 1998. [51](#)
- MR Osborne. Some special nonlinear least squares problems. *SIAM Journal on Numerical Analysis*, 12(4):571–592, 1975. [46](#)
- John Paisley and Lawrence Carin. Nonparametric factor analysis with beta process priors. In *Proceedings of the 26th annual international conference on machine learning*, pages 777–784, 2009. [23](#), [24](#)
- Candido Pomar, Ines Andretta, and Aline Remus. Feeding strategies to reduce nutrient losses and improve the sustainability of growing pigs. *Frontiers in Veterinary Science*, 8:742220, 2021. [64](#)
- Simona MC Porto, Claudia Arcidiacono, Umberto Anguzza, and Giovanni Cascone. A computer vision-based system for the automatic detection of lying behaviour of dairy cows in free-stall barns. *Biosystems Engineering*, 115(2):184–194, 2013. [4](#)
- Steve Pressé and Ioannis Sgouralis. *Data Modeling for the Sciences: Applications, Basics, Computations*. Cambridge University Press, 2023. [46](#)
- Jacquelyn M Prestegard-Wilson, VL Daley, Tiffany A Drape, and Mark D Hanigan. A survey of united states dairy cattle nutritionists’ practices and perceptions of reducing crude protein in lactating dairy cow diets. *Applied Animal Science*, 37(6): 697–709, 2021. [65](#)
- Lawrence R Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989. [1](#)
- K Ramani, S Ganapathy, and R Srinivasan. A fourier transform approach to voigt profile analysis and its application to nuclear magnetic resonance. *Journal of Magnetic Resonance (1969)*, 24(2):231–237, 1976. [19](#)

- LW Reeves and KN Shaw. Nuclear magnetic resonance studies of multi-site chemical exchange. i. matrix formulation of the bloch equations. *Canadian Journal of Chemistry*, 48(23):3641–3653, 1970. [61](#)
- Christian Robert, George Casella, Christian P Robert, and George Casella. Metropolis–hastings algorithms. *Introducing Monte Carlo Methods with R*, pages 167–197, 2010. [111](#)
- Christian P Robert, George Casella, and George Casella. *Monte Carlo statistical methods*, volume 2. Springer, 1999. [16](#)
- Denis V Rubtsov and Julian L Griffin. Time-domain bayesian detection and estimation of noisy damped sinusoidal signals applied to nmr spectroscopy. *Journal of Magnetic Resonance*, 188(2):367–379, 2007. [19](#), [20](#), [41](#)
- Gordon S Rule and T Kevin Hitchens. *Fundamentals of protein NMR spectroscopy*, volume 5. Springer Science & Business Media, 2006a. [7](#), [12](#), [13](#), [18](#), [25](#), [47](#), [49](#)
- Gordon S Rule and T Kevin Hitchens. *NMR spectroscopy*. Springer, 2006b. [3](#), [11](#), [12](#), [45](#)
- Ernest Rutherford and Nevill Francis Mott. Discussion on the structure of atomic nuclei. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 136(830):735–762, 1932. [8](#)
- Nicolas Schmid, Simon Bruderer, F Paruzzo, G Fischetti, Giuseppe Toscano, Dominik Graf, Michael Fey, Andreas Henrici, V Ziebart, B Heitmann, et al. Deconvolution of 1d nmr spectra: A deep learning-based approach. *Journal of Magnetic Resonance*, 347:107357, 2023. [13](#), [19](#)
- Ioannis Sgouralis, Shreya Madaan, Franky Djutanta, Rachael Kha, Rizal F Hariadi, and Steve Pressé. A bayesian nonparametric approach to single molecule forster resonance energy transfer. *The Journal of Physical Chemistry B*, 123(3):675–688, 2018. [3](#), [20](#)

- Diana M Sima and Sabine Van Huffel. Separable nonlinear least squares fitting with linear bound constraints and its application in magnetic resonance spectroscopy data quantification. *Journal of computational and applied mathematics*, 203(1): 264–278, 2007. [46](#)
- Devinderjit Sivia and John Skilling. *Data analysis: a Bayesian tutorial*. OUP Oxford, 2006. [24](#)
- Ionel Solomon. Relaxation processes in a system of two spins. *Physical Review*, 99(2):559, 1955. [44](#)
- Alois Steigel and Hans W Spiess. *Dynamic NMR spectroscopy*, volume 15. Springer Science & Business Media, 2012. [44](#), [45](#)
- Meysam Tavakoli, Sina Jazani, Ioannis Sgouralis, and Steve Presse. Bayesian nonparametrics for fluorescence methods. *Biophysical Journal*, 116(3):39a, 2019. [20](#)
- Meysam Tavakoli, Sina Jazani, Ioannis Sgouralis, Wooseok Heo, Kunihiro Ishii, Tahei Tahara, and Steve Pressé. Direct photon-by-photon analysis of time-resolved pulsed excitation data using bayesian nonparametrics. *Cell reports physical science*, 1(11), 2020a. [3](#), [20](#)
- Meysam Tavakoli, Sina Jazani, Ioannis Sgouralis, and Steve Presse. Single molecules dynamics learned from single photons-flim and fcs with bayesian nonparametrics. *Biophysical Journal*, 118(3):313a–314a, 2020b. [3](#), [20](#)
- Meysam Tavakoli, Sina Jazani, Ioannis Sgouralis, Omer M Shafraz, Sanjeevi Sivasankar, Bryan Donaphon, Marcia Levitus, and Steve Pressé. Pitching single-focus confocal data analysis one photon at a time with bayesian nonparametrics. *Physical review X*, 10(1):011021, 2020c. [20](#)
- Quincy Teng. *Structural biology: practical NMR applications*. Springer Science & Business Media, 2012. [19](#)

- Suhas Tikole, Victor Jaravine, Vladimir Rogov, Volker Dötsch, and Peter Güntert. Peak picking nmr spectral data using non-negative matrix factorization. *BMC bioinformatics*, 15(1):1–7, 2014. [20](#)
- Richard Chace Tolman. *The principles of statistical mechanics*. Courier Corporation, 1979. [10](#)
- Rens van de Schoot, Sarah Depaoli, Ruth King, Bianca Kramer, Kaspar Märtens, Mahlet G Tadesse, Marina Vannucci, Andrew Gelman, Duco Veen, Joukje Willemsen, et al. Bayesian statistics and modelling. *Nature Reviews Methods Primers*, 1(1):1, 2021. [24](#), [41](#)
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017. [2](#)
- Udo Von Toussaint. Bayesian inference in physics. *Reviews of Modern Physics*, 83(3):943, 2011. [24](#)
- Shmuel Winograd. On computing the discrete fourier transform. *Mathematics of computation*, 32(141):175–199, 1978. [22](#)
- Jin Yang, William J Bruno, William S Hlavacek, and John E Pearson. On imposing detailed balance in complex reaction mechanisms. *Biophysical journal*, 91(3):1136–1141, 2006. [51](#)
- Ying Yin, Mengyu Wu, Lejla Zubcevic, William F Borschel, Gabriel C Lander, and Seok-Yong Lee. Structure of the cold-and menthol-sensing ion channel trpm8. *Science*, 359(6372):237–241, 2018. [25](#)
- Chan Zheng, Xunmu Zhu, Xiaofan Yang, Lina Wang, Shuqin Tu, and Yueju Xue. Automatic recognition of lactating sow postures from depth images by deep learning detector. *Computers and electronics in agriculture*, 147:51–63, 2018. [4](#)

Nicolai Zientek, Clément Laurain, Klas Meyer, Matthias Kraume, Gisela Guthausen, and Michael Maiwald. Simultaneous ^{19}F – ^1H medium resolution nmr spectroscopy for online reaction monitoring. *Journal of magnetic resonance*, 249:53–62, 2014. [18](#)

Appendix A

Bayesian nonparametric statistical learning for resonance identification and characterization in ^{19}F -NMR

Detailed model description

Magnetization

An FID measurement made at time t , during an NMR experiment, probes the transverse components of the *total magnetization* of the NMR specimen. In the lab's stationary frame, this is represented by the sum

$$\vec{M}(t) = \sum_{k=1}^K \vec{M}^k(t). \quad (\text{A.1})$$

Here, $\vec{M}^k(t)$ is the *net magnetization* of the k^{th} spin species in the specimen. This is the resultant magnetization that adds up contributions by all individual spins of

the species. I use superscripts $k = 1, \dots, K$ to label the species. All individual spins contributing to $\vec{M}(t)$ have the same atomic nuclei; individual spins contributing to each $\vec{M}^k(t)$ are partitioned into species according to the local chemical environment of the nuclei.

Ignoring chemical exchange, each species magnetization $\vec{M}^k(t)$ follows its own *Bloch equation* Bloch [1946], which reads

$$\frac{d}{dt}M^k(t) = \gamma(1 - \sigma^k)(\vec{M}^k(t) \times \vec{B}_0^\#(t)) + \begin{bmatrix} R_2^k & 0 & 0 \\ 0 & R_2^k & 0 \\ 0 & 0 & R_1^k \end{bmatrix} \left(M_0^k \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \vec{M}^k(t) \right) \quad k = 1, \dots, K$$

The gyromagnetic ratio γ and external magnetic field $\mathbf{B}_0^\#(t)$ apply to all spin species; however, the shielding constant σ^k , the relaxation rates R_1^k, R_2^k and the equilibrium magnetization M_0^k are species dependent.

The *detection phase* begins at time $t_{\min}$ and ends at time $t_{\max}$. Between these times, the external magnetic field is independent of time and given by

$$\mathbf{B}_0^\#(t) = \begin{pmatrix} 0 \\ 0 \\ B_0^\# \end{pmatrix}, \quad t_{\min} \leq t \leq t_{\max} \quad (\text{A.2})$$

Accordingly, the general solution of the Bloch's equations, during detection phase, reads

$$\vec{M}^k(t) = \begin{pmatrix} e^{-R_2^k(t-t_{\min})} \left(G_1^k \cos \left(\gamma(1 - \sigma^k) B_0^\#(t - t_{\min}) \right) \right. \\ \left. + G_2^k \sin \left(\gamma(1 - \sigma^k) B_0^\#(t - t_{\min}) \right) \right) \\ e^{-R_2^k(t-t_{\min})} \left(G_2^k \cos \left(\gamma(1 - \sigma^k) B_0^\#(t - t_{\min}) \right) \right. \\ \left. - G_1^k \sin \left(\gamma(1 - \sigma^k) B_0^\#(t - t_{\min}) \right) \right) \\ M_0^k + e^{-R_2^k(t-t_{\min})} G_3^k \end{pmatrix}$$

$t_{\min} \leq t \leq t_{\max}$

Here G_1^k, G_2^k, G_3^k are the constants of integration. These are related to the initial conditions at the beginning of the detection phase. For this reason, the total magnetization reads

$$\vec{M}(t) = \begin{pmatrix} \sum_{k=1}^K e^{-R_2^k(t-t_{\min})} \left(G_1^k \cos \left(\gamma(1 - \sigma^k) B_0^\#(t - t_{\min}) \right) \right. \\ \left. + G_2^k \sin \left(\gamma(1 - \sigma^k) B_0^\#(t - t_{\min}) \right) \right) \\ \sum_{k=1}^K e^{-R_2^k(t-t_{\min})} \left(G_2^k \cos \left(\gamma(1 - \sigma^k) B_0^\#(t - t_{\min}) \right) \right. \\ \left. - G_1^k \sin \left(\gamma(1 - \sigma^k) B_0^\#(t - t_{\min}) \right) \right) \\ \sum_{k=1}^K M_0^k + e^{-R_2^k(t-t_{\min})} G_3^k \end{pmatrix},$$

$t_{\min} \leq t \leq t_{\max}$

Prior to acquisition, the total magnetization is transformed electronically to a *rotating frame*. By convention, this corresponds to the transformation $\mathbf{M}'(t) = \mathbb{R}_z^{\Omega t} \mathbf{M}(t)$

where the rotation matrix is given by

$$\mathbb{R}_z^{\Omega t} = \begin{bmatrix} +\cos(\Omega t) & -\sin(\Omega t) & 0 \\ +\sin(\Omega t) & +\cos(\Omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (\text{A.3})$$

Under this convention, the magnetization in the rotating frame reads

$$\mathbf{M}'(t) = \left\{ \begin{array}{l} \sum_{k=1}^K e^{-R_2^k(t-t_{\min})} \left(G_1^k \cos((\omega^k - \Omega)(t - t_{\min})) \right. \\ \qquad \qquad \qquad \left. + G_2^k \sin((\omega^k - \Omega)(t - t_{\min})) \right) \\ \sum_{k=1}^K e^{-R_2^k(t-t_{\min})} \left(G_2^k \cos((\omega^k - \Omega)(t - t_{\min})) \right. \\ \qquad \qquad \qquad \left. - G_1^k \sin((\omega^k - \Omega)(t - t_{\min})) \right) \\ \sum_{k=1}^K M_0^k + e^{-R_2^k(t-t_{\min})} G_3^k \end{array} \right\}, \quad t_{\min} \leq t \leq t_{\max}$$

Here, Ω is the frequency of the rotating frame and $\omega^k = \gamma(1 - \sigma^k)B_0^\#$.

Data acquisition

During the detection phase of the NMR experiment, measurements are acquired at time levels t_n that are arranged in strictly increasing order

$$t_{\min} \leq t_1 < t_2 < \dots < t_N \leq t_{\max}. \quad (\text{A.4})$$

Here, the times $t_{\min}$ and $t_{\max}$ mark the onset and conclusion of the detection phase.

I consider FID data collected from a 1D NMR experiment, $z_{1:N}$ where z_n is the magnetization sampled from the spectrometer at time t_n . Since, the FID results from quadrature sampling of the magnetization, each z_n is a complex signal consisting of

a real and an imaginary part. I represent this complex signal as a two dimensional vector (x_n, y_n) where $z_n = x_n + iy_n$.

Due to the spectrometer's sampling electronics, I assume that these datapoints are independent from each other. I equally note that the signal recorded by the spectrometer consists of the true FID measurements as predicted by the Bloch's equations accompanied with some noise. I represent the magnetization predicted by Bloch's equations as $(M'_x(t_n), M'_y(t_n))$.

The measurement (x_n, y_n) acquired at the n^{th} time level is bivariate and probes the total magnetization like this:

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} P \\ Q \end{pmatrix} + \lambda \begin{pmatrix} \frac{1}{\tau} \int_{t_n - \frac{\tau}{2}}^{t_n + \frac{\tau}{2}} dt M'_x(t) \\ \frac{1}{\tau} \int_{t_n - \frac{\tau}{2}}^{t_n + \frac{\tau}{2}} dt M'_y(t) \end{pmatrix} + \zeta_n, \quad n = 1, \dots, N. \quad (\text{A.5})$$

Here, τ is the exposure time, P, Q are the offsets in the two channels, λ is the gain, and ζ_n denotes additive noise.

- The short-term approximation reads

$$\begin{aligned} \frac{1}{\tau} \int_{t_n - \frac{\tau}{2}}^{t_n + \frac{\tau}{2}} dt M'_x(t) &\approx M'_x(t_n) & n = 1, \dots, N \\ \frac{1}{\tau} \int_{t_n - \frac{\tau}{2}}^{t_n + \frac{\tau}{2}} dt M'_y(t) &\approx M'_y(t_n) & n = 1, \dots, N \end{aligned}$$

- Gaussian noise reads

$$\zeta_n \sim \text{Normal}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, S \right) \quad n = 1, \dots, N$$

Here, S is a positive definite 2×2 covariance matrix. Gathering these conventions together, our likelihood reads

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} \sim \text{Normal}_2 \left(\begin{pmatrix} P + \lambda M'_x(t_n) \\ Q + \lambda M'_y(t_n) \end{pmatrix}, S \right) \quad n = 1, \dots, N$$

Notation

To simplify the notation, I cast the likelihood in this form

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} \sim \text{Normal}_2 \left(\begin{pmatrix} P + \sum_{k=1}^K p^k e^{-r^k T_n} \cos(w^k T_n) - q^k e^{-r^k T_n} \sin(w^k T_n) \\ Q + \sum_{k=1}^K q^k e^{-r^k T_n} \cos(w^k T_n) + p^k e^{-r^k T_n} \sin(w^k T_n) \end{pmatrix}, S \right) \quad n = 1, \dots, N$$

I further suppose S is a diagonal matrix that represents white noise with the same variance in both channels. The magnetization in both independent channels can both be simplified as:

$$x_n \sim \text{Normal} \left(P + \sum_{k=1}^K p^k e^{-r^k T_n} \cos(w^k T_n) - q^k e^{-r^k T_n} \sin(w^k T_n), v \right) \quad (\text{A.6})$$

$$y_n \sim \text{Normal} \left(Q + \sum_{k=1}^K q^k e^{-r^k T_n} \cos(w^k T_n) + p^k e^{-r^k T_n} \sin(w^k T_n), v \right) \quad (\text{A.7})$$

The new parameters are given by

$$\begin{aligned} T_n &= t_n - t_{\min}, & n &= 1, \dots, N \\ p^k &= G_1^k, & k &= 1, \dots, K \\ q^k &= G_2^k, & k &= 1, \dots, K \\ r^k &= R_2^k, & k &= 1, \dots, K \end{aligned} \quad (\text{A.8})$$

In this likelihood, the unknowns are: $P, Q, v_{xx}, v_{yy}, p^{1:K}, q^{1:K}, r^{1:K}, w^{1:K}$.

Derivation of frequency domain spectra

To derive the Fourier spectrum, I consider a signal

$$A(t) = M'_x(t) + iM'_y(t) \quad (\text{A.9})$$

Using the formulation in appendix A above, I can write

$$A(t) = \begin{cases} 0, t < t_{\min} \\ \sum_{k=1}^K (p^k + iq^k) e^{(iw^k - r^k)t}, t_{\min} \leq t \end{cases} \quad (\text{A.10})$$

The *spectral representation* $B(\nu)$ of the signal $A(t)$, is given by the Fourier transform as follows

$$B(\nu) = \int_{-\infty}^{+\infty} dt A(t) e^{-2\pi i \nu t} = \int_{t_{\min}}^{+\infty} dt A(t) e^{-2\pi i \nu t} \quad (\text{A.11})$$

$$= \int_{t_{\min}}^{+\infty} dt \sum_{k=1}^K (p^k + iq^k) e^{(iw^k - 2\pi i \nu - r^k)t} \quad (\text{A.12})$$

$$= \sum_{k=1}^K (p^k + iq^k) \int_{t_{\min}}^{+\infty} dt e^{(iw^k - 2\pi i \nu - r^k)t} \quad (\text{A.13})$$

$$= \sum_{k=1}^K (p^k + iq^k) \frac{e^{-(r^k + i(2\pi \nu - w^k))t_{\min}}}{r + i(2\pi \nu - w^k)} \quad (\text{A.14})$$

Bayesian model without model selection

If K is known, the following is a parametric Bayesian model that can be used to estimate the FID signal measurements.

$$\begin{aligned}
w^k &\sim \text{Beta}(E, F, w_{\min}, w_{\max}) & k = 1, \dots, K \\
r^k &\sim \text{Beta}(E', F', r_{\min}, r_{\max}) & k = 1, \dots, K \\
p^k &\sim \text{Normal}(0, v) & k = 1, \dots, K \\
q^k &\sim \text{Normal}(0, v) & k = 1, \dots, K \\
P &\sim \text{Normal}(0, V) \\
Q &\sim \text{Normal}(0, V) \\
\tau &\sim \text{Gamma}(\alpha, \beta) \\
x_n | \dots &\sim \text{Normal} \left(P + \sum_{k=1}^K p^k e^{-r^k T_n} \cos(w^k T_n) - q^k e^{-r^k T_n} \sin(w^k T_n), \frac{1}{\tau} \right) & n = 1, \dots, N \\
y_n | \dots &\sim \text{Normal} \left(Q + \sum_{k=1}^K q^k e^{-r^k T_n} \cos(w^k T_n) + p^k e^{-r^k T_n} \sin(w^k T_n), \frac{1}{\tau} \right) & n = 1, \dots, N
\end{aligned}$$

Bayesian model with model selection

Since K is generally unknown, a non-parametric Bayesian model can be used to estimate the FID signal measurements.

$$\begin{aligned}
b^k &\sim \text{Bernoulli}\left(\frac{\alpha}{K}\right) & k = 1, \dots, K \\
w^k &\sim \text{Beta}(E, F, w_{\min}, w_{\max}) & k = 1, \dots, K \\
r^k &\sim \text{Beta}(E', F', r_{\min}, r_{\max}) & k = 1, \dots, K \\
p^k &\sim \text{Normal}(0, v) & k = 1, \dots, K \\
q^k &\sim \text{Normal}(0, v) & k = 1, \dots, K \\
P &\sim \text{Normal}(0, V) \\
Q &\sim \text{Normal}(0, V) \\
\tau &\sim \text{Gamma}(\alpha, \beta) \\
x_n | \dots &\sim \text{Normal} \left(\begin{pmatrix} P + \sum_{k=1}^K b^k p^k e^{-r^k T_n} \cos(w^k T_n) \\ -b^k q^k e^{-r^k T_n} \sin(w^k T_n), \frac{1}{\tau} \end{pmatrix} \right) & n = 1, \dots, N \\
y_n | \dots &\sim \text{Normal} \left(\begin{pmatrix} Q + \sum_{k=1}^K b^k q^k e^{-r^k T_n} \cos(w^k T_n) \\ +b^k p^k e^{-r^k T_n} \sin(w^k T_n), \frac{1}{\tau} \end{pmatrix} \right) & n = 1, \dots, N
\end{aligned}$$

Detailed description of the inference framework

Prior probability distributions

To guarantee a range for both positive and negative real numbers, I place Normal distribution priors on p^k, q^k, P, Q .

In a similar fashion, I choose Beta priors to model w^k and r^k since the frequency and rates will typically take positive values within a certain range.

Since the load term b^k takes values in $\{0, 1\}$, I select a Bernoulli prior for it. Finally, I choose a Gamma distribution for the prior on the precision parameter τ due its conjugacy with the Normal distribution.

The Beta-Bernoulli process

As explained in the main text, estimating how many resonances are present in an FID is equivalent to estimating how many of those resonances are active (i.e. $b^k = 1$).

In the non-parametric formulation for my model, I use a Bernoulli prior of weight q^k with a conjugate Beta hyperprior to make sure that each load b^k takes only values 0 or 1.

$$b^k | q^k \sim \text{Bernoulli}(q^k) \quad (\text{A.15})$$

$$q^k \sim \text{Beta}(m, n) \quad (\text{A.16})$$

where m, n are parameters from the Beta distribution we need to estimate

Note that the marginal distribution $p(b^k)$ is computed by the following:

$$\begin{aligned} p(b^k) &= \int_0^1 p(b^k | q^k) p(q^k) dq^k \\ &= \int_0^1 \text{Bernoulli}(b^k; q^k) \text{Beta}(q^k; m, n) dq^k \end{aligned}$$

$$\text{Beta}(q^k; m, n) = \frac{1}{B(m, n)} (q^k)^{m-1} (1 - q^k)^{n-1},$$

where $B(m, n)$ is the Beta function:

$$B(m, n) = \int_0^1 (q)^{m-1} (1-q)^{n-1} dq = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}.$$

We note that since b^k is a Bernoulli random variable,

$$p(b^k) = p(b^k = 1) + p(b^k = 0)$$

Case 1: $b^k = 1$

$$\begin{aligned} P(b^k = 1) &= \int_0^1 q^k \cdot \frac{1}{B(m, n)} (q^k)^{m-1} (1-q^k)^{n-1} dq^k \\ &= \frac{1}{B(m, n)} \int_0^1 (q^k)^m (1-q^k)^{n-1} dq^k \\ &= \frac{B(m+1, n)}{B(m, n)} \\ &= \frac{m}{m+n} \end{aligned}$$

Case 2: $b^k = 0$

$$\begin{aligned} P(b^k = 0) &= \int_0^1 (1-q^k) \cdot \frac{1}{B(m, n)} (q^k)^{m-1} (1-q^k)^{n-1} dq^k \\ &= \frac{1}{B(m, n)} \int_0^1 (q^k)^{m-1} (1-q^k)^n dq^k \\ &= \frac{B(m, n+1)}{B(m, n)} \\ &= \frac{n}{m+n} \end{aligned}$$

Therefore, $b^k \sim \text{Bernoulli}\left(\frac{m}{m+n}\right)$

Physically this means that the probability a resonance is active is $\frac{m}{m+n}$.

Since, $K \rightarrow \infty$, we desire this probability to be small to guarantee that a sparse amount of resonances are active. Simultaneously, we would like $\sum_{k=1}^K \mathbb{E}[b^k]$ to be finite

as $K \rightarrow \infty$ since we have a finite expected number of resonances in our spectrum. An infinite expectation does not agree with the physical meaning of the spectrum.

Now $\sum_{k=1}^K \mathbb{E}[b^k] = \frac{Km}{m+n}$ So, I pick m, n so that the following conditions are met:

- $\frac{Km}{m+n}$ is finite as $K \rightarrow \infty$
- $\frac{m}{m+n}$ is small (it suffices to ensure that the denominator depends on K)

I make choices of $m = \frac{\alpha}{K}, n = \frac{\beta}{K}(K-1)$ so that the process is defined by the following hierarchical model:

$$\begin{aligned} b^k | q^k &\sim \text{Bernoulli}(q^k) \\ q^k &\sim \text{Beta}\left(\frac{\alpha}{K}, \frac{\beta}{K}(K-1)\right) \end{aligned} \quad k = 1, \dots, K \quad (\text{A.17})$$

By taking the limit as the number of components $K \rightarrow \infty$, this sequence of finite models converges to a Beta-Bernoulli process [Broderick et al. \[2012\]](#), which is a non-parametric prior. This specific parametrization is a common way to construct a finite approximation of the Indian Buffet Process (IBP) [Griffiths and Ghahramani \[2011\]](#), where α and β are hyperparameters that control the expected number of features and their sharing across observations in the limiting infinite process.

Computational implementation

The joint probability distribution of my model is given by

$$p(\tau, P, Q, p^{1:K}, q^{1:K}, r^{1:K}, w^{1:K}, b^{1:K} \mid z_{1:N}) \quad (\text{A.18})$$

I generate posterior samples according to a Gibbs sampling scheme [Gelman et al. \[2013\]](#). This is done by sampling a random variable conditioned on all other random variables and the FID data measurements $z_{1:N}$.

Conceptually, the steps in the generation of each posterior sample are:

1. Using direct sampling [Brooks et al. \[2011\]](#), update the precision parameter matrix τ and simultaneously update jointly the indicators b^k for all spin species. In particular, τ and loads for all spin species $b^{1:K}$ are updated from the distribution

$$p(\tau, b^{1:K} | P, Q, p^{1:K}, q^{1:K}, r^{1:K}, w^{1:K}, z_{1:N}). \quad (\text{A.19})$$

2. Using Metropolis-Hastings algorithm [Robert et al. \[2010\]](#), update jointly the offsets P , Q and simultaneously update jointly the indicators b^k , and initial constants p^k , q^k for all spin species. In particular, P , Q , p^k , q^k and loads for all spin species $b^{1:K}$ are updated from the distribution

$$p(P, Q, p^{1:K}, q^{1:K} | \tau, r^{1:K}, w^{1:K}, z_{1:N}). \quad (\text{A.20})$$

3. Using elliptical slice sampling [Neal \[2003\]](#), [Murray et al. \[2010\]](#), simultaneously update jointly the initial constants p^k , q^k , the relaxation rate r^k and the frequencies ω^k for all spin species. In particular, p^k , q^k , r^k , w^k for all spin species are updated from the distribution

$$p(p^{1:K}, q^{1:K}, r^{1:K}, w^{1:K} | \tau, P, Q, z_{1:N}). \quad (\text{A.21})$$

Appendix B

Probing statistical learning with chemical exchange

Proposition B.1. *Let V be the $K \times K$ strictly lower-triangular rate matrix*

$$V = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ v^{1 \leftrightarrow 2} & 0 & 0 & \cdots & 0 \\ v^{1 \leftrightarrow 3} & v^{2 \leftrightarrow 3} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v^{1 \leftrightarrow K} & v^{2 \leftrightarrow K} & v^{3 \leftrightarrow K} & \cdots & 0 \end{bmatrix},$$

and let $\pi = (\pi_1, \dots, \pi_K)^\top$ be a positive probability vector. Define

$$D_\pi = \text{diag}(\pi), \quad \Lambda = (V + V^\top) D_\pi^{-1},$$

and let

$$G = \Lambda + \text{diag} \left(-\Lambda \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right)$$

be the column-stochastic generator obtained by shifting Λ so that each column sums to zero. Then

$$G \pi = 0 \quad \text{and} \quad G D_\pi = D_\pi G^\top,$$

i.e. π is the stationary distribution of G and G satisfies detailed balance with respect to π .

Proof. Write

$$\kappa = \text{diag} \left(-\Lambda \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right),$$

so $G = \Lambda + \kappa$. We check the two conditions in turn.

(i) Stationarity:

$$G \pi = (\Lambda + \kappa) \pi = \Lambda \pi + \kappa \pi.$$

Since $\Lambda = (V + V^\top) D_\pi^{-1}$, we have

$$\Lambda \pi = (V + V^\top) D_\pi^{-1} \pi = V D_\pi^{-1} \pi + V^\top D_\pi^{-1} \pi = V \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + V^\top \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},$$

and by definition

$$\kappa \pi = \text{diag} \left(-\Lambda \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right) \pi = -\text{diag} \left((V + V^\top) D_\pi^{-1} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right) \pi.$$

But for any diagonal D and vector s , $\text{diag}(D s) = D$, so

$$\kappa \pi = -V \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} - V^\top \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Hence

$$G \pi = \left(V \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + V^\top \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right) - \left(V \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + V^\top \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right) = 0.$$

(ii) Detailed balance:

$G D_\pi = (\Lambda + \kappa) D_\pi = \Lambda D_\pi + \kappa D_\pi = V + V^\top + \kappa D_\pi$, since $\Lambda D_\pi = (V + V^\top) D_\pi^{-1} D_\pi = V + V^\top$. On the other hand,

$$D_\pi G^\top = D_\pi (\Lambda + \kappa)^\top = D_\pi (\Lambda^\top + \kappa) = D_\pi (D_\pi^{-1} (V + V^\top) + \kappa) = V + V^\top + D_\pi \kappa.$$

Since both κ and D_π are diagonal, they commute and hence

$$G D_\pi = V + V^\top + \kappa D_\pi = V + V^\top + D_\pi \kappa = D_\pi G^\top,$$

completing the proof. □

Computational implementation

The joint probability distribution of my model is given by

$$p(p, r_{ref}, \omega^{1:K}, q^{1:K}, dS^{1:K}, dH^{1:K} | z_{T_1:T_{\text{num of temp}}, 1:N}) \quad (\text{B.1})$$

I sample from the posterior distribution using a Metropolis-within-Gibbs sampling scheme [Gelman et al. \[2013\]](#). This scheme iteratively updates each of the model’s parameters by sampling from its full conditional distribution, which is conditioned on the current state of all other parameters, as well as the FID data. I use custom metropolis steps for the update of each conditional distribution since many of them are not in standard form.

For strictly positive parameters in the model (e.g. $q^k, r^k, dS_{kl}, dH_{hl}$), I implement a multiplicative random walk with elliptical slice proposals [McCoy et al. \[2025\]](#). This sampling strategy is equivalent to a standard random walk in the log-space of the parameter. The proposals for this walk are generated using elliptical slice sampling which is highly robust for correlated parameters [Gelman et al. \[2013\]](#), [Murray et al. \[2010\]](#). For parameters (such as w^k, p), whose support include negative values, I implemented a Metropolis-Hasting algorithm with an adaptive proposal mechanism. This method implements slice sampling [Neal \[2003\]](#) to generate the proposal step.

Finally, to ensure efficient exploration of the posterior [Gelman et al. \[2013\]](#), I identify and jointly update highly correlated parameters (for e.g. dS_{hl} and dH_{hl}).

Appendix C

Development of precision feed AI systems

Table C.1: Comprehensive summary of **visit events** and feed consumption for all animals during the time course of this study. The reported values indicate mean $\pm$ one standard deviation.

Cow ID	Feed Consumed per visit (g)	Feed Delivered per visit (g)	Feed Leftover per visit (g)	Feed Adjusted per visit (g)	Number of Images per visit
1	3920.2 $\pm$ 1077.3	3431.3 $\pm$ 1035.8	0.0 $\pm$ 0.0	-488.8 $\pm$ 370.6	28 $\pm$ 17
2	1645.8 $\pm$ 775.6	2000.3 $\pm$ 608.3	0.0 $\pm$ 0.0	354.5 $\pm$ 522.5	16 $\pm$ 12
3	1529.2 $\pm$ 455.0	1359.0 $\pm$ 354.5	0.0 $\pm$ 0.0	-170.1 $\pm$ 338.8	28 $\pm$ 24
4	1300.4 $\pm$ 357.4	1337.0 $\pm$ 312.4	0.0 $\pm$ 0.0	36.7 $\pm$ 311.3	26 $\pm$ 20
5	1588.0 $\pm$ 520.1	1433.8 $\pm$ 424.9	0.0 $\pm$ 0.0	-154.1 $\pm$ 361.9	30 $\pm$ 21
6	1933.1 $\pm$ 530.4	1714.6 $\pm$ 512.1	0.0 $\pm$ 0.0	-218.5 $\pm$ 369.9	23 $\pm$ 13
7	2116.0 $\pm$ 655.4	1919.4 $\pm$ 533.8	0.0 $\pm$ 0.0	-196.6 $\pm$ 359.7	27 $\pm$ 15
8	1873.0 $\pm$ 465.9	1770.2 $\pm$ 374.3	0.0 $\pm$ 0.0	-102.8 $\pm$ 293.3	29 $\pm$ 18
9	2028.1 $\pm$ 551.6	2035.1 $\pm$ 426.8	0.0 $\pm$ 0.0	7.0 $\pm$ 317.7	28 $\pm$ 16
10	1769.8 $\pm$ 487.6	1486.4 $\pm$ 331.9	0.0 $\pm$ 0.0	-283.3 $\pm$ 378.0	28 $\pm$ 20
11	2134.8 $\pm$ 544.3	2143.4 $\pm$ 456.9	0.0 $\pm$ 0.0	8.6 $\pm$ 398.3	28 $\pm$ 23
12	2147.0 $\pm$ 535.0	2219.7 $\pm$ 362.3	0.0 $\pm$ 0.0	72.7 $\pm$ 406.5	26 $\pm$ 19
13	1676.7 $\pm$ 472.5	1655.7 $\pm$ 353.3	0.0 $\pm$ 0.0	-21.0 $\pm$ 387.2	22 $\pm$ 17
14	1321.9 $\pm$ 489.6	1440.3 $\pm$ 308.4	0.0 $\pm$ 0.0	118.5 $\pm$ 380.3	24 $\pm$ 16
15	1956.6 $\pm$ 529.5	1979.0 $\pm$ 536.4	0.0 $\pm$ 0.0	22.4 $\pm$ 311.1	35 $\pm$ 21
16	1427.1 $\pm$ 516.4	1407.5 $\pm$ 289.9	0.0 $\pm$ 0.0	-19.6 $\pm$ 462.5	21 $\pm$ 22
17	2103.4 $\pm$ 536.4	2096.8 $\pm$ 451.6	0.0 $\pm$ 0.0	-6.6 $\pm$ 296.0	26 $\pm$ 16
18	1727.5 $\pm$ 502.9	1841.6 $\pm$ 349.2	0.0 $\pm$ 0.0	114.1 $\pm$ 393.8	36 $\pm$ 20
19	1570.8 $\pm$ 541.6	1674.6 $\pm$ 420.7	0.0 $\pm$ 0.0	103.8 $\pm$ 392.1	32 $\pm$ 20
20	1669.2 $\pm$ 480.2	1804.5 $\pm$ 338.0	0.0 $\pm$ 0.0	135.4 $\pm$ 353.6	26 $\pm$ 16
21	2466.2 $\pm$ 462.7	2406.1 $\pm$ 407.0	0.0 $\pm$ 0.0	-60.1 $\pm$ 357.8	33 $\pm$ 18
22	2286.9 $\pm$ 633.9	2383.8 $\pm$ 521.3	0.0 $\pm$ 0.0	96.9 $\pm$ 337.3	28 $\pm$ 19
23	1890.4 $\pm$ 611.9	2093.8 $\pm$ 426.2	0.0 $\pm$ 0.0	203.4 $\pm$ 453.6	24 $\pm$ 14
24	2014.2 $\pm$ 559.9	1844.3 $\pm$ 441.0	0.0 $\pm$ 0.0	-169.9 $\pm$ 349.0	28 $\pm$ 16
25	1524.6 $\pm$ 357.7	1288.6 $\pm$ 255.0	0.0 $\pm$ 0.0	-236.0 $\pm$ 302.2	31 $\pm$ 20
26	2481.7 $\pm$ 576.6	2412.7 $\pm$ 405.5	0.0 $\pm$ 0.0	-68.9 $\pm$ 394.6	26 $\pm$ 16
27	1741.0 $\pm$ 519.9	2033.2 $\pm$ 371.0	0.0 $\pm$ 0.0	292.2 $\pm$ 344.5	25 $\pm$ 16
28	1841.3 $\pm$ 585.7	1659.8 $\pm$ 422.4	0.0 $\pm$ 0.0	-181.5 $\pm$ 416.8	30 $\pm$ 24
29	1844.1 $\pm$ 566.3	2031.8 $\pm$ 338.5	0.0 $\pm$ 0.0	187.7 $\pm$ 426.8	28 $\pm$ 15
30	2156.3 $\pm$ 659.4	2288.7 $\pm$ 589.6	0.0 $\pm$ 0.0	132.5 $\pm$ 422.9	26 $\pm$ 15
31	1809.0 $\pm$ 475.2	1693.6 $\pm$ 370.9	0.0 $\pm$ 0.0	-115.4 $\pm$ 435.9	32 $\pm$ 21
32	1627.2 $\pm$ 452.4	1570.5 $\pm$ 372.4	0.0 $\pm$ 0.0	-56.7 $\pm$ 300.2	35 $\pm$ 21
33	2167.4 $\pm$ 544.4	2315.7 $\pm$ 505.8	0.0 $\pm$ 0.0	148.3 $\pm$ 353.2	24 $\pm$ 19
34	1990.9 $\pm$ 544.8	2298.3 $\pm$ 435.8	0.0 $\pm$ 0.0	307.4 $\pm$ 429.1	22 $\pm$ 17
35	2065.4 $\pm$ 556.7	2189.3 $\pm$ 456.6	0.0 $\pm$ 0.0	123.9 $\pm$ 315.2	22 $\pm$ 16
36	2746.1 $\pm$ 607.7	2703.0 $\pm$ 477.3	0.0 $\pm$ 0.0	-43.1 $\pm$ 373.1	27 $\pm$ 22
37	1428.7 $\pm$ 458.4	1674.1 $\pm$ 205.5	0.0 $\pm$ 0.0	245.4 $\pm$ 394.9	23 $\pm$ 20
38	1708.7 $\pm$ 433.1	1438.2 $\pm$ 272.6	0.0 $\pm$ 0.0	-270.5 $\pm$ 326.7	30 $\pm$ 23
39	2327.0 $\pm$ 623.1	2404.5 $\pm$ 320.0	0.0 $\pm$ 0.0	77.5 $\pm$ 508.2	20 $\pm$ 15
40	1553.4 $\pm$ 550.0	1624.3 $\pm$ 438.5	0.0 $\pm$ 0.0	70.9 $\pm$ 343.3	24 $\pm$ 20
41	1458.5 $\pm$ 438.9	1767.2 $\pm$ 407.9	0.0 $\pm$ 0.0	308.7 $\pm$ 357.1	29 $\pm$ 20
42	1888.5 $\pm$ 601.8	1849.9 $\pm$ 464.7	0.0 $\pm$ 0.0	-38.5 $\pm$ 364.0	27 $\pm$ 18
43	1696.4 $\pm$ 577.6	1708.7 $\pm$ 461.3	0.0 $\pm$ 0.0	12.3 $\pm$ 333.0	34 $\pm$ 20
44	1729.0 $\pm$ 584.3	1939.2 $\pm$ 383.3	0.0 $\pm$ 0.0	210.1 $\pm$ 391.7	19 $\pm$ 11
45	1394.5 $\pm$ 452.9	1480.0 $\pm$ 336.8	0.0 $\pm$ 0.0	85.5 $\pm$ 335.2	23 $\pm$ 18
46	2214.1 $\pm$ 542.0	2358.4 $\pm$ 362.3	0.0 $\pm$ 0.0	144.4 $\pm$ 460.8	20 $\pm$ 18
47	1790.3 $\pm$ 461.4	2147.1 $\pm$ 250.2	0.0 $\pm$ 0.0	356.8 $\pm$ 371.8	32 $\pm$ 23
48	1922.4 $\pm$ 627.1	2277.4 $\pm$ 329.5	0.0 $\pm$ 0.0	355.0 $\pm$ 508.8	38 $\pm$ 20
49	1641.2 $\pm$ 467.9	1545.2 $\pm$ 346.0	0.0 $\pm$ 0.0	-96.1 $\pm$ 309.1	29 $\pm$ 17
50	1336.8 $\pm$ 329.9	1353.2 $\pm$ 177.1	0.0 $\pm$ 0.0	16.4 $\pm$ 319.8	35 $\pm$ 23
51	1608.5 $\pm$ 510.7	1552.8 $\pm$ 334.3	0.0 $\pm$ 0.0	-55.7 $\pm$ 400.4	29 $\pm$ 17
52	1811.0 $\pm$ 425.7	1755.2 $\pm$ 229.1	0.0 $\pm$ 0.0	-55.7 $\pm$ 343.2	31 $\pm$ 20
53	1310.1 $\pm$ 499.7	1441.2 $\pm$ 235.2	0.0 $\pm$ 0.0	131.0 $\pm$ 440.8	30 $\pm$ 19
54	2890.3 $\pm$ 626.3	2725.5 $\pm$ 498.6	0.0 $\pm$ 0.0	-164.9 $\pm$ 365.5	20 $\pm$ 15

Table C.2: Comprehensive summary of **daily events** and feed consumption for all animals during the time course of this study. The reported values indicate mean $\pm$ one standard deviation.

Cow ID	Feed Consumed per day (g)	Feed Delivered per day (g)	Feed Leftover per day (g)	Feed Adjusted per day (g)	Number of Images per day
1	9990.1 $\pm$ 2031.4	8744.4 $\pm$ 1666.3	2110.2 $\pm$ 1573.1	-1245.7 $\pm$ 878.6	71 $\pm$ 46
2	3981.7 $\pm$ 850.6	4839.3 $\pm$ 711.5	2909.3 $\pm$ 1697.1	857.6 $\pm$ 836.6	38 $\pm$ 27
3	5722.0 $\pm$ 1031.3	5085.4 $\pm$ 869.2	2497.6 $\pm$ 1605.7	-636.6 $\pm$ 706.0	106 $\pm$ 75
4	4865.9 $\pm$ 894.4	5003.1 $\pm$ 687.2	2915.0 $\pm$ 1568.6	137.2 $\pm$ 591.0	97 $\pm$ 58
5	5634.7 $\pm$ 1096.0	5087.8 $\pm$ 810.5	2301.1 $\pm$ 1202.5	-546.9 $\pm$ 680.5	107 $\pm$ 61
6	5986.3 $\pm$ 1318.6	5309.6 $\pm$ 938.7	2032.8 $\pm$ 1165.8	-676.7 $\pm$ 954.9	70 $\pm$ 37
7	6484.4 $\pm$ 1607.2	5881.9 $\pm$ 1310.0	1776.7 $\pm$ 963.3	-602.5 $\pm$ 732.7	83 $\pm$ 46
8	6827.4 $\pm$ 885.1	6452.8 $\pm$ 907.7	2247.4 $\pm$ 997.7	-374.6 $\pm$ 478.3	104 $\pm$ 51
9	5626.3 $\pm$ 1205.3	5645.7 $\pm$ 998.0	2148.4 $\pm$ 1454.6	19.4 $\pm$ 606.0	77 $\pm$ 46
10	7021.9 $\pm$ 1513.8	5897.8 $\pm$ 891.4	2308.7 $\pm$ 1082.4	-1124.1 $\pm$ 1176.1	110 $\pm$ 61
11	5715.8 $\pm$ 1384.4	5738.7 $\pm$ 1076.9	2136.8 $\pm$ 1368.4	22.9 $\pm$ 715.9	75 $\pm$ 48
12	6025.4 $\pm$ 1146.1	6229.5 $\pm$ 1051.5	2451.9 $\pm$ 1082.0	204.1 $\pm$ 598.5	74 $\pm$ 42
13	6274.1 $\pm$ 1037.4	6195.6 $\pm$ 795.2	3040.0 $\pm$ 1544.9	-78.4 $\pm$ 824.3	84 $\pm$ 47
14	5244.9 $\pm$ 1310.3	5714.8 $\pm$ 717.1	3640.7 $\pm$ 2194.5	470.0 $\pm$ 1071.0	94 $\pm$ 47
15	5491.1 $\pm$ 1231.4	5554.1 $\pm$ 989.6	2167.6 $\pm$ 1301.9	63.0 $\pm$ 581.0	97 $\pm$ 61
16	6261.0 $\pm$ 1377.0	6174.9 $\pm$ 628.9	3629.9 $\pm$ 1708.3	-86.1 $\pm$ 1109.8	91 $\pm$ 64
17	6513.8 $\pm$ 1031.0	6493.5 $\pm$ 1105.7	2422.8 $\pm$ 1465.9	-20.3 $\pm$ 581.9	79 $\pm$ 43
18	5182.5 $\pm$ 1142.2	5524.9 $\pm$ 996.8	2971.1 $\pm$ 1911.7	342.4 $\pm$ 678.7	109 $\pm$ 57
19	5168.5 $\pm$ 1025.0	5510.0 $\pm$ 829.4	2999.6 $\pm$ 1991.7	341.5 $\pm$ 842.4	104 $\pm$ 57
20	4792.2 $\pm$ 862.1	5180.8 $\pm$ 804.2	2836.4 $\pm$ 1445.7	388.6 $\pm$ 630.1	76 $\pm$ 48
21	6921.2 $\pm$ 1315.6	6752.6 $\pm$ 959.7	2118.2 $\pm$ 1111.6	-168.6 $\pm$ 622.8	93 $\pm$ 48
22	6270.6 $\pm$ 1454.0	6536.3 $\pm$ 1324.3	2794.5 $\pm$ 1975.1	265.7 $\pm$ 669.2	76 $\pm$ 53
23	5610.2 $\pm$ 1257.1	6213.7 $\pm$ 782.9	2819.1 $\pm$ 1624.5	603.5 $\pm$ 975.3	70 $\pm$ 37
24	7407.1 $\pm$ 1154.7	6782.2 $\pm$ 797.0	2534.4 $\pm$ 1569.2	-624.9 $\pm$ 666.1	103 $\pm$ 51
25	6540.9 $\pm$ 1062.1	5528.5 $\pm$ 645.0	2453.2 $\pm$ 1222.8	-1012.4 $\pm$ 781.0	134 $\pm$ 66
26	6484.3 $\pm$ 1447.5	6304.2 $\pm$ 1094.5	2001.4 $\pm$ 1176.7	-180.1 $\pm$ 737.4	67 $\pm$ 39
27	3007.2 $\pm$ 1129.8	3511.9 $\pm$ 1108.6	1639.7 $\pm$ 552.8	504.7 $\pm$ 470.9	43 $\pm$ 27
28	6830.8 $\pm$ 1485.3	6157.4 $\pm$ 1010.0	2356.6 $\pm$ 967.5	-673.4 $\pm$ 1013.6	110 $\pm$ 63
29	4937.3 $\pm$ 1120.6	5440.0 $\pm$ 1263.6	2762.6 $\pm$ 1892.2	502.7 $\pm$ 1008.5	75 $\pm$ 41
30	5773.2 $\pm$ 1203.9	6127.9 $\pm$ 1171.1	2509.1 $\pm$ 1408.4	354.7 $\pm$ 740.3	69 $\pm$ 33
31	5952.2 $\pm$ 970.7	5572.5 $\pm$ 803.7	2310.2 $\pm$ 1377.9	-379.7 $\pm$ 851.6	107 $\pm$ 53
32	5301.6 $\pm$ 1002.1	5116.8 $\pm$ 912.7	2488.2 $\pm$ 1542.0	-184.8 $\pm$ 448.2	113 $\pm$ 61
33	5313.7 $\pm$ 1390.7	5677.2 $\pm$ 1225.9	2358.0 $\pm$ 1295.5	363.5 $\pm$ 645.3	59 $\pm$ 46
34	3438.8 $\pm$ 1362.8	3969.8 $\pm$ 1227.4	1611.2 $\pm$ 662.6	531.0 $\pm$ 588.8	38 $\pm$ 32
35	6262.8 $\pm$ 1268.6	6638.6 $\pm$ 1099.5	2820.4 $\pm$ 1673.7	375.7 $\pm$ 505.0	65 $\pm$ 36
36	6378.1 $\pm$ 1049.1	6278.0 $\pm$ 996.2	2091.3 $\pm$ 1295.6	-100.1 $\pm$ 513.8	64 $\pm$ 47
37	3502.7 $\pm$ 1128.3	4104.3 $\pm$ 964.9	2495.9 $\pm$ 1537.9	601.6 $\pm$ 767.3	57 $\pm$ 37
38	7275.9 $\pm$ 1426.8	6124.1 $\pm$ 861.2	2201.6 $\pm$ 1258.1	-1151.9 $\pm$ 910.7	127 $\pm$ 76
39	3807.8 $\pm$ 1148.3	3934.7 $\pm$ 1138.5	1242.4 $\pm$ 533.3	126.9 $\pm$ 476.8	33 $\pm$ 25
40	5161.3 $\pm$ 1062.1	5396.9 $\pm$ 770.4	2965.5 $\pm$ 1788.2	235.6 $\pm$ 633.0	80 $\pm$ 49
41	4328.5 $\pm$ 871.6	5244.7 $\pm$ 821.2	3371.9 $\pm$ 1817.8	916.2 $\pm$ 751.1	85 $\pm$ 54
42	6640.1 $\pm$ 1335.3	6504.6 $\pm$ 944.2	2702.8 $\pm$ 1996.5	-135.5 $\pm$ 794.9	93 $\pm$ 53
43	5362.8 $\pm$ 1195.6	5401.6 $\pm$ 1067.4	2504.8 $\pm$ 1785.1	38.8 $\pm$ 616.1	109 $\pm$ 63
44	4406.2 $\pm$ 865.2	4941.7 $\pm$ 712.9	2724.6 $\pm$ 1650.3	535.5 $\pm$ 763.5	47 $\pm$ 29
45	4678.3 $\pm$ 989.2	4965.2 $\pm$ 779.2	2961.6 $\pm$ 1798.5	286.9 $\pm$ 634.7	77 $\pm$ 52
46	4428.1 $\pm$ 1157.5	4716.9 $\pm$ 857.1	1902.2 $\pm$ 965.9	288.8 $\pm$ 608.9	41 $\pm$ 33
47	3696.1 $\pm$ 871.0	4432.6 $\pm$ 921.6	2400.9 $\pm$ 1044.3	736.6 $\pm$ 546.8	66 $\pm$ 38
48	3782.9 $\pm$ 1026.8	4481.4 $\pm$ 790.5	2289.9 $\pm$ 1191.4	698.5 $\pm$ 870.7	74 $\pm$ 36
49	5772.6 $\pm$ 1403.9	5434.8 $\pm$ 1165.4	2354.6 $\pm$ 1498.8	-337.9 $\pm$ 698.1	103 $\pm$ 61
50	6209.4 $\pm$ 819.3	6285.8 $\pm$ 595.8	3654.3 $\pm$ 1944.7	76.3 $\pm$ 657.0	164 $\pm$ 89
51	5707.7 $\pm$ 1300.7	5510.1 $\pm$ 990.3	2738.1 $\pm$ 1338.8	-197.7 $\pm$ 910.2	102 $\pm$ 55
52	5491.4 $\pm$ 1355.2	5322.4 $\pm$ 1124.0	2385.0 $\pm$ 1572.2	-169.0 $\pm$ 624.1	95 $\pm$ 56
53	5409.5 $\pm$ 1354.1	5950.6 $\pm$ 670.4	3888.8 $\pm$ 2265.4	541.1 $\pm$ 1151.3	124 $\pm$ 71
54	6713.1 $\pm$ 1332.4	6330.2 $\pm$ 1205.4	1578.3 $\pm$ 1050.3	-382.9 $\pm$ 434.2	47 $\pm$ 40

Vita

Kelechi Akwataghibe was born and raised in Aba, Nigeria, where he completed his secondary education at Dority International Secondary School. In 2015, he moved to Jacksonville, Florida, to pursue a Bachelor of Science degree in Mathematics with a double major in Computer Science at Jacksonville University. During his undergraduate studies he discovered a passion for software engineering, and published two mobile applications to the iOS App Store. Motivated to deepen his expertise in mathematical modeling and its practical applications, Kelechi enrolled in the Doctor of Philosophy program in Mathematics (Applied Mathematics) at the University of Tennessee, Knoxville. Under the mentorship of Dr. Ioannis Sgouralis, he has explored the integration of Bayesian statistics, machine learning, and computational modeling to advance both NMR spectroscopy analysis and precision-agriculture robotics. Upon conferral of his degree in August 2025, Kelechi aims to a new role as a Data Scientist where he plans to continue developing data-driven solutions at the intersection of mathematics, computing, and the life sciences. He is grateful for the encouragement of his family, friends, mentors, and colleagues throughout this journey.