THE UNIVERSITY OF CHICAGO

COMPUTATIONAL METHODS FOR MATERIALS SCIENCE WITH APPLICATIONS IN NEXT-GENERATION COMPUTING ARCHITECTURES

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

COMMITTEE ON COMPUTATIONAL AND APPLIED MATHEMATICS

BY
ABIGAIL N. POTESHMAN

CHICAGO, ILLINOIS
JUNE 2026

© 2026 by Abigail N. Poteshman

ORCID iD: https://orcid.org/0000-0002-4873-4826

In memory of Sharon R. Poteshman

We can only see a short distance ahead, but we can see plenty there that needs to be done.

— A. M. Turing, "Computing Machinery and Intelligence"

TABLE OF CONTENTS

LIST OF FIGURES viii

LIST OF TABLES x

ACKNOWLEDGMENTS xi

ABSTRACT xii

1 INTRODUCTION 1
1.1 Motivation 1
1.1.1 Computing as a materials science problem 1
1.1.2 Energy and efficiency challenges in classical computing 2
1.1.3 Beyond classical computing: Next-generation computing architectures 3
1.1.4 Towards hybrid computing paradigms 5
1.2 Computational methods in materials discovery and characterization 6
1.2.1 High-throughput computational materials discovery 7
1.2.2 From discovery to deployment: Integrating computation and experiment 9
1.2.3 Complementarity with machine learning- and AI-based materials discovery 11
1.3 Thesis scope and summary of contents 13

2 HIGH-THROUGHPUT COMPUTATION OF ELECTRIC POLARIZATION IN SOLIDS VIA BERRY FLUX DIAGONALIZATION 16
2.1 Introduction 17
2.2 Theoretical formalism 20
2.2.1 Standard approach for computing polarization using DFT 21
2.2.2 Berry flux diagonalization 23
2.3 Heuristic for sufficient wavefunction overlap and gauge consistency 27
2.3.1 Artificial $BaTiO_3$ test system 28
2.3.2 Translations operations to minimize maximal atomic displacements . 29
2.3.3 Interpolated structures 31
2.4 Validation of the Berry flux diagonalization approach for computing effective polarization 33
2.5 High-throughput effective polarization calculations with Berry flux diagonalization 38
2.6 Discussion 40
2.7 Methods 43
2.7.1 Input cell parameters 43
2.7.2 Computational Details for DFT calculations 44
2.7.3 Computational details for implementation of Berry flux digaonalization 45
2.7.4 Manually-traced branches for validation with QUANTUM ESPRESSO 45

2.7.5 Convergence studies with respect to number of k-points for QUANTUM ESPRESSO . . . 48
2.7.6 DFT+U calculations for $BiFeO_3$ for QUANTUM ESPRESSO . . . 48
2.8 Additional Supporting Results . . . 49
2.8.1 Parallel results for exploring translation and interpolation heuristics in QUANTUM ESPRESSO . . . 49
2.8.2 High-throughput calculation results in VASP . . . 50

3 TRANS-DIMENSIONAL HAMILTONIAN MODEL SELECTION AND PARAMETER ESTIMATION FROM SPARSE, NOISY DATA . . . 62
3.1 Introduction . . . 63
3.2 Inverse Problem . . . 65
3.2.1 General formulation . . . 65
3.2.2 Hybridized MCMC-based Hamiltonian model selection and parameter estimation . . . 67
3.3 Algorithms . . . 69
3.3.1 Random walk Metropolis-Hastings (RWMH) for continuous and discrete domains . . . 70
3.3.2 Reverse Jump Markov Chain Monte Carlo (RJMCMC) . . . 71
3.3.3 Parallel Tempering (PT) . . . 72
3.3.4 Hybrid algorithm . . . 74
3.4 Example application: Recovering hyperfine coupling parameters of a nuclear spin bath of defects in semiconductors from coherence data . . . 75
3.4.1 Hamiltonian and forward models . . . 76
3.4.2 Likelihood model and priors . . . 77
3.4.3 Hyperparameter tuning and dependence . . . 79
3.5 Results . . . 82
3.5.1 Results from simulations . . . 82
3.5.2 Metrics for performance and recovery rates . . . 83
3.5.3 Validation on experimental data . . . 96
3.6 Discussion . . . 110

4 HIGH-THROUGHPUT CHARACTERIZATION OF SPIN-DEFECTS IN SEMICONDUCTORS FOR SCALE-UP OF QUANTUM SENSING . . . 116
4.1 Introduction . . . 117
4.2 Background . . . 121
4.2.1 Hamiltonian description of nuclear spin bath . . . 121
4.2.2 Dynamical decoupling experiments . . . 122
4.2.3 Relative decoherence parameter . . . 124
4.3 Detection limits imposed by noisy, sparse data . . . 126
4.3.1 Detection limits imposed by sparsity . . . 126
4.3.2 Detection limits imposed by shot noise . . . 128
4.3.3 Determining limits on detectability from sparse, noisy experimental data 130

4.4 Application of a hybrid MCMC algorithm for experimental design and nuclear spin bath recovery . . . 131
4.4.1 Guiding experimental design . . . 133
4.4.2 Application to experimental data . . . 136
4.5 Discussion . . . 138

5 ONGOING AND FUTURE WORK, OUTLOOK, AND CONCLUSION . . . 147
5.1 Ongoing and future work for Berry flux digaonalization . . . 147
5.1.1 Outlook for Berry flux diagonalization code . . . 147
5.1.2 Extensions to larger supercells and ML frameworks . . . 148
5.1.3 Integrating Berry flux diagonalization into ML workflows for polarization and dielectric response . . . 150
5.2 Ongoing and future work for spin-bath characterization of spin defects from coherence spectra . . . 152
5.2.1 Extending framework from NV centers in diamond to axial divacancies in silicon carbide . . . 152
5.2.2 Adaptive and online experimental design . . . 153
5.2.3 Adaptive pulse sequence optimization . . . 157
5.3 Concluding remarks . . . 158

LIST OF FIGURES

2.1 Schematic of automated workflow to compute effective polarization 20
2.2 Schematic of a Berry phase "plaquette" 24
2.3 Toy BTO system for minimum singular values and maximum eigenvalues vs atomic displacement 28
2.4 Numerical stability of effective polarization for different translations of standard ferroelectrics 30
2.5 Numerical stability of effective polarization for different numbers of interpolations of standard ferroelectrics 34
2.6 High-throughput results comparing automated Berry flux diagonalization with the standard interpolation-based approach for 176 candidate ferroelectrics . . . 41
2.7 Polarization branches for $BaTiO_3$, $PbTiO_3$, $KNbO_3$, and $LiNbO_3$ computed using QUANTUM ESPRESSO 46
2.8 Polarization branches for Bi_2MoO_6, $CrO_3^{\dagger}$, and $BiFeO_3$ computed using QUANTUM ESPRESSO 47
2.9 Convergence of effective polarization computed using Berry flux diagonalization with respect to the number of k-points 49
2.10 Toy BTO system with numerics computed using QUANTUM ESPRESSO 51
2.11 Numerical stability of effective polarization for different translations of standard ferroelectrics computed with QUANTUM ESPRESSO 60
2.12 Numerical stability of effective polarization for different numbers of interpolations of standard ferroelectrics computed with QUANTUM ESPRESSO 61

3.1 Schematics of hybrid MCMC workflow 66
3.2 Schematic of sampling algorithms 69
3.3 Error trajectories of 5 ensembles as a function of walker steps for varying R_{spin} . 79
3.4 Error trajectories of 5 ensembles as a function of walker steps for varying σ^2 . . 80
3.5 Recovery in sparse data limit 85
3.6 Recovery in noisy data limit 89
3.7 Recovery method on data with experimental parameters in machine learning regime 91
3.8 Recovery method accuracy versus hyperfine coupling perturbations 93
3.9 Validation of hybrid MCMC approach on experimental data 95

4.1 Schematic of inverse problem 120
4.2 Comparison of coherence signal reconstruction with and without the relative decoherence parameter λ 125
4.3 Hyperfine coupling frequency detection limits from frequency sampling limits and simulations 127
4.4 Accuracy of hybrid MCMC algorithm for recovering hyperfine couplings from simulated data. 142
4.5 Interdependence of recovery rates on experimental settings. 143
4.6 Reconstructions of experimental coherence signal data and nuclear spin positions for 10 different nitrogen-vacancy (NV) centers 144

4.7 Posterior distributions of the number and hyperfine couplings of spins from the hybrid algorithm applied to 10 NV centers in diamond 145
4.8 Minimum detectable hyperfine coupling as a function of time averaging 146

LIST OF TABLES

2.1 Validation of Berry flux diagonalization against the standard interpolation-based Berry phase approach 35
2.2 Number of interpolations used to compute effective polarization values reported in Table 2.1 36
2.3 Polar input cell information 43
2.4 Nonpolar input cell information 44
2.5 Dependence of effective polarization values for $BiFeO_3$ on effective U values. . . 48
2.6 Full results of high-throughput calculation of effective polarization using Berry flux diagonalization 59

3.1 Validation of Wasserstein weighting parameter ζ on experimental data 98
3.2 Comparison of experimentally measured and recovered hyperfine interaction values 105
3.3 Comparison of experimentally measured and recovered nuclear spin positions . . 107

4.1 Average uncertainty per data point 137

ACKNOWLEDGMENTS

First and foremost, I thank my advisor, Giulia Galli for her continuous support, encouragement, and scientific guidance. I am also deeply grateful to have been advised by F. Joseph Heremans at Argonne National Laboratory and Jeffrey B. Neaton at Lawrence Berkeley National Laboratory during the summers. During my time in grad school, I am grateful to have collaborated with Francesco Ricci, Mykyta Onizhuk, and Christopher Egerstrom. I thank Yongjin Shin, Vrindaa Somjit, Jonah Nagura, Daniel Sanz-Alonso, Benjamin Pingault, and Yuxin Wang for many useful discussions that directly guided and impacted the direction of the research presented in this thesis. I acknowledge support during my PhD from the DOE CSGF under Award Number DE-SC0022158, AFOSR Grant No. FA9550-22-1-0370, the University of Chicago Research Computing Center, NERSC, IPAM, IHPCSS, Q-NEXT, and the DSI for funding, computing resources, and conference travel support. Most of all, I salute my CCAM, CSGF, and Galli Group friends and colleagues! No words could even begin to capture how fortunate I feel to have met each and every one of you.

ABSTRACT

Computing performance is constrained by the properties of the materials used to store, transmit, and manipulate information. For decades, improvements in silicon semiconductor technologies enabled rapid growth in computing capability through device scaling. However, as device dimensions approach atomic length scales, quantum tunneling, leakage currents, variability, and thermal dissipation increasingly limit computing performance and energy efficiency. These physical limits motivate the development of alternative computing paradigms, such as low-power multiferroic spintronics and quantum computing, whose realizations depend on discovering and engineering materials with specific functional properties. Computational materials science using classical computing is a powerful tool for identifying and characterizing such materials over a combinatorially large space, but efficient computational screening and characterization methods must be accurate, scalable, and effective with limited or noisy data. This thesis develops physics-informed computational algorithms that enable high-throughput computational screening and experiment-coupled characterization of materials for next-generation computing hardware. First, I introduce an automated and scalable implementation of Berry flux diagonalization to compute spontaneous polarization in insulating materials. This method enables robust high-throughput computational screening of candidate ferroelectric materials for applications in multiferroic spintronics. Second, I develop and validate a data-efficient numerical inference framework to reconstruct nuclear spin environments surrounding spin-defects from sparse, noisy data, enabling rapid experimental characterization of spin-defects for use as qubits. Finally, I demonstrate how this framework can guide high-throughput experimental design and establish sensing limits for spin-defect characterization. Together, these contributions provide a set of computational tools that integrate simulation, inference, and experiment to accelerate materials discovery for emerging computing architectures.

CHAPTER 1

INTRODUCTION

1.1 Motivation

1.1.1 Computing as a materials science problem

Modern computing is often discussed in terms of algorithms, software, and architecture; yet, in practice, every computation is executed by a physical system. Bits are stored in charge distributions, spins, or polarization states; logic operations are implemented through material-dependent switching mechanisms; memory, data transfer, and communication rely on transport through materials with finite resistance, capacitance, and dissipation. As a result, computing performance is not determined solely by algorithms and abstract models of computation, but by the physical properties of the materials from which hardware is built.

For several decades, improvements in silicon-based semiconductor devices and fabrication methods advanced computing power. Device miniaturization reduced switching energy for logic operations, increased transistor density, and improved computing performance across a range of applications. However, as silicon-based semiconductor chips approach atomic length scales, intrinsic physical limits increasingly dominate behavior. Quantum tunneling, leakage currents, interfacial scattering, and variability arising from atomic-scale disorder degrade device performance. Thermal dissipation and interconnect losses further limit computation speed and energy efficiency. These phenomena arise from fundamental material parameters and cannot be eliminated through circuit design and fabrication engineering alone.

As a result, progress in computational science depends not only on improved architectures and algorithms, but also on identifying and engineering materials with new functional properties for new kinds of computing hardware. New computing mechanisms, whether based on quantum coherence, collective spin order, or spontaneous polarization, require materials that exhibit specific stability, controllability, and coupling characteristics. Ergo, the

development of next-generation computing hardware is inseparable from materials discovery, characterization, and design.

1.1.2 Energy and efficiency challenges in classical computing

Conventional complementary metal-oxide-semiconductor (CMOS) technologies have approached their limits in energy and economic efficiency. Historically, improvements in classical computing performance followed empirical scaling trends. Moore's law, proposed in 1965 by Gordon E. Moore, observed that the number of transistors on an integrated circuit doubles roughly every 18-24 months, leading to an exponential growth in computing capability and a drop in cost per logic operation [1]. Dennard scaling, proposed in 1974, stated that the power density of transistors remains constant as they shrink, so transistor density can increase without a proportional rise in energy cost [2, 3]. Together, these trends reflected simultaneous gains in computing performance, cost efficiency, and energy efficiency in semiconductor technologies from the 1960s to the mid-2000s.

In practice, these heuristic scaling laws began to fail in the mid-2000s as device dimensions began to approach fundamental physical limits. Transistors can become only so small before the gates begin "running out of atoms" and current leakage due to quantum tunneling of electrons leads to a notable reduction in overall chip power [4]. Dennard scaling began to break when further voltage reduction became impractical due to noise margins and reliability constraints, leading to increased power consumption and making energy efficiency a bottleneck in the design of high-performance computing (HPC) systems [5]. Recent advances in computing performance have largely relied on parallelism and architectural specialization rather than faster individual devices [6].

Energy consumption is now a central constraint in modern computing infrastructure. High-performance computing systems require substantial cooling and power delivery resources, with significant environmental and economic cost [7]. Data centers account for

a growing fraction of global electricity usage, and large-scale AI and machine learning workloads are especially energy intensive due to their computational scale and data movement requirements [8]. At the same time, increased computational capability is essential for addressing major scientific and societal challenges. These combined energy, economic, and physical constraints motivate exploration of computing approaches that move beyond conventional charge-based CMOS hardware.

1.1.3 Beyond classical computing: Next-generation computing architectures

Next-generation computing architectures aim to address the physical and energy limitations of classical CMOS systems by using fundamentally different mechanisms for information representation and processing. Rather than relying exclusively on electronic charge and von Neumann data movement models, these approaches exploit alternative physical degrees of freedom and computational primitives.

Quantum computing represents one such paradigm shift by exploiting the fundamentally different nature of quantum from classical information. Unlike classical bits, quantum bits (qubits) exist in coherent superpositions of states and can become entangled, enabling correlations that have no classical analogue. Quantum algorithms leverage interference and entanglement to explore exponentially large Hilbert spaces using resources that scale polynomially with system size for certain problems. Hence, quantum computing can provide algorithmic advantages for specific computational tasks, including integer factorization, quantum simulation, computational chemistry, and materials modeling [9–12]. In particular, quantum computers offer a natural framework for simulating quantum many-body systems, whose classical description requires computational resources that grow exponentially with particle number. In this sense, quantum computing does not merely accelerate existing algorithms; it provides access to an entirely different class of problems that are intractable on classical architectures.

However, realizing these advantages requires hardware that differs radically from conventional electronics. Quantum information is fragile: coherent quantum states must be preserved in the face of environmental noise, thermal fluctuations, and material defects. Qubits must simultaneously be isolated from their environment to maintain coherence and strongly controllable for initialization, manipulation, and readout. These competing requirements impose stringent materials constraints, including long coherence times, precise control of spin or superconducting degrees of freedom, and reproducible fabrication at scale [13]. As a result, progress in quantum computing is deeply tied to materials discovery and defect engineering, where understanding microscopic noise sources and their interactions with quantum states is essential for overcoming decoherence.

In parallel, multiferroic spintronics offers a complementary route toward overcoming the energy-efficiency limitations of classical computing. Conventional CMOS logic dissipates energy primarily through resistive charge transport and repeated charging and discharging of capacitive elements. In contrast, spintronic devices encode information in electron spin or magnetic order parameters, allowing information storage and manipulation without sustained charge currents [14]. Multiferroic materials, that is, systems in which ferroelectric and magnetic order coexist and are coupled, enable electrical control of magnetism, providing a pathway to switch magnetic states using electric fields rather than current-driven magnetic torques [15]. Because electric-field switching can occur with minimal charge flow, multiferroic spintronic devices have the potential for orders-of-magnitude reductions in switching energy relative to CMOS technologies [16, 17].

Beyond simple energy savings, multiferroic spintronics addresses structural inefficiencies inherent to von Neumann architectures. Magnetic states are intrinsically nonvolatile, meaning information can be retained without continuous power, enabling memory and logic to coexist within the same physical device. This blurs the traditional separation between computation and memory, reducing costly data movement and enabling in-memory or neuromor-

phic computing paradigms. Such architectures directly target one of the central challenges facing modern HPC systems: the disproportionate energy cost of moving data compared to processing it. By co-locating storage and computation and exploiting collective material degrees of freedom, spintronic systems offer a materials-based pathway toward low-power, highly parallel information processing.

Across the emerging architectures of quantum computing and multiferroic spintronics, materials science plays a central and enabling role. Quantum devices require materials that suppress decoherence while enabling precise quantum control, whereas spintronic technologies demand materials with robust, switchable ferroic order parameters and strong cross-coupling between electric, magnetic, and structural degrees of freedom. In both cases, performance is ultimately limited not by abstract algorithms but by microscopic material properties: defects, interfaces, lattice distortions, and electronic structure. Practical realization for widespread deployment therefore depends on identifying, characterizing, and optimizing materials with highly specific functional properties and compatibility with scalable fabrication. Developing computational and experimental tools to discover and understand such materials is thus a key prerequisite for testing and deploying next-generation computing technologies at scale—and forms the central motivation for this thesis.

1.1.4 Towards hybrid computing paradigms

Computational materials science plays a central role in the development of next-generation computing hardware by enabling the identification, characterization, and optimization of candidate materials before experimental realization. At present, these computational methods are executed almost entirely on classical computing platforms, including high-performance computing systems. Classical computing remains the entry point for discovering and validating the materials needed to build alternative computing architectures. Looking forward, once emerging computing architectures reach sufficient maturity, computing ecosystems are likely

to become heterogeneous [18–20]. This trend is already evident in modern high-performance computing and machine learning systems, where graphics processing units (GPUs), originally designed for parallel graphics and vector processing, enable highly efficient large-scale linear algebra and model training [21]. Classical, quantum, and spintronic platforms may be used in combination, with subproblems assigned to the architecture best suited to their algorithmic structure, accuracy requirements, or energy cost. Such hybrid computing workflows could substantially improve efficiency for simulation, optimization, and inverse-design tasks that arise in materials discovery and device engineering. In turn, these improved workflows could accelerate the discovery of still better materials for computing hardware and other technologies [22]. Establishing this positive feedback cycle requires robust and efficient classical computational methods as a foundation, enabling the initial discovery and characterization steps that make hybrid computing paradigms possible.

1.2 Computational methods in materials discovery and characterization

Computational methods play multiple complementary roles in modern materials science, spanning discovery, interpretation, and prediction. At the most basic level, simulation provides a framework for hypothesis generation by enabling controlled exploration of how microscopic structure and composition determine macroscopic properties. Electronic structure calculations, atomistic simulations, and continuum models allow researchers to test candidate materials and mechanisms in silico before committing to experimental synthesis. This capability is especially valuable when exploring large compositional or structural design spaces, where exhaustive experimental search would be impractical. In particular, high-throughput computational screening frameworks make it possible to evaluate large material spaces in a systematic and automated way, trading per-material depth for statistical coverage and speed [23]. Once candidate materials with target properties have been identified, computational

workflows are key for practical scale-up and deployment of novel computing technologies. When coupled with experimental workflows, computational modeling can reduce synthesis overhead, prioritize promising candidates, and guide measurement strategies [24]. In this thesis, I present computational methods for applications both in high-throughput computational screening of materials for applications in multiferroic spintronics and in experiment-coupled workflows for experimental characterization of candidate qubits based on their coherence properties.

1.2.1 High-throughput computational materials discovery

High-throughput computational materials discovery has emerged as a powerful strategy for exploring the combinatorially large space of possible compounds using first-principles simulations. By combining electronic-structure methods with automated structure generation, workflow management, and database infrastructure, it becomes possible to evaluate materials properties across thousands or even millions of candidate systems in a systematic and reproducible manner. These approaches have enabled rapid screening for target properties such as band gaps, magnetic ordering, defect energetics, thermodynamic stability, and ferroelectric polarization, transforming computational materials science from a tool for studying individual compounds into a platform for large-scale materials exploration [25–28].

The central advantage of high-throughput methodologies is their ability to operate at scale. Rather than optimizing calculations for a single material, workflows are designed to apply consistent computational protocols across broad chemical and structural spaces, enabling statistical discovery and efficient elimination of unsuitable candidates early in the design process. This paradigm is particularly well suited to ferroelectric materials discovery, where identifying promising candidates requires searching across vast families of compounds with diverse chemistries, crystal symmetries, and structural distortions. Because only a small fraction of materials exhibit switchable spontaneous polarization, scalable screening methods

are essential for narrowing the search space prior to detailed experimental or theoretical investigation.

Achieving such scale, however, requires careful algorithmic design. High-throughput workflows must balance computational cost, numerical robustness, and physical accuracy. Calculations are typically performed at a uniform level of theory to ensure consistency and reproducibility, emphasizing reliable comparative trends rather than maximal accuracy for any individual material. Consequently, the efficiency of each property evaluation becomes a critical factor: even modest reductions in computational cost per material translate into substantial gains when applied across large datasets. Developing algorithms that preserve rigorous physical definitions while reducing computational overhead is therefore a central challenge in high-throughput materials screening.

One aspect of this thesis focuses specifically on addressing this challenge for first-principles polarization calculations, with the goal of enabling automated screening for candidate ferroelectric materials for application in multiferroic spintronics. Computing spontaneous polarization within the modern theory of polarization is inherently more complex than evaluating many other materials properties, as it traditionally requires constructing interpolation pathways between crystal structures and carefully tracking polarization branches [29]. Previous efforts toward automation have largely concentrated on systematizing these interpolation-based workflows or automating symmetry analysis to identify appropriate centrosymmetric reference structures [30, 31]. While these advances enabled the first generation of high-throughput polarization studies, the underlying methodology remains computationally inefficient and prone to numerical inaccuracies.

The work presented in this thesis builds on these foundations by developing an alternative framework to compute spontaneous polarization designed explicitly for high-throughput application. The emphasis is not only on automation, but on reducing the computational cost per material while maintaining rigorous numerical definitions and accuracy guarantees.

By reformulating polarization evaluation to avoid expensive interpolation procedures, the methods developed in this thesis enable robust polarization calculations suitable for screening large materials spaces, thereby expanding the practical scale at which first-principles ferroelectric screening can be performed.

1.2.2 From discovery to deployment: Integrating computation and experiment

Rather than viewing theory and experiment as sequential stages—where computation predicts materials and experiments later validate them—many emerging research areas benefit from iterative feedback between measurement and modeling [32]. Computational methods can interpret incomplete or indirect experimental observations, guide measurement design, and rapidly test competing physical hypotheses, while experiments provide the real-world constraints necessary to refine and validate theoretical models [24]. In these settings, computation serves not only as a predictive tool but as an essential component of experimental characterization itself.

Solid-state spin defects provide a compelling example of this paradigm. Point defects in wide-bandgap materials, such as nitrogen–vacancy (NV) centers in diamond, are leading candidates for applications in quantum sensing, quantum communication, and distributed quantum information processing [33]. Their suitability arises from a combination of properties: optical initialization and readout, long spin coherence times at room temperature, and strong coupling to nearby nuclear spins that can serve as auxiliary quantum registers [33–35]. However, the performance of individual spin defects is not determined solely by the defect's intrinsic electronic structure. Instead, coherence and control are strongly influenced by the surrounding spin environment—a complex bath of nuclear and electronic spins arising from isotopes, impurities, and lattice defects [36–38]. Because this environment varies from defect to defect, understanding and controlling the local spin bath becomes a central challenge for

scalable quantum technologies.

Experimentally, the spin bath can be probed indirectly through measurements such as dynamical decoupling and correlation spectroscopy, which encode information about environmental couplings in time-domain signals. Interpreting these measurements poses a challenging inverse problem: many distinct microscopic spin configurations can produce similar experimental signatures. Direct structural characterization at the atomic scale is not possible at scale, making computational inference key for extracting physically meaningful reconstructions of the magnetic noise environment.

One aspect of this thesis focuses on developing computational methods that integrate directly with experimental workflows to characterize spin baths surrounding individual NV centers. The approach taken in this thesis treats spin-environment reconstruction as a probabilistic inference problem, combining physical modeling of spin dynamics with scalable numerical algorithms capable of exploring high-dimensional parameter spaces. By leveraging hybrid Markov chain Monte Carlo techniques and physics-informed priors, the methods developed here infer candidate nuclear spin configurations consistent with experimental coherence measurements while quantifying uncertainty in the reconstruction. Crucially, these algorithms are designed for iterative use alongside experiments, enabling rapid updates as new measurements become available.

In this framework, computation and experiment operate as components of a unified feedback loop. Experimental data constrain possible spin environments; computational inference identifies the most probable microscopic configurations; and subsequent experiments may be adaptively designed to resolve remaining ambiguities or optimize quantum control strategies. Efficient computational turnaround is therefore essential, as faster inference directly accelerates experimental progress. By enabling data-efficient reconstruction of otherwise inaccessible microscopic environments, the work in this thesis lays the foundation for integrating computational–experimental workflows to transform characterization into an active

process of learning and control.

1.2.3 Complementarity with machine learning- and AI-based materials discovery

Machine learning (ML) has rapidly emerged as a powerful paradigm in materials science, enabling fast property prediction, surrogate modeling, and accelerated exploration across vast chemical and structural design spaces. By learning complex structure–property relationships from data, ML models can dramatically reduce the computational cost associated with evaluating candidate materials once trained. Consequently, data-driven approaches now play a central role in modern materials discovery workflows, supporting large-scale screening, inverse design, and automated analysis pipelines [39]. Recent developments in foundation models—large, general-purpose models trained on diverse materials datasets and designed to transfer across multiple tasks—further extend this paradigm by enabling transferable representations that can be adapted to new properties, chemistries, or simulation regimes with comparatively limited additional data, suggesting a pathway toward more unified and scalable materials prediction frameworks [40].

Despite these successes, ML methods remain fundamentally dependent on the availability of large, high-quality training datasets [41]. In computational materials science, generating such datasets is itself a demanding task, typically requiring extensive first-principles simulations or carefully controlled experimental measurements. In many cases, data generation—not model training—constitutes the dominant computational and time expense. Furthermore, ML models may lose predictive reliability when applied outside the distribution represented in their training data, particularly when encountering new chemistries, structural motifs, or physical mechanisms. These challenges highlight an important limitation of purely data-driven approaches and motivate the continued development of complementary computational strategies that do not rely primarily on large pre-labeled datasets.

Physics-based and numerical methods therefore remain indispensable components of the materials discovery ecosystem. First-principles simulations and structured numerical inference techniques provide reliable ground-truth data used to train, validate, and interpret ML models. Equally importantly, they offer direct access to underlying physical mechanisms, enabling hypothesis testing and interpretability in ways that purely statistical models often cannot [42]. Such methods are especially valuable in regimes where data are sparse, noisy, or expensive to obtain, including experiment-coupled characterization and inverse problems. By incorporating known physical laws and uncertainty quantification directly into the computational framework, physics-based algorithms can maintain stability and predictive consistency even when extrapolating beyond previously observed data.

Scalable non-ML algorithms, such as the ones presented in this thesis, are thus not competitors to machine learning but complements. Efficient physics-driven methods strengthen ML-enabled research by producing trustworthy datasets, guiding model development, and enabling controlled exploration of new regions of materials space. At the same time, incorporating domain knowledge, such as symmetry constraints, conservation laws, locality principles, and known functional relationships, can significantly improve both physics-based and ML approaches [43]. Embedding such structure reduces the effective hypothesis space, improves robustness and generalization, and enhances interpretability.

This thesis adopts a physics-informed, data-efficient perspective aimed at complementing machine learning rather than replacing it. The methods developed here emphasize fast numerical characterization, scalable inference, and high-throughput property evaluation, particularly in settings where experimental feedback is closely coupled to computation or where data remain limited. Beyond their direct application to materials discovery and characterization, these approaches also accelerate the generation of reliable training data and provide rigorous validation pathways for ML predictions.

1.3 Thesis scope and summary of contents

The central goal of this thesis is to develop computational methods for identifying, characterizing, and predicting materials with applications in next-generation computing hardware. As discussed in the preceding sections, continued advances in computing performance and efficiency increasingly depend on discovering and engineering materials with specific functional properties, including electronic, magnetic, ferroelectric, and quantum-coherent behavior. Computational materials science provides a primary pathway for accelerating this discovery process, but doing so at scale requires algorithms and numerical methods that are not only accurate, but also efficient, robust, and compatible with high-throughput and experiment-coupled workflows. This thesis focuses on numerical methods and computational frameworks that are (i) scalable to large materials sets and parameter spaces, (ii) data-efficient in regimes where labeled data or measurements are limited, and (iii) compatible with automated screening pipelines and rapid experimental feedback. Rather than emphasizing black-box prediction, the methods developed here are designed to integrate with first-principles calculations, inverse modeling, and experimental characterization.

In this thesis, I develop computational methods for application in two next-generation computing architectures: multiferroic spintronics and quantum computing. In Chapter 2, which is based on work in [44], I develop and test heuristics relevant to automate and deploy an algorithm known as “Berry flux diagonalization" to compute spontaneous polarization for insulating materials. Spontaneous polarization is a key quantity in characterizing a class of materials known as ferroelectrics, which are promising candidates for spintronic devices and low-energy computing architectures. Berry flux diagonalization circumvents many of the numerical challenges of the traditional methods to compute polarization, including reducing the number of expensive density functional theory (DFT) calculations necessary to obtain polarization values by an order of magnitude. This work represents the first time Berry flux diagonalization has been implemented and tested as a viable alternative method to

the traditional method of computing spontaneous polarization at scale, and we demonstrate that our implementation of this method is numerically robust and maintains accuracy over a diverse set of insulating compounds.

Next, I turn to computational method development for the scale-up and deployment of spin-defects as candidate qubits in quantum computers. In Chapter 3, I present and validate a computational method to resolve the inverse problem of reconstructing the nuclear spin bath of a spin-defect based on work in [45]. Understanding the coherence of individual spin defects requires characterizing the surrounding spin bath, which governs how long quantum information can be retained before interactions with the local environment lead to decoherence. Beyond providing insight into coherence limitations, knowledge of the spin bath may also enable the design of optimized control pulse sequences and allow nearby nuclear spins to be harnessed as local quantum memory units. In particular, I show that the computational method I develop performs as well as comparable machine learning methods, but at a much lower computational cost. Furthermore, I demonstrate that this computational method works even with sparse, noisy data, which is important when trying to characterize many spin-defects quickly. After presenting and validating this computational method, in Chapter 4 I demonstrate how this computational method can be used to guide experimental design for high-throughput experimental characterization of spin-defects while maintaining physically and mathematically rigorous bounds on sensing limits based on work from [46]. In this chapter, we apply this method to demonstrate the utility of this approach for high-throughput experimental characterization of individual spin-defects. I conclude in Chapter 5 by exploring extensions of the work presented in Chapters 2-4. Taken together, these contributions support a unified objective: to develop computational tools that accelerate the discovery and characterization of materials for next-generation computing architectures. By emphasizing scalable, physics-informed, and data-efficient algorithms, this thesis aims to strengthen the computational foundation needed to develop new materials platforms and,

ultimately, new forms of computing hardware.

CHAPTER 2

HIGH-THROUGHPUT COMPUTATION OF ELECTRIC POLARIZATION IN SOLIDS VIA BERRY FLUX DIAGONALIZATION

Electric polarization in the absence of an externally applied electric field is a key property of polar materials, but the standard interpolation-based *ab initio* approach to compute polarization differences within the modern theory of polarization presents challenges for automated high-throughput calculations. Berry flux diagonalization [J. Bonini *et al.*, Phys. Rev. B **102**, 045141 (2020)] has been proposed as an efficient and reliable alternative, though it has yet to be widely deployed. In this chapter, we assess Berry flux diagonalization using *ab initio* calculations of a large set of materials, introducing and validating heuristics that ensure branch alignment with a minimal number of intermediate interpolated structures. Our automated implementation of Berry flux diagonalization succeeds in cases where prior interpolation-based workflows fail due to band-gap closures or branch ambiguities. Benchmarking with *ab initio* calculations of 176 candidate ferroelectrics, we demonstrate the efficacy of the approach on a broad range of insulating materials and obtain accurate effective polarization values with fewer interpolated structures than prior automated interpolation-based workflows. Our real-space heuristics that can predict gauge stability *a priori* from ionic displacements enable a general automated framework for reliable polarization calculations and efficient high-throughput screening of chemically and structurally diverse polar insulators. These results establish Berry flux diagonalization as a robust and efficient method to compute the effective polarization of solids and to accelerate the data-driven discovery of functional polar materials.

2.1 Introduction

Polar crystals are solids that exhibit a nontrivial electric polarization in the absence of an externally applied electric field. Materials with this property play a central role in a range of electronic, electromechanical, and energy-related applications. For example, ferroelectrics are aclass of polar materials in which the polarization is switchable under an external electric field, with applications in non-volatile memory devices, sensors, and energy storage [47, 48]. Multiferroics, which feature ferroelectricity coexisting with magnetic order, further expand the scope of applications to include magnetoelectric devices, spintronic devices, and multifunctional sensors [49].

A key step in first principles studies of insulating polar materials is the computation of the effective polarization. Within the modern theory of polarization, the effective polarization is defined as the change in formal polarization along an adiabatic, insulating path connecting a polar material to a nearby nonpolar reference structure [50]. The formal polarization of an insulator is computed via the Berry phase of occupied electronic states. The formal polarization is an inherently multi-valued quantity, since it is defined modulo a quantum of polarization set by the unit cell volume and lattice vectors. Although the Berry phase is now a routine output of first-principles electronic structure codes, obtaining the effective polarization requires taking differences between multi-valued formal polarizations and tracking the correct branch of the Berry phase across the chosen path. As a result, automated high-throughput calculations of effective polarizations remain challenging, particularly in cases where the effective polarization is comparable in magnitude to the polarization quantum [51].

Neither the formal nor the effective polarization are quantities that can be measured directly. Experimental studies typically characterize spontaneous polarization, which can be inferred from the integrated switching current that results when an external electric field reverses the polarization of a material. The spontaneous polarization is related to half of

the charge transported during a switch between two polar enantiomorphs. The nonpolar reference structure used in the theoretical definition of effective polarization is not intended to represent a physical configuration that the material visits during switching but is instead a practical construction indicating a halfway point between two polar enantiomorphs along a single symmetric path [29]. Effective polarizations obtained in this way can be compared to experimental spontaneous polarizations, although experimental values may depend on non-bulk factors such as surface terminations and interfaces with electrodes. While several considerations arise in selecting an appropriate nonpolar reference structure that can be connected to a give polar structure along an insulating path, here we focus on the challenge of computing differences in formal polarization along the proper polarization branch in a consistent, efficient, and automatic manner. For discussion of automated strategies for identifying suitable nonpolar reference structures, we refer the reader to Ref. [31].

The standard approach to computing effective polarization relies on linear interpolation of atomic coordinates and repeated Berry-phase evaluations to compute formal polarizations along an insulating, adiabatic path [30, 31]. Although robust for simple systems, this procedure is computationally costly and can be unreliable when automated. The required number of intermediate configurations to resolve polarization branches depends on system-specific parameters, such as the polarization quantum and the unknown value of effective polarization itself. If too few interpolation structures are used, the branch cannot be properly traced, but adding more such structures increases the computational cost. Moreover, intermediate structures may exhibit spurious band gap closures, leading to metallic states for which the Berry phase is ill defined. Born effective charges, obtained either via density functional perturbation theory (DFPT) or finite differences calculations, have previously been used to guide the choice of the correct branch or as a direct proxy to the polarization itself, but they come at greater computational cost and often with lesser accuracy [52, 53]. Adaptive interpolation schemes (e.g., the nudged elastic band (NEB) method [54]), though viable, can

be even more computationally expensive due to the need for repeated *ab initio* calculations and Berry phase evaluations.

An alternative approach to computing effective polarization, known as Berry flux diagonalization, was recently proposed by Bonini *et al.* [55] and can circumvent these challenges. Rather than interpolating along a real-space structural path between two structures, Berry flux diagonalization directly computes the differences in formal polarization along the proper branch between two insulators [55]. Due to the underlying periodicity of Bloch wavefunctions in crystalline solids, there is a direct correspondence between branch-hopping in polarization and non-zero winding numbers of Berry phase differences [29]. By constructing a gauge that aligns wavefunctions across k-points in the Brillouin zone and structures, formal polarization differences can be computed without explicit interpolations of atomic positions.

Initial implementations of Berry flux diagonalization [55] were limited to single-direction effective polarization calculations and small test sets, and lacked numerical heuristics to guarantee gauge stability. Consequently, the method, while conceptually elegant, has yet to be broadly adopted or validated for diverse material classes, including systems involving large displacements. In this work, we extend, systematize, and automate an implementation of Berry flux diagonalization to enable robust and efficient effective polarization calculations across a wide range of polar insulators. We generalize the implementation of Berry flux diagonalization to compute polarization automatically along all three directions and support spin-polarized systems. Additionally, we provide a generalizable and extensible workflow that, while currently demonstrated for VASP and Quantum ESPRESSO, can be compatible with any plane-wave based density functional thoery (DFT) code. Crucically, we introduce a physical heuristic relating atomic displacements to wavefunction overlap singular values, providing an *a priori* criterion for when adiabatic gauge alignment is likely to succeed. We benchmark our automated implementation of the Berry flux diagonalization method with the developed heuristics against the standard interpolation-based approach on

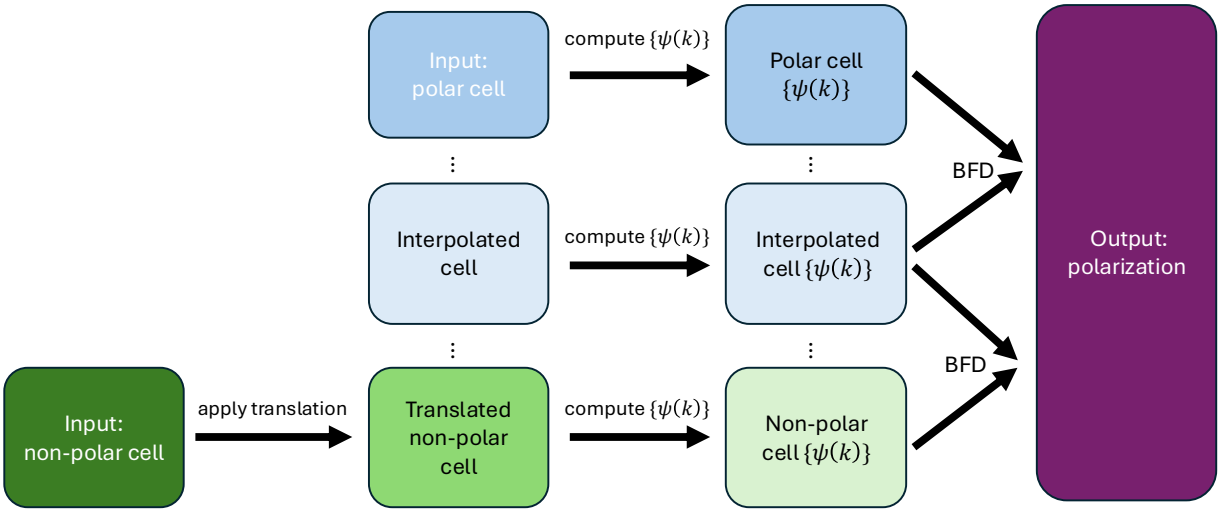

Figure 2.1: Schematic of our automated workflow to compute effective polarization. First, a rigid translation is applied to the nonpolar reference cell to minimize maximal atomic displacement relative to the polar cell. If the maximal atomic displacement between the polar and translated nonpolar reference cell is larger than 0.3 Å, intermediate structures are created by linearly interpolating the atomic coordinates. Then, Bloch wavefunctions for a uniform k-grid are computed using DFT. The Berry flux diagonalization algorithm is directly applied to wavefunctions corresponding to the polar and nonpolar cells (and any intermediates) to obtain a value for polarization.

176 candidate ferroelectric materials that were previously identified in [31], demonstrating that our implementation of Berry flux diagonalization accelerates the calculation of effective polarization while keeping the same accuracy as standard approaches and providing a foundation for large-scale screening and discovery of new polar materials.

2.2 Theoretical formalism

We begin provide a brief summary of the standard Berry phase approach to computing effective polarization within the modern theory of polarization, and we provide a review of the Berry flux diagonalization method to compute effective polarization. In our presentation of the standard Berry phase approach method to compute effective polarization and the Berry flux diagonalization formalism, we emphasize the mathematical connections between

the two approaches, and we highlight the equations that inform the heuristics we develop for our implementation of an automated workflow using Berry flux diagonalization to compute effective polarization. For detailed background and pedagogy, we refer the reader to Refs. [52, 53, 56] for the modern theory of polarization and to Ref. [55] for Berry flux diagonalization.

2.2.1 Standard approach for computing polarization using DFT

In the modern theory of polarization, the formal polarization ($\mathbf{P}$) of a crystalline solid, consisting of both ionic and electronic contributions, can be written as

$$\mathbf{P} = \mathbf{P}_{\text{elec}} + \mathbf{P}_{\text{ion}} \mod \mathbf{P}_{\text{quanta}}, \tag{2.1}$$

where $\mathbf{P}_{\text{quanta}}$ is the polarization quantum, an integer multiple of ${}^{e\mathbf{R}}/_{\Omega_{\text{cell}}}$, which is related to the periodicity of the wavefunction and ionic coordinates under translation by a lattice vector. The polarization quanta for a lattice is given by

$$\mathbf{P}_{\text{quanta}} = \sum_{i \in a,b,c} n_i \frac{e\mathbf{R}_i}{\Omega_{\text{cell}}}, \tag{2.2}$$

where Ω_{cell} is the volume of the unit cell, $\mathbf{R}_i$ are primitive lattice vectors, and $n_i \in \mathbb{Z}$. The formal polarization $\mathbf{P}$ of a crystalline solid is therefore a lattice of values and any one value can only be defined modulo a quantum. In practice, only differences in polarization ($\Delta\mathbf{P}$) between related structures are meaningful [52], and as long as the difference in structures are small enough one can always select a common reference quantum, or branch, that yields a single well-defined $\Delta\mathbf{P}$. We will return to this issue later.

In practice, the ionic contribution, $\Delta\mathbf{P}_{\text{ion}}$, can be computed from the difference in atomic coordinates as

$$\Delta\mathbf{P}_{\text{ion}} = \frac{e}{\Omega_{\text{cell}}} \sum_i Z_i \Delta\mathbf{r}_i, \tag{2.3}$$

where for each atom indexed by i, Z_i is the ionic charge, defined as the total atomic number of the atom minus the number of its electrons explicitly treated as valence by the pseudopotentials in the calculation, and $\Delta\mathbf{r}_i$ is the displacement of the atom between the polar and nonpolar reference structures.

The electronic contribution to the formal polarization for a given atomic arrangement is given by the Berry phase γ_n of occupied Bloch states, namely

$$\mathbf{P}_{\text{elec}} = \frac{e}{2\pi} \sum_n \gamma_n, \tag{2.4}$$

where n is over occupied states, and the electronic contribution to effective polarization is computed as the difference between the formal polarization of the polar (pol) and nonpolar (np) structures:

$$\Delta\mathbf{P}_{\text{elec}} = \frac{e}{2\pi}\mathbf{\Phi} = \mathbf{P}_{\text{elec}}^{\text{pol}} - \mathbf{P}_{\text{elec}}^{\text{np}}. \tag{2.5}$$

The Berry phase γ_n is typically computed as the integral of the Berry connection $\mathbf{A}_n(\mathbf{k})$ along a closed path $\mathcal{C}$ in reciprocal space, that is

$$\gamma_n = \oint_{\mathcal{C}} \mathbf{A}_n(\mathbf{k}) \cdot d\mathbf{k}, \tag{2.6}$$

where

$$\mathbf{A}_n(\mathbf{k}) = \text{Im}\langle u_{n\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n\mathbf{k}} \rangle. \tag{2.7}$$

Here, $|u_{n\mathbf{k}}\rangle$ is the cell-periodic part of a Bloch wavefunction and $\nabla_{\mathbf{k}}$ denotes the gradient with respect to the crystal momentum $\mathbf{k}$. For a discrete set of $\mathbf{k}$-points in the Brillouin zone, Eq. 2.6 becomes

$$\gamma_n = \mathrm{Im} \ln \prod_{m=1}^{N-1} \langle u_{n\mathbf{k_{m+1}}} | u_{n\mathbf{k_m}} \rangle. \tag{2.8}$$

The multi-valued nature of formal polarization originates with γ_n in Eq. 2.8 being itself multi-valued, due to the fact that the imaginary part of the logarithm of a phase is defined only modulo 2π. This multi-valuedness is a consequence of the periodicity of the wavefunction under translations by a lattice vector.

The effective polarization is finally determined by subtracting the Berry phase of the nonpolar structure from that of the polar one (Eq. 2.5). However, due to the multi-valued nature of the Berry phase, directly taking this difference can be ambiguous. To ensure the subtraction occurs on a consistent branch of the Berry phase, where both values of the formal polarization in the difference refer to the same quantum, it is common practice to compute the formal polarizations for intermediate structures by interpolating atomic positions between the polar and nonpolar configurations [30, 51]. Alternatively, another practice is to estimate the effective polarization via

$$\Delta \mathbf{P} \cong \frac{1}{\Omega} \sum_i Z_i^* \Delta \mathbf{u}_i, \tag{2.9}$$

where the sum runs over all atoms in the unit cell, i indexes an atom, and $\{Z_i^*\}$ are the Born effective charges of, e.g., the nonpolar or polar phase. This estimate, which is approximate but becomes exact in the limit of small $\Delta \mathbf{u}$, can guide the manual choice of a consistent branch for both phases [52, 53].

2.2.2 Berry flux diagonalization

Berry flux diagonalization computes the Berry phase differences (denoted ϕ_n^p) directly from the Bloch wavefunctions of the two structures (see Fig. 2.1), as opposed to the standard procedure of computing the Berry phases for each structure independently and then taking their difference. To compute the electronic contribution to effective polarization using Berry

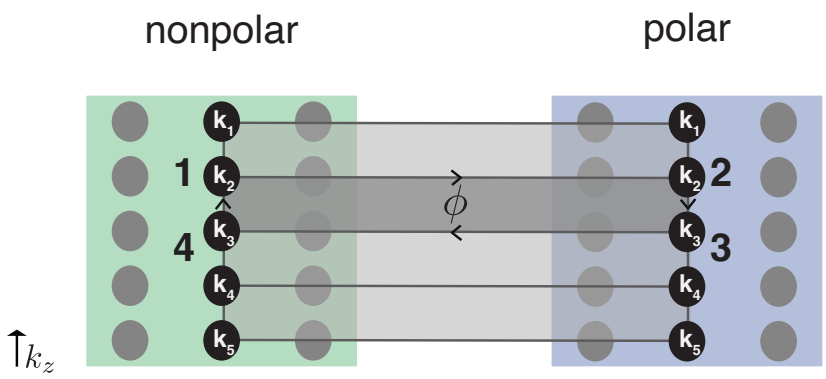

Figure 2.2: Schematic of a Berry phase "plaquette" used in the BFD calculation of Berry phase differences ϕ_n^p along the z-direction. Each plaquette is defined by a closed k-point loop $\mathcal{P}$ involving adjacent k-points from both the polar and nonpolar structures. To compute ϕ for a single plaquette, we compute overlap matrices between Bloch wavefunctions along the path $\mathcal{P} = (1)\,\mathbf{k}_2^{\text{np}} \to (2)\,\mathbf{k}_2^{\text{pol}} \to (3)\,\mathbf{k}_3^{\text{pol}} \to (4)\,\mathbf{k}_3^{\text{np}} \to (1)\,\mathbf{k}_2^{\text{np}}$. That is, we compute overlap matrix $M^{\langle 1,2\rangle}$ corresponding to the overlap between Bloch wavefunctions at k-points labeled (1) $\mathbf{k}_2^{\text{np}}$ and (2) $\mathbf{k}_2^{\text{pol}}$, and then we compute overlap matrix $M^{\langle 2,3\rangle}$ corresponding to the overlap between Bloch wavefunctions at k-points labeled (2) $\mathbf{k}_2^{\text{pol}}$ and (3) $\mathbf{k}_3^{\text{pol}}$ in the polar cell, and likewise for $M^{\langle 3,4\rangle}$ and $M^{\langle 4,1\rangle}$. The Berry phase difference ϕ corresponding to this plaquette is computed via Eqs. 2.12-2.17, and the Berry phase differences corresponding to all plaquettes defined by the k-points are summed to obtain the total Berry phase difference $\mathbf{\Phi}$.

flux diagonalization [55], the total phase difference $\mathbf{\Phi}$ (as in Eq. 2.5) is written as a sum over Berry phase "plaquettes":

$$\mathbf{\Phi} = \sum_p \sum_n \phi_n^p, \tag{2.10}$$

where n indexes the occupied bands and p labels the plaquettes.

Each plaquette is defined by a closed loop in k-space involving two adjacent k-points in both the polar (pol) and nonpolar (np) structures:

$$\mathcal{P}(\text{np, pol}) = (1)\,\mathbf{k}_j^{\text{np}} \to (2)\,\mathbf{k}_j^{\text{pol}} \to (3)\,\mathbf{k}_{j+1}^{\text{pol}} \to (4)\,\mathbf{k}_{j+1}^{\text{np}} \to (1)\,\mathbf{k}_j^{\text{np}}, \tag{2.11}$$

where $\mathbf{k}_i^{\text{pol}}$ and $\mathbf{k}_i^{\text{np}}$ denote k-points in the polar and nonpolar structures, respectively (see

Fig. 2.2). The points $\mathbf{k}_j$ and $\mathbf{k}_{j+1}$ are adjacent along the string of k-points in the direction of the polarization component being computed. Similar to the standard Berry phase approach, this process is repeated along strings of k-points, and the effective polarization is obtained by averaging over strings oriented along the complementary reciprocal lattice directions.

To compute the Berry phase ϕ_n^p for a given band n and plaquette p, we first define overlap matrices between Bloch wavefunctions at adjacent vertices of the plaquette. The Bloch states $|u_n(\mathbf{k}_a)\rangle$ are labeled by band index n and k-point $\mathbf{k}_a$, where a is a k-point along the path of $\mathcal{P}$ (Eq. 2.11). The overlap matrix between two adjacent points $\mathbf{k}_a$ and $\mathbf{k}_b$ on the path $\mathcal{P}$ is

$$M_{mn}^{\langle a,b\rangle} = \langle u_m(\mathbf{k}_a)|u_n(\mathbf{k}_b)\rangle, \tag{2.12}$$

where m and n are band indices.

Each overlap matrix $M^{\langle a,b\rangle}$ is then decomposed using a singular value decomposition (SVD) as:

$$M^{\langle a,b\rangle} = V^{\langle a,b\rangle}\Sigma^{\langle a,b\rangle}W^{\langle a,b\rangle\dagger}, \tag{2.13}$$

where $V^{\langle a,b\rangle}$ and $W^{\langle a,b\rangle}$ are unitary matrices, and $\Sigma^{\langle a,b\rangle}$ is a diagonal matrix of singular values. We then compute the unitary approximant of each overlap matrix from its SVD by discarding the singular values, expressed as

$$\mathcal{M}^{\langle a,b\rangle} = V^{\langle a,b\rangle}W^{\langle a,b\rangle\dagger}. \tag{2.14}$$

The total unitary matrix $U_{\mathcal{P}}$ that describes evolution around plaquette p is constructed as the ordered product of unitary approximants along the four edges of the plaquette path $\mathcal{P}$, namely

$$U_{\mathcal{P}} = \mathcal{M}^{\langle 1,2\rangle}\mathcal{M}^{\langle 2,3\rangle}\mathcal{M}^{\langle 3,4\rangle}\mathcal{M}^{\langle 4,1\rangle}. \tag{2.15}$$

The eigenvalues of this unitary matrix encode the geometric phase acquired by each band,

and can be expressed as

$$\mathrm{eig}(U_{\mathcal{P}}) = \left\{ e^{i\phi_n^p} \right\}; \tag{2.16}$$

the phases themselves are extracted as

$$\phi_n^p = \mathrm{Im} \ln \mathrm{eig}(U_{\mathcal{P}}). \tag{2.17}$$

Importantly, this approach requires both intra-structure overlap matrices (e.g., $\mathcal{M}^{\langle 2,3 \rangle}$ and $\mathcal{M}^{\langle 4,1 \rangle}$) and inter-structure ones (e.g., $\mathcal{M}^{\langle 1,2 \rangle}$ and $\mathcal{M}^{\langle 3,4 \rangle}$). The inter-structure overlaps are not present in standard Berry phase calculations, but they are essential here to evaluate the differences in formal polarizations directly and consistently across branches.

Berry flux diagonalization can circumvent branch-hopping so long as the singular values of the overlap matrices (Eq. 2.13) is greater than zero and each $\phi_i^p \ll \pi$. The singular value requirement ensures that the phase in Eq. 2.17 is not extracted from a complex number whose magnitude is zero. The eigenvalue requirement is equivalent to ensuring that the phase in the eigenvalues $e^{i\phi_n^p}$ do not wrap around the unit circle (i.e., they have a non-zero winding number). In the context of the modern theory of polarization, which is formulated within the adiabatic approximation, this eigenvalue condition corresponds to ensuring that the wavefunctions evolve smoothly along an adiabatic path in parameter space. That is, wavefunctions at adjacent k-points or corresponding k-points across the polar and nonpolar structure, must start sufficiently close to each other so that they are adiabatically connected. To ensure that both of these conditions are satisfied in practice, we develop a real-space heuristic based on the difference in ionic positions, $\Delta \mathbf{r}_i$, from Eq. 2.3, in polar and nonpolar structures. In the next section, we connect the real-space displacement of ions, the singular values of the overlap matrices (Eq. 2.13), and the eigenvalues of the plaquette evolution matrix (Eq. 2.16). This analysis enables us to formulate a criterion to ensure the Berry phase contributions are numerically stable, branch-consistent, and avoid the breakdown of

adiabaticity.

2.3 Heuristic for sufficient wavefunction overlap and gauge consistency

To compute the difference in Berry phase along the correct branch of polarization using Berry flux diagonalization, singular values of the overlap matrices (Eq. 2.13) must be greater than zero and the eigenvalues of the Berry flux plaquettes (Eq. 2.17) must be less than π. When the singular values of the overlap matrices are zero, we are in a numerically unstable regime in which we are extracting phases from complex numbers with zero magnitude. When eigenvalues of the Berry flux plaquettes reach or exceed π, the Berry flux diagonalization formalism fails and is equivalent to jumping between polarization branches. Bonini et al [55] suggested that maintaining a minimum singular value > 0.15 in the overlap matrices (Eq. 2.13) is sufficient to preserve proper branch alignment. Here, we propose a complementary real-space heuristic that requires only atomic coordinates: a maximal atomic displacement of ≈ 0.3 Å between polar and nonpolar structures reliably yields minimum singular values > 0.15 and maximum eigenvalues $\ll \pi$ and thus ensures proper numerical stability and gauge continuity. We systematically test this heuristic using a system with artificially displaced atoms, translations of atomic coordinates for a set of standard ferroelectrics, and a deterministic number of interpolations to reduce atomic displacements in cases where atomic displacements are large (> 0.3 Å). The results in this section are all presented for calculations using VASP, but the same calculations are also performed in Quantum ESPRESSO, and we present those results in Sec. 2.8.1.

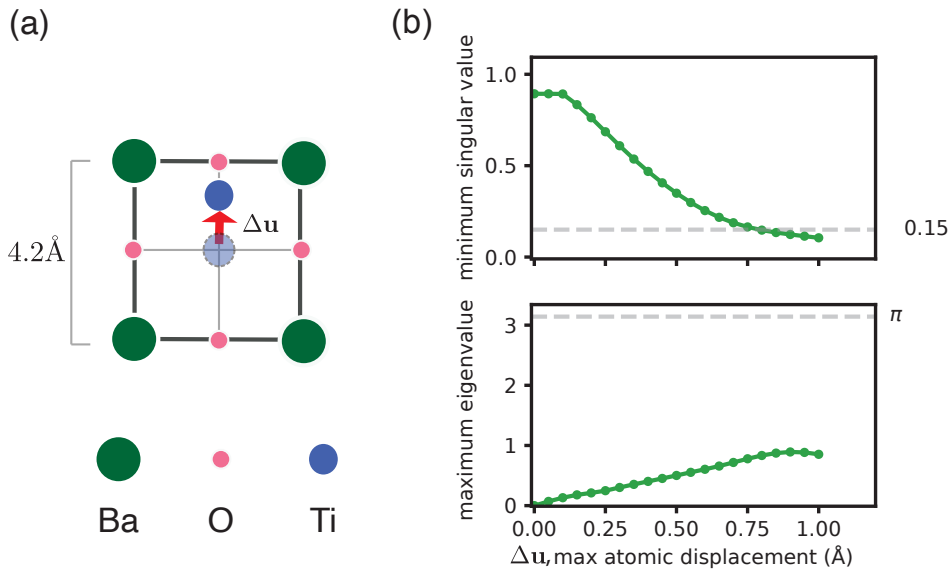

Figure 2.3: (a) Schematic of cubic perovskite $BaTiO_3$ test system with fixed Ba and O atoms and a Ti atom that is artificially displaced along the [001] direction (red arrow). The displacement of the Ti atom, $\mathbf{\Delta u}$, is the difference between its artificial position and its centrosymmetric position. (b) Minimal singular values of the overlap matrices $M^{\langle i,i+1\rangle}$ and maximal eigenvalues of the unitary evolution matrices U_P corresponding to the effective polarization calculated using Berry flux diagonalization for the $BaTiO_3$ with the artificially displaced Ti atom.

2.3.1 Artificial $BaTiO_3$ test system

To illustrate the relationship between nuclear displacement, the minimum singular values of Eq. 2.13, and maximum eigenvalues of Eq. 2.16, we construct a cubic perovskite $BaTiO_3$ system in which we initialize two identical nonpolar configurations fixed with Ba and O atoms. In one of the nonpolar configurations, we artificially displace only the Ti atom from its centrosymmetric position in the unit cell up to 1 Å along the [001] direction (see Fig. 2.3a). We compute overlap matrices between the nonpolar cell and the cell with the displaced Ti atom using Berry flux diagonalization, and we report the minimal singular values over all overlap matrices (Eq. 2.13) and maximal loop eigenvalues over all unitary evolution matrices (Eq. 2.16) in Fig. 2.3b.

We observe a monotonic decrease in the minimal singular values with increasing dis-

placement of the Ti atom relative to its centrosymmetric position in the cubic perovskite structure. The threshold singular value of 0.15 corresponds to a Ti displacement of approximately 0.75 Å. This set of calculations establishes that a quantitative correspondence between ionic displacement in real space, the minimum singular values of the overlap matrices, and the behavior of the maximal eigenvalues of the unitary evolution matrices. We also see that the 0.3 Å heuristic for maximal atomic displacement and the 0.15 heuristic for the minimal singular value is well below the corresponding threshold for which the maximal eigenvalue approaches π.

2.3.2 Translations operations to minimize maximal atomic displacements

To ensure sufficient overlap between wavefunctions of the polar and nonpolar structures in the Berry flux diagonalization approach, we introduce a pre-processing step where the translation between the structures is optimized to minimize the maximal atomic displacement. Specifically, we search for a translation vector $T = (T_a, T_b, T_c)$ such that the maximal atomic displacement along any single direction, $\max_i ||\Delta \mathbf{r}_i||_\infty$, is minimized over all atoms i in the primitive cell. This pre-processing step, performed before any DFT calculation, preserves the cell parameters and interatomic spacings but alters the origin of the nonpolar structure to reduce the distances between corresponding atoms in the polar and nonpolar configurations, reducing wavefunction mismatch.

Figure 2.4 illustrates the impact of this pre-processing translation step on the minimal singular values and maximal eigenvalues of the overlap matrices for $BaTiO_3$, $KNbO_3$, $PbTiO_3$, and $LiNbO_3$ computed directly from the wavefunctions corresponding to the polar and nonpolar structures (see Tables 2.3 & 2.4 for the cell parameters of the polar and nonpolar structures). For $BaTiO_3$, $KNbO_3$, and $PbTiO_3$, optimal translations result in sufficiently small maximal atomic displacements, leading to high minimal singular values (above 0.15) and low maximal eigenvalues (below π), ensuring proper gauge alignment and

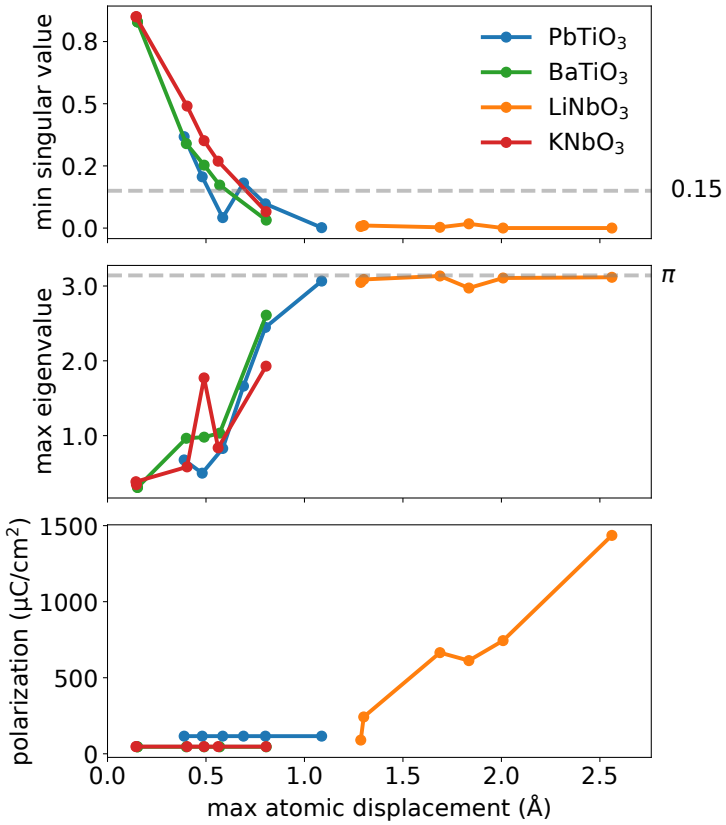

Figure 2.4: Minimal singular values (top), maximal eigenvalues (middle), and effective polarization (bottom) computed from overlap matrices for standard ferroelectrics under different translations of the nonpolar structure relative to the polar structure. For $BaTiO_3$, $KNbO_3$, and $PbTiO_3$, the translations that minimize the maximal atomic displacement ensure sufficient overlap between wavefunctions at each k-point, with minimal singular values remaining above the threshold of 0.15 and maximal eigenvalues below π. These conditions guarantee proper branch alignment and accurate effective polarization calculations. For $LiNbO_3$, even the translation minimizing maximal atomic displacement results in insufficient overlap, with singular values approaching zero and eigenvalues near π, leading to improper resolution of polarization branches and inaccurate values of effective polarization. This can be addressed by adding interpolated structures, see main text. Note: polarization values for $BaTiO_3$ (45.8 $\mu C/\text{cm}^2$) are almost identical to those of $KNbO_3$ (47.3 $\mu C/\text{cm}^2$), so the green curve is obscured by the red one in the effective polarization plot.

accurate effective polarization calculations (see Fig. 2.4a). Furthermore, we observe that increases in the maximal atomic displacements result in decreases in the minimal singular values and increases in maximal eigenvalues. In contrast, for $LiNbO_3$, even the translation that minimizes maximal atomic displacement results in values of minimal singular values

of the overlap matrices approaching zero and maximal eigenvalues of the unitary evolution matrix near π (see Fig. 2.4b), resulting in a polarization of $\sim 90\ \mu C/\text{cm}^2$, which is higher than the reference value of 84.5 $\mu C/\text{cm}^2$ computed from the standard interpolation-based approach. Therefore, this case clearly demonstrates a need to use interpolated structures to further reduce maximal atomic displacement, restore gauge alignment, and recover accurate effective polarizations when a single translation does not result in sufficiently small ionic displacements. We find that adding intermediate interpolated structures for $LiNbO_3$ indeed resolves these numerical instability issues and results in a polarization value (83.4 $\mu C/\text{cm}^2$) in close agreement with the reference value (84.5 $\mu C/\text{cm}^2$).

2.3.3 Interpolated structures

For materials where the minimum atomic displacement between polar and nonpolar structures remains relatively large (> 0.3 Å), we can introduce intermediate (interpolated) structures along the path connecting the two states, along the lines of the standard Berry phase approach. By decreasing the maximal atomic displacement between adjacent structures below the heuristic value of 0.3 Å, we can increase the minimum singular values above 0.15 and obtain maximum eigenvalues that are below π. In cases where a single translation does not sufficiently reduce the maximum atomic displacement, linearly interpolating atoms between the polar and nonpolar structures and computing the Berry phase along this path can minimize structural differences and ensure gauge alignment for Berry flux diagonalization.

We modify the Berry flux diagonalization formalism [55] to compute the electronic contribution to polarization using intermediate, interpolated structures by considering the set of $n-1$ structures {nonpolar (i_0), interpolation$_1$ (i_1), interpolation$_2$ (i_2),..., interpolation$_i$ (i_i), ..., polar (i_n)}. We can re-write the sum in Eq. 2.10 as a sum over the plaquettes formed by adjacent sets of interpolated structures, namely

$$\mathbf{\Phi} = \sum_{i=0}^{n-1} \sum_{p(i,i+1)} \sum_{n} \phi_n^{p(i,i+1)}, \tag{2.18}$$

where the plaquettes $p(i, i+1)$ are defined by the path:

$$\mathcal{P}(\text{i, i+1}) = (1)\,\mathbf{k}_j^{i_i} \rightarrow (2)\,\mathbf{k}_j^{i_{i+1}} \rightarrow (3)\,\mathbf{k}_{j+1}^{i_{i+1}} \rightarrow (4)\,\mathbf{k}_{j+1}^{i_i} \rightarrow (1)\,\mathbf{k}_j^{i_i}, \tag{2.19}$$

where i_i are the indices over the interpolated structures. In this modified definition of Eq. 2.11, $\mathbf{k}_j$ and $\mathbf{k}_{j+1}$ are still adjacent k-points along a string in the direction of the polarization component being computed, and n is still the index of the occupied bands.

Figure 2.5 demonstrates the impact of using additional interpolated structures on numerical stability for three materials, CrO_3, $LiNbO_3$, and HfO_2, for which we observe poor numerical stability after applying the translation that minimizes maximal nuclear distances across polar and nonpolar structures. In CrO_3, the translation that minimizes the maximum atomic displacement between the polar and nonpolar structures still yields a low minimum singular value (0.020) and a high maximum eigenvalue (3.11) for a displacement of 1.31 Å (Fig. 2.5). Introducing four interpolated structures reduces the maximum atomic displacement between adjacent structures to 0.26 Å, which is below our heuristic value of 0.3 Å. This leads to a significant improvement in numerical stability: the minimum singular value increases to 0.69, and the maximum eigenvalue decreases to 0.62, and we obtain the correct effective polarization value of 122.9 $\mu C/\text{cm}^2$. Similarly, for $LiNbO_3$, we observe that large maximal atomic displacements (1.30 Å) result in small minimal singular values (0.0063) high maximal eigenvalues (3.05) and an incorrect effective polarization (90.2 $\mu C/\text{cm}^2$). As we add an increasing number of interpolated structures, the numerical stability improves (the minimal singular values increase and the maximal eigenvalues decrease), and the effective polarization value converges to the reference value. In contrast, even under poor numerical conditions (large eigenvalues and small singular values), the effective polarization for HfO_2

is computed correctly, indicating that the branch alignment is preserved despite numerical instability. Nevertheless, there are no guarantees that this will always be the case.

These effective polarization values computed from different numbers of interpolated structures demonstrate that while our heuristic—requiring maximum atomic displacement below 0.3 Å—is more stringent than strictly necessary for some materials, it may be crucial for others. The heuristic provides a conservative guideline for ensuring numerical stability and branch continuity in Berry phase calculations without requiring any DFT input, but our automated workflow will automatically detect and warn users if the maximum eigenvalue is close to π or if the minimum singular value is too small (< 0.15), regardless of the maximum atomic displacement. Crucially, this heuristic is useful for estimating the number of interpolations required *before* performing any DFT calculations.

2.4 Validation of the Berry flux diagonalization approach for computing effective polarization

We validate the Berry flux diagonalization algorithm by computing effective polarization as the difference in formal polarization between polar and nonpolar structures for a benchmark set of well-known ferroelectrics ($BaTiO_3$, $PbTiO_3$, $KNbO_3$, $LiNbO_3$) as well as a set of proposed polar and known multiferroic materials (Bi_2MoO_6, CrO_3, and $BiFeO_3$) that posed challenges for previous implementations of automated interpolation-based approaches (Table 2.1) [30, 31]. For each material, wavefunctions are obtained from self-consistent DFT calculations using VASP or QUANTUM ESPRESSO for polar and nonpolar structures (see Tables 2.3 & 2.4 for cell parameter information). The Berry flux diagonalization method is applied directly to these wavefunctions (with the heuristics explored in the previous sections). For comparison, we compute effective polarization values using the standard interpolation-based approach approach, either manually resolving polarization branches with QUANTUM ESPRESSO [53, 57, 58] or via the ferroelectric workflow in ATOMATE2 [30, 31] for VASP.

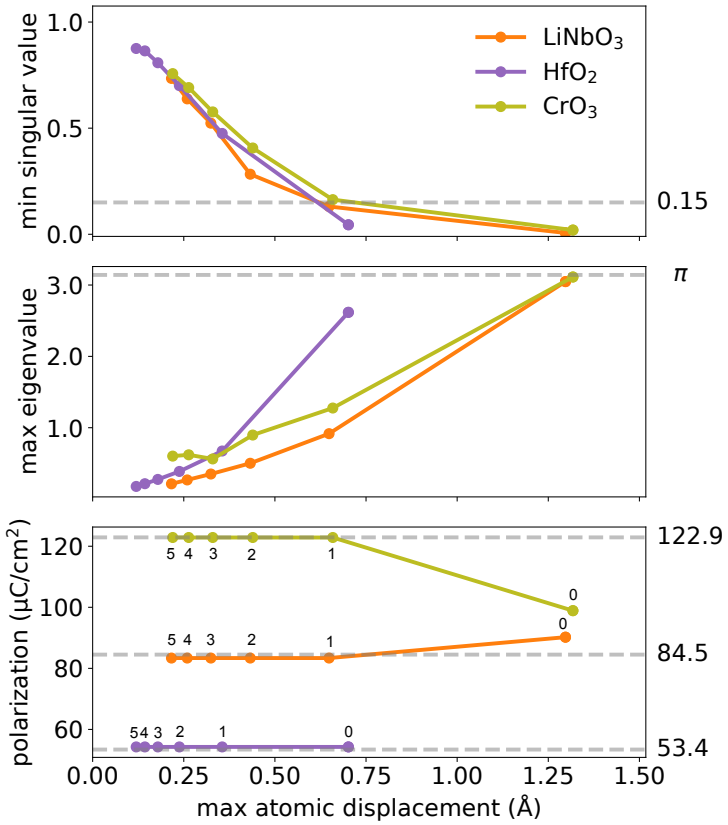

Figure 2.5: Calculated minimum singular values (top), maximal eigenvalues (middle) and effective polarization (bottom) as functions of the maximum atomic displacement for CrO_3, $LiNbO_3$, and HfO_2 computed from different numbers of interpolated structures. Results are presented for varying numbers of interpolations, where the points with the largest maximal atomic distance correspond to the polarizations computed directly from the polar and nonpolar reference structures (with the optimal translation applied to the nonpolar structure), and each successive point with a smaller maximal atomic displacement corresponds to one additional interpolated structure. The number of interpolations used in the calculation for each data point is labeled in the bottom panel. For points corresponding to calculations with more than one interpolation, the numbers plotted are the maxima of the maximum atomic displacements across the sets of adjacent interpolated structures used in the intermediate steps of the computation, and likewise for the maxima of the maximum eigenvalues and minima of the minium singular values.

The effective polarization values reported in Table 2.1 are calculated using the PBE functional [59] with ONCV SG15 pseudopotentials for calculations from wavefunctions using Quantum ESPRESSO and with PAW potentials [60] for calculations using VASP. Despite these detailed methodological differences, our results obtained using the Berry flux diago-

Material	effective polarization ($\mu C/\text{cm}^2$)			
	BFD method (QE)	interpolation method (QE, present work, manual branch tracking)	BFD method (VASP)	interpolation method (VASP, automated branch tracking [30, 31])
$BaTiO_3$	45.9	45.8	45.8	45.9
$KNbO_3$	48.0	49.3	47.3	48.7
$PbTiO_3$	115.8	118.6	116.0	116.8
$LiNbO_3$	83.8	83.7	83.4	84.5
Bi_2MoO_6	76.9	80.6	75.7	7.7* (78.0†)
CrO_3	122.7	123.3	122.9	78.2* (122.9§)
$BiFeO_3$	93.3	95.8	93.8	95.1‡

Table 2.1: Summary of validation of Berry flux diagonalization (BFD) algorithm with the standard interpolation-based Berry-phase approach computed with QUANTUM ESPRESSO (columns 2 and 3) and VASP (columns 4 and 5). The interpolation values in the column 5 are reported directly from prior Materials Project (MP)-based ferroelectric (FE) databases [30, 31]. We use the same cell parameters for all effective polarization calculations based on the relaxed polar and nonpolar phases as were used in the MP FE database. The lattice directions and cell parameters for all materials are reported in Tables 2.3 & 2.4. *Values of polarization reported in Ref. [30] due to incorrect resolution of the polarization branches. †Value of effective polarization calculated with the same method as [30] but with double the number of interpolated structures to resolve the polarization branch as reported in [31].§Value of effective polarization as reported in [30] by manually resolving the polarization branch.‡Value of effective polarization as reported in [51], since Ref. [30] failed to compute a value of effective polarization due to spurious metallic intermediates. The numbers of interpolations used to obtain these effective polarization values are reported in Table 2.2. Manually resolved polarization branches for the interpolation-based method for QUANTUM ESPRESSO (column 3) are shown in Sec. 2.7.4, and convergence studies with respect to the number of k-points used are reported in Sec. 2.7.5.

nalization algorithm in combination with our real-space heuristics show excellent agreement (with an average percent error of 1.86% ± 0.63% using QUANTUM ESPRESSO and 1.34% ± 0.45% using VASP) with the interpolation-based reference values, demonstrating that Berry flux diagonalization reliably resolves polarization branches, including in cases where prior automated workflows face significant challenges.

For Bi_2MoO_6 and CrO_3, the standard interpolation-based approach to computing effec-

Material	# of interpolations		
	BFD method (QE & VASP)	interpolation method (QE, present work, manual branch tracking)	interpolation method (VASP, automated branch tracking [30, 31])
$BaTiO_3$	0	3	10
$KNbO_3$	0	3	10
$PbTiO_3$	1	3	10
$LiNbO_3$	4	3	10
Bi_2MoO_6	2	9	20
CrO_3	4	7	20
$BiFeO_3$	1	5	7*

Table 2.2: Number of interpolations used to compute the effective polarization values reported in Table 2.1. The number of interpolations required for Berry flux diagonalization (column 2) is calculated according to our heuristics. The number of interpolations for polarization calculations using QUANTUM ESPRESSO (column 3) is determined by manually tracing branches, where we began with 3 interpolations and increased the number of interpolations as necessary. These manually-resolved branches are reported in Sec. 2.7.4. The number of interpolations for polarization calculations using VASP is drawn from previous implementations of an automated workflow to compute polarization using the interpolation-based approach [30, 31], where the default setting for the number of interpolations is 10, and this number was increased to 20 for CrO_3 and Bi_2MoO_6, since the default 10 interpolations resulted in an incorrect polarization value. This may be viewed as an upper bound on the number of interpolations, instead of a minimum number. *The reference value for polarization for $BiFeO_3$ in the antiferromagnetic configuration using the standard interpolation-based approach is from [51], and we report the number of interpolations that were used in this calculation.

tive polarization has proven sensitive to the number of intermediate structures used; computing the effective polarization in this manner can lead to incorrect results due to unresolved branch structure. In previous work, insufficient interpolation led to substantial errors. For instance, CrO_3 was reported to have an effective polarization of 78.2 $\mu C/\text{cm}^2$ instead of a value of 122.9 $\mu C/\text{cm}^2$, the latter of which was reported after manual resolution of the branches [30]. Similarly, Bi_2MoO_6 was initially reported to have an effective polarization of 7.7 $\mu C/\text{cm}^2$ instead of 78.0 $\mu C/\text{cm}^2$, the latter of which was reported after calculation using the standard approach with denser interpolation scheme [31]. In contrast, the Berry

flux diagonalization method together with our heuristics enables calculation of effective polarization for these materials using a predetermined (often smaller–see Table 2.2) number of interpolations between the polar and nonpolar structures. This greatly reduces the need for multiple interpolated geometries and manual intervention to track branch continuity, while still yielding effective polarization values that match the reference results (See Table 2.1).

$BiFeO_3$ is a multiferroic material that exhibits a G-type antiferromagnetic (AFM) ordering in its ground state structure with spacegroup R3c (161) [51]. Ref. [30] did not report the effective polarization of this phase due to erroneous band gap closure and metallic intermediate states for some of interpolated structures. But, it reported a large effective polarization of 124.5 $\mu C/\text{cm}^2$ for another polar structure of $BiFeO_3$ with spacegroup R3m (167) in the Materials Project database with ferromagnetic ordering, computed using the cubic nonpolar reference phase with spacegroup Pm-3m (221) and the PBE functional. In better agreement with experiment and prior calculations, Ref. [51] reported an effective polarization of 95.1 $\mu C/\text{cm}^2$ for the R3c phase in the AFM ordering, using the standard nonpolar reference with spacegroup R-3c and LDA+U functional to correctly access its semiconducting nature. In this work, we use the same polar and nonpolar phases spacegroups in the same AFM ordering and we find a value that agreed within 2 $\mu C/\text{cm}^2$ with the value reported in Ref. [51] using both QUANTUM ESPRESSO and VASP (see Table 2.1), with differences attributable to variations in exchange-correlation functional (PBE + U in our work for both codes, versus LSDA + U in [51]). (The functional can also impact optimized structural parameters, and to obtain comparisons across codes and materials, we keep the structural parameters, reported in Tables 2.3 & 2.4, fixed across calculations.) For variation of the effective polarization with U values for QUANTUM ESPRESSO, see Sec. 2.7.6. For VASP, we use a $U_{\text{eff}} = 5.3$ eV, which is consistent with the Material Project's recommendation and is the default ATOMATE2 [61] setting for the $BiFeO_3$.

This validation demonstrates that the Berry flux diagonalization algorithm reliably re-

solves polarization branches and can avoid failures due to metallic intermediates, making it a robust and efficient alternative to interpolation-based approaches. These advantages are particularly critical for prior automated high-throughput workflows, where 149 out of 413 cases were reported in Ref. [30] to have failed due to such challenges.

2.5 High-throughput effective polarization calculations with Berry flux diagonalization

We implement a fully automated workflow for Berry flux diagonalization with our real-space heuristics to apply the appropriate translation and number of interpolations using VASP and ATOMATE2 [61], and we benchmark the automated Berry flux diagonalization workflow on 176 candidate ferroelectrics from Ref. [31]. For this comparison, we use the same relaxed polar and nonpolar structures used for the reference effective polarization values and the same functionals (PBE v.54). Across the dataset, agreement with the standard interpolation-based approach is excellent: 154 of the 176 materials differ by less than 1 % in effective polarization, and 172 of 176 materials differ by less than 10 %. The root-mean-square difference in effective polarization between the two methods is 0.91 $^{\mu C}/_{cm^2}$, demonstrating their overall consistency. Furthermore, our implementation of Berry flux diagonalization requires at most 6 interpolations for any material, with 115 out of 176 materials requiring 1 or fewer and 174 out of 176 requiring 4 or fewer. In contrast, the standard interpolation-based approach used a fixed 10 interpolations by default, and used 20 interpolation in 17 cases when the branch detection algorithm failed with only 10 interpolations. This reduction represents a substantial reduction in computational effort without loss of accuracy. We provide a full table of this high-throughput comparison in Table 2.6.

Only 6 of the 182 candidate ferroelectric materials identified in Ref. [31] could not be calculated automatically. For $CdBiO_3$, $CuReO_4$, and $K_2SrCdSb_2$, we are unable to obtain a large enough band gap for the nonpolar reference structure (using ATOMATE2 settings

for Brillouin zone integrations, a 0.2 eV Gaussian smearing parameter). Increasing the k-point density and using the tetrahedron method, or using a hybrid functional (e.g., HSE), may recover the correct insulating state, and while our workflow flags this situation, future workflows could automatically switch between these strategies when it encounters a gap closure. For $LiGaO_2$, H_7ClO_3, and $RbBe_2F_5$, the automated workflows failed due to parsing issues with the VASP WAVECAR files, which arose due to memory constraints and large file sizes in our calculations. This is due to the fact that our current implementation of Berry flux diagonalization requires wavefunctions over the full Brillouin zone. A Berry flux diagonalization implementation with wavefunctions computed over only the irreducible Brillouin zone, which will result in smaller VASP WAVECAR files would address this issue; we reserve this for future work. Alternatively, one could modify the existing workflow to read in wavefunctions a few k-points at a time and compute the overlap matrices for those specified k-points independently to circumvent this memory issue.

Figure 2.6 summarizes the performance of the automated heuristics for selecting the number of interpolations in effective polarization calculations. Among 176 materials, only four exhibit errors greater than 10 %: $K_4Ba_2SnBi_4$ (12.9 %), Rb_2Se_3 (15.5 %), $Ba_2InSbSe_5$ (15.8 %), and Ba_2CdAs_2 (58.3 %). Each of these cases is automatically flagged by the workflow due to small minimum singular values (≤ 0.15), indicating numerical instability. The user is then prompted to refine the k-point mesh, increase the number of interpolations, or validate the results against independent references such as Born effective charges or standard interpolation-based approaches.

For $K_4Ba_2SnBi_4$, increasing the number of interpolations from 3 to 9 enhances numerical stability (minimum singular value from 0.10 to 0.19) and reduces the effective polarization error to 1.5 %. In contrast, for $Ba_2InSbSe_5$, Rb_2Se_3, and Ba_2CdAs_2, increasing the number of interpolations alone does not significantly improve stability or effective polarization accuracy. Instead, a denser k-point mesh is required for the self-consistent DFT calculation. The high-

throughput workflow employs the default static k-point settings from ATOMATE2 [61], which are sufficient for most materials but undersample in these three cases, leading to minimum singular values of the overlap matrices less than the heuristic value of 0.15.

Doubling the k-point density in each direction improves the minimum singular values from 0.12 to 0.19 for $Ba_2InSbSe_5$, from 6.2×10^{-4} to 0.40 for Rb_2Se_3, and from 5.1×10^{-6} to 0.34 for Ba_2CdAs_2, accompanied by improved effective polarization accuracy—percent errors decrease to 8.1 %, 0.51 %, and 31.1 %, respectively, indicating more stringent covergence criteria beyond the default settings can explain the discrepancy. Figure 2.6 reports the results obtained using the automated heuristics and default k-point meshes, while both the original and recalculated values are listed in the high-throughput results table in Sec. VI of the Supplementary Information.

These outliers highlight the value of automated flagging based on the intermediate singular values and eigenvalues of the overlap matrices. This capability, combined with an *a priori* estimate of the required number of interpolations based solely on ionic coordinates, makes Berry flux diagonalization both more reliable and more efficient than standard interpolation-based methods, which lack predictive heuristics and automated error detection.

2.6 Discussion

Berry flux diagonalization represents an advance in the automated computation of effective polarization and can address limitations of interpolation-based approaches in both computational efficiency and reliability. Previous automated workflows for effective polarization calculations rely on interpolation schemes that assume a fixed number of interpolated structures between polar and nonpolar states, regardless of material-specific characteristics [30, 31]. This framework leads to inefficiencies: in cases where fewer interpolated structures suffice to resolve the polarization path, redundant *ab initio* calculations are performed. Conversely, for materials requiring finer resolution, insufficient interpolation can result in branch

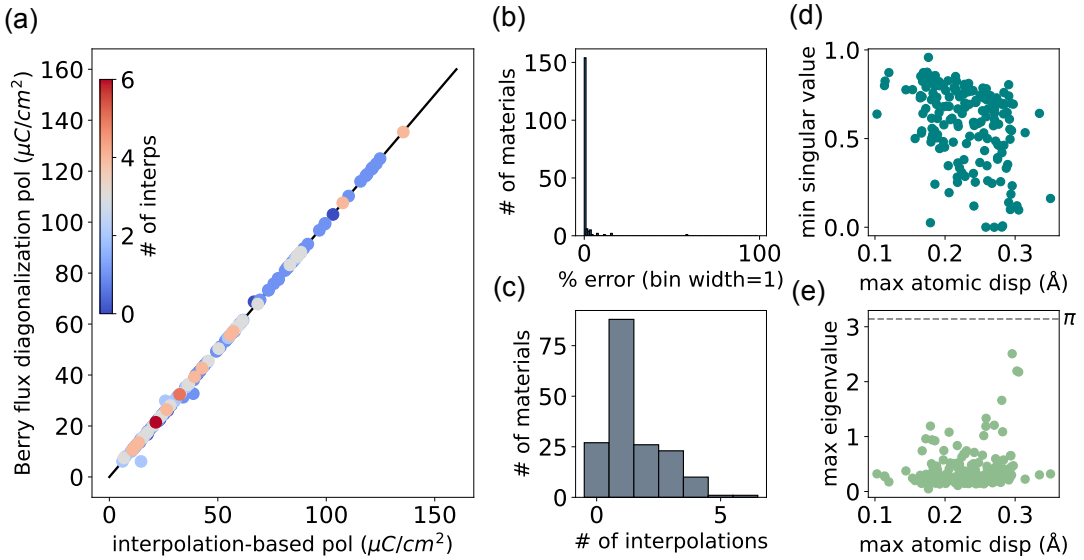

Figure 2.6: High-throughput calculation results comparing our automated workflow using Berry flux diagonalization with the standard interpolation-based approach for 176 candidate ferroelectric materials. (a) Effective polarization computed with Berry flux diagonalization versus the standard interpolation-based approach. (b) Percent error of the polarization computed with Berry flux digaonalization compared to the standard interpolation-based approach, and (c) the number interpolations used in the Berry flux diagonalization calculation (the standard interpolation-based approach uses 10 or 20 interpolations). (d) Minimum singular values and (e) maximum eigenvalues from the overlap matrices used in the Berry flux diagonalization calculation. When there were multiple interpolation structures, maxima and minima over all structures corresponding to a single material are reported.

misalignment and erroneous effective polarization values. The Berry flux diagonalization method, combined with a set of heuristics based on the ionic coordinates of the polar and nonpolar reference structures presented in this work, overcomes these issues. These heuristics determine the minimal number of interpolations necessary, ensuring both efficiency and accuracy. By monitoring the singular values of the overlap matrices and the eigenvalues of the unitary evolution matrix at each step, the algorithm identifies instances of numerical instability or a failure to enforce gauge alignment, indicating a failure to resolve the polarization branch correctly. In such cases, our workflow halts the calculation and alerts the user, avoiding the propagation of incorrect results. This advance addresses a key challenge in traditional interpolation-based workflows, where branch-jumping behavior is difficult to

detect automatically. By explicitly verifying gauge alignment throughout the calculation, this method can provide robust and trustworthy effective polarization values across a wide range of materials.

However, challenges remain for cases where a translation cannot sufficiently minimize the maximal atomic displacement between polar and nonpolar structures. In such cases, additional linear interpolations may fail, similar to interpolation-based Berry phase methods, particularly if the interpolated structures exhibit band gap closures. These limitations highlight the importance of exploring complementary strategies to further enhance the robustness of effective polarization calculations.

Our results indicate that while Berry flux diagonalization can accurately compute polarization for a broad class of materials at a comparable accuracy to the interpolation-based approach, there exists a small class of materials for which the method does not yield stable polarization values, even when the number of interpolations is increased. Crucially, these cases can be automatically identified through monitoring of maximum eigenvalues and minimum singular values, allowing unreliable results to be flagged and excluded. Built-in stability checks ensure that it does not return spurious effective polarization values, making the method reliable and efficient for large-scale screening.

Looking ahead, several improvements can extend the utility of the Berry flux diagonalization method. Leveraging the symmetry of wavefunctions at different k-points could reduce the memory requirements associated with saving wavefunctions across the k-mesh for each structure, streamlining high-throughput applications. Additionally, expanding the method's compatibility with other plane-wave-based density functional theory (DFT) codes through the development of parsers for formats beyond Quantum ESPRESSO and VASP will broaden its applicability and adoption. These ongoing developments position the Berry flux diagonalization algorithm as a powerful and versatile tool for advancing ferroelectric materials discovery.

2.7 Methods

2.7.1 Input cell parameters

The structure parameters (and Materials Project (MP) id's) of the polar and nonpolar cells we used as input for the DFT calculations for validation of the Berry flux diagonalization approach in Sec. 2.4 are reported in Tables 2.3 & 2.4. When possible, we used the same structures that were used in the calculations performed in [30] and [31], which are available through the Materials Project [62].

For $BiFeO_3$, we follow the procedure from [31] and obtain the polar structure for the antiferromagnetic configuration by starting from the ferromagnetic structures and space group reported in the Materials Project (mp-24932). This structure was relaxed in VASP with PBE according to the criterion reported in [30]. Then, we imposed an antiferromagnetic configuration and re-relaxed the structure with the correct antiferromagnetic configuration, and these cell parameters are reported in Table 2.3. For the nonpolar cell, we use a nonpolar cell generated from the polar cell using PSEUDO, and then relaxed the nonpolar structure from PSUEDO, and these cell parameters are reported in Table 2.4.

material	polar cell source	polar (a, b, c) (Å)	polar (α, β, γ) (°)
$BaTiO_3$	mp-5986	(4.002, 4.002, 4.216)	(90.0, 90.0, 90.0)
$KNbO_3$	mp-4342	(4.029, 4.029, 4.219)	(90.0, 90.0, 90.0)
$PbTiO_3$	mp-20459	(3.872, 3.872, 4.585)	(90.0, 90.0, 90.0)
$LiNbO_3$	mp-3731*	(5.223, 5.223, 14.129)	(90.0, 90.0, 120.0)
Bi_2MoO_6	mp-25708*	(5.600, 16.669, 5.625)	(90.0, 90.0, 90.0)
CrO_3	mvc-13134*	(5.001, 5.259, 7.129)	(90.0, 89.9, 90.0)
$BiFeO_3$	[31]**	(5.686, 5.686, 5.686)	(59.2, 59.2, 59.2)

Table 2.3: Polar input cell information and corresponding sources for the calculations in Tables 2.1 & 2.2. Cell parameters for mp-3731 have been updated since the publication of [30], and we report and use the same cell parameters as in [30] to maintain consistency in comparisons between effective polarization values. **$BiFeO_3$ from [31], and see Sec. 2.7.1 for how the structure was generated.

material	nonpolar cell source	nonpolar (a, b, c) (Å)	nonpolar (α, β, γ) (°)
$BaTiO_3$	mp-2998	(4.036, 4.036, 4.036)	(90.0, 90.0, 90.0)
$KNbO_3$	mp-935811	(4.057, 4.057, 4.057)	(90.0, 90.0, 90.0)
$PbTiO_3$	mp-19845	(3.970, 3.970, 3.970)	(90.0, 90.0, 90.0)
$LiNbO_3$	mp-552588*	(5.269, 5.269, 13.903)	(90.0, 90.0, 120.0)
Bi_2MoO_6	mp-567075*	(5.424, 16.949, 5.402)	(90.0, 90.0, 90.0)
CrO_3	mp-779986*	(5.852, 4.570, 9.571)	(90.0, 90.0, 90.0)
$BiFeO_3$	[31]**	(5.537, 5.37, 5.537)	(61.2, 61.2, 61.2)

Table 2.4: Nonpolar input cell information and corresponding sources for the calculations in Tables 2.1 & 2.2. *The cell parameters for have been updated since the publication of [30], and we report and use the same cell parameters as in [30] to maintain consistency in comparisons between effective polarization values. **$BiFeO_3$ from [31], and see Sec. 2.7.1 for how the structure was generated.

2.7.2 Computational Details for DFT calculations

The Berry flux diagonalization method employed in this work is an open-source postprocessing algorithm designed for analyzing wavefunctions in a plane-wave basis generated by first-principles density functional theory (DFT) codes. All *ab initio* calculations for generating wavefunctions were performed using QUANTUM ESPRESSO v7.3 [63] or VASP v6.4.2 with cell inputs obtained from the Materials Project database [62]. For QUANTUM ESPRESSO, we used SG15 ONCV pseudopotentials and the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional. Unless otherwise noted, calculations using QUANTUM ESPRESSO employed a plane-wave kinetic energy cutoff of 100 Ry and a $7 \times 7 \times 7$ k-point grid for self-consistent field (SCF) calculations. For non-self-consistent field (NSCF) calculations, we used a denser grid with 21 k-points along the g-direction and at least 10 k-points per string to compute reference polarizations using the "lberry" tag. For DFT+U calculations on $BiFeO_3$, a Hubbard parameter $U = 3$ eV was applied. For VASP, we used the PBE functional (v54) and the projected augmented wave (PAW) pseudopotentials. The standard interpolation-based approach is implemented in the workflow described in [30] and in the ATOMATE2 code [61]. We refer to those papers for more details on the standard interpolation-based workflow,

and we use the default DFT parameters for k-points and energy cutoffs from the Materials Project for the VASP calculations presented here.

2.7.3 Computational details for implementation of Berry flux digaonalization

Our implementation of the Berry flux diagonalization method is independent of the specific DFT code used to generate wavefunctions, relying only on the plane-wave format. The algorithm's implementation is agnostic to internal data structures of the codes and depends solely on PYMATGEN [64] for parsing and symmetry operations. In this work, the wavefunction parser for QUANTUM ESPRESSO was adapted from QESCHEMA [65], and similar parsers can be developed for other plane-wave-based DFT codes to interface with our framework seamlessly. The wavefunction parser for VASP uses PAWPYSEED to read PAW DFT wavefunctions and recover the full wavefunctions from the pseudowavefunctions based on their projector functions [66].

2.7.4 Manually-traced branches for validation with QUANTUM ESPRESSO

We benchmark the Berry flux diagonalization method for computing effective polarization in QUANTUM ESPRESSO in Sec. 2.4 against the conventional approach, which constructs polarization paths via linear interpolation between polar and nonpolar reference structures. In QUANTUM ESPRESSO, all polarization branches were manually traced to obtain effective polarization values, including the standard ferroelectric set ($KNbO_3$, $BaTiO_3$, $PbTiO_3$, and $LiNbO_3$), shown in Figure 2.7, as well as materials for which the automated workflow of [30] fails (Bi_2MoO_6, CrO_3, and $BiFeO_3$), shown in Figure 2.8.

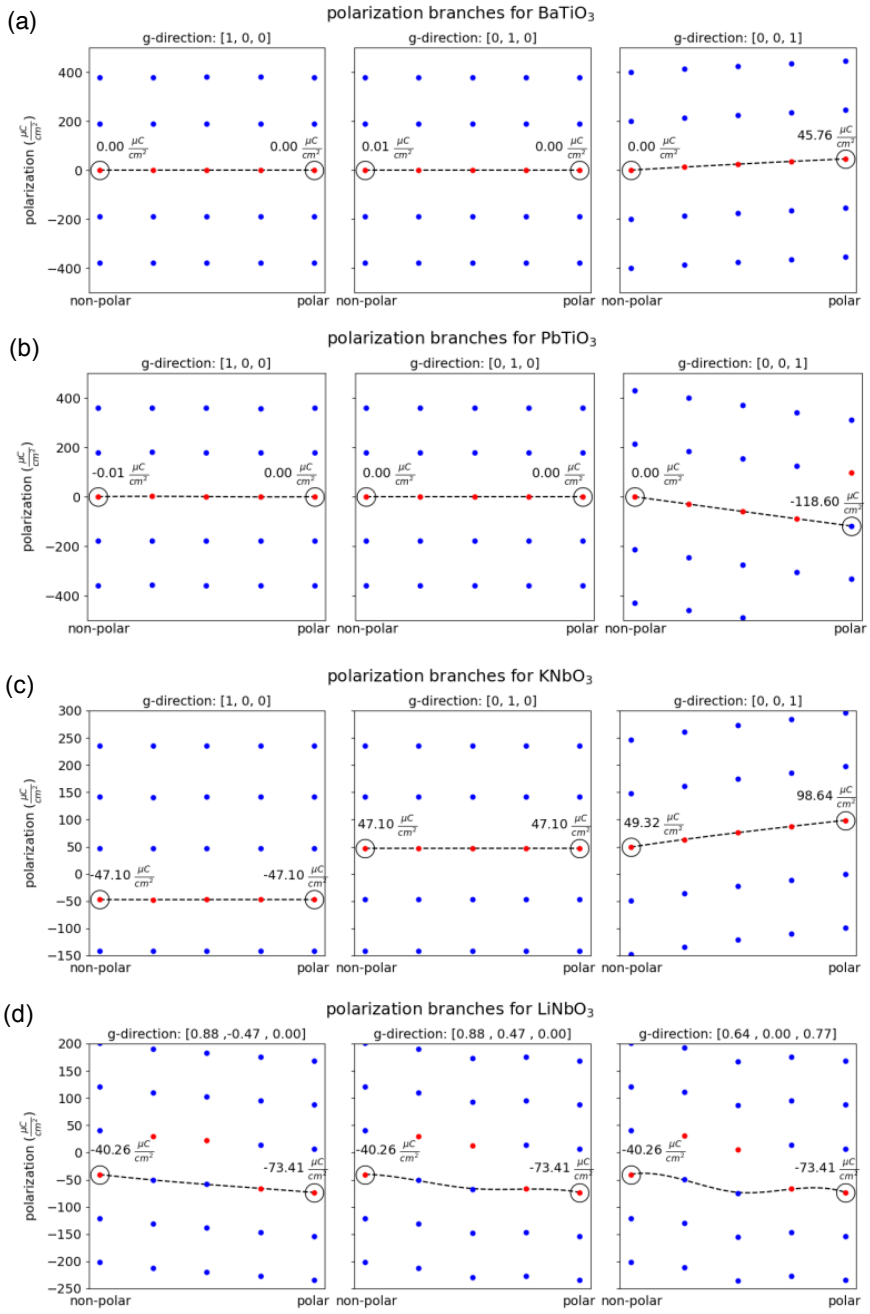

Figure 2.7: Polarization branches for (a) $BaTiO_3$, (b) $PbTiO_3$, (c) $KNbO_3$, and (d) $LiNbO_3$ calculated using the "lberry" tag in QUANTUM ESPRESSO. Polarization values output from QUANTUM ESPRESSO are in red, the polarization branches calculated using the polarization quanta are in blue, the values used for the effective polarization are circled and labeled, and the dashed line is to guide the eye along a proper branch of the polarization.

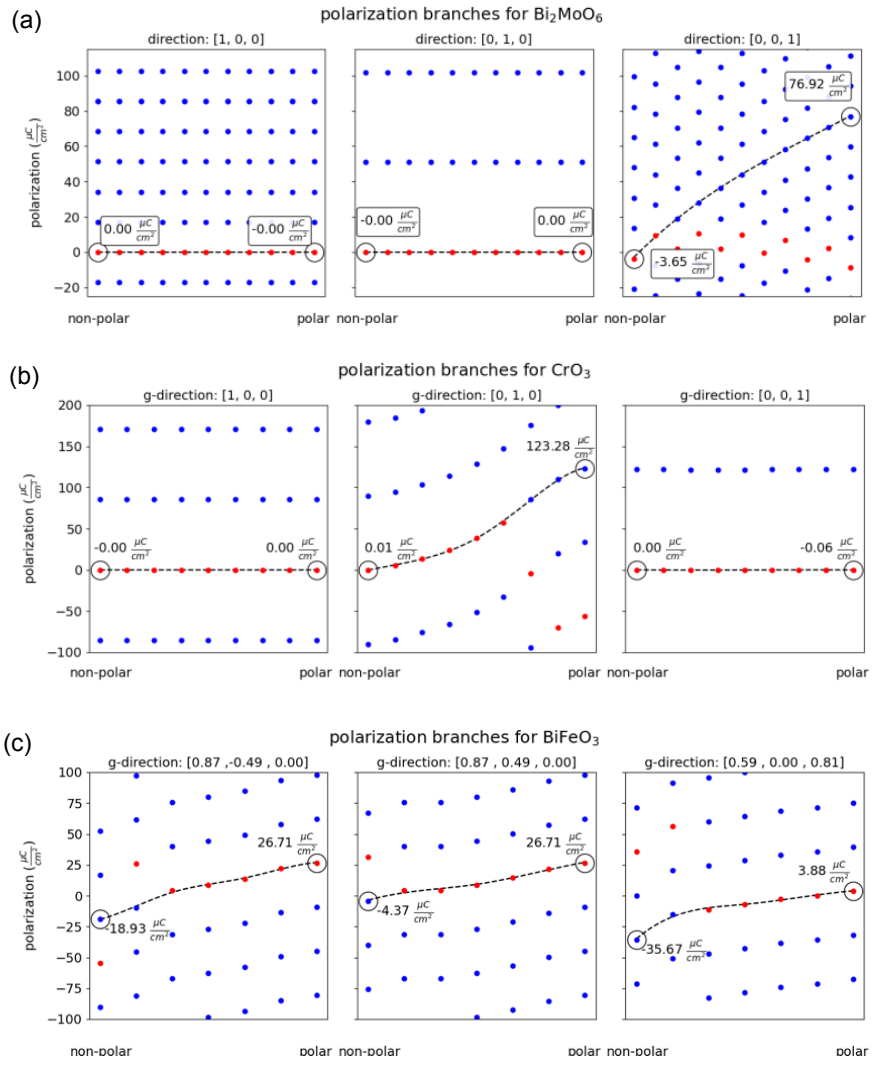

Figure 2.8: Polarization branches for (a) Bi_2MoO_6*, (b) $CrO_3^{\dagger}$, and (c) $BiFeO_3$ calculated using the "lberry" tag in QUANTUM ESPRESSO. Polarization values output from QUANTUM ESPRESSO are in red, the polarization branches calculated using the polarization quanta are in blue, the values used for the effective polarization are circled and labeled, and the dashed line is to guide the eye along a proper branch of the polarization. *We used a $9\times9\times9$ k-points in the scf calculation and $18\times9\times9$ k-points in the [1, 0, 0] g-direction (and likewise for the [0, 1, 0] and [0, 0, 1] g-directions) with 15 strings per k-point for Bi_2MoO_6. †We used $9\times9\times9$ k-points in the scf calculation and $18\times27\times18$ k-points with 15 strings per k-point for the [0, 1, 0] g-direction of CrO_3.

2.7.5 Convergence studies with respect to number of k-points for QUANTUM ESPRESSO

For the standard ferroelectrics, $BaTiO_3$, $KNbO_3$, and $PbTiO_3$, we see that the values of effective polarization are converged within 0.5 $^{\mu C}/_{cm^2}$, the minimal singular values from all SVD decompositions of the overlap matrices are converged within 0.02 and the maximum eigenvalues are converged within and for k-point meshes ranging from 125 k-points ($5\times5\times5$) to 1331 k-points ($11\times11\times11$) (see Fig. 2.9). While the maximum eigenvalues decrease with additional k-points, even for sparse k-point grids, the maximum eigenvalues remains below the threshold value of π to ensure that effective polarization is computed along the proper branch.

2.7.6 DFT+U calculations for $BiFeO_3$ for QUANTUM ESPRESSO

U (eV)	Effective polarization using BFD ($\frac{\mu C}{cm^2}$)
0	94.5
3	93.3
5	92.7

Table 2.5: Dependence of effective polarization values for $BiFeO_3$ on effective U values.

Due to the multiferroic nature of $BiFeO_3$ [51], we used DFT+U to calculate the wavefunctions in QUANTUM ESPRESSO, and the dependence of the effective polarization value on U is shown in Table 2.5. In [51], they use a U_{eff} = 2 eV with LSDA+U and report an effective polarization value of 95.05 $^{\mu C}/_{cm^2}$.

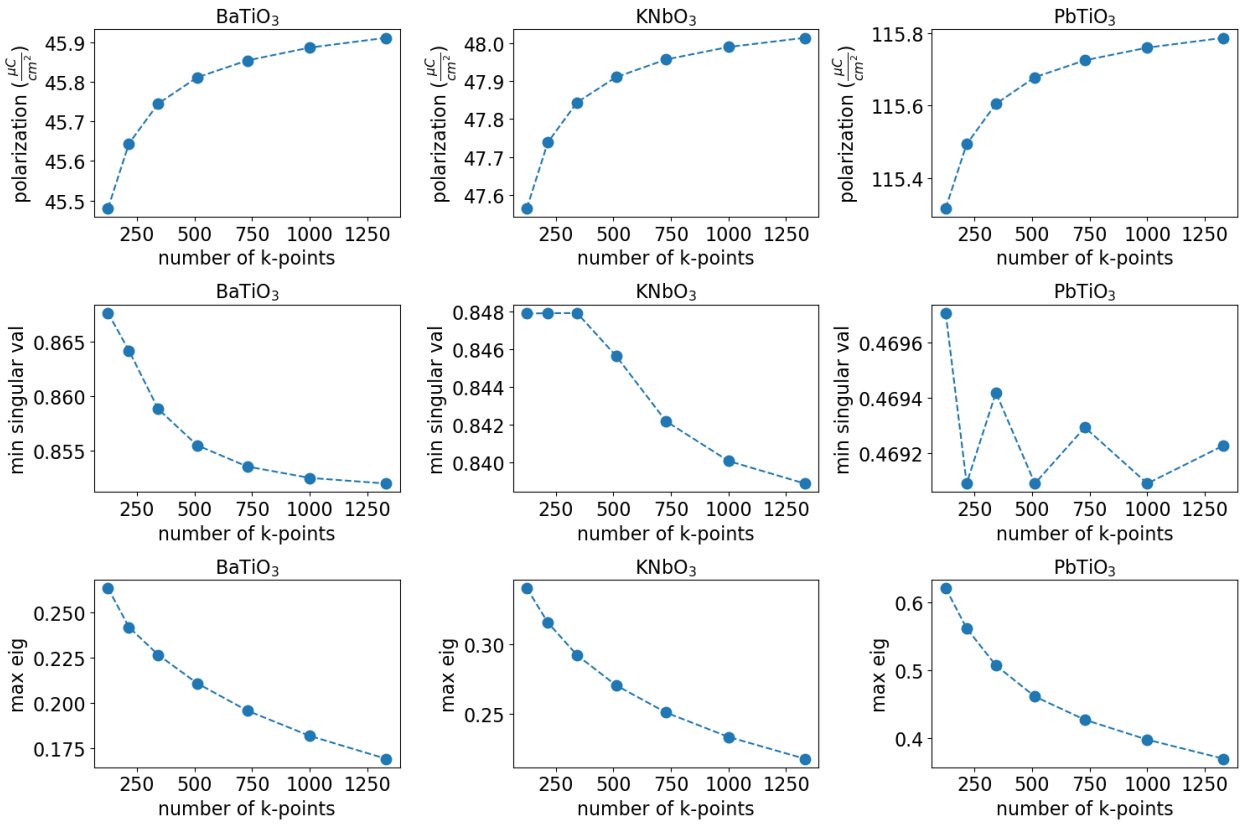

Figure 2.9: Convergence of effective polarization values computed using the Berry flux diagonalization method with respect to k-point grid. All calculations were carried out using the translation between polar structure and nonpolar reference structure that minimized the maximal atomic displacement, and we used k-grids ranging from $5 \times 5 \times 5$ to $11 \times 11 \times 11$, with k-grids increasing by an integer along each direction.

2.8 Additional Supporting Results

2.8.1 Parallel results for exploring translation and interpolation heuristics in Quantum ESPRESSO

We perform the same set of experiments in Quantum ESPRESSO as described in Sec. 2.3 and report the results here.

In Fig. 2.10, we reproduce the behavior observed in VASP for the artificial $BaTiO_3$ test system (Sec. 2.3.1). Specifically, increasing the ionic displacement of the central Ti atom leads to a decrease in the minimal singular value of the overlap matrices and an increase in

the maximal eigenvalues.

In Fig. 2.11, applying rigid translations to $BaTiO_3$, $KNbO_3$, $PbTiO_3$, and $LiNbO_3$ using QUANTUM ESPRESSO yields trends consistent with those obtained with VASP (2.3). For $BaTiO_3$, $KNbO_3$, and $PbTiO_3$, numerical stability is maximized for translations that minimize the maximal atomic displacement, and the effective polarization computed via Berry flux diagonalization agrees with the values obtained from manual interpolation-based calculations in QUANTUM ESPRESSO (see Sec. 2.7.4 for the branch calculations). For $LiNbO_3$, however, the translation that minimizes the maximal atomic displacement results in poor numerical stability and inaccurate effective polarization values, indicating the need to use interpolated structures to compute effective polarization along the proper branch.

Fig. 2.12 shows that increasing the number of interpolated structures improves stability for CrO_3 and $LiNbO_3$, increasing minimal singular values and reducing maximal eigenvalues. For $LiNbO_3$, sufficient interpolations are essential for accurate effective polarization, whereas CrO_3 converges even with poor stability. These results confirm that the heuristics we develop are conservative: they flag potentially inaccurate effective polarization values but are not strictly required to obtain correct results in all cases.

2.8.2 High-throughput calculation results in VASP

We present in Table 2.6 the effective polarization values computed using our automated implementation of Berry flux diagonalization with the reference effective polarization values computed with the same polar and nonpolar relaxed structures using the interpolation-based approach in [31]. All calculations were done in VASP v6.4.2 with the default settings for static calculations provided by the Materials Project [62] and ATOMATE2 [61], with the that exception we use PBE v.54 to maintain consistency with the polarization values computed in [31].

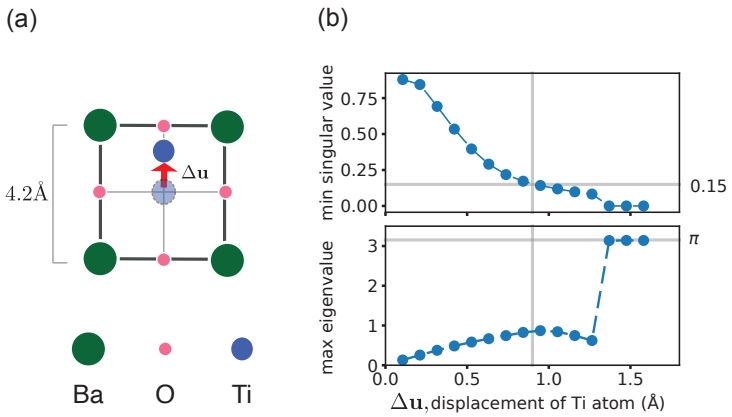

Figure 2.10: (a) Schematic of cubic perovskite $BaTiO_3$ system with fixed Ba and O atoms and a Ti atom that is artificially displaced along the [001] direction (red arrow). The displacement of the Ti atom, $\mathbf{\Delta u}$, is the difference between its artificial position and its centrosymmetric position. (b) Minimal singular values of the overlap matrices $M^{\langle i,i+1\rangle}$ and maximal eigenvalues of the unitary evolution matrices $U_{\mathcal{P}}$ corresponding to the effective polarization calculated using Berry flux diagonalization for the $BaTiO_3$ with the artificially displaced Ti atom.

material	mp-ID	FE DB effective polarization ($\mu C/cm^2$)	BFD effective polarization ($\mu C/cm^2$)	# interps (BFD)	min. singular value	max. eigenvalue
BeO	mp-2542	121.5	121.5	1	0.85	0.21
Y_3TlF_{10}	mp-1207844	55.7	55.6	4	0.64	0.33
$AlHO_2$	mp-625054	24.1	24.1	0	0.78	0.19
$BaNa(B_3O_5)_3$	mp-17864	11.7	11.7	1	0.39	0.23
H_2O	mp-697111	25.8	25.9	1	0.96	0.05
AsF_3	mp-28027	69.0	69.0	1	0.65	0.31
$Li_2Si_2O_5$	mp-4117	84.1	84.1	1	0.81	0.16
Li_2SiO_3	mp-5012	77.9	77.9	1	0.86	0.16
$CsSO_3$	mp-561681	10.9	10.9	2	0.73	0.14
$Na_5B_2P_3O_{13}$	mp-557836	42.7	42.7	4	0.35	0.26
$Ga_3Pb_5F_{19}$	mp-556117	16.6	16.5	2	0.71	0.24

material	mp-ID	FE DB effective polarization ($\mu C/cm^2$)	BFD effective polarization ($\mu C/cm^2$)	# interps (BFD)	min. singular value	max. eigenvalue
LiSiNO	mp-6015	110.3	110.2	1	0.87	0.22
$Zn(PO_3)_2$	mp-558930	26.4	26.4	4	0.37	0.15
$ScHO_2$	mp-625150	22.1	22.2	0	0.78	0.19
Na_3HfF_7	mp-34579	26.0	26.0	2	0.51	0.16
$KLiSi_2O_5$	mp-558102	69.5	69.5	1	0.64	0.30
$BaZnF_4$	mp-3881	12.2	12.2	1	0.67	0.16
$Sr_7Y_6OF_{30}$	mp-1218554	10.3	10.3	2	0.79	0.22
$K_2Si_2O_5$	mp-29825	24.2	24.2	3	0.79	0.23
$NaSb_3F_{10}$	mp-27573	50.4	50.4	3	0.72	0.14
RbH_3O_2	mp-28264	28.1	28.1	0	0.59	0.45
LiI	mp-570935	21.9	21.9	1	0.82	0.23
$CsAsF_4$	mp-5707	21.6	21.6	1	0.80	0.30
$NaHSeO_4$	mp-1220802	14.9	14.9	0	0.80	0.22
$NaAlO_2$	mp-9212	77.9	77.9	1	0.61	0.21
Li_2GeO_3	mp-15349	75.8	75.8	1	0.64	0.17
$TePb_3Cl_4O_3$	mp-1196950	32.8	32.7	3	0.77	0.48
$TaTlWO_6$	mp-1217830	14.1	14.0	0	0.55	0.39
Al_2PbO_4	mp-21892	96.8	96.8	1	0.59	0.23
Ga_2PbO_4	mp-20496	99.3	99.2	1	0.42	0.26
$LiTaO_3$	mp-3666	57.8	57.6	1	0.84	0.19
$Rb_5Nb_3OF_{18}$	mp-540995	11.2	11.0	2	0.70	0.19
$KNaGeO_3$	mp-558085	39.7	39.7	0	0.14	0.35
Al_2CdBr_8	mp-1214910	32.4	32.4	5	0.57	0.17
$Cd(GaCl_4)_2$	mp-1213982	36.2	36.2	3	0.26	0.17

material	mp-ID	FE DB effective polarization ($\mu C/cm^2$)	BFD effective polarization ($\mu C/cm^2$)	# interps (BFD)	min. singular value	max. eigenvalue
$Ta_2Pb_2O_7$	mp-755663	40.6	40.3	0	0.61	0.38
$KNbWO_6$	mp-1223396	25.5	25.2	1	0.87	0.20
Li_2MgGeO_4	mp-1210797	86.6	86.6	1	0.50	0.18
$Rb_2Cd(IBr)_2$	mp-1103709	20.4	20.4	3	0.53	0.22
$KNbSi_2O_7$	mp-560407	22.5	22.0	0	0.82	0.18
$Ca_4YB_3O_{10}$	mp-1214020	28.1	28.0	3	0.72	0.18
$TlSbWO_6$	mp-1216623	11.2	11.2	0	0.70	0.37
$Ca_2Nb_2O_7$	mp-13979	22.6	22.1	1	0.75	0.28
$ZnBi_6P_2O_{15}$	mp-645375	35.1	35.0	1	0.50	0.26
$Bi_4Br_2O_5$	mp-23544	24.6	24.5	1	0.59	0.19
$BaTi(IO_3)_6$	mp-1227524	10.4	10.4	2	0.77	0.24
Li_3AsO_4	mp-9197	89.6	89.6	1	0.84	0.23
$BaCa(GaO_2)_4$	mp-561446	85.5	85.5	1	0.43	0.22
$CaTa_2Bi_2O_9$	mp-556697	40.9	40.9	1	0.59	0.21
$Ba_5Re_3NO_{18}$	mp-555095	23.3	23.4	3	0.80	0.17
$KYSiS_4$	mp-867328	10.1	10.1	1	0.84	0.20
$AgIO_3$	mp-27384	19.1	19.1	1	0.61	0.22
$TiBi_2O_5$	mp-752676	35.1	35.1	1	0.72	0.19
$BaAl_4S_7$	mp-8258	45.2	45.2	1	0.82	0.30
$NaNbSe_2O_7$	mp-1190450	14.0	13.6	0	0.65	0.17
$RbSbMoO_6$	mp-1219606	28.5	28.2	0	0.80	0.30
$NaLi_2AsO_4$	mp-9066	81.0	81.0	1	0.68	0.22
$Rb_2Mo_3SeO_{12}$	mp-1198721	68.4	67.8	3	0.77	0.36
TiI_2O_7	mp-1205304	12.9	12.9	0	0.87	0.17

material	mp-ID	FE DB effective polarization ($\mu C/cm^2$)	BFD effective polarization ($\mu C/cm^2$)	# interps (BFD)	min. singular value	max. eigenvalue
$NaGaO_2$	mp-3338	73.3	73.3	1	0.45	0.20
$SrMgH_4$	mp-643009	18.2	18.2	1	0.74	0.21
$CsGeCl_3$	mp-22988	22.5	22.4	0	0.31	0.63
Na_3VO_4	mp-755436	69.0	68.9	1	0.57	0.71
$SeOF_2$	mp-27367	85.7	85.6	2	0.75	0.32
Na_3PSO_3	mp-542136	41.4	41.5	1	0.85	0.27
Li_3VO_4	mp-19219	86.6	86.7	1	0.76	0.45
$LiAlS_2$	mp-1106183	60.3	60.3	1	0.83	0.32
GaClO	mp-1212662	50.0	49.9	1	0.48	0.16
Mg_2PN_3	mp-3933	91.5	91.4	1	0.85	0.34
BrF_5	mp-27987	25.3	25.3	1	0.79	0.31
$CsGeBr_3$	mp-1068340	20.0	19.9	0	0.38	0.69
$Nb_3Al(PO_4)_6$	mp-1220763	88.3	88.2	3	0.70	0.25
$TiSb(PO_4)_3$	mp-1239394	83.2	83.2	3	0.72	0.24
$Sr_2Nb_2O_7$	mp-3870	38.9	38.1	1	0.64	0.19
$NbTlTeO_6$	mp-1220473	25.2	24.7	0	0.55	0.30
Sc_2TiO_5	mp-753401	35.3	35.2	2	0.77	0.16
$KYGeS_4$	mp-867334	11.0	10.9	1	0.81	0.23
Ga_2S_3	mp-539	60.9	60.7	2	0.25	0.41
Al_2Se_3	mp-11674	54.7	54.6	2	0.76	0.41
$Ti_3P_6WO_{24}$	mp-775198	87.8	87.8	3	0.70	0.24
Tl_3AsO_4	mp-15573	43.9	43.9	1	0.77	0.23
$TiNb(PO_4)_3$	mp-1216943	87.0	87.0	3	0.69	0.25
$TaTi(PO_4)_3$	mp-1244815	86.5	86.5	3	0.69	0.25

material	mp-ID	FE DB effective polarization ($\mu C/cm^2$)	BFD effective polarization ($\mu C/cm^2$)	# interps (BFD)	min. singular value	max. eigenvalue
$TlSnPS_4$	mp-6057	15.1	15.1	1	0.74	0.20
BrF_3	mp-23297	30.7	30.7	2	0.77	0.15
$PSeCl_9$	mp-662547	13.0	13.0	2	0.80	0.34
$BiIO_4$	mp-1191266	77.7	77.6	1	0.73	0.43
Li_7BiO_6	mp-38487	15.1	15.2	2	0.81	0.14
$KLiZn_3O_4$	mp-1223469	58.2	58.3	3	0.38	0.44
$SnPS_3$	mp-13923	32.0	31.9	1	0.73	0.30
$K_2Sn_2Hg_3S_8$	mp-18115	45.7	45.5	3	0.50	0.40
$SrNb_2Bi_2O_9$	mp-23614	57.7	57.1	1	0.68	0.30
$BaGa_4S_7$	mp-4309	43.9	43.9	1	0.52	0.39
$LiAlSe_2$	mp-7117	54.1	54.0	1	0.74	0.45
$RbMoIO_6$	mp-554787	55.6	55.5	2	0.78	0.25
$LiGaS_2$	mp-3647	59.8	59.7	1	0.48	0.51
Sr_2SnS_4	mp-30294	58.0	58.0	3	0.46	0.31
Bi_2WO_6	mp-23480	53.4	53.3	1	0.78	0.24
$MgGeN_2$	mp-7798	123.1	122.9	1	0.52	0.34
$LiInS_2$	mp-1188392	54.6	54.5	1	0.68	0.41
ClF_5	mp-1213763	31.2	31.2	1	0.73	0.39
$Bi_{10}(MoO_8)_3$	mp-558215	10.6	10.6	4	0.74	0.29
$Li_2Mo_3S_4$	mp-675779	14.4	14.4	1	0.70	0.44
$TaZn_2N_3$	mp-1029387	118.9	118.9	1	0.50	0.40
KNa_2CuO_2	mp-545359	22.8	22.9	1	0.58	0.11
$SnPSe_3$	mp-570370	24.7	24.6	0	0.64	0.57
$Ga_2Sn_2S_5$	mp-14280	21.4	21.5	6	0.26	0.39

material	mp-ID	FE DB effective polarization ($\mu C/cm^2$)	BFD effective polarization ($\mu C/cm^2$)	# interps (BFD)	min. singular value	max. eigenvalue
Bi_2TeO_5	mp-23334	59.6	59.5	3	0.76	0.22
Li_2CdGeO_4	mp-7688	82.3	82.2	1	0.35	0.19
Ag_2HgSI_2	mp-556866	11.7	11.6	1	0.73	0.24
$RbGeBr_3$	mp-28558	11.2	11.1	0	0.10	0.65
$NbZn_2N_3$	mp-1029422	115.8	116.0	1	0.51	0.40
$NbTlBr_4O$	mp-551826	17.8	17.4	0	0.78	0.37
$GaSeBr_7$	mp-23376	26.4	26.2	3	0.16	0.32
Mg_2SbN_3	mp-1029378	103.1	103.0	0	0.69	0.47
CdS	mp-672	51.2	51.1	1	0.68	0.32
VZn_2N_3	mp-1029262	120.9	120.8	1	0.48	0.57
$YTl(PSe_3)_2$	mp-1105136	13.5	13.5	4	0.68	0.28
Li_2CdSnS_4	mp-1188784	55.4	55.3	1	0.61	0.52
$LiGaSe_2$	mp-11582	53.9	53.7	1	0.36	0.52
$SrTaNO_2$	mp-1101327	42.0	41.7	0	0.69	0.46
$TlVTeO_5$	mp-639714	61.4	61.4	3	0.66	0.30
$Sr_{14}(SnAs_4)_3$	mp-1192393	20.0	20.0	2	0.69	0.46
Ca_5NbN_5	mp-675101	39.3	39.3	4	0.70	0.20
$TiZnN_2$	mp-1016108	121.1	121.0	0	0.19	0.65
Li_2CdGeS_4	mp-1189383	58.3	58.3	1	0.30	0.51
$CaSnN_2$	mp-1029633	88.1	88.0	1	0.61	0.33
Ag_2GeS_3	mp-9900	49.3	49.3	1	0.49	0.60
$NbBr_2O$	mp-550070	23.4	22.7	0	0.80	0.28
NbI_2O	mp-549720	17.5	16.7	0	0.64	0.32
$NbCl_2O$	mp-1025567	26.9	26.1	0	0.82	0.27

material	mp-ID	FE DB effective polarization ($\mu C/cm^2$)	BFD effective polarization ($\mu C/cm^2$)	# interps (BFD)	min. singular value	max. eigenvalue
$CsGeI_3$	mp-28377	17.6	17.7	0	0.29	0.69
$Ba_2In_2O_5$	mp-20546	26.9	26.9	1	0.58	0.29
MgTe	mp-1039	42.7	42.7	1	0.67	0.43
Li_2ZnGeS_4	mp-1210804	61.6	61.6	1	0.33	0.50
Mg_2VN_3	mp-1029368	99.5	99.5	1	0.74	0.74
Li_3SbS_4	mp-756316	56.6	56.5	1	0.58	0.75
$CsIO_3$	mp-28295	73.4	73.3	1	0.67	0.96
$Ca_{11}AlSb_9$	mp-1214215	17.6	17.5	3	0.51	0.53
$KCu(BiS_2)_2$	mp-558063	14.3	14.2	1	0.50	0.20
$LiVN_2$	mp-1029932	124.8	124.9	1	0.56	0.94
GeTe	mp-938	66.7	68.8	0	0.24	0.92
$Ba_2Ge_2Te_5$	mp-17885	11.5	11.6	4	0.25	0.40
$Hg(IO_3)_2$	mp-23407	36.6	36.4	3	0.80	0.34
Cd_2GaAgS_4	mp-6356	55.1	55.0	1	0.35	0.46
Ca_2CdP_2	mp-1078908	14.7	14.8	2	0.57	0.22
$RbHgSbTe_3$	mp-1219595	14.6	14.8	0	0.28	0.84
Ba_2InBiS_5	mp-864638	28.0	28.1	2	0.33	0.34
ClO_3	mp-22869	43.0	43.0	2	0.66	0.66
Cs_2Te_3	mp-505634	28.5	29.4	1	0.19	1.08
Ca_2CdAs_2	mp-1095180	15.5	15.6	2	0.25	0.28
$NaNbS_6$	mp-1210002	15.4	15.4	3	0.46	0.40
K_2Se_3	mp-7670	40.1	40.1	1	0.47	0.74
SrGaSiH	mp-979137	39.6	39.6	1	0.55	0.34
Li_5SiP_3	mp-685991	31.8	31.8	1	0.53	0.34

material	mp-ID	FE DB effective polarization ($\mu C/cm^2$)	BFD effective polarization ($\mu C/cm^2$)	# interps (BFD)	min. singular value	max. eigenvalue
SrAlGeH	mp-980057	40.0	40.1	1	0.61	0.33
$LiAsS_2$	mp-555874	81.9	81.9	1	0.03	1.19
Sr_2CdAs_2	mp-867203	11.7	11.8	2	0.12	0.34
Cs_2Se_3	mp-7449	34.0*	31.3	1	0.01	1.09
$LiB(SO_4)_2$	mp-1020106	57.1*	57.1	4	0.74	0.33
Li_3AlSiO_5	mp-1020023	85.4*	85.4	1	0.86	0.15
B_4PbO_7	mp-9747	118.2*	118.1	1	0.58	0.41
Li_3AlGeO_5	mp-6765	82.6*	82.7	1	0.53	0.15
Li_3GaSiO_5	mp-16996	84.6*	84.5	1	0.44	0.14
Li_3GaGeO_5	mp-17685	81.8*	81.8	1	0.54	0.16
$Ti(SO_4)_2$	mp-775149	107.6*	107.6	4	0.67	0.36
$Sb_6S_2O_{15}$	mp-27229	57.3*	57.2	2	0.81	0.15
Bi_2MoO_6	mp-23064	78.0*	77.4	1	0.72	0.51
$KAu(IO_3)_4$	mp-557441	135.5*	135.4	4	0.62	1.33
$Ba_2InBiSe_5$	mp-1105202	30.0*	29.7	2	0.23	2.51
$K_4Ba_2SnAs_4$	mp-1224063	6.1*	6.1	2	0.46	0.52
$K_4Ba_2SnSb_4$	mp-1224100	7.5*	8.1	2	0.00	1.66
$K_4Ba_2SnBi_4$	mp-1224198	6.7*	7.5 ($6.5^{\dagger}$)	3 ($9^{\dagger}$)	0.10 ($0.19^{\dagger}$)	2.18 ($2.76^{\dagger}$)
Rb_2Se_3	mp-7447	38.7*	32.7 ($38.5^{\ddagger}$)	1	0.00 ($0.40^{\ddagger}$)	1.19 ($0.75^{\ddagger}$)
$Ba_2InSbSe_5$	mp-1189339	25.8*	29.9 ($27.9^{\ddagger}$)	2	0.12 ($0.19^{\ddagger}$)	2.19 ($2.16^{\ddagger}$)
Ba_2CdAs_2	mp-1079666	14.6	6.1 ($10.1^{\ddagger}$)	2	0.00 ($0.34^{\ddagger}$)	1.21 ($0.47^{\ddagger}$)

material	mp-ID	FE DB effective polarization ($\mu C/cm^2$)	BFD effective polarization ($\mu C/cm^2$)	# interps (BFD)	min. singular value	max. eigenvalue

Table 2.6: The FE DB effective polarization values were computed in [31] using a standard 10 interpolated structures unless denoted with *, which denotes that the reference value for the FE DB effective polarization using the standard interpolation-based approach was computed with 20 interpolated structures. The parenthetical effective polarization, minimum singular, and maximum eigenvalue values denoted with † correspond to corrected values of polarization computed with Berry flux diagonalization for a higher number of interpolations (also given parenthetically) than heuristic number of interpolations. The parenthetical effective polarization, minimum singular, and maximum eigenvalue values denoted with ‡ correspond to corrected values of effective polarization computed with twice the default number of k-points for static calculations in ATOMATE2 [61]. In the main text, the values computed with the original heuristics are plotted in Fig. 2.6, but both original and recomputed values are presented here.

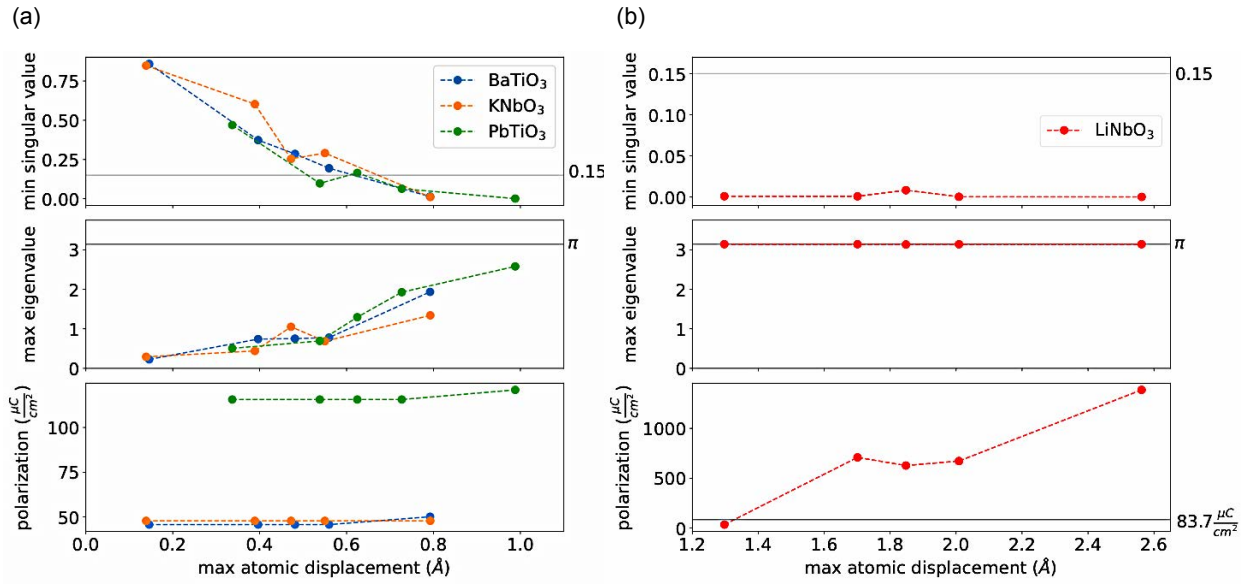

Figure 2.11: Minimal singular values (top), maximal eigenvalues (middle), and effective polarization (bottom) computed from overlap matrices for standard ferroelectrics under different translations of the nonpolar structure relative to the polar structure for (a) $BaTiO_3$, $KNbO_3$, and $PbTiO_3$, and (b) $LiNbO_3$. The translations that minimize the maximal atomic displacement ensure sufficient overlap between wavefunctions at each k-point, with minimal singular values remaining above the threshold of 0.15 and maximal eigenvalues below π. These conditions guarantee proper branch alignment and accurate effective polarization calculations. For $LiNbO_3$, even the translation minimizing maximal atomic displacement results in insufficient overlap, with singular values approaching zero and eigenvalues near π, leading to improper resolution of polarization branches and inaccurate values of effective polarization.

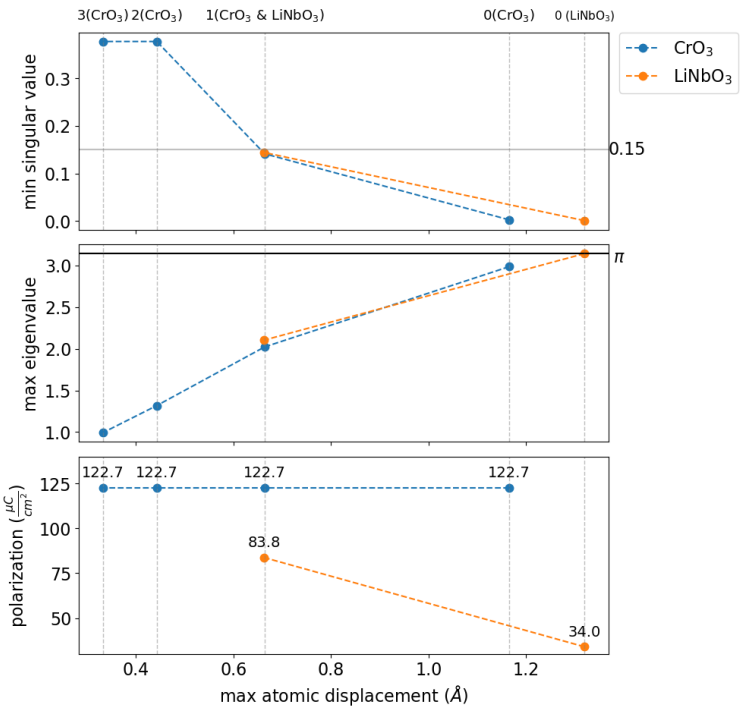

Figure 2.12: Calculated minimum singular values (top), maximal eigenvalues (middle) and effective polarization (bottom) as functions of the maximum atomic displacement for CrO_3, $LiNbO_3$, and computed from different numbers of interpolated structures, where the number of interpolations is labeled on the top x-axis. For points corresponding to calculations with more than one interpolation, the numbers plotted are the maxima of the maximum atomic displacements across the sets of adjacent interpolated structures used in the intermediate steps of the computation, and likewise for the maxima of the maximum eigenvalues and minima of the minium singular values.

CHAPTER 3

TRANS-DIMENSIONAL HAMILTONIAN MODEL SELECTION AND PARAMETER ESTIMATION FROM SPARSE, NOISY DATA

High-throughput characterization often requires estimating parameters and model dimension from experimental data of limited quantity and quality. Such data may result in an ill-posed inverse problem, where multiple sets of parameters and model dimensions are consistent with available data. This ill-posed regime may render traditional machine learning and deterministic methods unreliable or intractable, particularly in high-dimensional, nonlinear, and mixed continuous and discrete parameter spaces. To address these challenges, we present a Bayesian framework that hybridizes several Markov chain Monte Carlo (MCMC) sampling techniques to estimate both parameters and model dimension from sparse, noisy data. By integrating sampling for mixed continuous and discrete parameter spaces, reversible-jump MCMC to estimate model dimension, and parallel tempering to accelerate exploration of complex posteriors, our approach enables principled parameter estimation and model selection in data-limited regimes. We apply our framework to a specific ill-posed problem in quantum information science: recovering the locations and hyperfine couplings of nuclear spins surrounding a spin-defect in a semiconductor from sparse, noisy coherence data. We show that a hybridized MCMC method can recover meaningful posterior distributions over physical parameters using an order of magnitude less data than existing approaches, and we validate our results on experimental measurements. More generally, our work provides a flexible, extensible strategy for solving a broad class of ill-posed inverse problems under realistic experimental constraints.

3.1 Introduction

Estimation problems from sparse and noisy data are often ill-posed inverse problems, where multiple sets of parameters or models may be consistent with observed data, making unique reconstruction of ground-truth quantities nearly impossible. In such cases, the goal is not to recover a single “true" parameter set, but to produce a posterior distribution of probable models and parameters, given a set of data.

The central motivating example for this work arises in the characterization of local spin environments in optically active spin defects in solid-state systems, such as divacancies in silicon carbide [67] and nitrogen-vacancy (NV) centers in diamond [68]. These defects are key candidates for quantum information technologies due to their long coherence times [28, 69] and ability to couple coherently to nearby nuclear spins [70, 71]. The surrounding nuclear spin bath plays a dual role: it is both the primary source of decoherence and a potential quantum resource for memory or register applications, depending on the ability to accurately resolve hyperfine couplings and spatial configurations [34, 72–75]. As scalable fabrication techniques allow for the creation of billions of such spin defects [76, 77], high-throughput, data-efficient methods to characterize the nuclear spin environment of each defect are a prerequisite for the practical deployment of spin-defect-based quantum devices. Beyond characterizing the intrinsic spin environment of defects, related techniques are increasingly being applied to image external spin systems, including molecules or materials placed near a spin-defect based quantum sensor, as part of a broader effort to achieve nanoscale magnetic resonance imaging [78].

Current approaches to estimating the parameters of the Hamiltonian describing the interaction between a central spin and neighboring nuclear spins, such as correlation spectroscopy, offer high-resolution insights but require long acquisition times and extensive manual tuning [35, 79], making them infeasible at scale. Recently proposed machine learning methods offer some automation [80, 81], but rely on experimental datasets that are expensive to obtain and

rely on a large amount of computing resources. Hence, a framework that can operate in the sparse, noisy data regime and still extract physically meaningful Hamiltonian parameters is desirable. Furthermore, since the number of nuclear spins interacting with the spin-defect is unknown *a priori* and is itself a quantity to be estimated, the inverse problem at hand is a model dimension selection problem in addition to a Hamiltonian parameter estimation problem.

Existing approaches to similar inverse problems aim to design experimental protocols or to optimize measurement strategies to escape the ill-posed regime by collecting more informative data [82–85]. In contrast, here we assume a fixed, potentially sparse dataset, and we focus on robust inference under these constraints. Bayesian estimation techniques provide a natural language for this setting, allowing one to incorporate prior knowledge, account for measurement uncertainty, and explore the full posterior landscape of plausible models. However, performing this kind of inference efficiently—particularly in the presence of both discrete and continuous parameters, unknown model dimensionality, and high-dimensional loss landscapes riddled with local minima—remains a significant technical challenge.

To address these challenges simultaneously, we present a hybridized Markov chain Monte Carlo (MCMC) framework that combines several specialized techniques to address separately the mathematical challenges that arise. Discrete and continuous random-walk Metropolis-Hastings (RWMH) samplers are used to explore parameter spaces with both continuous and discrete parameters. Trans-dimensional model selection is handled through the reversible-jump MCMC (RJMCMC) algorithm [86], which allows for sampling over models with differing numbers of parameters, a necessity in problems where the model dimension (e.g., the number of nuclear spins) is itself unknown. To better sample from complex and multimodal posteriors in high-dimensional spaces, we incorporate parallel tempering (PT) [87], which improves convergence and avoids trapping in local minima. Together, these methods form a generalized, flexible, and extensible framework to solve inverse problems that can operate

robustly in the data-scarce regime. We demonstrate that our hybridized MCMC framework produces posterior distributions that describe simulated nuclear spin baths in the sparse, noisy data regime, and we validate our approach with experimental data to demonstrate that it achieves an accuracy comparable to much more computationally and experimentally intensive procedures, including direct machine learning approaches [80, 81], but using an order of magnitude fewer data points.

The remainder of this chapter is organized as follows. In Sec. 3.2, we formalize the class of inverse problems we target and discuss the key challenges. In Sec. 3.3, we describe the MCMC algorithms and how they are hybridized to form a cohesive inference engine. In Sec. 3.4, we apply the method to a specific example: recovering the structure and parameters of a spin environment from sparse coherence data. In Sec. 3.5, we assess performance on both simulated and experimental data. Finally, in Sec. 3.6, we reflect on the capabilities and limitations of this framework and suggest directions for future work.

3.2 Inverse Problem

3.2.1 General formulation

We consider some physical process that is governed by a family of Hamiltonians $\mathcal{H} \in \{\mathcal{H}_{k_1}(\mathbf{a}_{k_1}), \mathcal{H}_{k_2}(\mathbf{a}_{k_2}), \ldots, \mathcal{H}_{k_j}(\mathbf{a}_{k_j})\}$, where each Hamiltonian $\mathcal{H}_{k_i}$ is parametrized by k_i variables given by $\mathbf{a}_{k_i} = (a_{k_1}, a_{k_2}, \ldots, a_{k_i})$. We consider a sparse set of data $\mathbf{d}_d \in \mathbb{R}^d$ that is generated by a forward model $\mathbf{d}_d = f(\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c}) + \epsilon_d$ where the forward model $f(\cdot)$ depends on the parameterized Hamiltonian $\mathcal{H}_k(\mathbf{a}_k)$, unknown parameters $\mathbf{b}$, known parameters $\mathbf{c}$, and ϵ_d is a noise model. Examples of unknown parameters represented by $\mathbf{b}$ include parameters that are not directly captured by Hamiltonian terms but are nevertheless crucial to reproduced experimental data, such as experimental setup-dependent calibration and sensitivity effects or emergent effects that may not be directly accounted for in the Hamiltonian.

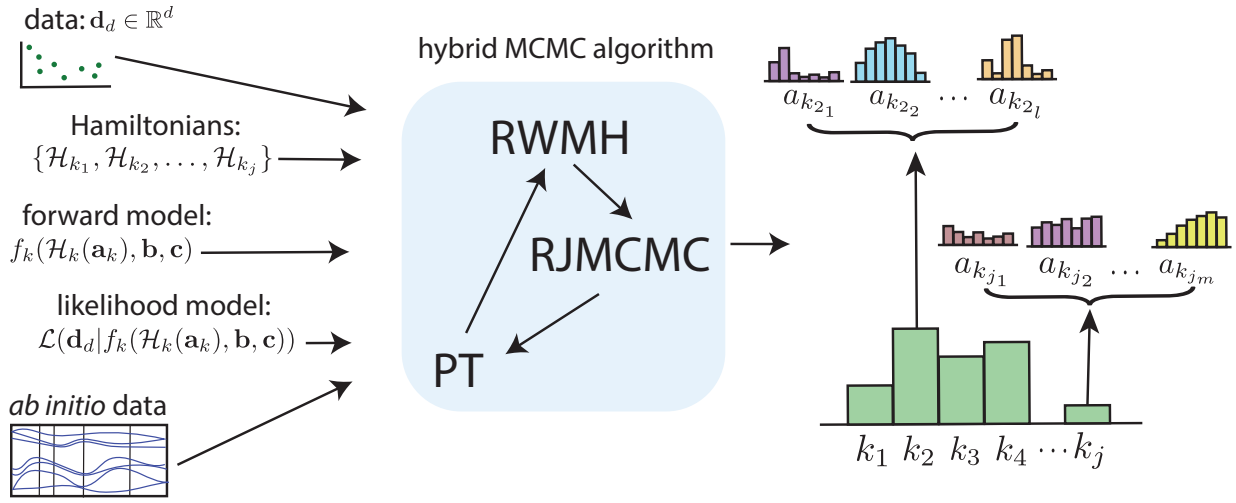

Figure 3.1: **Schematic of hybrid MCMC workflow.** The input (left) to the hybrid MCMC algorithm is a fixed set of data $\mathbf{d}_d \in \mathbb{R}^d$, a family of candidate Hamiltonians $\{\mathcal{H}_{k_1}, \mathcal{H}_{k_2}, \ldots, \mathcal{H}_{k_j}\}$ which are parameterized by $\mathbf{a}_{k_i}$ where the number of parameters for different Hamiltonians $\mathcal{H}_{k_i}$ and $\mathcal{H}_{k_j}$ can be different. We also take in a forward model $f_k(\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c})$ that generates a set of data from candidate Hamiltonian $\mathcal{H}_k$, and a likelihood model $\mathcal{L}(\mathbf{d}_d | f_k(\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c}))$ that quantifies the probability of observing the data $\mathbf{d}_d$ given a set of model parameters, and any priors, which may come in the form of *ab initio* data. The output for the hybrid MCMC algorithm is a probability distribution over the different Hamiltonian models, and for each Hamiltonian model we have a posterior distribution over the values of the parameters.

Examples of known parameters modeled by $\mathbf{c}$ may include known experimental settings, such as magnetic field or temperature. Formally, the inverse problem we seek to resolve is:

$$\arg \min_{k, \mathbf{a}_k, \mathbf{b}} ||f(\mathcal{H}_k(\mathbf{a}_k)), \mathbf{b}, \mathbf{c}) - \mathbf{d}_d||_{N(\epsilon_d)}. \tag{3.1}$$

That is, we seek to find the candidate model $\mathcal{H}_k$, indexed by k, and the parameters $\mathbf{a}_k$ and $\mathbf{b}$ that minimize the norm of the residual between some forward model f applied to the Hamiltonian of dimension k with parameters $\mathbf{a}_k$ and $\mathbf{b}$ and some $d-$dimensional data, where the norm $N(\epsilon_d)$ depends on underlying assumptions on the noise model ϵ_d. The assumptions on the noise model impact the appropriate likelihood function $\mathcal{L} = \mathbb{P}(\mathbf{d} | \mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c})$, which is the probability of observing data d given $\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}$, and $\mathbf{c}$. Without loss of generality, we

assume that the dimension of $\mathbf{a}_k$ is variable and the dimension of $\mathbf{b}$ is fixed, but we note that the same strategies that are applied to account for the variable number of dimensions of $\mathbf{a}_k$ may be applied to $\mathbf{b}$.

Furthermore, we acknowledge that "sparse" and "noisy" are quantities that are context-dependent, and we generally consider data $\mathbf{d}_d$ to be "sparse" and noisy if there are multiple parameterizations k_i, k_j and $(\mathbf{a}_{k_i}, \mathbf{b}_i), (\mathbf{a}_{k_j}, \mathbf{b}_j)$ such that

$$\min_{k_i, \mathbf{a}_{k_i}} ||f(\mathcal{H}_{k_i}(\mathbf{a}_{k_i}), \mathbf{b}_i, \mathbf{c}) - \mathbf{d}_d||_{N(\epsilon_d)} \approx \min_{k_j, \mathbf{a}_{k_j}} ||f(\mathcal{H}_{k_j}(\mathbf{a}_{k_j}), \mathbf{b}_j, \mathbf{c}) - \mathbf{d}_d||_{N(\epsilon_d)} \tag{3.2}$$

for a fixed set of data $\mathbf{d}_d$. That is, the sparsity and noise regimes we consider here are with respect to the ill-posedness of a specific inverse problem.

Since we will focus on the regime where the inverse problem is ill-posed, we take a Bayesian approach, as we know that we will not be able to exactly recover k, $\mathbf{a}_k$, and $\mathbf{b}$, but we will instead recover posterior distributions over k, $\mathbf{a}_k$, and $\mathbf{b}$. That is, instead of a single k and $\mathbf{a}_k$, we will recover a set of $\{k_l, k_m, \ldots, k_n\}$ with corresponding probabilities $\{p_{k_l}, p_{k_m}, \ldots, p_{k_n}\}$ (where $\sum_j p_{k_j} = 1$), with the interpretation the data $\mathbf{d}_d$ is generated by $\mathcal{H}_{k_i}$ with probability p_i. For each model class k_i, we will also obtain posterior distributions over the parameters $\{\mathbf{a}_{k_{i_1}}, \mathbf{a}_{k_{i_2}}, \ldots \mathbf{a}_{k_{ip}}\}$ with probabilities $\{p_{k_{i_1}}, p_{k_{i_2}}, \ldots, p_{k_{ip}}\}$ and $\sum_j p_{k_{i_j}} = 1$, and likewise for $\mathbf{b}$ (see Fig. 3.1).

3.2.2 Hybridized MCMC-based Hamiltonian model selection and parameter estimation

Markov chain Monte Carlo (MCMC) algorithms provide a robust framework for exploring complex, high-dimensional parameter spaces by generating a sequence of samples that approximate a target probability distribution. In general, these algorithms initialize a set of walkers across the domain of a parameter space and propose updates to the parameters based

on a specified transition rule. The acceptance or rejection of a proposed move is governed by whether the likelihood (i.e., the probability of observing some data given a set of parameters) increases for an updated set of parameters. After an initial burn-in period, the sequence of walker positions represents samples from the posterior distribution, which provides the most probable parameter values given the observed data.

As shown in Fig. 3.1, the hybridized MCMC algorithm takes as input a fixed dataset $\mathbf{d}_d \in \mathbb{R}^d$, a set of candidate Hamiltonian models $\{\mathcal{H}_{k_1}, \mathcal{H}_{k_2}, \ldots, \mathcal{H}_{k_j}\}$, each parameterized by a potentially different number of parameters $\mathbf{a}_{k_i}$, and a forward model $f_k(\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c})$ that simulates data for a given Hamiltonian. The algorithm also requires a likelihood function $\mathcal{L}(\mathbf{d}_d \mid f_k(\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c}))$ to quantify how well the simulated data from a candidate model match the measured data, as well as any available prior information, such as constraints from *ab initio* calculations. The algorithm produces as outputs posterior distributions over the candidate Hamiltonian models and, for each model, a posterior distribution over its associated parameter values.

We address the challenge of recovering parameters that may be drawn from different domains (i.e., the set or distribution from which parameters are sampled) by partitioning the parameter space into subsets, $(\mathbf{a}_k, \mathbf{b}) = (a_1, \ldots, a_i | a_{i+1}, \ldots, a_k | b_1, \ldots, b_j)$, where different partitions can follow either discrete distributions based on precomputed or experimental datasets or continuous distributions suited for a broader search, and subsets of variables can also have constant or varying dimensions. This partitioning enables efficient sampling by tailoring MCMC strategies to different types of parameters. The hybrid algorithm includes different MCMC strategies appropriate for each subset of parameters. Since the approach accommodates arbitrary likelihood functions and arbitrary forward models (see Fig. 3.1), our framework is designed for extensibility across a variety of physical systems and experimental setups; although a few key MCMC algorithms for dealing with varying dimensionality and efficient exploration of non-linear, non-convex landscapes have already been implemented,

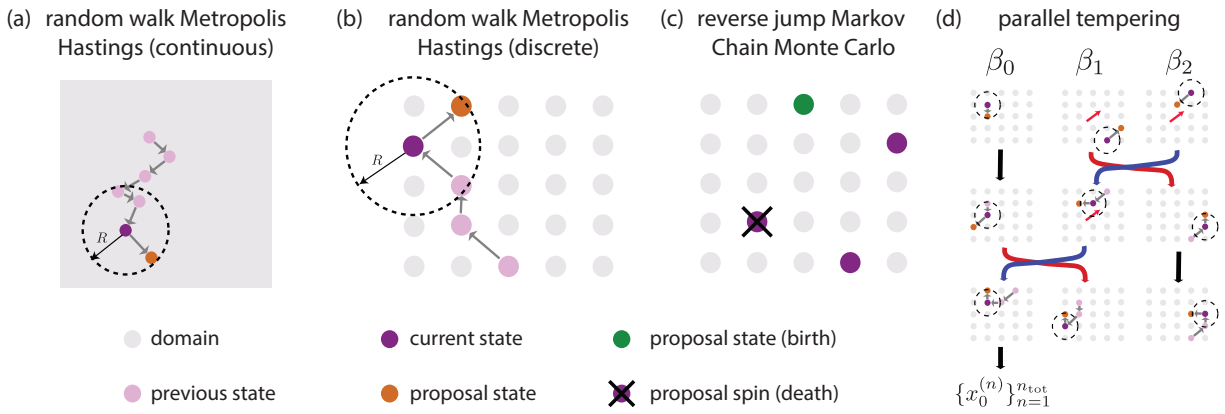

Figure 3.2: **Schematic of sampling algorithms** (a) random walk Metropolis Hastings in over a continuous domain (Alg. 1), (b) random walk Metropolis Hastings over a discrete domain (Alg. 1), (c) reverse jump Markov chain Monte Carlo (Alg. 2), and (d) parallel tempering (Alg. 3).

additional MCMC-based algorithms to evolve subsets of given parameters may be combined and implemented.

3.3 Algorithms

In this section, we elaborate on details of the implementation of the hybrid algorithm and provide pseudocode for the sub-algorithms of the hybridized MCMC procedure (Fig. 3.2). For details on these Monte-Carlo based methods, we refer the reader to [88, 89]. While many of these algorithms are standard within the MCMC literature, we present them here with consistent notation to demonstrate how they can be hybridized with each other for different types of continuous and discrete parameters and how they can be used for specific sampling tasks.

Algorithm 1 Random walk Metropolis Hastings (RWMH)

Input: data $\mathbf{d}_d$, partition of a fixed-length set of parameters to update $\mathbf{p} = \{p_1, \ldots, p_m\}$ and fixed parameters $\mathbf{q} = \{q_1, \ldots, q_s\}$, with $\mathbf{p} \cup \mathbf{q} = \{a_1, \ldots, a_k, b_1, \ldots, b_j\}$ and $\mathbf{p} \cap \mathbf{q} = \emptyset$ and their initial values $\mathbf{p}^{(0)} = (p_1^{(0)}, p_2^{(0)}, \ldots, p_m^{(0)})$ and $\mathbf{q}^{(0)} = (q_1^{(0)}, q_2^{(0)}, \ldots, q_s^{(0)})$, known parameters $\mathbf{c}$, fixed Hamiltonian model $\mathcal{H}_k(\mathbf{a}_k)$ indexed by k, forward model $f_k(\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c}) = f_k(\mathbf{p}, \mathbf{q}, \mathbf{c})$, likelihood model $\mathcal{L}(\mathbf{d}_d | f_k(\mathbf{p}, \mathbf{q}, \mathbf{c}))$, domain $\mathcal{D}$, proposal kernel $r_{R,\mathcal{D}}(x, z)$, radius R, number of steps N_{step}

1: **for** $n = 0, \ldots N_{\text{step}} - 1$ **do**

2: **for** $l = 1, \ldots, m$ **do**

3: **Proposal step**: Sample $p_l^* \sim r_{R,\mathcal{D}}(p_l^{(n)}, \cdot)$;
$\mathbf{p}^* = \{p_1^{(n)}, \ldots, p_{l-1}^{(n)}, p_l^*, p_{l+1}^{(n)}, \ldots, p_m^{(n)}\}$

4: **Accept/reject step**: Update

$$\mathbf{p}^{(n+1)} = \begin{cases} \mathbf{p}^* & \text{with probability } \alpha(\mathbf{p}^{(n)}, \mathbf{p}^*) = \min\{1, \frac{\mathcal{L}(\mathbf{d}_d | f_k(\mathbf{p}^*, \mathbf{q}, \mathbf{c}))}{\mathcal{L}(\mathbf{d}_d | f_k(\mathbf{p}^{(n)}, \mathbf{q}, \mathbf{c}))} \frac{r_{R,\mathcal{D}}(p_l^*, p_l^{(n)})}{r_{R,\mathcal{D}}(p_l^{(n)}, p_l^*)}\} \\ \mathbf{p}^{(n)} & \text{with probability } 1 - \alpha(\mathbf{p}^{(n)}, \mathbf{p}^*) \end{cases}$$

$\mathbf{q}^{(n+1)} = \mathbf{q}^{(n)}$

Output: Sequence of updated parameters: $\{\mathbf{p}^{(n)}\}_{n=1}^{N_{\text{step}}} = (\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \ldots \mathbf{p}^{(N_{\text{step}})})$ and sequence of fixed parameters: $\{\mathbf{q}^{(n)}\}_{n=1}^{N_{\text{step}}} = (\mathbf{q}^{(1)}, \mathbf{q}^{(2)}, \ldots \mathbf{q}^{(N_{\text{step}})}) = (\mathbf{q}^{(0)}, \mathbf{q}^{(0)}, \ldots, \mathbf{q}^{(0)})$

3.3.1 Random walk Metropolis-Hastings (RWMH) for continuous and discrete domains

We employ a random walk Metropolis-Hastings (RWMH) algorithm to sample a subset of unknown parameters with fixed dimension, allowing for both continuous and discrete domains (see Alg. 1). Specifically, we partition the set of unknown parameters into $\mathbf{p} = \{p_1, \ldots, p_m\}$, which are updated during the sampling, and $\mathbf{q} = \{q_1, \ldots, q_s\}$, which remain fixed. At each iteration, a single parameter p_l^* is proposed from a domain-aware proposal distribution $r_{R,\mathcal{D}}(p_l^{(n)}, \cdot)$, where the domain $\mathcal{D}$ and the radius R define the step size and allowable transitions. The proposal mechanism is flexible, permitting updates in either a discrete or continuous setting, depending on the nature of the domain $\mathcal{D}$ from which the parameters are sampled (see Fig. 3.2a & b). For continuous variables, this can be a closed set

over a fixed interval, in which case there is a standard interpretation of a random walk. For a discrete domain $\mathcal{D}$, there may be cases, such as selecting different configurations of spins over lattice sites that have real-space coordinates, where there is still a standard interpretation of a random walk in real space. For discrete sets without parameters that have a natural mapping to real space, such as possible energy transitions or interaction strengths, one may need a problem-specific interpretation of a random walk to use RWMH. The proposed state $\mathbf{p}^*$ (containing p_l^*) is then accepted or rejected based on the Metropolis criterion, ensuring detailed balance. The likelihood function $\mathcal{L}$ governs the acceptance probability, leveraging a forward model $f(\mathbf{p}, \mathbf{q}, \mathbf{c})$ to compare simulated and observed data. Although this is a standard MCMC algorithm, we emphasize the practical importance of the built-in partitioning of parameters in the implementation described in Alg. 1; such partitioning is particularly relevant in a Bayesian approach to Hamiltonian parameter estimation problems for which prior information on certain parameters may be available from *ab initio* calculations or experimental data.

3.3.2 Reverse Jump Markov Chain Monte Carlo (RJMCMC)

In the previous algorithm (RWMH), the parameters are partitioned into a set to be updated and a set to remain fixed, but the dimensions of the updated set are fixed. To account for scenarios in which the number of parameters must be estimated in addition to the parameters themselves, we use a MCMC-based technique known as "reverse jump" Markov chain Monte Carlo, which is a method for trans-dimensional Bayesian model selection (see Alg. 2) [90]. The RJMCMC method for trans-dimensional model selection has been applied to inverse problems across the physical and biological sciences, including estimating and studying the number of black holes from gravitational wave data [91, 92], seismic inversion problems that arise in geophysics [93, 94], and inferring phylogenetic trees from genomic data [95, 96]. This approach enables simultaneous trans-dimensional model selection and parameter

estimation, allowing for a systematic exploration of the parameter space in cases where standard regularization techniques, such as sparsity-inducing regularization, may not be physically meaningful.

When using RJMCMC, we partition the set of parameters in $\mathbf{p}$ and $\mathbf{q}$, where $\mathbf{p}$ is a set of parameters to be estimated using RJMCMC where the number of parameters $\mathbf{p}$ is unknown, and $\mathbf{q}$ is a set of unknown parameters that remains fixed. RJMCMC extends traditional MCMC by allowing proposals between models of different dimensionality, enabling the simultaneous exploration of both parameter values and model structures through dimension-changing proposal moves that ensure detailed balance and reversibility. In the discrete case, RJMCMC operates as a birth-death process where new parameters (e.g., additional particles or interactions) are "born" into the model or "die" based on probabilistic moves that maintain detailed balance (see Fig. 3.2c) [88]. In the continuous case, dimension matching is ensured by introducing auxiliary variables that transform proposals between parameter spaces of different dimensions, preserving reversibility by maintaining a one-to-one mapping with a Jacobian determinant correction [86].

3.3.3 Parallel Tempering (PT)

Parallel tempering, also known as replica exchange Monte Carlo, is a computational technique used to efficiently sample complex energy landscapes, particularly in systems with many local minima. The method involves running multiple copies (replicas) of the system in parallel, each at a different temperature. Higher temperature replicas explore the energy landscape more freely, overcoming barriers between local minima, while lower temperature replicas focus on finding the global minimum. At regular intervals, configurations between replicas are exchanged with a probability determined by the Metropolis criterion (see Fig. 3.2d). This exchange allows for low-temperature replicas to escape local minima by occasionally adopting the high-temperature configuration. The exchanges are designed

Algorithm 2 Reverse Jump Markov Chain Monte Carlo (RJMCMC) for discrete parameters

Input: data $\mathbf{d}_d$, partition of variable parameters with variable length $\mathbf{p} = \{p_1, \ldots, p_m\}$ and fixed parameters with fixed length $\mathbf{q} = \{q_1, \ldots, q_s\}$ and $\mathbf{p} \cup \mathbf{q} = \{a_1, \ldots, a_k, b_1, \ldots, b_j\}$ and $\mathbf{p} \cap \mathbf{q} = \emptyset$, known parameters $\mathbf{c}$, family of Hamiltonians $\{\mathcal{H}_k\}_{k=1}^{k_{\max}}$, family of forward models $\{f_k(\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c})\}_{k=1}^{k_{\max}}$, with $f_k(\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c}) = f_k(\mathbf{p}, \mathbf{q}, \mathbf{c})$, initial values $\mathbf{p}^{(0)} = (p_1^{(0)}, p_2^{(0)}, \ldots, p_m^{(0)})$, $\mathbf{q}^{(0)} = (q_1^{(0)}, q_2^{(0)}, \ldots, q_s^{(0)})$, $k^{(0)}$, likelihood model $\mathcal{L}(\mathbf{d}_d | f_k(\mathbf{p}, \mathbf{q}, \mathbf{c}))$, domain $\mathcal{D}$, number of steps N_{step}, dimension-changing kernel $\gamma(k, \cdot)$

1: **for** $n = 0, \ldots N_{\text{step}} - 1$ **do**

2: **Proposal step**: sample $k^* \sim \gamma(k^{(n)}, \cdot)$

3: **if** $k^* > k^{(n)}$ **then**

$$
\begin{aligned}
k^* &= k^{(n)} + 1 \\
p_{k+1}^* &\sim \mathcal{D} \\
\mathbf{p}^* &= (p_1^{(n)}, p_2^{(n)}, \ldots, p_k^{(n)}, p_{k+1}^*)
\end{aligned}
$$

4: **else**

$$
\begin{aligned}
k^* &= k^{(n)} - 1 \\
l &\sim \{1, 2, \ldots, k^{(n)}\} \\
\mathbf{p}^* &= (p_1^{(n)}, \ldots, p_{l-1}^{(n)}, p_{l+1}^{(n)}, \ldots, p_k^{(n)})
\end{aligned}
$$

5: **Accept/reject step**: Update

$$
\begin{aligned}
k^{n+1}, \mathbf{p}^{(n+1)} &= \begin{cases} k^*, \mathbf{p}^* & \text{with probability } \alpha(k^{(n)}, \mathbf{p}^{(n)}; k^*, \mathbf{p}^*) \\ k^{(n)}, \mathbf{p}^{(n)} & \text{with probability } 1 - \alpha(k^{(n)}, \mathbf{p}^{(n)}; k^*, \mathbf{p}^*) \end{cases} \\
\mathbf{q}^{(n+1)} &= \mathbf{q}^{(n)} \\
\alpha(k^{(n)}, \mathbf{p}^{(n)}; k^*, \mathbf{p}^*) &= \min\{1, \frac{\mathcal{L}(\mathbf{d}_d | f_{k^*}(\mathbf{p}^*, \mathbf{q}, \mathbf{c})}{\mathcal{L}(\mathbf{d}_d | f_{k^{(n)}}(\mathbf{p}^{(n)}, \mathbf{q}, \mathbf{c})} \frac{\gamma(k^*, k^{(n)})}{\gamma(k^{(n)}, k^*)}\}
\end{aligned}
$$

Output: $\{k^{(n)}\}_{n=1}^{N_{\text{step}}}$, $\{\mathbf{p}^{(n)}\}_{n=1}^{N_{\text{step}}}$, $\{\mathbf{q}^{(n)}\}_{n=1}^{N_{\text{step}}}$

to maintain detailed balance, ensuring that the sampling process correctly represents the equilibrium distribution at each temperature. By allowing configurations to move between different temperature levels, parallel tempering enhances the ability to explore the energy landscape, reducing the risk of getting trapped in local minima. This leads to more efficient and accurate sampling, particularly for systems with rugged energy landscapes. PT is an acceleration technique for a fixed number of parameters in highly non-linear non-convex

Algorithm 3 Parallel tempering (PT)

Input: Number of strands J, inverse temperatures $\beta_1 = 1 > \cdots > \beta_J$, data $\mathbf{d}_d$, partition of a fixed-length set of parameters to update $\mathbf{p} = \{p_1, \ldots, p_m\}$ and fixed parameters $\mathbf{q} = \{q_1, \ldots, q_s\}$, with $\mathbf{p} \cup \mathbf{q} = \{a_1, \ldots, a_k, b_1, \ldots, b_j\}$ and $\mathbf{p} \cap \mathbf{q} = \emptyset$ and their initial values $\mathbf{p}^{(0)} = (p_1^{(0)}, p_2^{(0)}, \ldots, p_m^{(0)})$ and $\mathbf{q}^{(0)} = (q_1^{(0)}, q_2^{(0)}, \ldots, q_s^{(0)})$, known parameters $\mathbf{c}$, fixed Hamiltonian model $\mathcal{H}_k(\mathbf{a}_k)$ indexed by k, likelihood model $\mathcal{L}(\mathbf{d}_d | f_k(\mathbf{a_k}, \mathbf{b}, \mathbf{c}))$, domain $\mathcal{D}$, number of steps N_{step}

1: **for** $n = 0, \ldots N_{\text{step}} - 1$ **do**

2: **Within-strand step**: Generate $\tilde{\mathbf{p}}_j^{(n+1)}$ from $\mathbf{p}_j^{(n+1)}$ by one step of (Alg. 1) for each strand $j \in [1, \ldots, J]$ using the corresponding likelihood model

$$\mathcal{L}_j(\mathbf{d}_d | f_k(\mathbf{p}, \mathbf{q}, \mathbf{c})) = (\mathcal{L}(\mathbf{d}_d | f_k(\mathbf{p}, \mathbf{q}, \mathbf{c})))^{\beta_j}$$

3: **Draw strands to attempt swap**: Sample $a, b \sim \{1, \ldots J\}$ with $a \neq b$ uniformly at random

4: **Evolve non-swapped chains forward**: For $j \notin \{a, b\}$, set

$$\mathbf{p}_j^{(n+1)}, \mathbf{q}_j^{(n+1)} = \tilde{\mathbf{p}}_j^{(n+1)}, \mathbf{q}_j^{(n)}$$

5: **Swap proposal step**: Attempt a swap of states between strands a and b:

$$(\mathbf{p}_a^{(n+1)}, \mathbf{p}_b^{(n+1)}) = \begin{cases} (\tilde{\mathbf{p}}_b^{(n+1)}, \tilde{\mathbf{p}}_a^{(n+1)}) & \text{with probability } \alpha_{\text{PT}}(a, b) \\ (\tilde{\mathbf{p}}_a^{(n+1)}, \tilde{\mathbf{p}}_b^{(n+1)}) & \text{with probability } 1 - \alpha_{\text{PT}}(a, b) \end{cases}$$

$$\mathbf{q}_a^{(n+1)}, \mathbf{q}_b^{(n+1)} = \mathbf{q}_a^{(n)}, \mathbf{q}_b^{(n)}$$

$$\alpha_{\text{PT}} = \min\{1, \frac{\mathcal{L}_a(\mathbf{d}_d | \tilde{\mathbf{p}}_b^{(n+1}, \mathbf{q}, \mathbf{c})}{\mathcal{L}_a(\mathbf{d}_d | \tilde{\mathbf{p}}_a^{(n+1}, \mathbf{q}, \mathbf{c})} \frac{\mathcal{L}_b(\mathbf{d}_d | \tilde{\mathbf{p}}_a^{(n+1)}, \mathbf{q}, \mathbf{c})}{\mathcal{L}_b(\mathbf{d}_d | \tilde{\mathbf{p}}_b^{(n+1)}, \mathbf{q}, \mathbf{c})}\}$$

Output: Sequences corresponding to the $j = 1, \beta_j = 1$ chain: $\{\mathbf{p}_{j=1}^{(n)}\}_{n=1}^{N_{\text{step}}}$, $\{\mathbf{q}_{j=1}^{(n)}\}_{n=1}^{N_{\text{step}}}$

energy landscapes.

3.3.4 Hybrid algorithm

Hybridizing Markov Chain Monte Carlo (MCMC) algorithms can enhance sampling efficiency and model selection for complex inverse problems. A robust approach combines the Random Walk Metropolis-Hastings (RWMH) algorithm for exploring continuous parameter spaces with Reversible Jump MCMC (RJMCMC) for transdimensional moves that enable discrete model selection. To improve within-model fits, particularly in highly non-linear and non-

convex likelihood landscapes, parallel tempering (PT) based on RWMH can be incorporated for a fixed number of discrete parameters, allowing for enhanced exploration of local minima and better posterior sampling. This multi-tiered strategy balances global model selection with efficient local exploration, mitigating issues like poor mixing and mode trapping, which are common in high-dimensional and multi-modal inverse problems.

3.4 Example application: Recovering hyperfine coupling parameters of a nuclear spin bath of defects in semiconductors from coherence data

While the trans-dimensional Hamiltonian model selection and parameter estimation problem in the sparse, noisy regime is general, we validate our implementation and illustrate the utility of a hybrid MCMC-based approach on a specific characterization problem of interest to the quantum information community: the recovery of hyperfine couplings of a nuclear spin bath of a single spin defect in a semiconductor from sparse, noisy coherence data. This example problem is a trans-dimensional Hamiltonian selection problem, since the number of nuclear spins in the bath is not known *a priori*; in addition, it is a problem for which there is interest in accelerating recovery of spin environments for purposes of device scale-up when using spin defects in seminconductors as units of quantum information processing, with nuclear spins as auxiliary qubits for quantum algorithms and memory. This problem further illustrates the ability of a hybrid MCMC approach to deal with parameters for which strong priors via *ab initio* calculations can be incorporated, as well as parameters that are drawn from a continuous distribution.

3.4.1 Hamiltonian and forward models

We consider a nuclear spin bath consisting of ^{13}C surrounding a nitrogen vacancy (NV) center in diamond. We seek to characterize the number of spins in the ^{13}C nuclear spin bath and the strength of each spin's hyperfine coupling $(A_\perp, A_\parallel)$. We consider the Hamiltonians indexed by the variable of the number of nuclear spins k with the Hamiltonian given by:

$$\mathcal{H}_k = D\mathcal{S}_z^2 + \gamma_e B_z \mathcal{S}_z + \sum_{i=1}^{k} \gamma_n B_z \mathcal{I}_{z,i} + \sum_{i=1}^{k} \mathcal{S}_z (A_{\parallel,i} \mathcal{I}_{z,i} + A_{\perp,i} \mathcal{I}_{x,i}) \tag{3.3}$$

where the known quantities are $D = 2\pi \times 2.87$ GHz, B_z is the strength of an externally applied magnetic field along the axis of the NV center, γ_e is the gyromagnetic ratio of an electron, γ_n is the nuclear gyromagnetic ratio of the nuclei, and $\mathcal{S} = (\mathcal{S}_x, \mathcal{S}_y, \mathcal{S}_z)$ and $\mathcal{I} = (\mathcal{I}_x, \mathcal{I}_y, \mathcal{I}_z)$ are the spin operators for the electronic and nuclear spins, respectively. We neglect nuclear–nuclear interactions in the Hamiltonian, as we expect these interactions to be negligible for the relatively short τ (< 10 μs) and relatively low pulse number (≤ 32 CP pulses) in the experiments we consider for high-throughput characterization. We seek to recover the number of spins k, which determines the dimension of the Hamiltonian, and for each spin, we aim to recover the parallel and perpendicular components $(A_\parallel, A_\perp)$ of the hyperfine coupling tensor, which characterizes the coupling strength between the nuclei and the NV center.

The forward model $f_k(\mathcal{H}_k(\mathbf{a}_k), \mathbf{b}, \mathbf{c})$ for a set of k nuclear spins with hyperfine components $\mathbf{A}_\parallel, \mathbf{A}_\perp$ under dynamical decoupling experiment with N CP-pulses measured at inter-pulse

spacing τ_j in external applied magnetic field B_z is:

$$f_k(\mathcal{H}_k(\mathbf{A}_\parallel, \mathbf{A}_\perp, B_z), \lambda, N, \tau_j) = (\frac{1}{2}(1 + \prod_{i=1}^{k} M_i(A_{\parallel,i}, A_{\perp,i}, N, B_z, \tau_j)))^{-\frac{\tau_j}{\lambda}} \tag{3.4}$$

$$M_i(A_{\parallel,i}, A_{\perp,i}, N, B_z, \tau) = (1 - m_{i,x}^2 \frac{(1-\cos\alpha_i)(1-\cos\beta)}{1+\cos\alpha_i\cos\beta - m_{i,z}\sin\alpha_i\sin\beta} \sin\frac{N\phi_i}{2}^2) \tag{3.5}$$

$$\cos\phi_i = \cos\alpha_i\cos\beta - m_{i,z}\sin\alpha_i\sin\beta \tag{3.6}$$

$$m_{i,z} = \frac{A_{\parallel,i} + \omega_L}{\tilde{\omega}_i} \tag{3.7}$$

$$m_{i,x} = \frac{A_{\perp,i}}{\tilde{\omega}_i} \tag{3.8}$$

$$\tilde{\omega}_i = \sqrt{(A_{\parallel,i} + \omega_L)^2 + A_{\perp,i}^2} \tag{3.9}$$

$$\alpha_i = \tilde{\omega}_i \tau_j \tag{3.10}$$

$$\beta = \omega_L \tau_j \tag{3.11}$$

$$\omega_L = -\gamma_n B_z, \tag{3.12}$$

where we can consider our partition of parameters to be $\mathbf{a}_k = \{\mathbf{A}_\parallel, \mathbf{A}_\perp\}$, $\mathbf{b} = \{\lambda\}$, and $\mathbf{c} = \{N, B_z, \tau_j\}$ [34].

3.4.2 Likelihood model and priors

Given that coherence signals are measured with photon counts on the order of thousands to tens of thousands, we approximate the noise as Gaussian rather than Poisson-distributed. Accordingly, we define the likelihood function as:

$$\mathcal{L}(\mathbf{d}_d | f_k(\mathcal{H}_k(\mathbf{a_k}), \mathbf{b}, \mathbf{c})) = \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{d}(\mathbf{d}_i - f_k(\mathcal{H}_k(\mathbf{A}_\parallel, \mathbf{A}_\perp, B_z), \lambda, N, \tau_i))^2\right), \tag{3.13}$$

where σ^2 is a hyperparameter that relates to assumptions on the variance of the noise. The partition of parameters is $\mathbf{a}_k = (A_{\parallel,1}, A_{\perp,1}, A_{\parallel,2}, A_{\perp,2}, \ldots, A_{\parallel,k}, A_{\perp,k})$, $\mathbf{b} = (\lambda)$, and $\mathbf{c} = (N, \tau_j, B_z)$. The set $\mathbf{a}_k$ corresponds to the hyperfine parameters of the nuclear spins in the bath; since we do not know *a priori* how many nuclear spins are present, we do not know *a priori* how many hyperfine parameters $(A_{\parallel}, A_{\perp})$ we need to recover. The continuous parameter $\mathbf{b} = (\lambda)$ that we restrict to the interval $[0, 1]$ accounts for relative decoherence effects. This parameter is crucial for experimental consistency and likely reflects interactions between the NV center and unaccounted lattice impurities, such as nitrogen substitutional defects, and the broader, distant spin bath, which causes and overall decoherence decay. The known parameters $\mathbf{c} = (N, B_z, \tau_j)$ are the externally applied magnetic field B_z, the number of laser pulses N required in the experimental protocol to obtain a coherence signal, and the amount of time τ_j in between sequential laser pulses.

To enhance the efficiency of parameter estimation and model selection, we integrate data from *ab initio* calculations as physical constraints. Specifically, we restrict possible nuclear positions to specific lattice sites of diamond, leveraging computational predictions of hyperfine couplings at these sites based on the Fermi contact and point-dipole approximations [97]. By using lattice site constraints, we address two main issues. First, in high-dimensional spaces, random walks are inefficient, so by mapping the possible values of the parameters of interest to a discrete set, we can sample from such a discrete set via a random walk rather than allowing the unknown parameters to take on any value in a specified closed domain. Second, because lattice position determines the strength of the hyperfine coupling, and there are lattice sites that, by symmetry, have identical hyperfine couplings, we have further constraints on the number of allowable nuclear spins with identical hyperfine couplings, provided that two walkers are not allowed to simultaneously occupy the same lattice site.

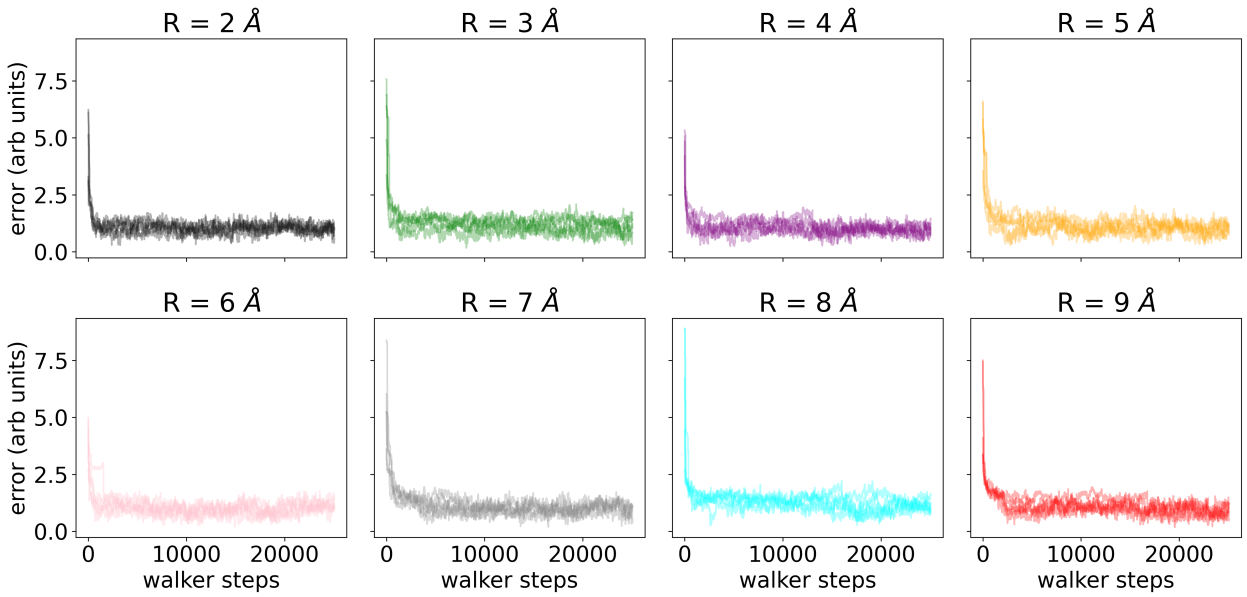

Figure 3.3: **Error trajectories of 5 ensembles as a function of walker steps for varying R_{spin}.** The recovery algorithm was applied to the same coherence signal generated using $N = 16$ CP pulses, $\varepsilon = 0.001$ shot noise, hyperfine perturbation 0.001, $\tau_{\max} = 0.008$ ms, and 250 interpulse times sampled from 15 simulated spins with $\sigma^2 = 0.1$.

3.4.3 Hyperparameter tuning and dependence

There are a variety of hyperparameters in the hybrid algorithm, including the walker radius of the lattice sites (R_{spin}), the walker radius over λ values (R_{T_2}), the number of strands for parallel tempering, the inverse temperatures used for parallel tempering, and the noise variance parameter (σ^2) in the likelihood function. While we do not tune every hyperparameter for optimal performance (and we make no claim that the hyperparameters selected for this work are optimal), we note that we get sufficient performance even without having to go through the computationally intensive steps to systematically and exhaustively tune the hyperparameters.

We find that the hybrid algorithm presented is highly robust to R_{spin}, but the performance of the hybrid algorithm is sensitive to σ^2. In Fig. 3.3, we find that the variance, mean, and convergence properties of the error trajectories are uncorrelated with the size of the walker radius over the lattice sites R_{spin}. That is, the probabilistic selection rules

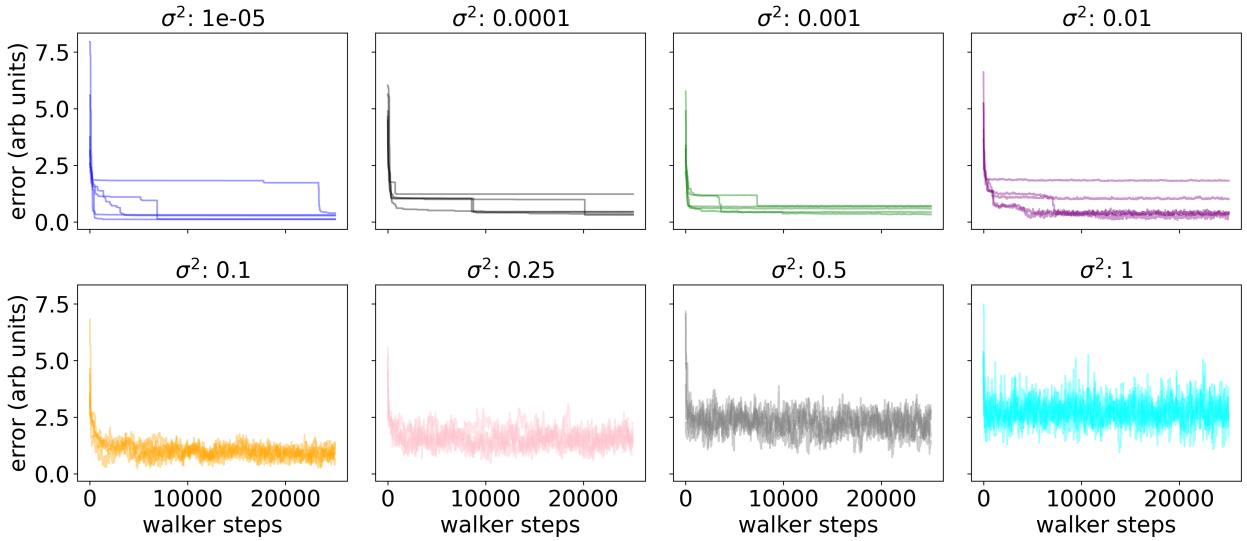

Figure 3.4: **Error trajectories of 5 ensembles as a function of walker steps for varying σ^2.** The recovery algorithm was applied to the same coherence signal generated using $N = 16$ CP pulses, $\varepsilon = 0.001$ shot noise, hyperfine perturbation 0.001, $\tau_{\max} = 0.008$ ms, and 250 interpulse times sampled from 15 simulated spins with $R_{\text{spin}} = 5$ Å.

introduced by the hybrid algorithm bias the walkers toward the same solution, regardless of the size of the neighborhood by which each lattice walker is constrained. Notably, there is a physical limit on how small the walker radius can be, which is the distance of nearest neighbors between lattice sites. For diamond, the distance between nearest neighbors is around 1.54 Å. Unless otherwise stated, all simulations and experimental data recovery were performed with $R_{\text{spin}} = 5$ Å.

In contrast, the performance of the hybrid algorithm is highly sensitive to the value of σ^2 in the likelihood. We find that for values of σ^2 that are too small, even though the variance in the error is small and the mean errors tend to be low, ensembles are highly likely to be trapped in local minima, whereas for values of σ^2 that are too large, the mean error and variance both increase (see Fig. 3.4). Unless stated otherwise, we use $\sigma^2 = 0.1$, which represents the best tradeoff between low variance, low mean, and minimal trapping in local minima.

We note that hyperparameters need not be strictly optimal: beyond a certain point, fur-

ther optimization yields diminishing returns, and the cost of hyperparameter optimization should not exceed the cost of the inference itself. Our primary requirement for hyperparameter values is reliable convergence of the MCMC walkers to the posterior distribution on an accessible timescale given available computational resources. In practice, tuning the hybrid algorithm involves balancing exploration and convergence across both discrete and continuous parameter spaces, although one should also account for the trade-off between optimizing hyperparameters, which can lead to faster convergence, versus the time-consuming and computationally intensive process of optimizing the hyperparameters.

For continuous parameters, adjusting proposal step sizes or covariance matrices can help achieve an acceptance rate that promotes efficient mixing while avoiding local trapping. For discrete parameters, such as lattice site selection, the choice of neighborhood radius should respect physical constraints (e.g., nearest-neighbor distances) but can otherwise be broad without degrading convergence. Parallel tempering parameters, including the number of strands and inverse temperatures, can be selected to ensure sufficient exchange between chains, particularly in multimodal or rugged likelihood landscapes. Finally, the noise variance σ^2 in the likelihood should be tuned to balance model flexibility against overconfidence: too small values risk trapping the walkers in local minima, whereas too large values reduce the algorithm's ability to accurately recover underlying parameters. In many cases, adaptive strategies provide a practical approach to systematically improve performance without requiring exhaustive hyperparameter searches [98]; such as monitoring acceptance rates and effective sample sizes during burn-in and adjusting proposal distributions or tempering schedules accordingly.

3.5 Results

3.5.1 Results from simulations

Experimentally, the coherence signals are obtained by applying a series of N Carr-Purcell (CP) pulses separated by varying inter-pulse spacings τ_j. Hence, to obtain each additional data point at a new τ_j, the CP experiment must be repeated many times and the results averaged to mitigate noise from photon detection, and each experiment for each data point takes $\propto 2\tau_j N$. We assume we are working with data in a noise and sparsity regime where the inverse problem is ill-posed, and we seek to understand how sparsity and noise impact the accuracy of the recovery procedure by studying the accuracy on simulated data.

For each simulation, we generated coherence signals based on a set of simulated ^{13}C that determine the number of spins and the hyperfine couplings for specified experimental parameters and relative decoherence values. We hybridize the algorithms discussed in Sec. 3.3 as outlined in Alg. 4, and we initialized 5 ensembles and let each ensemble evolve for 25,000 steps, and we consider the first 10,000 steps to be burn-in time. We cycle between 50 steps of RJMCMC to determine the number of spins in the system, 100 steps of PT to obtain a better fit of the hyperfine coupling Hamiltonian parameters for a fixed number of spins, and 25 steps of RWMH for the relative decoherence parameter λ.

For the lattice sites and sets of hyperfine parameters computed at each lattice site, we use data from [97] and define a cutoff of lattice sites by considering only lattice sites with hyperfine couplings with either component of $A_{\parallel}$ or $A_{\perp}$ greater than 5 kHz. For RJMCMC, we consider a maximum of $k_{\max} = 50$ nuclear spins to be recovered. For PT, we use $J = 10$ strands and inverse temperatures $\beta_j = 2^{-j}$. The discrete random walk over lattice sites uses a walker radius of $R = 5$ Å, and the continuous RWMH for D over $[0, 1]$ with reflected boundary conditions uses a walker radius of $R = 0.01$. Unless stated otherwise, the default values for the simulated experiments are 250 uniformly sampled τ_j values between 0 and 8

μs, $N = 16$ dynamical decoupling pulses, an external applied magnetic field of 311 G, and $\epsilon \sim \mathcal{N}(0, 0.001)$ noise added to each data point of the simulated coherence signal.

3.5.2 Metrics for performance and recovery rates

We measured the recovery rate of hyperfine parameters by examining the fraction of simulated spins that appeared in the posterior distribution across the entire spin configuration space. A recovery rate of 1 indicates that a given spin appears in every spin configuration, while a rate of 0 signifies that the spin does not appear in any posterior configuration, and we properly account for duplicate and triplicate hyperfine couplings. In the simulations, where both lattice positions and hyperfine couplings are known, we impose a maximum allowed mismatch of $|\Delta A_{\parallel}|, |\Delta A_{\perp}| \leq 0.1$ kHz for a nuclear spin to be detected. This tolerance is not intended to relax the hyperfine constraint, but rather to account for small numerical errors in the DFT-computed hyperfine couplings between symmetry-related lattice sites. More precisely, let $B^{(j)}$ denote the j-th posterior spin-bath sample, with $j = 1, \ldots, M$, and define an indicator function $\mathbb{I}_i^{(j)}$ for the presence of a given simulated (or experimentally-confirmed) spin $s_i \in S$ posterior sample $B^{(j)}$, where $\mathbb{I}_i^{(j)}$ is 1 if $s_i \in B^{(j)}$ and 0 otherwise. Here, spin identity is determined by matching hyperfine parameters within the mismatch tolerance. The detection rate for an individual spin s_i is defined as

$$R_i = \frac{1}{M} \sum_{j=1}^{M} \mathbb{I}_i^{(j)}, \tag{3.14}$$

and the overall detection rate for a given simulated (or experimentally-confirmed) spin-bath S of size n is

$$R = \frac{1}{n} \sum_{i=1}^{n} R_i. \tag{3.15}$$

We then define the overall recovery rate by averaging these individual rates across all spins in the bath. When multiple simulated baths are considered (e.g., with different bath sizes),

the reported values are further averaged across all simulated baths.

Additionally, we assessed the accuracy of recovering the number of spins by comparing the modes of the posterior distribution with the number of simulated spins in each hyperfine coupling grouping. We report the discrepancy in the number of spins as |mode of posterior distribution of k - number of simulated spins|, so a discrepancy of 0 indicates that the mode of the posterior distribution matches the number of spins that were simulated, and the greater the discrepancy the worse the hybrid MCMC procedure selects the correct Hamiltonian model dimension. We quantified false positive rate by considering the modal spin distribution corresponding to the modal model dimension of the posterior distribution as the spins that are recovered from a given coherence signal. Let S^* denote the set of spins in the modal spin-bath configuration of size n, and let $B^{(j)}$ denote the j-th posterior spin-bath, with $j = 1, \ldots, M$. We define a sample-averaged false absence rate as

$$FP = \frac{1}{nM} \sum_{s \in S^*} \sum_{j=1}^{M} \mathbb{I}_{s \notin B^{(j)}} \tag{3.16}$$

where $\mathbb{I}_{s \notin B^{(j)}}$ is a non-occurrence indicator function for the presence of a spin s in nuclear spin bath $B^{(j)}$. This sampled-averaged false absence rate measures the fraction of posterior samples in which spins belonging to the modal configuration are absent. We note that this is a sample-averaged measure rather than a strict false-positive rate in the classical classification sense. All results were grouped by hyperfine coupling magnitude, which we define as $\sqrt{A_{\parallel}^2 + A_{\perp}^2}$, and averaged 16 different simulated spin-bath configurations containing between 5 and 20 simulated spins.

Performance in sparse data limit

To evaluate the performance of our recovery method in the limit of sparse data, we varied the number of τ_j sampled from the coherence signals from a simulated spin bath of ten

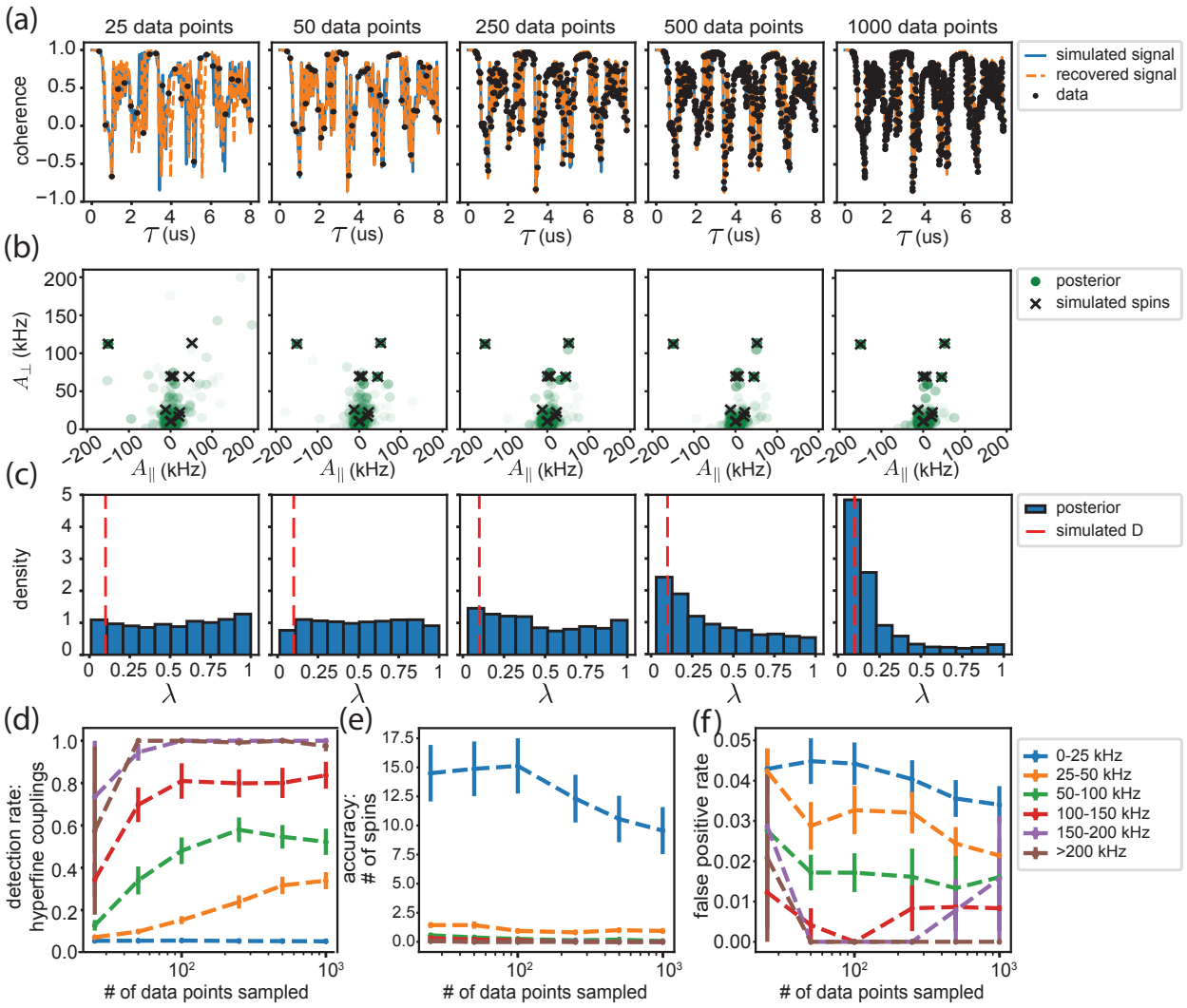

Figure 3.5: **Recovery in sparse data limit**. **(a)** We simulate a system of ten ^{13}C surrounding a nitrogen vacancy (NV) center in diamond, and plot the resulting coherence signal for a 16-pulse dynamical decoupling experiment in an external magnetic field of 311 G. We sample a varying number of data points with noise $\epsilon \sim \mathcal{N}(0, 0.001)$, and we plot the coherence signal from the nuclear spin parameters recovered from the specified number of sampled data points. Posterior distribution of the hyperfine components of the nuclear spins **(b)** and relative decoherence parameter (λ) **(c)** recovered from the coherence signal with varying number of data points sampled corresponding to the coherence signal in **(a)**. In **(d)**, **(e)**, and **(f)**, we plot the average results for detection rate of hyperfine couplings, the discrepancy the number of recovered spins in the bath with the number of spins that were simulated, and false positive rate for 16 simulated nuclear spin configurations containing between 5-20 nuclear spins, with the rates plotted by hyperfine magnitude ($\sqrt{A_\perp^2 + A_\parallel^2}$) of the spins.

nuclear spins (Fig. 3.5a). We then assessed whether the posterior distributions of the spin configurations (Fig. 3.5b) and relative decoherence (Fig. 3.5c) obtained from the hybrid algorithm described in Sec. 3.3.4 successfully captured the simulated nuclear spin configurations and relative decoherence parameter λ. To ensure these results generalize across spin configurations, we analyzed the recovery rate for the hyperfine couplings, the detection rate for the number of spins, and the false positive rate for nuclear detection for 16 different simulated spin baths containing 5 to 20 spins, with a hyperfine magnitude greater than 5 kHz (Fig. 3.5d).

For the sample bath consisting of 10 nuclear spins and varying number of τ_j, we observe, as expected, that the fidelity of the recovered signal improves with an increased number of sampled τ_j (Fig. 3.5a). However, what we are interested in is not necessarily to best recover the underlying signal but to recover the underlying nuclear spin bath configuration that gives rise to the measured signal. We find that even for signals sampled as sparsely as 50 τ_j, the posterior distribution of hyperfine couplings of spins contains all ten of the simulated spins (Fig. 3.5b), and the signal sampled as sparsely as 25 τ_j recovers the nuclear spin with the strongest hyperfine magnitude. As expected, the more sparsely sampled signals result in more fictitious spins in the posterior compared to more densely sampled spins. The density of the time-point sampling is related to the ill-posedness of the inverse problem; from this sample bath, we also observe that the ability to resolve hyperfine couplings from sparse signals is highly correlated with the strength of the hyperfine magnitude. That is, the strongly coupled spins are clearly resolved in the posterior distributions in Fig. 3.5b, whereas we are unable to clearly resolve the more weakly coupled spins are unable, even at relatively dense samplings of the signals. We emphasize that the failure to resolve weakly coupled spins is related to the inherence ill-posedness of the inverse problems. Weakly coupled spins lead to smaller dips in the fringes of the coherence signal; as a result, many different combinations of weakly coupled spins can give rise to identical patterns of peaks in the coherence signal, making it

impossible to resolve individual weakly coupled spins from a single sparse coherence signal originating from a dynamical decoupling experiment.

While we are able to recover the stronger nuclear spin hyperfine couplings, even from sparse time-point samplings, we observe that relatively higher resolutions of data are required to recover the relative decoherence parameter, λ, as seen in Fig. 3.5c. For samplings below 250 τ, we obtain a relatively uniform posterior distribution of λ, whereas for simulated data with more than 250 τ, the recovered value of λ (that is, the mode of the posterior distribution) matches the simulated value. Since λ is a decay parameter, extremal peak values of the coherence signal are required to capture the decay accurately. For sparsely sampled data, there is a lower probability that extremal peak values will be sampled, resulting in a relatively uniform distribution of relative decoherence values (Fig. 3.5c). In contrast, dense samplings of the data will result in many extremal peak values, resulting in a non-uniform posterior distribution of relative decoherence values centered on the simulated value.

To generalize the results beyond a single spin bath, we look at average values of detection rate (Fig. 3.5d), discrepancy in the number of recovered spins (Fig. 3.5e), and false positive rates (Fig. 3.5f). We find that the recovery rate of hyperfine couplings improves with increasing τ, but plateaus beyond approximately 100 sampled time points. Additional sampling beyond this threshold does not significantly enhance the recovered information, indicating a limit on the impact of increasing the rate of sampling in terms of improving the recoverable information content from specific experimental parameters. That is, beyond around 100 sampled τ_j, the ability to recover hyperfine couplings from this coherence signal is limited not by the data resolution but by the experimental values itself, such as the number of CP pulses N, the external magnetic field B_z, or the maximum τ_j sampled. When considering the discrepancy in the number of spins, we observe a similar trend: recovery rates increase with the number of sampled time points, eventually saturating beyond 500 time points.

The accuracy of recovery rates is notably lower for weakly coupled nuclei (< 25 kHz) across multiple metrics, including detection rate, number of recovered spins, and false positive rate, as seen in Fig. 3.5d-f. This behavior arises from the inherent ill-posedness of the problem in the sparse data regime: distinct configurations of weakly coupled spins can generate identical coherence signals with comparable likelihoods, limiting the ability to resolve individual weakly-coupled nuclear spins. This limitation is a fundamental consequence of the ill-posedness of this inverse problem in the sparse data regime rather than a failure of the hybrid algorithm to accurately resolve the simulated nuclear spin configuration.

In contrast, for nuclear spins with hyperfine couplings exceeding 150 kHz, the hybrid MCMC recovery method exhibits consistently high accuracy. For spins with intermediate coupling strengths, accuracy improves with increased sampling, particularly in the range of 25 to 100 sampled time points (Fig. 3.5d-f). However, beyond 100 sampled points, the performance of the hybrid algorithm is stable, with detection rates, recovered spin counts, and false positive rates remaining largely unchanged up to 1000 sampled points. This hybrid recovery algorithm is robust even when applied to coherence signals sampled as sparsely as 100 time points.

Performance in noisy data limit

Since experimentally, coherence signals are measured by averaging photon counts across repeats of a single given CP experimental pulse, the shot noise per data point can be decreased by a quadratically longer time average (i.e., reducing shot noise by half requires repeating an experiment four times as many repeats). Since averaging takes time, we are interested in how the accuracy of the recover procedure is impacted by the amount of noise. Remarkably, our approach demonstrated high robustness to shot noise across various magnitudes of hyperfine couplings. In an example simulated bath containing seventeen nuclear spins with varying amounts of noise added to each data point, we observe that even for significant amounts of

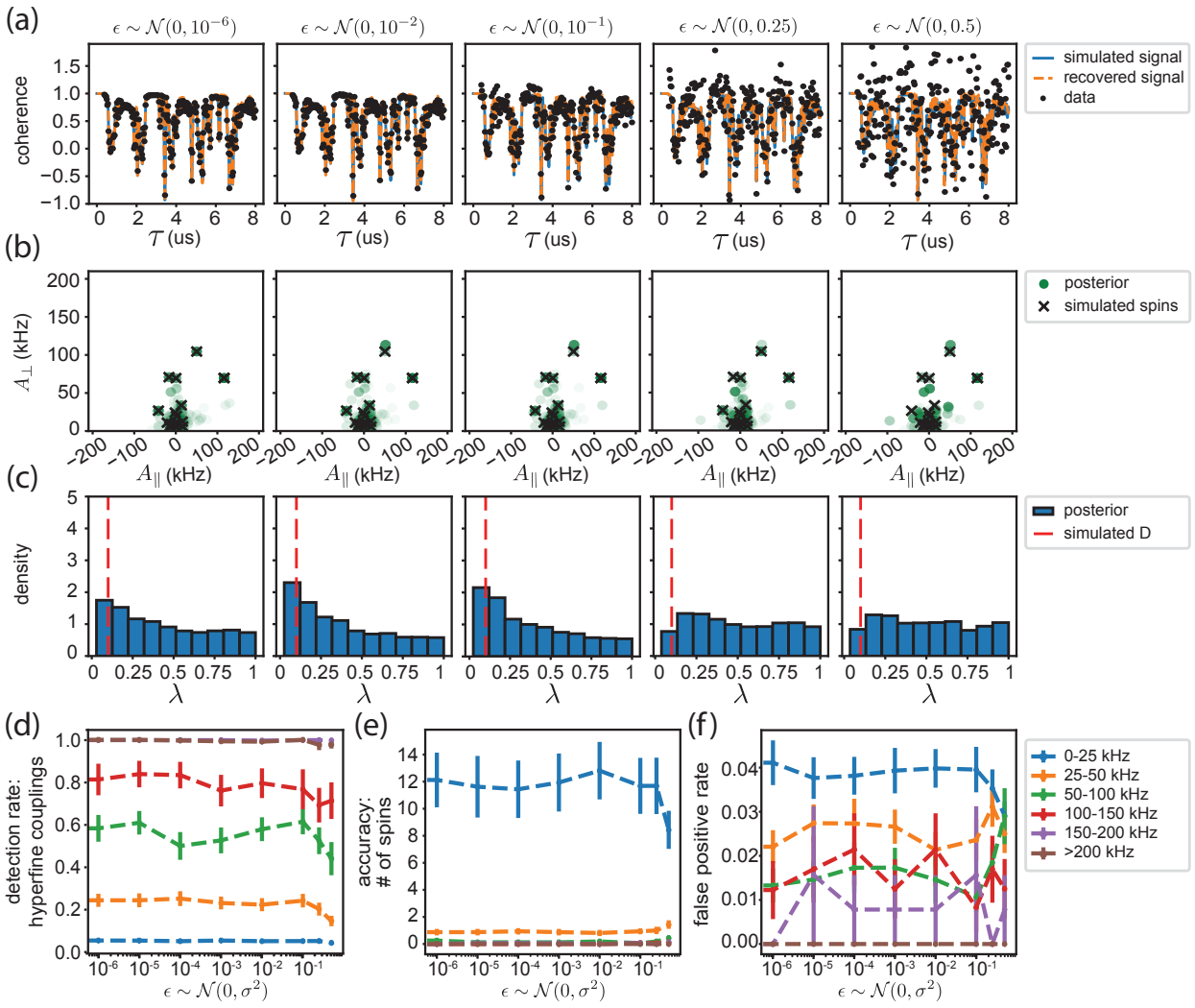

Figure 3.6: **Recovery in noisy data limit**. **(a)** We simulate a system of seventeen ^{13}C surrounding a nitrogen vacancy (NV) center in diamond, and plot the resulting coherence signal sampled uniformly for 250 τ_j for a 16-pulse dynamical decoupling experiment in an external magnetic field of 311 G. We add varying amounts of noise to each data point, and we plot the coherence signal from the nuclear spin parameters recovered from the specified number of sampled data points. We plot the posterior distribution of the hyperfine components of the nuclear spins **(b)** and relative decoherence parameter (λ) **(c)** recovered from the coherence signal with varying number of data points sampled corresponding to the coherence signal in **(a)**. In **(d)**, **(e)**, and **(f)**, we plot the average results for detection rate of hyperfine couplings, the discrepancy the number of recovered spins in the bath with the number of spins that were simulated, and false positive rate for 16 simulated nuclear spin configurations containing between 5-20 nuclear spins, with the rates plotted by hyperfine magnitude ($\sqrt{A_\perp^2 + A_\parallel^2}$) of the spins.

noise ($\epsilon \sim \mathcal{N}(0, 0.25)$), we are able to recover both the coherence signal (Fig. 3.6a) and the strongly coupled nuclear spins are well-resolved in the posterior distribution of the recovered hyperfine couplings (Fig. 3.6b). Recovery performance only showed significant degradation when the average shot noise per data point exceeded 0.25, which is substantial given that the coherence signal is bounded between -1 and 1. Even with average noise per data point up to 0.1 , the algorithm was able to recover hyperfine couplings and spin numbers, especially for spins with stronger hyperfine couplings. Furthermore, we find that, similar to decreasing the number of sampled time points, adding an increasing amount of noise results in a more uniform distribution of the posterior distribution of the relative decoherence parameter, λ (Fig. 3.6c). Since the relative decoherence relies on accurate values of the coherence at its peaks, adding an increasing amount of noise results in inaccurate peak values, leading to an inability to recover λ.

Over a range of simulated nuclear spin baths, we find that this hybrid recovery approach is remarkably stable to noise added to the simulated coherence data up to $\epsilon \sim \mathcal{N}(0, 0.1)$ in terms of multiple metrics, including the detection rate of the hyperfine couplings (Fig. 3.6d), discrepancy in number of spins recovered (Fig. 3.6e), and false positive rate (Fig. 3.6f). These results suggest that our method is well-suited for noisy data and does not require a separate de-noising step, which is a common requirement in machine learning-based approaches [80]. Furthermore, the algorithm's ability to recover nuclear spins across different magnitudes for average shot noise per data point ranging from 10^{-6} to 10^{-1} indicates that the ability to recover nuclear spins is limited by the information inherent in short, sparsely sampled dynamical decoupling experiments, rather than by shot noise obscuring the hyperfine coupling information present in the coherence signal.

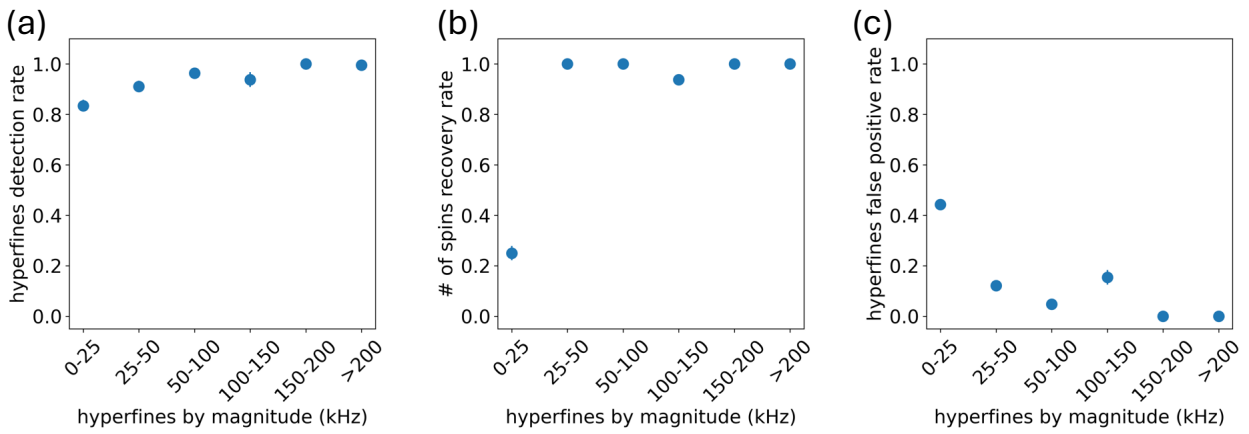

Figure 3.7: **Recovery method on data with experimental parameters in machine learning regime**. **(a)** Detection rate for hyperfine couplings of the posterior distribution, **(b)** detection rate for the number of spins in the modal configurations of the posterior distribution, and **(c)** false positive rate for hyperfine couplings for spins in the modal spin configuration over the posterior distribution for data generated with 15,000 time points for N=32 and N=256 CP pulses at 404 G magnetic field.

Performance of hybrid algorithm in dense, low noise limit

We demonstrate that the proposed method achieves high accuracy, even for weakly coupled spins, with over 80% accuracy for spins in the 0-25 kHz frequency range when applied to coherence data simulated in the large data limit commonly used in machine learning approaches. When the method is applied to experimental coherence signals that are in the resolution limit of the data used in recent machine learning studies (e.g., [80, 81]), as shown in Figure 3.7, it achieves low false positive rates and near-perfect accuracy for recovering the number of spins when the coupling frequencies exceed 25 kHz.

Despite these favorable trends, performance in the 0–25 kHz regime remains substantially degraded in terms of accuracy of the number of spins detected ($\tilde{3}0\%$) and false positive rate ($\tilde{4}0\%$) as seen in Figure 3.7, even in the limit of long dynamical decoupling sequences, large numbers of dynamical decoupling pulses, and dense samplings of interpulse spacings. This behavior should not be interpreted as an algorithmic failure, but rather as a manifestation of the fundamentally ill-posed nature of the inverse problem associated with weakly coupled

nuclear spins. In this regime, individual spins induce only minimal modulations of the NV coherence signal, and the resulting signal variations are often comparable in magnitude to experimental noise [46]. Thus, many distinct spin configurations, differing in both number and hyperfine couplings, produce coherence signals that are effectively indistinguishable within the accessible measurement resolution.

The ill-posedness is further exacerbated by the large number of nuclear spins at small hyperfine couplings, which leads to a combinatorial proliferation of nearly degenerate solutions in parameter space. Even with extensive data, including large numbers of sampled time points and high pulse numbers, the available measurements do not uniquely constrain the underlying spin configuration. In this setting, additional data primarily reinforce existing ambiguities rather than resolve them, reflecting a saturation of the information content of the coherence signal under fixed experimental parameters.

Within our Bayesian framework, this intrinsic non-identifiability is made explicit through the structure of the posterior distribution, which exhibits broad, multi-modal support in the weak-coupling regime. The strict detection criteria employed here, requiring consistency between detected spins and lattice-resolved hyperfine couplings, favor conservative reconstructions at the expense of reduced apparent detection rates for extremely weakly coupled spins. Importantly, this trade-off reflects a principled response to an underdetermined inverse problem rather than a limitation of the inference procedure itself.

These observations indicate that improving performance in the 0–25 kHz regime cannot be achieved solely through algorithmic refinement or increased data density under otherwise fixed experimental conditions. Instead, resolving such weakly coupled spins requires additional, complementary experimental information that alters the structure of the inverse problem itself, for example by combining measurements performed under different magnetic fields. Hence, the reduced performance in the weak-coupling regime reflects a fundamental information limit of the measurement setting, rather than a deficiency of the Bayesian

inference framework.

Dependence on hyperfine couplings

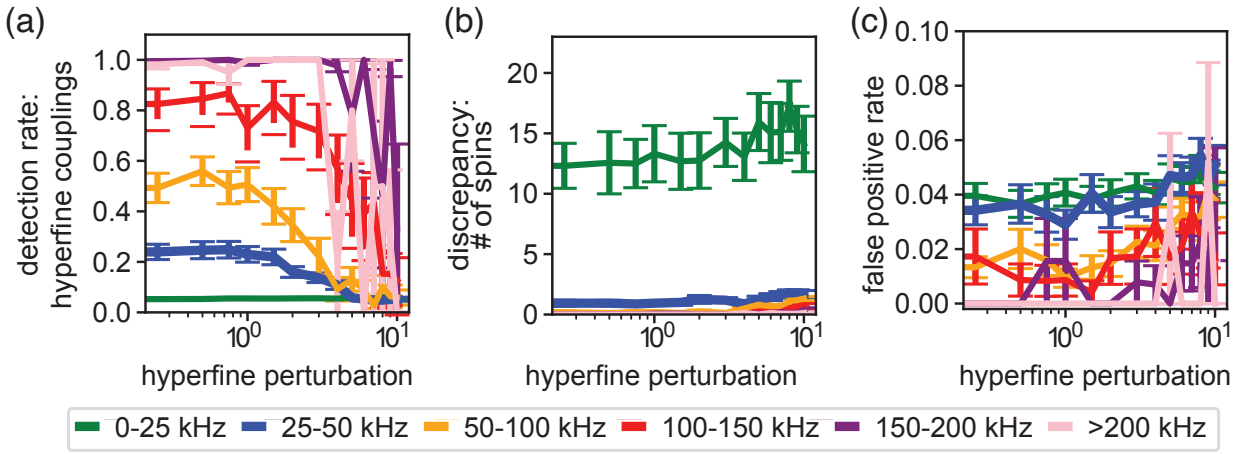

Figure 3.8: **Recovery method accuracy versus hyperfine coupling perturbations**. **(a)** Detection rate for hyperfine couplings of the posterior distribution, **(b)** discrepancy between the number of spins recovered and the number of spins simulated, and **(c)** false positive rate for hyperfine coupling perturbations of different magnitudes. The recovery algorithm was applied to the same coherence signal generated using $N = 16$ CP pulses, $\varepsilon = 0.001$ shot noise, $\tau_{\max} = 0.008$ ms, and 250 interpulse times and $\sigma^2 = 0.1$ averaged over simulated nuclear spin baths ranging from 6 to 20 spins.

Because our model currently constrains the possible hyperfine couplings to those computed from *ab initio* density functional theory for each lattice site, it is essential to understand how the recovery method performs when these input couplings are inaccurate. We investigated the hybrid algorithm's robustness to perturbations in hyperfine couplings, focusing on the ability to recover the hyperfine parameters $A_\perp$ and $A_\parallel$ from noisy input hyperfine data associated with each lattice site. Simulated inaccuracies in hyperfine couplings, introduced by perturbations $A_\perp \pm \delta$ and $A_\parallel \pm \delta$, were considered, where we varied the perturbation (denoted as δ) to assess recovery accuracy. We applied our hybrid recovery method to simulations with a 0.001 shot noise per data point, 250 τ sampled, a 0.008 ms longest τ, 16 CP pulses, and 311 G magnetic field.

Our results indicate that the algorithm remains robust to hyperfine perturbations up to approximately 1 kHz. When the input hyperfine couplings were accurate to the closest kHz, the algorithm performed well, with recovery rates close to 1 for the highest magnitude of hyperfine couplings ($>$ 150 kHz). However, when the perturbations exceed 1 kHz, the recovery rates for both hyperfine couplings and number of spins decreases, suggesting that the method is less effective when hyperfine couplings are known with low precision (Fig. 3.8). This threshold suggests that the method is particularly suited for systems where hyperfine couplings can be known to within a kilohertz. Given that state-of-the-art *ab initio* calculations based on density functional theory can achieve relative mean errors of around 1% for hyperfine couplings at all distances for an NV center in diamond [97], and that we seek to detect nuclear spins in the weakly coupled regime (up to 200 kHz), with the majority of spins considered here (over 99%) below 100 kHz, this level of accuracy falls within the algorithm's robustness threshold and is promising for practical applications.

However, based on these results for more strongly coupled spins (between 100-500 kHz), we expect a large drop off in accuracy for detection rates for individual spins, since the relative mean error results in a much larger absolute error for more strongly coupled spins. A 1% mean absolute error for a 250 kHz results in a 2.5 kHz hyperfine perturbation, which is in the regime where the detection rate is much lower than for hyperfine perturbations below 1 kHz and smaller. Nevertheless, the discrepancy between the number of spins for simulated groups of strongly coupled nuclear spins and the recovered number of spins remains relatively low and constant, even for simulations with highly perturbed hyperfine couplings. Taken together, these simulations indicate that this hybrid MCMC approach remains a powerful tool for screening individual samples based on the number of spins of varying magnitudes, even as we expect a decrease in detection accuracy when using theoretical values for nuclear spins with more strongly coupled hyperfine couplings.

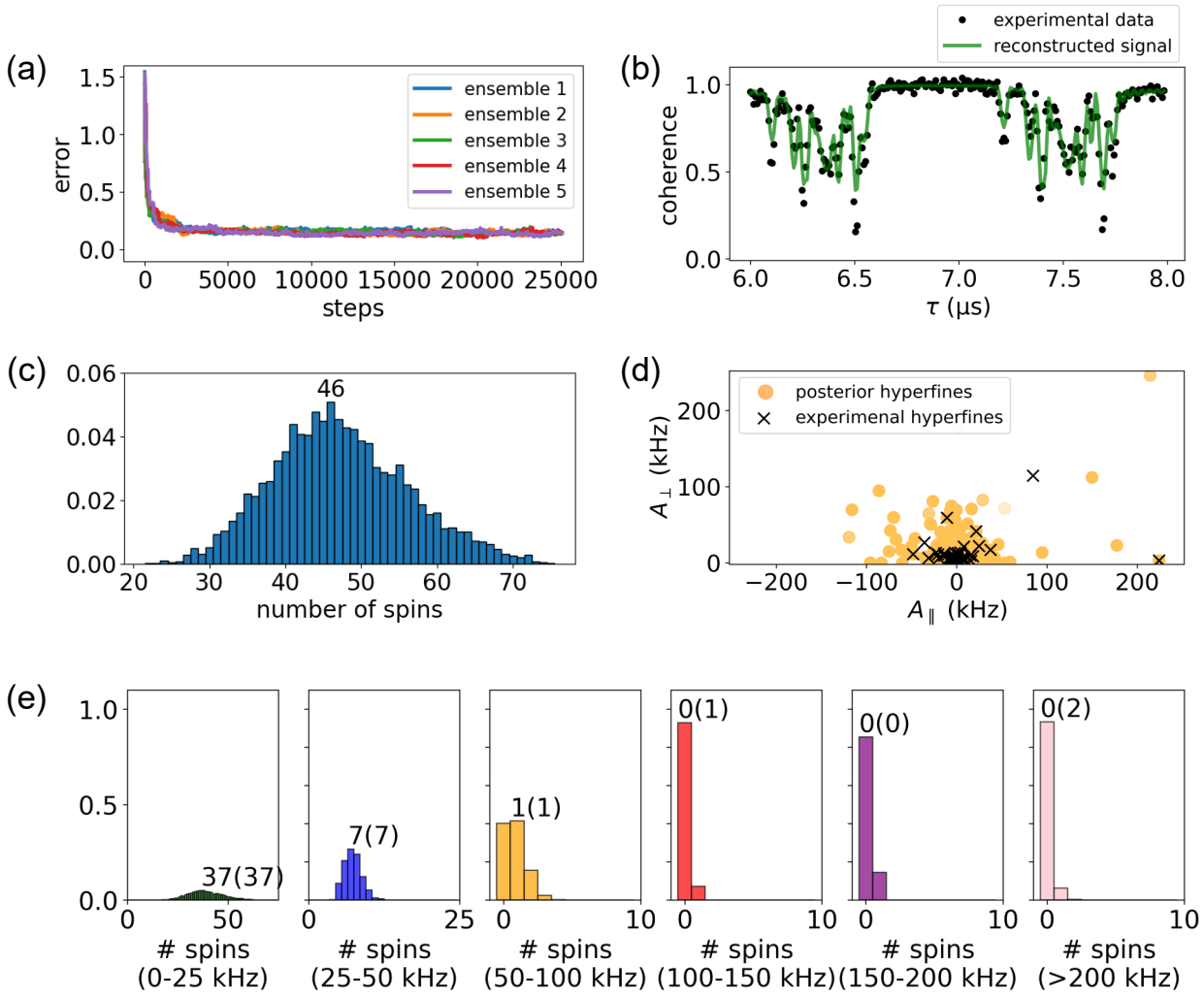

Figure 3.9: **Validation of hybrid MCMC approach on experimental data:** (a) Average error per data point as a function of the number of MCMC steps, with five random ensembles initialized for 25,000 steps each; the first 10,000 steps are treated as burn-in, and the remaining 15,000 steps are the posterior spin configurations. (b) Experimental coherence data (250 data points from 6 μs to 8 μs of a 32-pulse CPMG sequence) to we applied the hybrid MCMC algorithm, and the reconstructed signal using the best-fit spin configuration from the posterior distribution. (c) Posterior distribution of the total number of spins in the recovered spin configurations. (d) Posterior distribution of hyperfine couplings, with the 50 reported nuclear spin hyperfine couplings from the reference study overlaid. (e) Posterior distribution of the number of spins by spin magnitude ($\sqrt{A_\perp^2 + A_\parallel^2}$), with the mode displayed and the reference value reported in parentheses.

3.5.3 Validation on experimental data

Experimental data

We validate the hybrid MCMC approach by applying it to coherence data collected from a single NV center at a temperature of 3.7 K using a 32-pulse CPMG pulse sequence, with τ ranging from 6 to 8 μs. CPMG data were measured on a single, naturally occurring NV center in a high-purity chemical-vapor-deposition homoepitaxial diamond (type IIa, 1.1% natural abundance of ^{13}C) with $\langle 111 \rangle$ crystal orientation. A solid immersion lens with an aluminum-oxide anti-reflection coating (atomic-layer deposition) was used for enhanced photon collection. Microwave fields for electron spin manipulation were delivered via on-chip lithographically-defined strip lines, and a static magnetic field of $B_z \approx 403$ G was applied along the NV-axis using a room-temperature neodymium magnet, with stability < 3 mG and alignment within 0.07° (thermal echo calibration). The electron spin Rabi frequency was 14.31(3) MHz, with Hermite pulse shapes for efficient MW pulses, and phase alternation of π-pulses was performed using the XY-8 scheme. Experiments were conducted at 3.7 K in a commercial closed-cycle cryostat (Montana Cryostation), achieving single-shot NV spin-state readout fidelity of 94.5% via spin-selective resonant excitation. The electron spin relaxation time was $T_1 > 1$ h, natural dephasing time $T_2 = 4.9(2)$ μs, spin-echo coherence time $T_2 = 1.182(5)$ ms, and multipulse dynamical decoupling coherence time $T_{2,\mathrm{DD}} > 1$ s with optimized inter-pulse delay 2τ. This NV center has well-characterized hyperfine interactions and lattice positions of surrounding nuclear spins, determined from high-resolution multidimensional spectroscopy and correlated sensing experiments [35, 99]. These experimentally determined values serve as reference values.

Computational procedure for experimental validation

We use the hyperfine couplings and lattice positions calculated in [97] as theoretical reference values. The hybrid MCMC algorithm (Alg. 4) was initialized with five independent ensembles, each comprising 25,000 MCMC steps. All ensembles converged to the same error regime within 5,000 steps (Fig. 3.9a), and the initial 10,000 steps were discarded as burn-in. The posterior distribution was constructed from the remaining 15,000 steps.

To incorporate additional geometric information from the experimental coherence decay data, we modified the likelihood function from a standard squared-error form (Eq. 3.13) to a weighted likelihood function that balances point-wise residuals with global shape differences in the signal. Specifically, we introduce a regularized form that incorporates the squared 2-Wasserstein distance between the normalized empirical data distribution and the predicted signal distribution:

$$\mathcal{L}_{\text{mod}}(\mathbf{d}_d|\theta) = (1-\zeta)\cdot\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{d}\left(\mathbf{d}_i - f_k(\mathcal{H}_k(\theta),\tau_i)\right)^2\right) - \zeta\cdot W_2^2\left(f_k(\mathcal{H}_k(\theta)),\mathbf{d}\right) \tag{3.17}$$

where W_2^2 denotes the squared 2-Wasserstein distance between the empirical data distribution $\mathbf{d}$ and the predicted model distribution $f_k(\mathcal{H}_k(\theta))$, and ζ is a weighting parameter.

This hybrid formulation enables the inference algorithm to capture both local point-wise agreement and global distributional similarity. The squared-error term ensures that the model fits the observed decay values at each delay time τ_i, while the Wasserstein term penalizes global mismatches in the overall structure of the coherence decay curve—an important feature when dealing with noisy, underdetermined, or aliased experimental data [100]. This additional regularization is not required when testing our hybrid MCMC approach on simulated data generated from known hyperfine parameters. In the simulation case, the model is sampled from the same distribution as the simulated data, and there is no error due to

mismatches between the model and physical reality. Therefore, a standard likelihood based on squared residuals is sufficient for an accurate reconstruction, in the case of the simulated data.

In contrast, when analyzing experimental data, we know that the true hyperfine couplings differ systematically from the theoretical values used to define the sampling space [97]. As a result, directly comparing the predicted and experimental signals point-wise can lead to overfitting or biased inference. Incorporating the Wasserstein distance allows the inference process to remain robust to such mismatches by focusing on the agreement in the shape and distribution of the coherence signal, rather than on the exact point-wise correspondence. The weight $\zeta = 0.5$ was selected after benchmarking several values, where we show that this value of ζ balances local fit and global structure matching.

Table 3.1: Validation of Wasserstein weighting parameter ζ on experimental data. The number of nuclear spins detected from the posterior distribution is shown for various values of ζ.

ζ	Number of detected spins
0	20
0.25	18
0.5	22
0.75	12
1.0	2

As shown in Table 3.1, the value $\zeta = 0.5$ yielded the best balance between signal fidelity and structural regularization, resulting in the recovery of the most physically plausible number of spins (22). At $\zeta = 0$, the inference relies purely on the pointwise squared error, which can overfit noise and result in over-detection. Conversely, large values of ζ (e.g., $\zeta = 1$) heavily penalize small-scale mismatch and underfit the data, suppressing recovery.

This weighting is not needed for simulations using ground-truth hyperfine parameters, since both data generation and inference use consistent models (i.e., are drawn from the same distribution). For experimental data, however, uncertainties in measured hyperfine

couplings necessitate a regularization term to improve robustness and mitigate overfitting to systematic discrepancies.

Detection and recovery rates for experimental validation

We present the detection rates for hyperfine couplings (Table 3.2) and spatial positions (Table 3.3) of nuclear spins, computed from the posterior distribution sampled using the hybrid MCMC algorithm applied to 250 data points from an $N = 32$ CPMG experiment with τ ranging from 6 to 8 μs. These results are compared with experimentally reported values in Refs. [35, 99]. We align the computed lattice sites and hyperfine couplings, which were computed in Ref. [97], using the HSE06 functional with a supercell containing 1728 atoms, with the experimentally reported values in Ref. [99]. If an experimental nuclear spin position is reported with low precision, meaning that multiple lattice sites fall within the range defined by the standard deviation, we identify the "recovered" lattice site as the single lattice position within this range with the effective spin frequency (A^{-}_{rec}) that most closely matches the experimentally determined effective spin frequency (A^{-}_{exp}). There are two nuclear spins, C37 and C49, for which we could not find a good match between the experimentally reported positions and effective spin frequencies with the computed positions and effective spin frequencies. Similarly, while the posterior distribution contains additional hyperfine values corresponding to lattice sites not matched to experimental spins, only the subset of posterior hyperfines that can be associated with experimentally observed spins are reported in Table 3.2.

As noted in Ref. [97], there is a known sign inconsistency between the experimentally reported parallel components of hyperfine values and the values computed theoretically. To maintain consistency, we use the sign convention of the experimental data. The hyperfine components computed with the HSE06 hybrid functional in Ref. [97], which we also use in this work, were validated against three separate experimental datasets: Refs. [34, 99, 101].

One of these datasets is the same as the one we use for validation here. Because the HSE06 functional was tested on multiple datasets, we consider it a reliable choice for this study and are not overly concerned about overfitting to a single dataset. In Fig. 3.9(d), the yellow points correspond to all posterior hyperfines sampled by the MCMC algorithm, whereas the crosses indicate experimentally observed spins. The crosses plotted in Fig. 3.9(d) correspond to columns 7 and 8 (rather than columns 4 and 5) in Table 3.2, which was chosen to better visualize the correspondence between spins that could be recovered and the posterior hyperfine distribution. That is, all plotted crosses correspond to DFT-computed hyperfines, and in Tables 2 and 3 we provide the subset of spins for which we consider the DFT-computed hyperfines and experimentally measured hyperfines to be equivalent. The matching criterion uses both lattice positions and hyperfine couplings to identify corresponding spins.

We note that the 50 spins associated with this NV center were characterized using different types of experimental measurements, and therefore different amounts of hyperfine information are available across the set. All spins have reconstructed lattice coordinates (Table 3), which we use as the primary criterion to associate experimentally observed spins with DFT lattice sites for hyperfine comparison. For spins 1–27, both $m_s = \pm 1$ transition frequencies were measured, allowing direct extraction of $A_{\parallel}$ and $A_{\perp}$ (Table 2). For more information on the experimental procedures for spins 1-27, see [35]. For spins 28–50, only partial spectroscopic data are available (the $m_s = -1$ frequency and, for some spins, a local shift Δ_i), which provides information for $A_{\parallel}$ based on the local shift ($\Delta_i \approx A_{\parallel}$), but not for $A_{\perp}$. For more information on the experimental procedures for spins 28-50, see [99], and see the caption of Table 3.2 for more details on which approximations were used to compute hyperfine couplings for specific spins. This variation in available experimental data highlights the practical motivation for inference methods that reduce the need for time-consuming full hyperfine characterization measurements.

We calculate the baseline probabilities of nuclear spins appearing in a random selection

of 50 spins using the hypergeometric probability distribution. Specifically, we determine the probability that a spin from the set of 3,518 considered lattice sites is included in a random group of 50 spins without replacement, while taking into account the number of symmetrically equivalent lattice sites for each spin with respect to its hyperfine couplings. This approach allows us to compare the observed detection rate of each spin in the posterior distribution with the probability of its random occurrence.

Using the hypergeometric distribution, we determine baseline probabilities of 0.0420 for a site with three symmetrically equivalent sites, 0.0824 for six, 0.1210 for nine, and 0.1581 for twelve. For nuclear spins with identical hyperfine couplings (C8 and C50, C17 and C27, and C42 and C45), we calculate the baseline probability for a second spin based on the probability of a spin appearing at least twice. This results in probabilities of 0.0029 for a site with six symmetrically equivalent sites and 0.0119 for a site with twelve. Using these baseline thresholds, we find that 22 of the nuclear spins reported in [99] appear in the posterior distributions sampled by the hybrid MCMC method more frequently than expected by chance.

Some recovered spins have detection rates of zero (see Table 3.2), meaning they are absent in all posterior samples. This occurs in the presence of sparse, noisy experimental data and inaccurate priors, highlighting the ill-posedness of the inverse problem: certain spin configurations cannot be reliably resolved given the limited information content. Importantly, the goal of this Bayesian inference approach is not exact, high-precision reconstruction but rather to provide informative posterior distributions that quantify uncertainty and identify plausible nuclear spin configurations. Hence, the method is particularly useful for rapid, high-throughput screening of spin environments, guiding experimental focus and helping prioritize spins for further investigation, rather than producing perfect reconstructions from limited data.

We do not provide explicit guidelines for a minimum detection rate required to consider

a prediction "reliable," because this threshold is highly context dependent. It can vary based on the specific experimental system, the quality and quantity of available data, and the application or target observable of interest. For example, in regimes with sparse or noisy data, even a moderate detection rate may provide useful information for screening or identifying likely spin candidates, whereas in high-precision applications, only spins with very high detection rates may be considered reliable. Similarly, the acceptable threshold may depend on whether one is interested in qualitative identification of spin presence, quantitative recovery of hyperfine parameters, or downstream tasks such as simulating spin-bath dynamics. In practice, the detection rate should be interpreted alongside other metrics, such as the baseline probability, posterior uncertainty, and consistency with prior knowledge, to assess the confidence in a given prediction.

	A^-_{exp} (kHz)	A^-_{rec} (kHz)	$A_{\parallel,\text{exp}}$ (kHz)	$A_{\perp,\text{exp}}$ (kHz)	$A_{\parallel,\text{rec}}$ (kHz)	$A_{\perp,\text{rec}}$ (kHz)	detection rate
C1	452.83(2)	452.34	-20.72(1)	12(1)	-20.65	11.18	0.0797
C2*	455.37(2)	455.01	-23.22(1)	13(1)	-23.27	13.16	0.8760
C3*	463.27(2)	462.73	-31.25(2)	8(2)	-31.14	6.38	0.0498
C4*	446.23(4)	445.73	-14.07(2)	13(1)	-14.01	12.42	0.0564
C5	447.234(1)	446.31	-11.346(2)	59.21(3)	-10.81	59.27	0.0583
C6	480.625(1)	480.08	-48.58(2)	9(2)	-48.39	11.51	0.0383
C7	440.288(6)	439.88	-8.32(1)	3(5)	-8.33	1.08	0.0163
C8	441.77(1)	441.37	-9.79(2)	5(4)	-9.79	4.51	0.0683
C9	218.828(1)	207.39	213.154(1)	3.0(4)	224.18	3.06	0.0603
C10	414.407(1)	413.76	17.643(1)	8.6(2)	17.88	8.75	0.0679
C11*	417.523(4)	416.54	14.548(3)	10(1)	15.13	9.85	0.1585
C12	413.477(1)	412.07	20.569(1)	41.51(3)	21.57	41.50	0.0036
C13	424.499(1)	424.15	8.029(1)	21.0(4)	7.92	21.04	0.0

C14	451.802(1)	451.38	-19.815(3)	5.3(5)	-19.79	6.19	0.0593
C15*	446.01(5)	445.71	-13.961(3)	9(1)	-14.07	9.43	0.1108
C16*	436.67(5)	436.26	-4.66(3)	7(4)	-4.65	7.44	0.1025
C17*	437.61(1)	437.24	-5.62(1)	5(2)	-5.68	1.33	0.0991
C18	469.02(1)	468.14	-36.308(1)	26.62(4)	-35.83	26.69	0.0
C19	408.317(1)	407.58	24.399(1)	24.81(4)	24.57	22.12	0.0035
C20	429.403(4)	429.02	2.690(4)	11(1)	2.67	10.96	0.0462
C21*	430.937(3)	430.58	1.212(5)	13(1)	1.16	12.82	0.0966
C22*	424.289(3)	423.83	7.683(4)	4(3)	7.75	5.37	0.1092
C23	435.143(7)	436.19	-3.177(5)	2(4)	-4.61	5.23	0.1276
C24*	436.183(3)	436.84	-4.225(4)	0(6)	-5.29	1.04	0.1256
C25	435.829(5)	436.89	-3.873(5)	0(4)	-5.33	1.66	0.1519
C26	435.547(2)	435.23	-3.618(5)	0(2)	-3.58	9.59	0.0982
C27*	435.99(3)	437.24	-4.039(5)	0(3)	-5.68	1.33	0.0035
C28	440.9(1)	440.51	-8.91547(3)†	4.65†	-8.48	5.90	0.0287
C29*	434.3(1)	433.83	-2.1857(1)†	14.86†	-2.09	13.13	0.0890
C30	427.1(1)	426.64	4.87111(4)†	3.08†	4.94	5.43	0.1100
C31	428.3(1)	423.09	-3.66‡	0‡	8.48	4.32	0.0538
C32*	431.6(1)	415.60	-0.36‡	0‡	15.96	2.36	0.0835
C33*	439.0(1)	438.58	7.04‡	0‡	-7.00	4.76	0.1137
C34	437.3(1)	436.88	5.34‡	0‡	-5.32	3.12	0.0703
C35	427.4(1)	424.82	4.59133(4)†	5.17†	6.74	3.33	0.1213
C36	434.4(1)	365.41	-2.2142(8)†	14.00†	84.52	114.42	0.0
C37	429.1(1)	-	-2.86‡	0‡	-	-	-
C38*	434.0(1)	433.44	2.04‡	0‡	-1.83	7.03	0.0996
C39	432.5(1)	431.38	-0.450(5)†	8.82†	0.20	5.15	0.0267

C40*	433.3(1)	433.24	-1.173(5)†	12.03†	-1.66	5.75	0.0979
C41*	434.1(1)	430.69	-2.1893(6)†	0†	0.90	6.15	0.1749
C42	434.8(1)	437.08	2.84‡	0‡	-5.52	3.58	0.1462
C43*	432.2(1)	427.83	-0.2701(6)†	0†	3.94	13.63	0.0904
C44	433.9(1)	435.32	-1.8828(6)†	7.05†	-3.74	5.46	0.0903
C45*	436.2(1)	437.08	-4.174(1)†	7.59†	-5.52	3.58	0.0176
C46	434.8(1)	430.43	2.84‡	0‡	1.15	5.31	0.0332
C47*	429.4(1)	423.98	2.5878(3)†	4.89†	7.51	2.56	0.1065
C48*	431.0(1)	394.41	-0.96‡	0‡	37.51	17.04	0.9588
C49	428.3(1)	-	-3.66‡	0‡	-	-	-
C50	436.2(1)	207.39	-4.2270(4)†	3.37†	224.19	3.06	0.0026

Table 3.2: **Comparison of experimentally measured and recovered hyperfine interaction values.** This table compares experimentally measured and recovered hyperfine interaction values for nuclear spins, following the nuclear spin labeling convention in [35, 99]. The experimental data in this Table combines data from two different kinds of experimental measurements. The data for nuclear spins C1-C27 is from [35], in which the hyperfine components $(A_{\parallel,\text{exp}}, A_{\perp,\text{exp}})$ are fit from direct measurements of the ^{13}C spin precession frequencies for the $m_s = \pm 1$ electron spin projections (see Sec. I of the SI of [35] for more information about the experimental measurements and the fitting procedures). The data for nuclear spins C28-C50 is from [99], in which only the spin frequencies (A^-_{exp}) corresponding to the $m_s = -1$ electron projection are measured, and for a subset of these nuclear spins (denoted by †), a hyperfine shift Δ_i is measured (see Supplementary Note I of [99] for more details on experimental measurements and a derivation of the following approximations). For the nuclear spins denoted by †, we approximate $A_{\parallel,\text{exp}} \approx \Delta_i$, and we recover $A_{\perp,\text{exp}}$ based on $A^-_{\text{exp}} = \sqrt{(\omega_L - A_{\parallel,\text{exp}})^2 + A^2_{\perp,\text{exp}}}$, and any imaginary numbers are reported as $0^\dagger$. For the spins (denoted with ‡) where Δ_i was not measured (due to low polarization, coherence, or readout contrast), we approximate $A_{\parallel,\text{exp}} \approx \omega_L - A^-_{\text{exp}}$, where we have assumed that $A_{\parallel,\text{exp}}$ is small and $A_{\perp,\text{exp}} \approx 0$. These experimental values are compared with the hyperfine components $(A_{\parallel,\text{rec}}, A_{\perp,\text{rec}})$ recovered using a hybrid MCMC algorithm (see Sec. 3.5.3). We report A^-_{rec} as $A^-_{\text{rec}} = \sqrt{(\omega_L - A_{\parallel,\text{rec}})^2 + A^2_{\perp,\text{rec}}}$. Detection rates of individual nuclear spins are determined based on the frequency of their appearance in spin configurations sampled from the posterior distribution. Spins can only be identified up to a degree of symmetry determined by the number of lattice sites with identical hyperfine couplings; however, we properly account for the multiplicity of spins with identical hyperfine couplings but different lattice sites when computing detection rates. Spins with detection rates above their baseline probabilities are marked with an asterisk (*).

	Experimental values			Recovered values			# of
	x (Å)	y (Å)	z (Å)	x (Å)	y (Å)	z (Å)	symmetric spins
C1	0.0	0.0	0.0	0.0	0.0	0.0	6
C2*	2.52	2.91	-0.51	2.52	2.91	-0.52	6
C3*	3.78	0.73	-0.51	3.78	0.73	-0.52	3
C4*	-1.26	2.18	0.0	-1.26	2.18	-0.01	3
C5	0.0	4.37	-6.18	0.0	4.36	-6.18	6

C6	5.04	-1.46	-2.06	5.04	-1.46	-2.06	3
C7	5.04	-1.46	5.66	5.04	-1.46	5.66	3
C8	7.57	1.46	3.6	7.56	1.45	3.60	6
C9	7.57	-4.37	-10.81	7.56	-4.37	-10.82	6
C10	0.0	8.74	-12.36	0.0	8.73	-12.36	6
C11*	6.31	9.46	-12.87	6.30	9.46	-12.88	6
C12	11.35	0.73	-14.42	11.35	0.73	-14.42	6
C13	12.61	2.91	-6.69	12.61	2.91	-6.70	6
C14	5.04	-2.91	-22.65	5.04	-2.91	-22.66	6
C15*	1.26	3.64	-22.65	1.26	3.64	-22.66	6
C16*	2.52	8.74	-23.17	2.52	8.74	-23.17	6
C17*	6.31	-2.18	-29.34	6.30	-2.19	-29.35	6
C18	0.0	-1.46	-19.05	0.0	-1.46	-19.06	6
C19	3.78	-9.46	-8.75	3.78	-9.46	-8.76	3
C20*	3.78	10.92	-4.63	3.78	10.92	-4.64	3
C21*	-5.04	5.82	-4.12	-5.05	5.82	-4.12	6
C22*	16.39	-3.64	-8.24	16.39	-3.64	-8.24	6
C23	13.88	0.73	5.66	12.61	1.45	3.60	12
C24*	1.26	-0.73	9.78	1.26	-0.73	8.23	6
C25	7.57	1.46	9.78	6.30	2.18	7.72	12
C26	12.61	-5.82	-0.51	12.61	-5.83	-0.52	9
C27*	1.26	3.64	-31.4	1.26	2.18	-29.35	6
C28	-1.26	2.18	-24.71	-1.26	2.18	-24.72	3
C29*	-6.31	0.72	-19.05	-6.31	0.73	-19.05	6
C30	12.62	10.19	-14.93	12.61	10.19	-10.30	12
C31	11(4)	15(4)	-11(4)	7.56	11.65	-12.88	6

C32*	6(3)	12(2)	-3(9)	5.04	10.19	-14.94	6
C33*	-2.52	-1.46	4.12	-2.53	-1.46	4.11	6
C34	-2.52	0.0	6.18	-2.53	0.0	6.17	6
C35	20(3)	-1(7)	-8(1)	17.65	1.45	-8.76	12
C36	-8(7)	-0(7)	-23(9)	-1.26	0.73	-14.42	6
C37	13(2)	-12(1)	-18(1)	-	-	-	-
C38*	12.61	-8.73	0.0	12.61	-8.74	-0.01	6
C39	16(5)	-8(5)	-24(4)	16.39	-8.01	20.60	3
C40*	11(3)	-14(3)	-26(4)	10.09	-11.65	-22.66	6
C41*	8(4)	-12(4)	2(6)	8.82	-13.84	-2.58	12
C42	10.09	-7.28	-28.83	8.82	-5.10	-27.29	12
C43*	11(3)	7(3)	-19(3)	12.61	5.82	-16.48	6
C44	0.0	11.65	-26.77	0.0	8.74	-24.72	9
C45*	6.3	6.56	-29.35	5.04	5.82	-27.29	12
C46	14(4)	-5(4)	-24(5)	17.65	-5.83	-19.05	6
C47*	21(4)	-7(7)	-14(7)	17.65	0.0	-12.36	6
C48*	16(4)	-4(4)	-21(9)	12.61	0.0	-12.36	6
C49	23(4)	-4(8)	-12(3)	-	-	-	-
C50	11(7)	5(7)	-2(9)	5.04	4.36	-10.82	6

Table 3.3: **Comparison of experimentally measured and recovered nuclear spin positions.** We present the spatial coordinates (x, y, z) of nuclear spins determined experimentally [35, 99] alongside the recovered coordinates obtained using a hybrid MCMC algorithm. The lattice used in the hybrid MCMC algorithm was aligned to be consistent with the experimentally reported nuclear lattice positions and the experimental values of effective nuclear spin frequencies and hyperfine couplings. The column labeled "# of symmetric spins" indicates the number of lattice sites with identical hyperfine couplings. Of these symmetric lattice sites, the position closest to the experimentally reported value is shown.

The reconstructed coherence signal (Fig. 3.9b) captures many of the experimentally observed features, particularly the smaller peaks, but underestimates the depths of the strongest peaks. This is consistent with the known limitations in the accuracy of theoretical hyperfine calculations, particularly for strongly coupled nuclear spins.

The posterior distribution of the total number of nuclear spins (Fig. 3.9c), with a modal value of 46, shows excellent agreement with the experimental reference value of 50. We further evaluated the posterior distribution of hyperfine couplings (Fig. 3.9d). A total of 45 out of the 48 experimentally identified nuclear spins are detected in the posterior distribution, with 22 of these spins exhibiting detection rates exceeding their baseline probabilities (Table 3.2). One strongly coupled spin (labeled C36 in Table 2 with $(A_{\parallel}, A_{\perp}) = (84.52, 114.42)$ kHz) is not detected in the posterior distribution (see Fig. 3.9d). This illustrates the combined effects of sparse experimental data and known discrepancies between DFT-computed and experimentally measured hyperfine couplings: when the theoretical hyperfine values deviate from experiment, the posterior detection probability can be reduced even for otherwise well-resolved spins. As shown in Sec. 3.5.2, recovery rates for both spin number and hyperfine couplings remain high when the input couplings are accurate to within ~ 1 kHz, while larger discrepancies between the DFT-computed and experimentally-measured hyperfine couplings, as in the case of C36, lead to decreased detection.

Our approach offers significant advantages over prior methods in terms of data efficiency, computational cost, and adaptability. A machine learning method proposed in Ref. [80] identified 23 nuclear spins in the same experimental dataset but required 14 hours of computation using both CPU and GPU resources (an Intel Core i9-9920X with 12 cores and 2× NVIDIA RTX 2080 Ti), along with 11,000 data points for a 32-pulse experiment (and 7,500 for a 256-pulse one). Another ML-based method for this same inverse problem [81], which to the best of our knowledge has not yet been experimentally validated, uses 3.6 million synthetic training samples, each of which requires a forward model evaluation of the spin

coherence model. Although they do not report computational time, training data generation alone involves significantly more model evaluations than our hybrid approach, which requires approximately 750,000 forward evaluations in total (25,000 MCMC walker steps per ensemble with an additional approximately 125,000 evaluations per ensemble during parallel tempering, with 5 ensembles run in total). Furthermore, their model must be retrained when experimental conditions, such as magnetic field or pulse sequences, are changed.

In contrast, our method achieves comparable performance using only 250 data points from a single 32-pulse experiment and runs in approximately 8 hours using only 5-cores of a single CPU node (Intel Xeon E5-2680 v4), without requiring pre-training or additional tuning when experimental conditions are changed. The experimental dataset used here was originally acquired for different studies [35, 99]; we sub-sampled the available time points to test the inference framework in a regime currently inaccessible to prior inverse-problem techniques, demonstrating that accurate spin detection can be achieved with substantially fewer measurements. Typical acquisition of a single 32-pulse coherence signal with 250 time points requires roughly 5–10 hours under standard laboratory conditions, including NV selection, alignment, and calibration, making the measurement time comparable to the runtime of the inference algorithm. Compared to another non-ML approach proposed in Ref. [102], which identified 14 experimental spins using high-resolution data (around 10,000 time points of the coherence signal) across multiple ^{13}C Larmor periods, our method identifies more spins using a much sparser dataset. The tradeoff of working with sparse data is the persistence of ill-posedness; however, our Bayesian inference framework reduces the number of plausible lattice sites from 3,518 to 230 by leveraging posterior detection probabilities.

Finally, the hybrid MCMC method accurately characterizes the distribution of nuclear spins across different hyperfine magnitude regimes (Fig. 3.9e). Specifically, the modes of the posterior distributions for weakly coupled spins ($<$100 kHz) align exactly with reference values, while the number of strongly coupled spins is slightly underestimated. This outcome

is consistent with the known limitations of the theoretical hyperfine values.

3.6 Discussion

Hybridized Markov Chain Monte Carlo (MCMC) methods offer a flexible and powerful framework for addressing trans-dimensional inverse problems, particularly under ill-posed conditions. Although such hybrid approaches, where different MCMC strategies are combined to enhance sampling efficiency, have found broad utility in fields such as ecology [103], evolutionary biology [104], and geophysics [105, 106], their application to quantum systems presents unique advantages. This is largely due to the discrete nature of quantum states and the hierarchical structure of prior knowledge that often accompanies physical parameters, whether derived from experimental measurements or from first-principles calculations. When prior knowledge strongly constrains the possible values of certain parameters, such as nuclear spin quantum numbers or symmetry-allowed configurations, it is advantageous to represent these parameters as discrete, thereby restricting the search space to physically admissible values and reducing computational cost. Conversely, for parameters where the uncertainty is significant, a continuous representation may be more appropriate, allowing the model to explore a broader and less constrained space informed by physically motivated probability distributions. This dual treatment of model parameters—as either discrete or continuous depending on the level of prior information—provides a systematic means of integrating *a priori* knowledge into the sampling process, balancing computational efficiency with physical rigor.

Crucially, the hybrid MCMC framework does not depend on the availability of DFT-derived hyperfine tensors and remains fully operational even in their absence. When the crystal lattice is known but detailed electronic-structure information is unavailable, the same hybrid sampling strategy can be applied over discrete lattice sites using magnetic dipolar hyperfine interactions computed analytically and directly from the fixed lattice geometry. This

lattice-constrained formulation captures the dominant distance- and orientation-dependent physics governing the nuclear–electron interaction while substantially reducing the dimensionality of the inference problem. A continuous RWMH scheme can then relax the hyperfine coupling parameters associated with each lattice site around their approximate dipolar values, without relaxing the underlying lattice positions themselves. This hybrid inference strategy establishes the approach as a broadly applicable and practical technique for reconstructing nuclear spin environments across a wide class of spin-defect systems, including semiconductors for which reliable DFT hyperfine calculations are unavailable or impractical to compute. More generally, in systems where even lattice-resolved prior information is unavailable, the same Bayesian framework can be implemented in a fully continuous trans-dimensional setting using a continuous formulation of RJMCMC [90] or continuous RWMH sampling over both the number and spatial coordinates of nuclear spins without requiring any discrete lattice constraints.

A central challenge in applying Markov Chain Monte Carlo (MCMC) methods to quantum systems lies in the high computational cost associated with forward models, which map parameters of interest to observable quantities. In quantum materials and molecular systems, such forward models are often based on *ab initio* electronic structure methods, such as density functional theory (DFT), whose computational cost typically scales as $O(N^3)$ with system size N. For certain physical properties that require even more computationally expensive methods, such as coupled-cluster or multireference methods, the scaling becomes substantially worse—up to $O(N^6)$ or beyond—rendering them impractical for large systems or for problems requiring extensive sampling. In trans-dimensional Bayesian inference, MCMC algorithms often require tens of thousands of likelihood evaluations to adequately sample posterior distributions. If each of these evaluations involves a costly first principles calculation, the cumulative computational burden becomes prohibitive. Thus, the development of surrogate models, reduced-order methods, or data-driven approximations that retain

quantum mechanical fidelity while significantly lowering computational overhead is essential for enabling scalable inference of the properties of quantum systems.

One promising strategy to address the computational bottleneck is the use of machine learning (ML) to construct surrogate models for forward simulations. Rather than attempting to learn the inverse map from observables to model parameters—which is fundamentally ill-posed and highly sensitive to noise—ML approaches can be employed to approximate the forward model itself, which is well-posed and grounded in the underlying physics. Surrogate models trained on a limited set of high-fidelity quantum mechanical calculations can reproduce system observables with sufficient accuracy while drastically reducing the computational cost of individual likelihood evaluations. This approach enables efficient sampling within MCMC frameworks, without compromising the interpretability or uncertainty quantification inherent to Bayesian inference. Recent studies have demonstrated the effectiveness of surrogate ML models in capturing structure–property relationships in a range of quantum systems, including the prediction of spectroscopic signatures, electronic properties, and energy landscapes [107–112]. The integration of such models into inference workflows provides a scalable path forward for the characterization of the properties of quantum materials and molecular systems, especially in regimes where direct simulation remains computationally prohibitive.

A notable strength of the Bayesian framework, and hybridized MCMC approaches in particular, lies in their systematic improvability and conceptual flexibility. Enhancements in the accuracy or efficiency of forward models—whether through advances in numerical algorithms or the introduction of machine-learned surrogates—translate directly into improved sampling performance and more accurate inference, without requiring fundamental changes to the underlying methodology. Moreover, physical constraints, conservation laws, and symmetries can be naturally and rigorously incorporated into the likelihood function, allowing the inference process to remain firmly grounded in the physics of the problem. Modifying

the forward model or refining the likelihood to reflect improved understanding or additional experimental input incurs minimal overhead, enabling a modular and extensible workflow. This stands in contrast to direct machine learning approaches, which typically require retraining on new datasets and often involve nontrivial engineering of loss functions to enforce physical consistency. In this sense, Bayesian inference provides a transparent and adaptable framework for the characterization of many properties of quantum systems, capable of evolving alongside theoretical and computational advances.

A potential limitation of the hybrid MCMC approach arises when the true Hamiltonian does not lie within the predefined class of candidate models. However, the framework is naturally compatible with established techniques for model selection and averaging, which can mitigate issues related to model misspecification. Bayesian model averaging [113], mixture models [114], and ensemble strategies [115] such as bagging offer systematic approaches to incorporate uncertainty over model structure. In particular, bagging methods—originally developed in the context of machine learning—construct ensembles by training models on bootstrapped subsets of the data and combining their predictions through voting, averaging, or regularized aggregation to improve robustness; such procedures are adopted in sparse data regimes where small variations in the dataset may lead to significant shifts in the inferred posterior [116]. By enabling principled treatment of both parameter and model uncertainty, such extensions could preserve the interpretability and physical grounding of the Bayesian approach while enhancing its stability and generalizability. Nonetheless, if the true model lies far outside the chosen candidate set, the current framework cannot generate a fully data-driven model; some choice of forward model or candidate class is always required. This feature distinguishes our approach from purely data-driven or machine learning–based methods. At the same time, leveraging substantial prior information about the physical system, such as Hamiltonian structure and DFT-based priors, is precisely what allows us to extract meaningful information from this highly ill-posed inverse problem.

Another potential limitation of MCMC-based strategies is their computational cost, which can become significant for high-dimensional models. This is an inherent feature of sampling-based approaches: the walkers require a burn-in period to explore the posterior distribution before producing reliable samples, and each likelihood evaluation can be expensive, particularly when forward models involve complex physics or first-principles calculations. While surrogate models [117] and reduced-order approximations [118] can mitigate the computational burden, the fundamental requirement for a sufficiently long burn-in and adequate sampling cannot be eliminated. High-dimensional problems require more computational resources, and careful attention must be paid to balancing model complexity, sampling efficiency, and available computational power.

Future developments of hybrid MCMC methods for ill-posed inverse problems will benefit from several complementary advances. Adaptive hybridization schemes that dynamically toggle between discrete and continuous parameter spaces in response to data quality and uncertainty could significantly enhance sampling efficiency and robustness. The integration of surrogate models—potentially guided by active learning protocols—promises not only to accelerate convergence but also to enable real-time Bayesian updating, laying the foundation for autonomous, data-driven experimental design and continuous data assimilation. Moreover, realizing the full potential of these approaches will require efficient, scalable implementations of parallel and distributed MCMC algorithms capable of exploiting modern high-performance computing architectures. Coupled with user-friendly software frameworks that facilitate hybridization across diverse MCMC strategies, such advances will be crucial to the widespread adoption of hybrid MCMC for rapid and scalable quantum characterization across a broad range of physical quantum systems.

Algorithm 4 Hybridized MCMC algorithm for recovering local nuclear spin environment

Input: data $\mathbf{d}_d$, total number of steps N_{tot}, number of RWMH steps N_{RWMH}, number of RJMCMC steps N_{RJMCMC}, number of PT steps N_{PT}, discrete parameter corresponding to nuclear spins $\mathbf{a}_k = \{p_{d_1}, \ldots, p_{d_k}, k\}$, continuous parameters corresponding to relative decoherence $\mathbf{b} = \{\lambda\}$, fixed parameters in the forward model $\mathbf{c} = \{N, \tau, B_z\}$, fixed Hamiltonian model $\mathcal{H}_k(\mathbf{a}_k)$ indexed by k, likelihood model $\mathcal{L}(\mathbf{d}_d | f_k(\mathbf{a_k}, \mathbf{b}, \mathbf{c}))$, and hyperparameters for the individual MCMC algorithms

for $n = 0, \ldots N_{\text{tot}} - 1$ **do**
  $n_{\text{RWMH}} \leftarrow 0$
  $n_{\text{RJMCMC}} \leftarrow 0$
  $n_{\text{PT}} \leftarrow 0$
  for $n_{\text{RWMH}} = 0, \ldots, N_{\text{RWMH}} - 1$ **do**
    RWMH (Alg. 1) with
    Input: $\mathbf{p} = \{\lambda\}$, $\mathbf{q} = \{\mathbf{a}_k, k\}, \mathbf{p}^{(0)} = \{\lambda^{n+n_{\text{RWMH}}-1}\}, \mathbf{q}^{(0)} = \{\mathbf{a}_k^{n+n_{\text{RWMH}}-1}, k^{n+n_{\text{RWMH}}-1}\}$
    Output: $\{\mathbf{p}^n\}_{n=n}^{n+N_{\text{RWMH}}} = \{\lambda^n\}_{n=n}^{n+N_{\text{RWMH}}}$, $\{\mathbf{q}^n\}_{n=n}^{n+N_{\text{RWMH}}} = \{\mathbf{a}_k^n, k^n\}_{n=n}^{n+N_{\text{RWMH}}}$
    $n \leftarrow n + n_{\text{RWMH}}$
  for $n_{\text{RJMCMC}} = 0, \ldots, N_{\text{RJMCMC}} - 1$ **do**
    RJMCMC (Alg. 2) with
    Input: $\mathbf{p} = \{\mathbf{a}_k\}$, $\mathbf{q} = \{\lambda\}, \mathbf{p}^{(0)} = \{\mathbf{a}_k^{n+n_{\text{RJMCMC}}-1}\}, \mathbf{q}^{(0)} = \{\lambda^{n+n_{\text{RJMCMC}}-1}\}$
    Output: $\{\mathbf{p}^n\}_{n=n}^{n+N_{\text{RJMCMC}}} = \{\mathbf{a}_k^n\}_{n=n}^{n+N_{\text{RJMCMC}}}$, $\{\mathbf{q}^n\}_{n=n}^{n+N_{\text{RJMCMC}}} = \{\lambda^n\}_{n=n}^{n+N_{\text{RJMCMC}}}$, $\{k^n\}_{n=n}^{n+N_{\text{RJMCMC}}}$
    $n \leftarrow n + n_{\text{RJMCMC}}$
  for $n_{\text{PT}} = 0, \ldots, N_{\text{PT}} - 1$ **do**
    PT (Alg. 3) with
    Input: $\mathbf{p} = \{\mathbf{a}_k\}$, $\mathbf{q} = \{\lambda, k\}, \mathbf{p}^{(0)} = \{\mathbf{a}_k^{n+n_{\text{PT}}-1}\}, \mathbf{q}^{(0)} = \{\lambda^{n+n_{\text{PT}}-1}, k^{n+n_{\text{PT}}-1}\}$
    Output: $\{\mathbf{p}^n\}_{n=n}^{n+N_{\text{PT}}} = \{\mathbf{a}_k^n\}_{n=n}^{n+N_{\text{PT}}}$, $\{\mathbf{q}^n\}_{n=n}^{n+N_{\text{PT}}} = \{\lambda^n, k^n\}_{n=n}^{n+N_{\text{PT}}}$
    $n \leftarrow n + n_{\text{PT}}$

Output: append all $\mathbf{a}_k$, λ, and k to obtain $\{\mathbf{a}_k^n\}_{n=0}^{N_{\text{tot}}}$, $\{\lambda^n\}_{n=0}^{N_{\text{tot}}}$, and $\{k^n\}_{n=0}^{N_{\text{tot}}}$

CHAPTER 4

HIGH-THROUGHPUT CHARACTERIZATION OF SPIN-DEFECTS IN SEMICONDUCTORS FOR SCALE-UP OF QUANTUM SENSING

Detailed knowledge of the local environments of spin-defects in semiconductors, such as nitrogen vacancy (NV) centers in diamond or divacancies in silicon carbide, is crucial for optimizing control and entanglement protocols in quantum sensing and information applications. However, at present a direct experimental characterization of individual defect environments is not scalable, as conventional spin-bath measurements are time-consuming and difficult to automate. Achieving high-throughput characterization requires short experiments to probe the spin-bath. However, with fewer and noisier measurements, the inverse problem of recovering spin-bath properties from measured data becomes ill-posed, with multiple spin-baths having a high likelihood of yielding the same data. In this work, we present a set of computational tools to resolve the ill-posed inverse problem of recovering the atomic positions and hyperfine couplings of random nuclei surrounding spin-defects from sparse, noisy experimental coherence data, which can be obtained in hours. We use a trans-dimensional Bayesian approach that incorporates *ab initio* data to yield full posterior distributions over nuclear spin environments, enabling robust recovery from limited data. We also provide practical tools and guidelines to determine the limits of detectability for hyperfine couplings under specific dynamical decoupling sequences and sampling conditions. In addition, we demonstrate how the tools developed here, in combination with *ab initio* simulations of spin-baths, can guide the design of efficient experimental protocols for application-specific high-throughput screening. To showcase the utility of our approach, we apply it to design fast dynamical decoupling experiments to characterize the spin-baths of 10 individual NV centers in diamond. While the primary focus is on accelerating spin-bath characterization

of spin-defects, this Bayesian approach also lays the foundation for digital twin studies of spin-defects, where a virtual model of the spin-defect system evolves in real time with ongoing experimental measurements. Together, the set of tools we designed and applied paves the way for scalable deployment of spin-defects in semiconductors for quantum sensing and information applications.

4.1 Introduction

Optically active spin defects in solid-state systems, such as phosphorus donors in silicon [119], divacancies in silicon carbide [67], and nitrogen-vacancy (NV) centers in diamond [68], have emerged as promising platforms for quantum information science due to their long coherence times [28, 69] and their coupling to local nuclear spin environments [70, 71]. While long coherence times are critical for quantum applications, the presence of naturally occurring nuclear spin isotopes in these systems introduces both challenges and opportunities. On the one hand, nuclear spins are one of the main mechanisms of decoherence, necessitating precise characterization of the local environment to optimize control sequences and enhance performance [72]. On the other hand, these same nuclear spins may serve as auxiliary quantum resources, enabling functionalities such as local quantum registers or quantum memories, provided their hyperfine couplings and spatial configurations are known [34, 73–75].

Recent advances have enabled scalable fabrication of optically-active spin-defects in semiconductors (e.g., billions of NV centers in diamond can be created within a few days [76, 77]), prompting a need for high-throughput characterization tools capable of quickly screening large numbers of candidate defects for nuclear spin bath properties for a target quantum information application. While scalable techniques exist to assess optical and spin properties [120], there is currently no efficient approach to determine the local nuclear spin environment of individual defects at scale. High-resolution methods such as correlation spectroscopy pro-

vide access to detailed hyperfine and spatial information but require hours of integration time per data point and manual tuning of pulse sequences, rendering them impractical for large-scale deployment [35, 79]. Machine learning approaches have shown promise for extracting nuclear spin information from coherence data [80, 81], but they depend on high-quality, high-resolution datasets that are infeasible to collect in high-throughput settings.

In this work, we address the ill-posed, trans-dimensional inverse problem of identifying the number, positions, and hyperfine couplings of nuclear spins interacting with a single spin-defect from sparse, noisy coherence data. We focus on a data regime in which experimental coherence signals can be measured rapidly but may be limited in resolution and precision. We refer to this regime as as the sparse, noisy regime, which enables rapid initial screening of candidate defects but poses fundamental challenges for accurate characterization. The inverse problem is ill-posed because the same experimental data may be consistent with multiple distinct nuclear spin configurations, each with different numbers, positions, and hyperfine couplings of nuclear spins.

Existing approaches for extracting information about the nuclear spin environment of single defects from low-resolution data often rely on fitting parameterized models, using off-the-shelf nonlinear optimization methods [76, 121]. These methods require specifying the number of nuclear spins *a priori*, necessitating a separate model selection step, and often ignore physical constraints such as spatial lattice geometry or bandwidth limits. As a result, they may produce unphysical or uninterpretable results. Classical statistical approaches favor lower dimensional models when errors between higher- and lower-dimensional parameterizations are comparable [122], but in this setting, there is no physical reason to prefer, for instance, a five-spin nuclear spin configuration over a ten-spin one if both lie within the the expected statistical distribution for a given isotope concentration.

To achieve efficient and scalable characterization of nuclear spin environments of single defects, we apply a Bayesian framework for recovering nuclear spin environments from sparse

coherence data acquired from dynamical decoupling experiments that we introduced in Ref. [45]. This method performs joint model selection and parameter estimation using reversible jump Markov chain Monte Carlo (RJMCMC) [123], enabling trans-dimensional model selection over the number of nuclear spin, in addition to parameter estimation of the hyperfine couplings. By incorporating lattice parameters and hyperfine couplings calculated from *ab initio* density functional theory (DFT) calculations and enforcing experimental bandwidth constraints, our approach yields physically meaningful posteriors over nuclear spin configurations.

Given the ill-posed nature of recovering detailed nuclear spin information from sparse coherence data, our approach does not aim to fully resolve precise hyperfine couplings or nuclear spin locations. Instead, it provides posterior distributions over possible hyperfine coupling and spin location configurations, identifying defects with characteristics that are promising for specific applications. Determining the limits of detectability of nuclear spins is critical when working with sparse, noisy coherence data, as it allows us to distinguish physically meaningful spin features from artifacts caused by measurement noise and limited sampling. By explicitly accounting for these detection limits, we ensure that inferences remain robust and physically meaningful while avoiding overinterpreting weak or ambiguous signals. Defects showing favorable hyperfine coupling or nuclear spin distributions can then be selected for further experimental characterization using more time-consuming techniques, while avoiding the allocation of both experimental and computational resources to defects unlikely to meet application-specific requirements.

After providing background in Sec. 4.2 on the NV center and dynamical decoupling experiments we study here, we begin in Sec. 4.3 by analyzing the fundamental limits of hyperfine coupling detectability in sparse, noisy regimes, and we use these limits to guide experimental design. In Sec. 4.4, we then apply a trans-dimensional Bayesian model selection and parameter estimation method that reconstructs the nuclear spin environment by

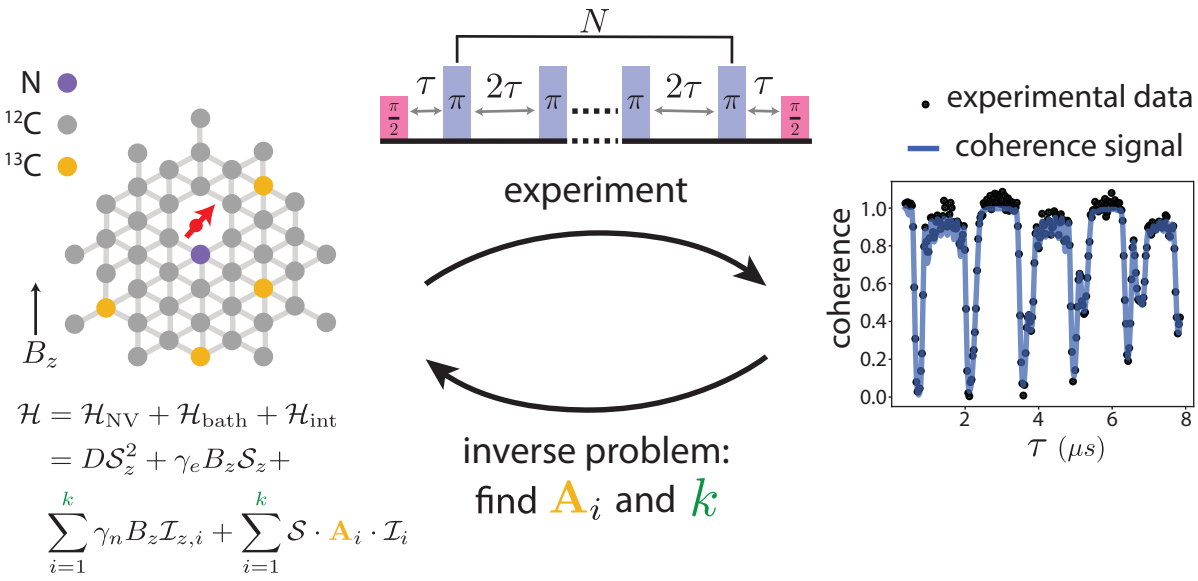

Figure 4.1: **Schematic of inverse problem**. (Left) Representation of a single NV center in diamond with an external magnetic field (B_Z) applied along the axis of the NV center, and there are k isotopic ^{13}C randomly distributed at lattice sites, interacting with the central electronic spin (red) via the hyperfine interaction ($\mathbf{A}$) (see text for the definition of Hamiltonian terms). (Center) Experimental probing by applying N dynamical decoupling π pulses with varying inter-pulse spacings (τ) to obtain (Right) a coherence signal. The inverse problem of interest is recovering the number of nuclear spins (k) and their hyperfine couplings ($\mathbf{A}$) from a sparse, noisy experimental coherence signal.

efficiently sampling over nuclear spin configurations with different numbers of nuclear spins. We apply a hybrid MCMC method in two steps: first, in Sec. 4.4.1, we demonstrate how we can use simulations of nuclear spin baths, within the fundamental limits on detectability, and we show how the hybrid MCMC method can guide experimental design for fast, scalable detection of nuclear spin bath properties; next, in Sec. 4.4.2, we apply the hybrid MCMC method to dynamical decoupling data from ten single NV centers in diamond with a natural abundance of ^{13}C nuclear spins. We conclude in Sec. 4.5 with a discussion on the capabilities, limitations, and possible extensions of the set of tools we provide. Altogether, our results provide a set of tools to design, analyze, and interpret dynamical decoupling experiments that can be performed rapidly for scalable characterization of nuclear spin environments of spin-defects in semiconductors.

4.2 Background

4.2.1 Hamiltonian description of nuclear spin bath

We consider a central electronic spin, the nitrogen-vacancy (NV) center in diamond, in an external magnetic field B_z interacting with a surrounding bath of nuclear spins, as shown in the schematic in Fig. 4.1. The dynamics of the system are governed by the Hamiltonian (in frequency units of $\mathcal{H}/\hbar$):

$$\mathcal{H} = \mathcal{H}_{\text{NV}} + \mathcal{H}_{\text{bath}} + \mathcal{H}_{\text{int}}, \tag{4.1}$$

where $\mathcal{H}_{\text{NV}}$ describes the electronic spin of the NV center, $\mathcal{H}_{\text{bath}}$ accounts for the nuclear spin bath, and $\mathcal{H}_{\text{int}}$ captures the hyperfine interaction between the NV center and the nuclear spins.

The NV center Hamiltonian is given by:

$$\mathcal{H}_{\text{NV}} = D\mathcal{S}_z^2 + \gamma_e B_z \mathcal{S}_z, \tag{4.2}$$

where D is the axial component of the zero-field splitting ($D = 2.87$ GHz for the NV center in diamond), and γ_e is the gyromagnetic ratio of the NV center. The nuclear spin bath is modeled as a collection of non-interacting spins:

$$\mathcal{H}_{\text{bath}} = \sum_{i=1}^{k} \gamma_n B_z \mathcal{I}_{z,i}, \tag{4.3}$$

where γ_n is the nuclear gyromagnetic ratio. Given the time and energy scales of the sparse, noisy experiments, we neglect interactions among the nuclear spins, since they are relatively weak, on the order of γ_n^2, compared to electron-nuclear spin interactions. The hyperfine

interaction is given by:

$$\mathcal{H}_{\text{int}} = \sum_{i=1}^{k} \mathcal{S} \cdot \mathbf{A}_i \cdot \mathcal{I}_i = \mathcal{S}_z(\sum_{i=1}^{k} A_{\parallel,i}\mathcal{I}_{z,i} + A_{\perp,i}\mathcal{I}_{x,i}) \tag{4.4}$$

where $\mathcal{A}$ represents the hyperfine tensor for each nuclear spin, and since the electronic level splitting is the dominant energy scale, we apply the secular approximation to describe the hyperfine interaction in terms of a parallel ($A_\parallel$) and a perpendicular ($A_\perp$) component with respect to the externally applied magnetic field B_z. We do not treat the central nitrogen explicitly in $\mathcal{H}_{\text{int}}$ because it is a fixed, well-understood degree of freedom rather than part of the unknown spin bath [124, 125].

4.2.2 Dynamical decoupling experiments

Dynamical decoupling (DD) sequences, such as the Carr-Purcell-Meiboom-Gill (CPMG) protocol, provide a powerful tool for probing the local spin environment by selectively filtering environmental noise at specific frequencies. These sequences are widely employed in experiments with nitrogen-vacancy (NV) centers in diamond and other solid-state spin qubits to extract information about nuclear spin dynamics and spin-spin interactions [126, 127].

In a CPMG-N experiment, a series of N equally spaced π-pulses with inter-pulse spacing 2τ are applied to the electronic spin (with τ time separating the initial and final π- and $\pi/2$-pulses), effectively modulating its interaction with the surrounding nuclear spin bath. The coherence of the electronic spin, measured as a function of the inter-pulse delay time τ, carries signatures of the nuclear spin bath, including the spectral density of its fluctuations and the spatial configuration of nuclear spins. The modulation function induced by the DD sequence acts as a spectral filter, enhancing the sensitivity to specific nuclear spin precession frequencies that match harmonics of the filter function peaks [128].

A widely used theoretical approach for modeling electronic spin coherence under DD se-

quences is the cluster-correlation expansion (CCE), which systematically accounts for many-body interactions within the nuclear spin bath [129, 130]. The CCE method provides a numerically controlled approximation to the coherence function by truncating the expansion at a given cluster order, where CCE-n includes interactions up to n-body nuclear spin correlations. For dilute nuclear spin baths where nuclear-nuclear interactions are weak, the lowest-order term, CCE-1, provides a quantitatively accurate description, capturing the dominant contribution to decoherence from independent nuclear spins [34, 36].In the specific case of NV centers in diamond, the combined impact of the sparsity of the ^{13}C nuclear spin bath, the relatively weak interaction strengths among ^{13}C, and the relatively short DD sequences studied results in a decoherence process that is dominated by independent contributions from single ^{13}C nuclei via the hyperfine interaction. As a result, CCE-1 is a sufficiently accurate description of the decoherence process for the ^{13}C bath of NV centers in diamond for the problem addressed in our work.

In this work, we focus on a regime where nuclear-nuclear interactions can be neglected, allowing for a closed-form analytical expression for the electronic spin coherence. This analytical formula, derived from the semiclassical treatment of the spin bath, is formally equivalent to the CCE-1 approximation and provides an exact description of the coherence decay under CPMG sequences in the absence of nuclear spin interactions for NV centers in diamond [34]. While CCE-1 is computationally efficient and yields accurate results for coherence calculations for NV centers in diamond, it is not universally applicable to all spin-defect systems. For example, for defects in SiC, coherence functions are well-converged only at the CCE-2 level of theory [131]. The level of CCE theory required for accurate coherence simulations depends both on the specific spin-defect system under study and on the experimental protocol being modeled. The Bayesian framework we employ, however, is general and can incorporate higher-order CCE calculations as required. As such, this framework is applicable to a broader class of spin-defect systems beyond the examples on NV centers we present here.

The coherence $\mathcal{L}$ for a known configuration of k nuclear bath spins $\{s_i\}_{i=1}^{k}$ for a CPMG-N experiment with inter-pulse spacing τ performed on an NV center with an external magnetic field B_z is given by:

$$\mathcal{L}(\{s_i\}_{i=1}^{k}, \tau, N, B_z) = \frac{1}{2}(1 + \prod_{i=1}^{k} M(A_{\parallel_i}, A_{\perp_i}, \tau, N, B_z))e^{-\frac{\tau}{\lambda}} \tag{4.5}$$

$$M(A_{\parallel,i}, A_{\perp,i}, \tau, N, B_Z) = (1 - m_{i,x}^2 \frac{(1-\cos\alpha_i)(1-\cos\beta)}{1+\cos\alpha_i\cos\beta - m_{i,z}\sin\alpha_i\sin\beta} \sin\frac{N\phi_i}{2}^2) \tag{4.6}$$

for $\cos\phi_i = \cos\alpha_i\cos\beta - m_{i,z}\sin\alpha_i\sin\beta$ with $m_{i,z} = \frac{A_{\parallel,i}+\omega_L}{\tilde{\omega}_i}$ and $m_{i,x} = \frac{A_{\perp,i}}{\tilde{\omega}_i}$ for $\tilde{\omega}_i = \sqrt{(A_{\parallel,i}+\omega_L)^2 + A_{\perp,i}^2}$, $\alpha_i = \tilde{\omega}_i\tau$, $\beta = \omega_L\tau$, and $\omega_L = \gamma_n \times B_z$ [34], and λ is a decay factor to account for dephasing of the electronic spin due to interactions that are not explicitly modeled here, such as interactions with lattice impurities other than ^{13}C that may be present in experimental samples or interactions with the distant spin bath, which can cause an overall decay in the coherence signal [80]. We find that including the decay factor λ to implicitly account for dephasing effects is crucial to ensure that our forward model (Eq. 4.5). In the next section, we provide more details on including the relative decoherence parameter λ for recovering nuclear spin baths from experimental data.

4.2.3 Relative decoherence parameter

In spin-based sensing experiments, decoherence arising from interactions with the surrounding environment becomes increasingly significant at longer evolution times. These effects are particularly relevant in sequences with a large number of dynamical decoupling pulses, such as Carr-Purcell (CP) sequences with 16 or more pulses, where the total experiment duration is extended and the NV center has more opportunity to interact with unmodeled or weakly coupled environmental spins, lattice phonons, or other sources of noise.

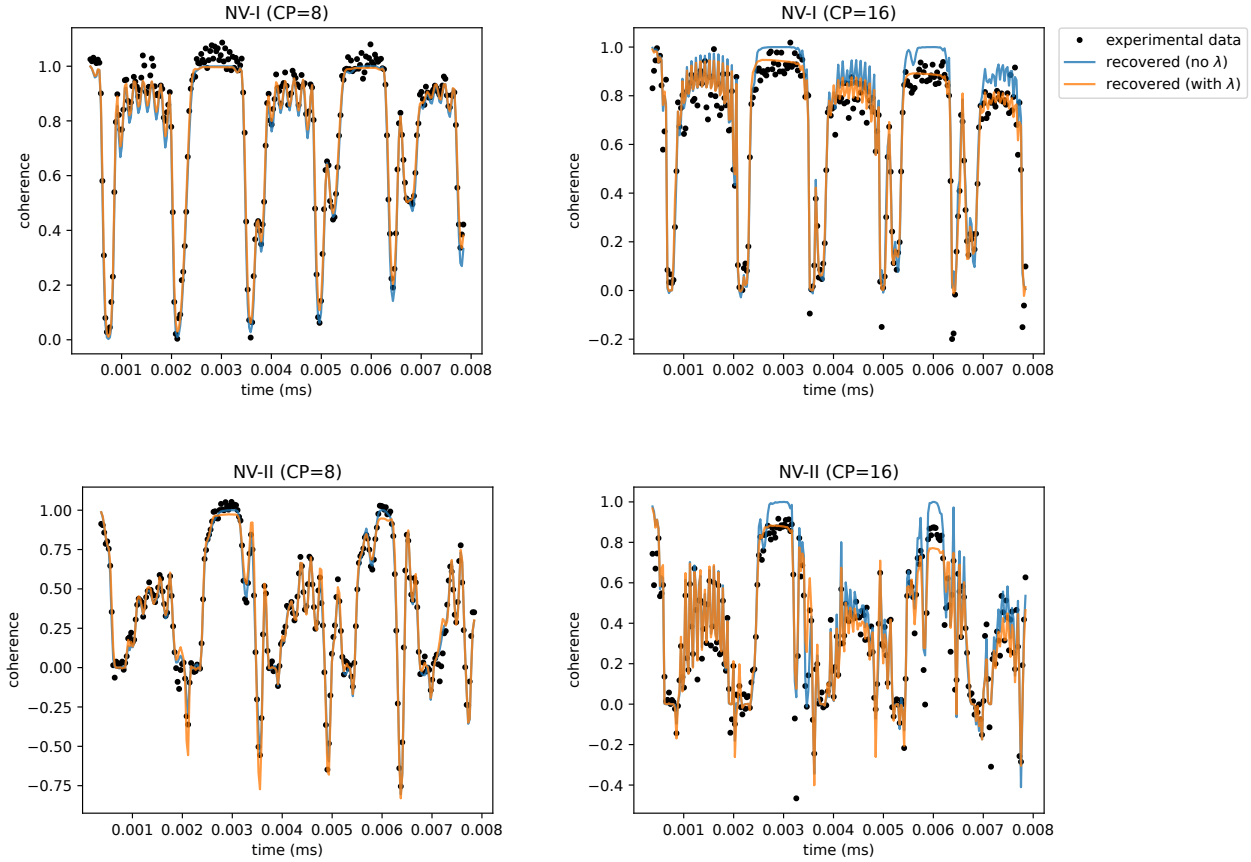

Figure 4.2: **Comparison of coherence signal reconstruction with and without the relative decoherence parameter** λ. The top row shows experimental data for NV-I with 8 and 16 pulses (CP=8, CP=16); the bottom row shows corresponding data for NV-II. In each panel, black dots denote experimental coherence data, while the blue and orange curves show reconstructed signals derived from the modal spin configuration of the posterior distribution. The blue curves correspond to reconstructions where the relative decoherence parameter λ is held fixed at 1, while the orange curves result from inference where λ is treated as a free parameter and jointly inferred along with the nuclear spin bath.

As shown in Fig. 4.2, neglecting decoherence from sources other than the explicitly-modeled magnetic noise arising from the nuclear spin bath leads to systematic overestimation of the coherence signal at larger values of the pulse interval τ, especially for CP=16. Reconstructions that do not account for these effects—such as those performed with a fixed λ parameter—fail to capture the experimentally observed decay. In contrast, incorporating decoherence phenomenologically by inferring an effective regularization parameter λ results

in reconstructed signals that more accurately reflect the measured coherence, suggesting that this approach can implicitly account for unresolved environmental contributions.

This highlights the importance of including or compensating for decoherence effects when interpreting long-time experimental data, particularly in cases where the model does not explicitly include all sources of environmental coupling.

4.3 Detection limits imposed by noisy, sparse data

Because we aim to recover hyperfine couplings from experimental coherence signals that are relatively short and sparsely sampled in terms of interpulse spacings τ, the resulting data are inherently noisy due to the photon detection process. This shot noise, combined with the limited frequency resolution dictated by the experimental parameters, constrains the extraction of meaningful hyperfine couplings. To systematically determine the regime where extracted hyperfine couplings remain physically significant rather than artifacts of shot noise, we analyze next the interplay between sensing limits, nuclear spin bath properties, and experimental averaging. These factors collectively determine how the chosen experimental parameters influence the physical significance of the extracted hyperfine couplings.

4.3.1 Detection limits imposed by sparsity

The fundamental sensing limits in dynamical decoupling experiments are governed by the Nyquist-Shannon sampling theorem and the maximum interpulse spacing ($\tau_{\max}$). The highest resolvable frequency is

$$f_{\max} = \frac{\text{number of sampled interpulse spacings}}{4 \times \tau_{\max}}, \tag{4.7}$$

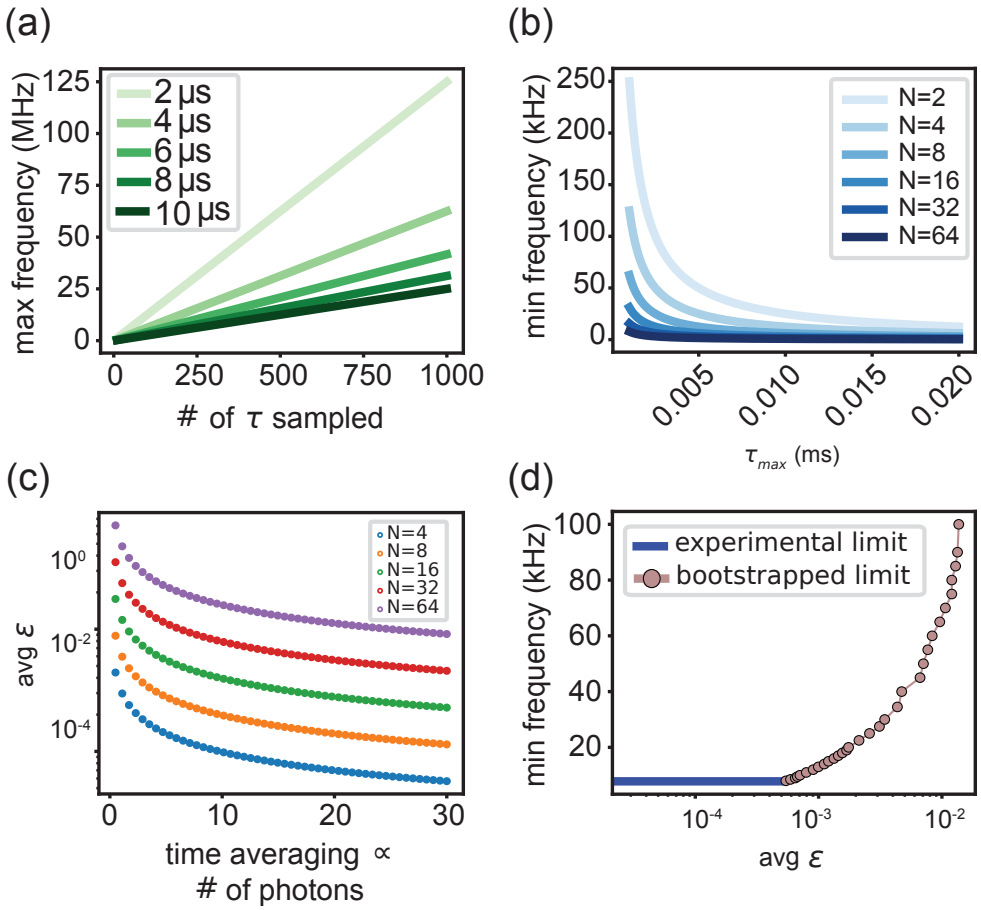

Figure 4.3: **Hyperfine coupling frequency detection limits from frequency sampling limits and simulations (a)** Maximum detectable frequencies as a function of the number of interpulse spacings (τ) sampled, shown for various maximum interpulse spacings ($\tau_{\max}$). **(b)** Minimum detectable frequencies as a function of $\tau_{\max}$ for different Carr-Purcell (CP) pulse numbers (N). **(c)** Schematic of average shot noise (ε) per data point for varying numbers of CP-pulses (N) with an arbitrary initial signal-to-noise ratio. **(d)** Minimum detectable hyperfine frequency, calculated as a function of the average ε per data point, for joint fits of an N=8 and N=16 CP pulse experiment with $\tau_{\max} = 8\mu$s, $B_{\text{ext}} = 311$G, and 250 τ sampled at the natural concentration (1.1%) of ^{13}C. The bootstrapped limit is shown for 0.95 confidence of detectability. The minimum frequency limit imposed by the experimental settings (7.8125 kHz).

while the lowest detectable frequency is limited by the duration of the experiment, given by

$$f_{\min} = \frac{1}{N \times 2 \times \tau_{\max}} \tag{4.8}$$

where N is the number of dynamical decoupling pulses and the total length of an experiment is $N \times 2 \times \tau$ for a given τ. These limits define the range of frequencies that can be extracted for a fixed magnetic field, B_z, given the window and number of interpulse spacings sampled (Fig. 4.3a & b).

For nuclear spin environments, the Zeeman energy splitting induced by B_{ext} determines the nuclear spins' Larmor precession frequency, thereby affecting the dynamical evolution of the system. Specifically, higher values of B_z result in faster Larmor precession, effectively compressing the characteristic timescales of the dynamics, while lower values extend these timescales. This relationship establishes a redundancy between τ_{max} and B_z, expressed as $\tau_{max} \propto \frac{1}{\gamma B_{ext}}$, where γ denotes the nuclear spins' gyromagnetic ratio.

In this work, we explore the impact of dynamical timescales under fixed B_z while varying τ_{max}. This choice reflects the experimental ease of modifying interpulse spacing, which only requires programmatic adjustments to the experimental control sequence, compared to altering B_z, a process that involves physical reconfiguration and fine-tuning of multiple hardware components. While increasing B_z shortens the timescales, it may also alter effective noise levels and potentially affect resolution limits due to environmental interactions. A detailed investigation of these noise and resolution effects is beyond the scope of this study but represents a direction for future work.

4.3.2 Detection limits imposed by shot noise

Each spin defect acts as a single-photon emitter, limiting the number of spin-sensitive photons that can be detected in a single measurement before the spin state decoheres or reinitializes. As a result, experimental coherence signals are acquired by counting single photons and averaging over repeated measurements, yielding data distributed according to Poisson statistics. In high-throughput settings, the requirement of rapid acquisition imposes a fundamental trade-off between measurement time and data quality. Here we investigate how

noise impacts our ability to recover hyperfine couplings from coherence data acquired in the sparse, noisy regime. In particular, we examine the extent to which weakly coupled nuclear spins can be reliably detected. This analysis establishes fundamental limits on the minimum detectable hyperfine coupling strength given a fixed amount of shot noise and informs experimental design choices for scalable characterization.

To systematically determine the crossover threshold at which hyperfine couplings of specific frequencies become undetectable, compared to shot noise for a fixed set of ^{13}C isotopes surrounding NV centers in experiments, we fixed key experimental parameters. These include the applied external magnetic field (B_{ext}), the number of dynamical decoupling pulses (N), the maximum interpulse spacing measured ($\tau_{\max}$), the number of interpulse spacings measured, the amount of shot noise ε, and the isotopic concentration of spins c. We then sampled hyperfine coupling constants for atomic positions within 30 Å of the NV center using *ab initio* density functional theory calculations [97] and using the dipole approximation for atoms between 30 Å and 40 Å. A 40 Å cutoff leads to well-converged results for the noise regimes and pulse sequences studied here; nuclear spins beyond this distance are so weakly coupled that their contributions are below the detection threshold imposed by shot noise.

Using a specified concentration (c), we generated 10,000 random configurations of lattice sites and, for each configuration, calculated the L_2 error E_T between the coherence signal $L(\tau)$ obtained with the full bath and the signal $L_T(\tau)$ resulting from omitting spins with hyperfine couplings below a specified frequency threshold T. For a fixed nuclear spin bath containing n spins $\mathbf{s} = \{s_i\}_{i=1}^{n}$ and a thresholded spin bath $\tilde{\mathbf{s}}_T = \{s_i | A_{\parallel,i} \text{ or } A_{\perp,i} \geq T\}$ with spins with one or both hyperfine couplings components below a threshold T removed we have:

$$L(\tau) \coloneqq \mathcal{L}(\mathbf{s}, \tau, N, B_z) \tag{4.9}$$

$$L_T(\tau) \coloneqq \mathcal{L}(\tilde{\mathbf{s}}_T, \tau, N, B_z) \tag{4.10}$$

$$E_T \coloneqq ||L(\tau) - L_T(\tau)||_2. \tag{4.11}$$

To assess the uncertainty in the errors E_T for each threshold value T, we employed a bootstrapping procedure with 10,000 iterations, resampling the calculated errors to obtain the 0.95 confidence level. Specifically, the 0.95 confidence value corresponds to the error level such that 95% of spin configurations yield smaller errors than the expected error from the shot noise. The averaged bootstrapped 95% confidence error was then normalized by the number of time points in the coherence signal (i.e., the number of interpulse spacings). This normalized value represents the experimental noise level for a given hyperfine threshold such that the perturbation induced by nuclear spins with small hyperfine couplings exceeds the noise with 95% confidence (Fig. 4.3c).

4.3.3 Determining limits on detectability from sparse, noisy experimental data

By combining fundamental limits on hyperfine frequency detectability with calculations of the crossover regime between intrinsic and external noise, we determined how the minimum detectable hyperfine frequency decreases with experimental averaging. This approach allows us to obtain the minimum detectable hyperfine frequency for a particular experiment with a desired amount of averaging. Notably, for the experimental parameters used in this study ($\tau_{\max} = 8\mu$s, $B_{\text{ext}} = 311$G, and 250 τ sampled for CP-8 and CP-16 pulses), we established a lower detectability limit of 7.8125 kHz, below which hyperfine couplings cannot be resolved regardless of experimental averaging (Fig. 4.3d).

Our combined approach achieves two objectives. First, given a specific experimental dataset, including the amount of time averaging, we can verify that the extracted hyperfine couplings are meaningful rather than extraneous fitting artifacts. Second, for targeted applications of spin defects in semiconductors, we can design optimal dynamical decoupling

experiments. By tailoring the number of time points sampled, the number of pulses measured, the applied magnetic field, the interpulse spacing, and the required time averaging, we can ensure efficient detection of hyperfine couplings within a desired range. This approach paves the way for high-throughput characterization strategies, enabling precise and resource-efficient evaluation of spin-defect properties.

4.4 Application of a hybrid MCMC algorithm for experimental design and nuclear spin bath recovery

Having established fundamental constraints on the detectability of hyperfine couplings that hold independently of the specific method used for parameter recovery, we now turn to the application of a concrete inference strategy designed to operate effectively in sparse and noisy experimental regimes. To this end, we employ a hybrid Markov chain Monte Carlo (MCMC) algorithm that integrates several complementary techniques: reverse jump MCMC (RJMCMC), parallel tempering (PT), and continuous and discrete versions of random walk Metropolis-Hastings (RWMH). Each of these components addresses a distinct challenge associated with the high-dimensional and trans-dimensional nature of the inverse problem under consideration. Our approach integrates *ab initio* information about the spin-defect system, including the crystal lattice geometry and hyperfine couplings computed from first principles [97]. Using this structural and electronic input as a prior, we define a probabilistic model over configurations of nuclear spins on lattice sites and their associated decoherence effects.

To obtain the posterior distribution over spin configurations and parameters, we employ a hybridized Markov chain Monte Carlo (MCMC) approach introduced in Ref. [45]. The hybridized algorithm is built around a discrete random walk Metropolis-Hastings (RWMH) algorithm, which samples lattice sites via a discrete random walk. Reverse jump MCMC (RJMCMC) addresses the trans-dimensional nature of the inference problem by proposing birth and death moves at lattice sites to vary the number of nuclear spin walkers [123].

Within a fixed-dimensional model, parallel tempering (PT) is employed to overcome the non-linear and non-convex structure of the likelihood landscape by running multiple chains of discrete RWMH at different effective temperatures, with exchanges between chains enhancing convergence and avoiding local minima. Finally, continuous RWMH updates are used to obtain posterior distributions for the real-valued relative decoherence scaling factor λ, which captures spin decoherence contributions beyond the dominant hyperfine couplings. This hybridized MCMC strategy provides a general and robust framework for extracting spin bath parameters across a wide range of experimental conditions and materials systems.

Using hyperfine couplings for all lattice sites near the central electronic spin, which were calculated using density functional theory (DFT) in Ref. [97], we use the hyperfine frequency detection limits to determine the number and location of lattice sites that correspond to the hyperfine coupling frequencies within the detection limits. We use a cut-off point of $s_{\max} = 50$ spins as the maximum number of spins considered for the NV center in diamond based on a dimensional analysis of the expected number of spins for the lattice given a natural abundance of ^{13} C within the cut-off radius of 40 Å. We randomly initialize a random number of walkers between 1 and $s_{\max}$ for 5 ensembles of walkers.

We then iterate through a hybrid MCMC algorithm that cycles through 50 steps of reverse jump Markov chain Monte Carlo (RJMCMC), 100 steps of parallel tempering (PT), and 25 steps of random walk Metropolis Hastings (RWMH) for 25,000 total steps. For pseudocode for each algorithm, see [45]. We consider the first 5,000 steps as burn-in, so the posterior distributions consist of the spin configurations and T_2 values corresponding to the last 20,000 steps, and spin configurations and T_2 values are combined across ensembles. Unless otherwise specified, we used $R_{\text{spin}} = 5$, $R_{T_2} = 0.05$, $\sigma^2 = 0.1$, 10 strands for the PT algorithm, and $\beta = 2^{-n}$ for strand $n \in [0, \ldots \text{num strands} - 1]$ for the PT algorithm. For more details on hyperparameters, see Sec. 3.4.3. While these hyperparameters are not necessarily optimal, they reliably result in convergence to stationary posterior distributions.

We demonstrate the utility of this approach in two settings. First, we demonstrate how this hybrid MCMC algorithm can be applied to guide the design of rapid, resource-efficient experiments targeting nuclear spins with hyperfine couplings in a specified range (Sec. 4.4.1). Second, we demonstrate how this hybrid MCMC approach enables high-throughput screening of spin-defect samples for desirable nuclear spin bath properties from dynamical decoupling data (Sec. 4.4.2).

4.4.1 Guiding experimental design

To determine which experiments are most effective for detecting nuclear spins with hyperfine couplings in a targeted range, we use simulations to systematically evaluate the performance of our hybrid MCMC recovery method across varied experimental conditions. This approach enables experimentalists to design and select protocols that maximize recovery success for their detection goals, even in the presence of constraints imposed by experimental setups, such as fixed magnetic fields, limited pulse numbers, or bounded total acquisition time.

For each simulation, we generate synthetic coherence signals based on model nuclear spin baths and simulate measurements under different experimental configurations. These configurations include variations in the number of Carr-Purcell (CP) pulses, the maximum interpulse spacing (τ), and the resolution of τ-sampling. By averaging recovery statistics across multiple simulated spin baths, each with a different number of simulated nuclear spins (between 5 and 20), we assess how well each experimental setting enables accurate inference of spin number and hyperfine parameters. Key performance metrics include the recovery rate of individual spins (i.e., how frequently each true spin appears in the posterior), the discrepancy between the inferred spin number (the modal number of spins in the posterior distribution) and the number of simulated spins, and the false positive rate of the nuclear spins in the modal spin configuration of the posterior distribution.

Increasing the maximum interpulse spacing ($\tau_{\max}$) allows observation of a greater number

of Larmor periods, and we observed that extended $\tau_{\max}$ increases recovery rates of hyperfine couplings for spins greater than 25 kHz (Fig. 4.4). To just resolve the number of spins in the system, however, we observe that discrepancy for the number of recovered spins compared to the number of simulated spins is stable and accurate (close to zero) for spins with magnitudes > 25 kHz, and we find that extending $\tau_{\max}$ results in a decreased discrepancy between the number of simulated and the number of recovered spins. This suggests that it is crucial to extend the length of the longest interpulse spacing to accurately resolve hyperfine couplings, especially for weakly coupled spins (< 25 kHz), but shorter experiments (on the order of $\tau_{\max} \approx 0.002$ ms for a $N = 16$ pulse CP experiment) may suffice for recovering just the number of spins with magnitudes > 25 kHz.

When varying the number of CP pulses, we observed that recovery rates stabilized after approximately 2^5 pulses for recovery rates of both hyperfine couplings and number of spins (Fig. 4.4). This plateau indicates that, after this point, increasing the number of pulses does not improve recovery rates unless the sampling resolution is increased to match the higher pulse count. This finding highlights the importance of optimizing the sampling resolution to fully exploit the information content provided by a higher pulse count.

Next, we explore the effect of varying the magnetic field strength on recovery performance. Lower magnetic fields (below 100 G) led to a significant drop in detection capability, with both hyperfine coupling recovery and the number of spins detected suffering considerably (Fig. 4.4). However, as the magnetic field strength increased, both the recovery rates for hyperfine couplings and the number of spins showed modest improvements, and the false positive rate decreased slightly. This suggests that higher magnetic fields improve the quality of the coherence signal and enhance the recovery performance, though the improvements are not dramatic beyond ~ 500 G for these sets of experimental settings.

Finally, we explore the interdependence of multiple experimental parameters on recovery rate by covarying mutliple experimental settings. Each CP pulse sequence defines a

specific filter function that interacts with the spectral noise, shaping how effectively different frequency components of the noise are suppressed [132]. As the number of pulses increases, the filter function typically broadens and shifts, enabling better suppression of low-frequency noise and leading to an extension of the system's coherence time. For a fixed maximum interpulse spacing, simulations show that recovery rates are comparable across hyperfine couplings of all magnitudes when comparing two sampling rates—250 and 500 time points—up to approximately 2^5 CP pulses (Fig. 4.5). Beyond this point, however, the recovery rate for the 250-time-point sampling plateaus, indicating that the additional information encoded by higher numbers of CP pulses cannot be accessed due to insufficient resolution. In contrast, at 500 time points, the recovery rate continues to improve even beyond 2^5 CP pulses, suggesting that the higher resolution allows for better extraction of information provided by the increased number of pulses. Additionally, increasing the number of CP pulses tends to extend the system's decoherence time by more effectively suppressing noise, which can further enhancing the recovery of hyperfine couplings (see main text, Sec.). Thus, it is important to optimize sampling resolution, the number of CP pulses, and the maximum interpulse spacing simultaneously, rather than treating each factor independently, to achieve the best performance in recovering hyperfine couplings.

Application of the hybrid MCMC method to simulated data demonstrates that the recovery fidelity depends on the interplay between experimental parameters. In particular, we find that hyperfine parameter recovery is correlated with the strength of the hyperfine coupling magnitude across all experimental regimes, where nuclear spins with stronger hyperfine couplings are recovered more accurately than more weakly coupled nuclear spins (see Fig. 4.4). In particular, we find that there is a saturation point beyond which increasing the length of the experiment (maximum τ) and increasing the number of dynamical decoupling pulses yields little additional benefit in terms of recovery, suggesting a regime of diminishing returns with increased experimental resources. For example, we observe a plateau in the

detection rate of hyperfine couplings in for hyperfine couplings between $25 - 100$ kHz for greater than 2^5 CP pulses, indicating that the recovery of hyperfine couplings is limited not by the number of pulses but by the length of longest interpulse space sampled or the number of time points sampled (see Fig. 4.4). We emphasize that this regime of diminishing returns is a result of the interdependence of multiple experimental parameters and can thus help experimentalists avoid inefficient dynamical decoupling sampling strategies.

Overall, these simulations enable targeted experimental planning: given a specific detection goal (e.g., recovering hyperfine couplings in the 10–50 kHz range) and hardware constraints, one can identify experimental protocols that will offer the highest likelihood of successful recovery using this hybrid MCMC method. Furthermore, experimental constraints such as fixed magnetic fields or pulse limits can be explicitly incorporated into the simulations, allowing users to determine what is achievable within their setup and to identify tradeoffs (e.g., more averaging vs. more resolution) before performing experiments. This capability is especially valuable in resource-constrained or high-throughput settings.

4.4.2 Application to experimental data

We further demonstrate the utility of this hybrid MCMC approach for high-throughput charaterization of spin-defects in semiconductors by applying it to coherence data obtained from ten distinct NV centers in diamond. The experimental parameters were selected to include two coherence signals—CP8 and CP16—each comprising 250 time points, with interpulse spacings up to 0.008 ms under a magnetic field of 311 G. Experimental noise was quantified using standard error propagation formulas, incorporating contributions from Poisson noise and photon counts during the transformation of raw photon count data into coherence data.

To reduce the effects of fluctuations among the data and calculate uncertainties, the raw photon counts from the bright (p_z) and dark (m_z) channels are normalized: $I_{\text{norm}} = \frac{p_z - m_z}{p_z + m_z}$.

NV	Avg. Exp. Uncertainty (CP=8)	Avg. Exp. Uncertainty (CP=16)	Lower Hyperfine Threshold from bootstrapping (kHz)
I	0.00266	0.00627	1.15
II	0.00211	0.00470	1.12
III	0.00279	0.00572	1.14
IV	0.00301	0.00620	1.15
V	0.00439	0.00850	1.22
VI	0.00369	0.00819	1.20
VII	0.00154	0.00128	1.04
VIII	0.00495	0.00997	1.25
IX	0.00290	0.00657	1.16
X	0.00257	0.00531	1.13

Table 4.1: **Average uncertainty per data point**. Columns corresponding to experimental uncertainties are calculated from photon counts from raw experimental data, and the lower hyperfine threshold is obtained numerically based on the averages of the experimental uncertainties.

The coherence is computed from the normalized bright and dark channels, and we propagate uncertainties based on standard error propagation formula for each manipulation of the raw data into coherence signals. We take the uncertainty of each channel to be the square root of the number of photons detected. As an increasing number of photons are detected, the uncertainty per data point decreases (Fig. 3.6), and as a result the minimum detectable hyperfine frequency decreases as an increased number of photons are collected, as shown in Fig. 4.8 for different values of statistical confidence.

Although the minimum hyperfine detection limit for the experimental data in the main text is limited by the total duration of the experiment rather than by shot noise, we report the average shot noise ε per data point for the ten samples of NV centers we applied this hybrid algorithm recovery approach to in the main text calculated from the photon counts from the raw experimental data in Table 4.1.

For each NV center, we obtained posterior distributions of spin configurations that have a high likelihood of producing the observed coherence signal data. Coherence signals generated from the modal spin configuration in these posterior distributions as well as the corresponding

lattice positions of the modal spin configuration are shown in Fig. 4.6. Posterior distributions of the hyperfine couplings and numbers of spins in each sample are shown in Fig. 4.7.

Based on the sensitivity analysis detailed in the previous section, we find that hyperfine couplings greater than 25 kHz can be reliably recovered. Lower-frequency couplings were less well-constrained, consistent with the limits of the experimental noise and hyperfine thresholds. For certain NV centers, the algorithm identified spin configurations that effectively recreate the experimental data. Notably, NV centers II, III, IV, V, IX, and X exhibit key features—such as a small number of strongly coupled spins—that make them particularly relevant as auxiliary qubits for quantum information applications, including local quantum memory units (Fig. 4.7).

In contrast, NV centers I, XI, XII, and XIII lack strongly coupled spins and are therefore less promising for quantum information applications (Fig. 4.7). These NVs could potentially be deprioritized in future experimental studies, allowing resources to be reallocated toward further characterization of more promising candidates.

Certain samples, such as NV-IX, presented systematic underfitting in their coherence signal reconstruction. This suggests the presence of strongly coupled spins not captured by the DFT-calculated hyperfine couplings. Possible explanations include the presence of substitutional nitrogen defects (P1 centers), interactions with nearby NV centers, or other lattice distortions that deviate from the idealized DFT model.

4.5 Discussion

Bayesian frameworks offers distinct advantages for characterizing quantum systems compared to machine learning (ML)-based approaches in the sparse data regime. While ML methods excel at pattern recognition and high-dimensional data processing, they often lack interpretability and explicit uncertainty quantification [133]. In contrast, Bayesian methods naturally provide posterior distributions that quantify uncertainty, enabling a probabilistic

understanding of the recovered spin environment. This interpretability is crucial for experimental design and decision-making, where understanding the confidence in recovered parameters can directly impact subsequent measurements and analyses.

The Bayesian framework introduced in this work offers a practical strategy for optimizing experimental design. By providing probabilistic estimates of spin configurations and their uncertainties, our approach enables experimentalists to identify measurement regimes that maximize information gain while minimizing acquisition time. In addition, the framework proposed in our work lays the groundwork for implementing digital twin approaches to quantum spin systems, where simulations continuously inform and are informed by experimental data. Digital twins have been successfully applied in areas such as photovoltaic characterization [134] and quantum dot synthesis [135], and similar strategies could enhance defect-based quantum technologies. Specifically, the method presented here could enable real-time data assimilation between a physical spin defect and a virtual model of its spin environment. The digital twin could incorporate additional spin dynamics using either high-fidelity ab initio simulations [136] or efficient surrogate models [109], supporting a closed-loop system where experiments refine the model, and the model, in turn, guides subsequent measurements.

Despite its strengths, the current implementation has limitations tied to the underlying physical models. The forward model relies on the cluster correlation expansion (CCE) to compute coherence signals from proposed nuclear spin baths, which accurately captures short-range interactions but becomes less reliable for long-range couplings [137]. While higher-order CCE can, in principle, capture more complex many-body correlations and long-range effects, the number of clusters grows rapidly with CCE order n and bath size k. Specifically, CCE-1 requires $\mathrm{O}(k)$ single-spin contributions, CCE-2 involves $\mathrm{O}(k^2)$ pairs, CCE-3 scales as $\mathrm{O}(k^3)$ triples, and so on, with the total number of n-spin clusters given by $\binom{k}{n}$. Although each fixed order scales polynomially with system size k, the growth with both n and k quickly becomes computationally prohibitive, which limits the practicality of going

beyond CCE-2 or CCE-3 in large-scale inference problems.

Additionally, the likelihood model assumes fixed hyperfine couplings derived from a single density functional theory (DFT) calculation on an idealized lattice, neglecting uncertainties due to lattice imperfections, unknown strain, or deviations in defect isolation that may arise when processing samples at scale. This idealized treatment excludes the influence of nearby paramagnetic impurities, such as P1 centers, or other types of lattice defects, which can significantly alter the coherence dynamics [37, 138]. However, if the presence of, e.g. P1 centers is known from experiments, their perturbation of the NV center's spin density and the resulting modification of the hyperfine couplings via the Fermi contact term could eventually be included in large scale DFT calculations. In such cases, one could first experimentally confirm the presence of NV-P1 pairs with double electron-electron resonance (DEER) experiments, which can isolate interactions between NV centers with electrons from those with nuclei [138]. Then, one could compute using DFT how the P1 electron perturbs the NV electronic density and alters the hyperfine couplings via the Fermi contact interaction. Finally, one could apply the hybrid MCMC framework to reconstruct the ^{13}C spin bath under these modified conditions.

In summary, we present a set of computational tools for scalable spin-defect characterization, with applications ranging from experimental data interpretation to the design of targeted experimental protocols. While we provide simulations and applications of NV centers in diamond, the approach we present is broadly applicable to other defect systems, such as phosphorus donors in silicon and divacancies in silicon carbide, as well as to broader classes of Hamiltonian learning problems. By addressing the challenges posed by high-throughput defect fabrication and the constraints of noisy experimental regimes, our work provides a pathway toward efficient integration of spin defects into quantum technologies.

Several avenues for future work can address these limitations and extend the applicability of this approach. First, this method could be extended to other systems, such as divacancies

in SiC [69, 139] or phosphorus donors in silicon [140]. A more robust likelihood model incorporating uncertainty in DFT-derived hyperfine couplings and lattice distortions could further improve accuracy. We could also reframe the problem to find the shortest dynamical decoupling experiment given a desired uncertainty bound for a target regime of hyperfine frequencies. Finally, extending the model to include full Hamiltonian learning, accounting for nuclear spin interactions or ensembles of defects such as NV-NV [141] or NV-P1 systems [37], represents an exciting challenge for future studies. These advancements would enhance the framework's predictive power and utility across a wide range of quantum systems.

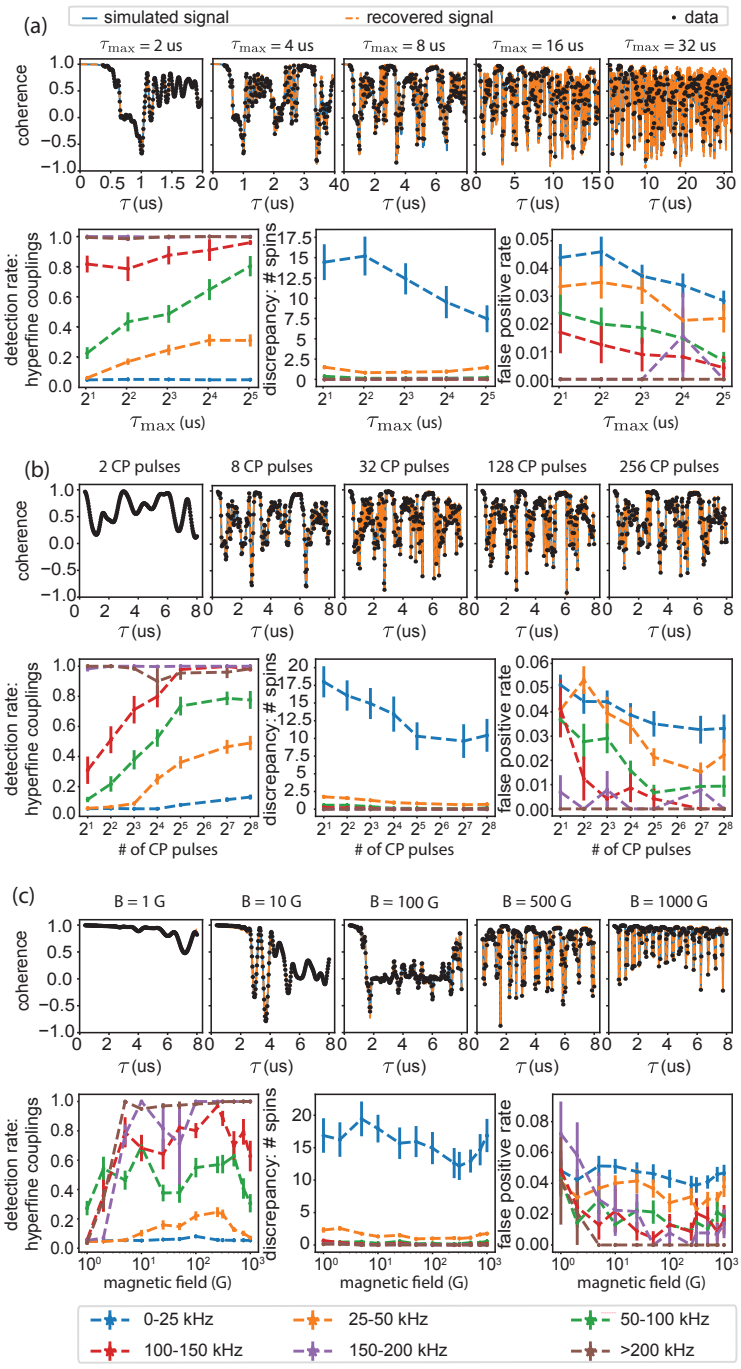

Figure 4.4: **Accuracy of hybrid MCMC algorithm for recovering hyperfine couplings from simulated data.** We applied our hybrid recovery method (see text) to simulations with 0.001 shot noise per data point, 250 τ sampled, a 0.008 ms longest τ, 16 CP pulses, and 311 G magnetic field, and varied each one of the parameters independently for (a) maximum interpulse length $\tau_{\max}$, (b) the number of CP pulses, and (c) magnetic field. For each of the parameters varied, we plot examples of simulated and recovered signals based on the data and detection rates, discrepancies in the number of recovered versus simulated spins, and false positive rates for the hybrid MCMC method averaged over random nuclear spin configurations consisting of 5 to 20 nuclear spins (for a total of 16 simulated nuclear spin configurations). Results are plotted by groupings of nuclear spins based on hyperfine magnitudes ($\sqrt{A_{\parallel}^2 + A_{\perp}^2}$) (see bottom colorbar).

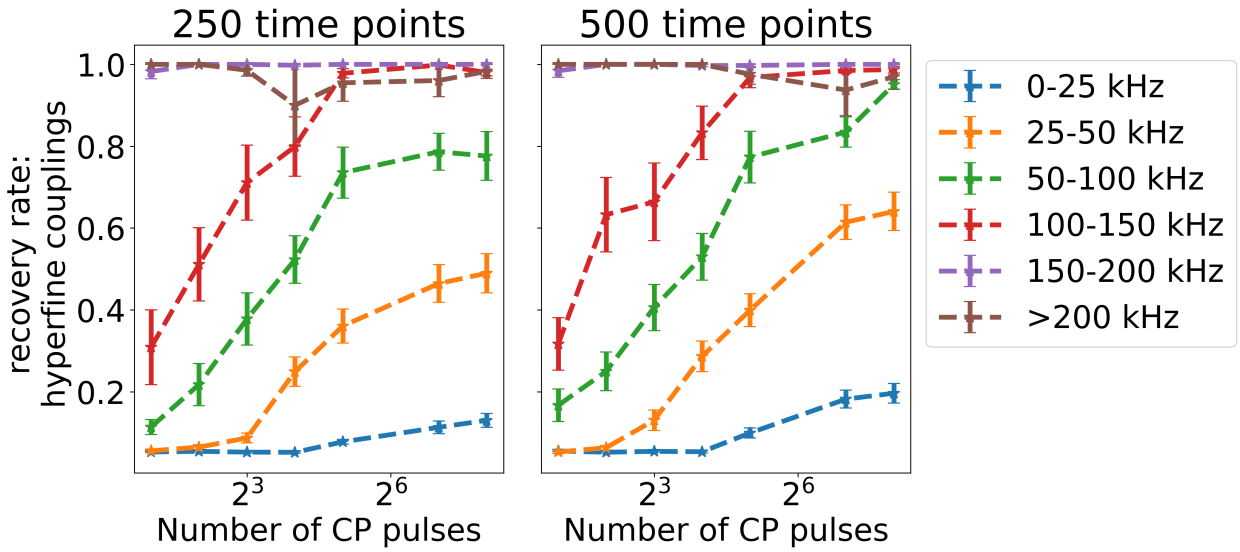

Figure 4.5: **Interdependence of recovery rates on experimental settings.** Recovery rate of hyperfine couplings as a function of the number of CP pulses for two different sampling resolutions: 250 time points (left) and 500 time points (right). Data are grouped by hyperfine coupling strength, ranging from 0–25 kHz to >200 kHz. Settings not specified are the same as Fig. 3 of the main text.

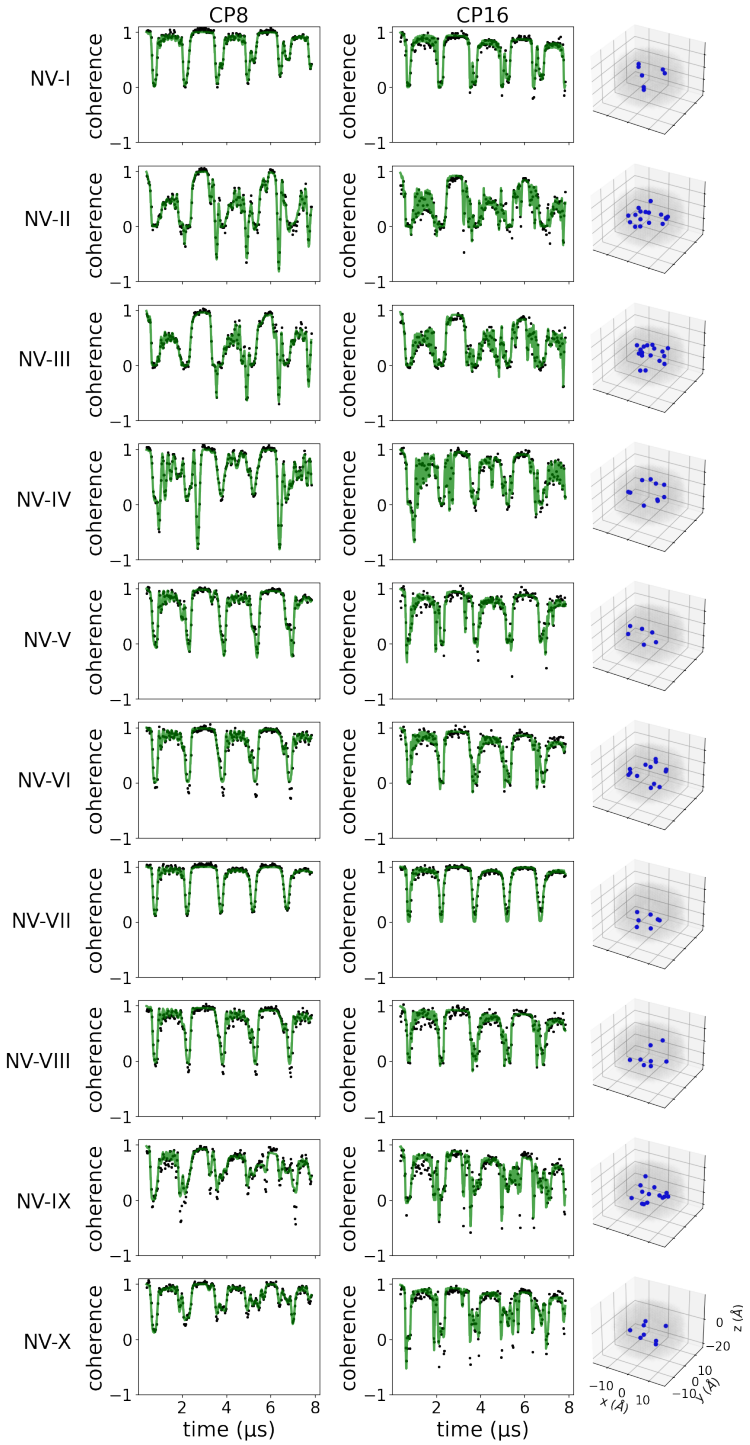

Figure 4.6: Experimental coherence signal data for 10 different nitrogen-vacancy (NV) centers under CP8 (left column) and CP16 (middle column) pulse sequences, overlaid with simulated coherence signals (green) generated from the modal spin configuration of the posterior distribution of nuclear spin configurations generated by our hybrid method to extract nuclear spin environments from experimental data. Each row corresponds to a distinct NV center. The right column shows the spatial distribution of nuclear spins in the modal configuration for each NV, plotted relative to the NV center (at the (0, 0, 0) position). Only lattice positions of nuclear spins with hyperfine magnitudes > 25 kHz are plotted. The hybrid method jointly analyzes CP8 and CP16 datasets to infer spin environments that best match the experimental data.

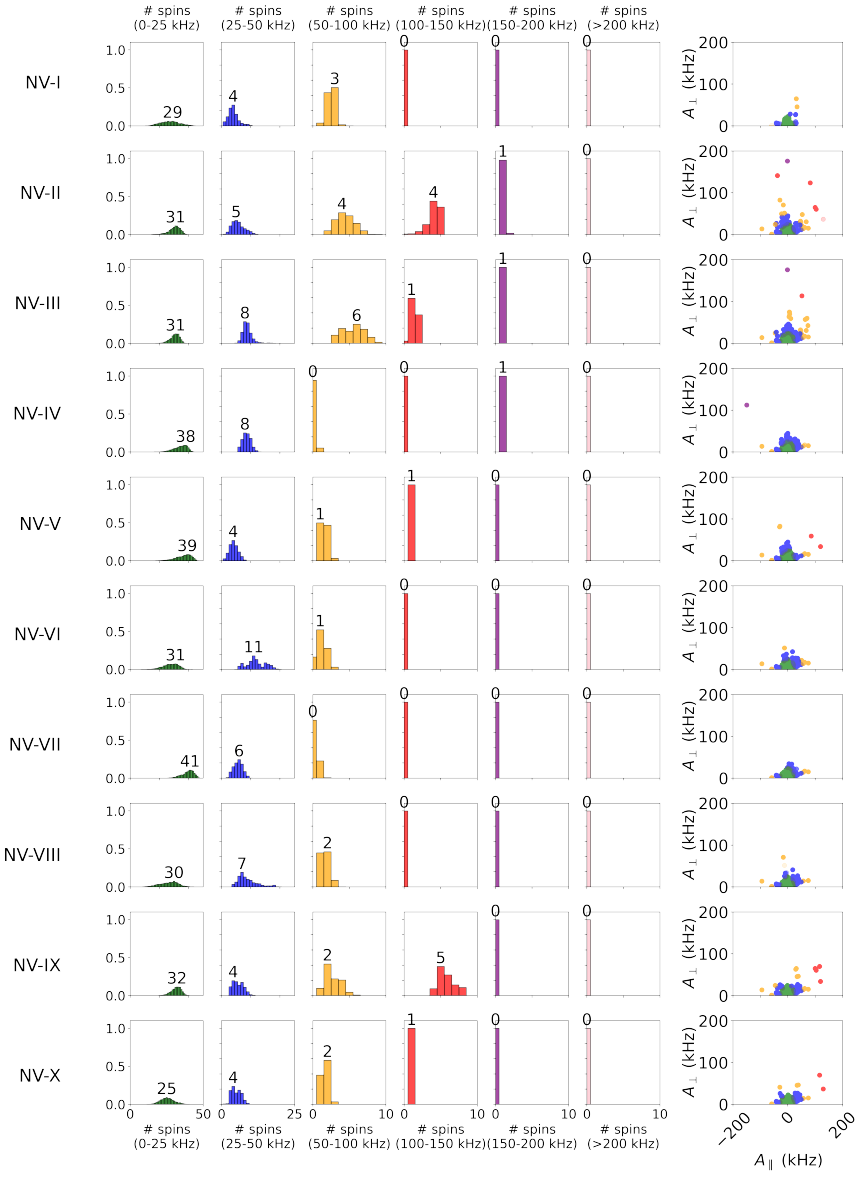

Figure 4.7: Posterior distributions of the number and hyperfine couplings of spins from the hybrid algorithm applied to ten NV centers in diamond, plotted by magnitude ($\sqrt{A_\perp^2 + A_\parallel^2}$) of hyperfine coupling.

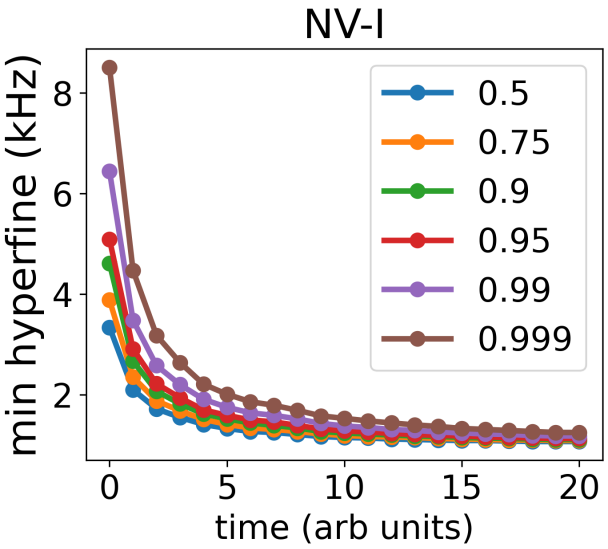

Figure 4.8: **Minimum detectable hyperfine coupling as a function of time averaging**. The time averaging corresponds to a number of raw photon counts that can be used to determine the average uncertainty per data point ε, which can be used as the threshold for the crossover regime in which small hyperfine couplings cannot be reliably distinguished from noise, given a specified statistical confidence.

CHAPTER 5

ONGOING AND FUTURE WORK, OUTLOOK, AND CONCLUSION

The previous chapters presented methods and results based on published work, establishing validated computational frameworks for high-throughput polarization calculations and spin-bath reconstruction. In this chapter, I discuss ongoing extensions of these methods, outline planned future directions, and provide an outlook on how these projects may evolve to further advance materials-driven approaches for next-generation computing hardware.

5.1 Ongoing and future work for Berry flux digaonalization

5.1.1 Outlook for Berry flux diagonalization code

An important near-term objective is the full integration of the existing Berry flux diagonalization implementation into the broader open-source materials simulation ecosystem. In particular, incorporating the method into PYMATGEN and releasing a corresponding ATOMATE2 workflow will make the approach accessible within established high-throughput infrastructures [61, 64]. The planned ATOMATE2 workflow will enable automated polarization calculations between two crystalline compounds while allowing users to select between the traditional interpolation-based Berry-phase workflow and the Berry flux diagonalization approach through a simple runtime flag. Such integration will lower the barrier to widespread adoption of this algorithm, facilitate further benchmarking across large materials datasets, and enable systematic comparisons between interpolation-dependent and interpolation-free polarization calculations in realistic computational studies.

Beyond workflow integration, an important direction involves improving computational efficiency. At present, the dominant bottleneck in the Berry flux diagonalization method

arises not from the diagonalization itself but from input/output overhead. The current implementation requires exporting plane-wave coefficients from a self-consistent density functional theory (DFT) calculation, storing them in intermediate files, and subsequently re-parsing these data within a Python environment. For large systems and high-throughput studies, this repeated serialization and reconstruction of wavefunction data introduces substantial runtime and storage costs, limiting scalability.

In the longer term, these considerations motivate embedding the algorithm directly within plane-wave DFT packages such as Quantum ESPRESSO, VASP, or related electronic-structure codes. Implementing Berry flux diagonalization in Fortran alongside existing Berry-phase routines would allow direct access to wavefunction coefficients already resident in memory during the self-consistent calculation. Eliminating intermediate file generation and parsing would remove the primary performance bottleneck, enabling significant speedups and reducing storage demands. More broadly, a native implementation would position Berry flux diagonalization as a first-class polarization method, opening the door to routine use in large-scale materials discovery workflows and enabling polarization calculations that are both robust and computationally efficient.

5.1.2 Extensions to larger supercells and ML frameworks

An extension of the Berry flux diagonalization method to compute spontaneous polarization developed in this thesis is its application to larger simulation cells and chemically complex materials, particularly those relevant to functional oxide design and defect engineering. In the current framework, spontaneous polarization is computed by relating a polar structure to a fictitious nonpolar reference structure, with structural heuristics introduced in Chapter 2. These heuristics rely primarily on real-space atomic displacements between corresponding structures. While effective for closely related unit cells, this strategy becomes increasingly fragile when applied to systems exhibiting significant mismatch in relaxed lattice parameters,

especially in enlarged supercells.

In practice, large mismatches between relaxed lattice parameters introduce ambiguities in defining atomic correspondences and displacement magnitudes. When structures are embedded in supercells, small differences in lattice vectors can produce artificially large displacements once atomic positions are compared relative to a shared real-space origin. These effects are not physically meaningful, but instead arise from the geometric mismatch between simulation cells. As a result, displacement-based metrics that function reliably at the primitive-cell level can lead to incorrect pathway identification or inflated estimates of structural change in supercell calculations.

To address this limitation, an important direction for future work is the development of algorithms that compute spontaneous polarization directly between two polar enantiomers, bypassing the need for an explicit centrosymmetric reference structure. By reformulating the polarization calculation in terms of symmetry-related polar states (with identical relaxed lattice constants), the method can avoid displacement ambiguities associated with mismatched lattice parameters while remaining consistent with the modern theory of polarization. Such an approach would also better reflect experimentally relevant switching processes, which often occur between symmetry-related polar configurations rather than through a fictitious nonpolar reference configuration.

Enabling robust calculations in large supercells is important for studying chemically realistic materials and design strategies. Many technologically relevant phenomena, including dopant ordering, defect complexes, strain engineering, and emergent multiferroicity, require supercell models to capture local symmetry breaking and long-range interactions [48, 142]. In particular, engineering perovskites with coupled functional properties demands simulations that incorporate compositional disorder and low-symmetry distortions beyond the primitive-cell description. Extending the present framework to these regimes will therefore allow systematic exploration of doping strategies aimed at stabilizing targeted polarization and

magnetic responses.

5.1.3 Integrating Berry flux diagonalization into ML workflows for polarization and dielectric response

A longer-term opportunity enabled by Berry flux diagonalization is its integration into emerging machine-learning (ML) frameworks for predicting electric response properties of materials. Recent advances in atomistic ML models have demonstrated that quantities traditionally considered computationally expensive—such as polarization, dielectric tensors, and Born effective charges—can be learned alongside energies and forces within unified workflows. Workflows based on large-scale automated density functional theory (DFT) datasets have begun to establish the feasibility of training transferable models for electric response, as demonstrated by recent data-driven polarization and dielectric modeling efforts [143]. Within this context, Berry flux diagonalization provides a computationally efficient and automation-friendly route for generating the high-quality training data required by such approaches.

A key advantage of the method is that it reformulates polarization calculations in a way that is compatible with high-throughput pipelines. Traditional Berry-phase workflows often rely on carefully constructed interpolation paths, manual branch tracking, and convergence checks that complicate automation across large chemical spaces. In contrast, Berry flux diagonalization eliminates the need for structural interpolation and enables direct evaluation of polarization differences between structures. This substantially reduces workflow complexity and failure modes in automated environments, making it well suited for integration into scalable infrastructures similar to those used in recent ML-driven materials discovery studies. As a result, the method can be deployed to generate large datasets of dielectric and ferroelectric properties spanning diverse chemistries, crystal symmetries, and distortion patterns.

Such datasets are particularly valuable in light of new ML architectures designed to

learn electric response quantities jointly with atomic energetics. The Allegro-pol framework demonstrated that electric enthalpy, Born effective charges, and polarizabilities can be incorporated into equivariant neural network models alongside energies and forces, yielding a unified representation of response to applied external electric fields [144]. Building on this paradigm, subsequent work trained a MACE-Field model on density-functional perturbation theory (DFPT) data from the Materials Project to predict dielectric and ferroelectric behavior across materials classes [145].

Berry flux diagonalization offers a complementary and potentially transformative data-generation pathway for these models. Unlike DFPT-based datasets, which are typically restricted to equilibrium configurations, the Berry flux diagonalization approach can compute polarization for structures away from equilibrium, including configurations with nonvanishing forces. Access to such out-of-equilibrium data is particularly valuable for ML training because modern interatomic potentials learn most effectively from diverse configurations sampling local distortions, thermal fluctuations, and switching pathways. Including polarization labels for distorted structures enables models to learn the local relationship between atomic geometry and electric response, improving transferability and stability in molecular dynamics simulations and finite-temperature studies.

Moreover, the computational efficiency gained by avoiding interpolation and minimizing post-processing overhead makes it feasible to generate polarization and dielectric datasets at scales comparable to existing energy–force databases. This opens the possibility of constructing large, systematically curated repositories of ferroelectric and dielectric responses analogous to current materials databases, but enriched with field-dependent observables. Such datasets could accelerate the discovery of functional materials by enabling rapid ML screening for target properties such as large spontaneous polarization, enhanced dielectric response, or strong electromechanical coupling.

Looking forward, integrating Berry flux diagonalization into ML-based workflows suggests

a unified computational pipeline: automated DFT calculations generate polarization and dielectric data using interpolation-free algorithms; these data train equivariant ML models capable of predicting electric response in real time; and the resulting models enable large-scale exploration of vast compositional and structural spaces.

5.2 Ongoing and future work for spin-bath characterization of spin defects from coherence spectra

5.2.1 Extending framework from NV centers in diamond to axial divacancies in silicon carbide

The trans-dimensional Bayesian inference framework presented in Chapters 3-4 reconstructs the Hamiltonians of single nitrogen-vacancy (NV) centers in diamond, but we can extend the framework to other materials. Of particular interest is extending this inference framework to the more complex case of axial divacancies in silicon carbide (SiC). SiC is a promising platform due to its telecom-band optical transitions and compatibility with scalable device architectures [70].

The extension to SiC introduces several significant methodological and computational challenges. First, the nuclear spin bath is more heterogeneous: in diamond, only one isotope (^{13}C) is relevant, whereas in SiC, both ^{29}Si and ^{13}C can be present. As a result,the existing inference algorithm can be augmented to simultaneously determine the isotope type and number of spins without incurring combinatorial scaling by associating isotope type with the lattice site, over which the MCMC walkers perform a discrete random walk. This modification retains the computational efficiency of the NV workflow while enabling isotope-specific inference.

Second, we must adapt our framework to the limited availability of validated DFT hyperfine datasets for SiC. For NV centers, we selected a hyperfine dataset computed with HSE06

based on benchmarking against experimental ^{13}C data [35, 97]; no such experimental benchmark datasets currently exist for SiC. While couplings can be computed with beyond-PBE functionals, the lack of experimental reference data adds uncertainty to the errors in the computed hyperfine couplings. To address this, each lattice site's parallel and perpendicular hyperfine tensor components can be modeled as drawn from a 2D Gaussian centered on PBE-computed values. The inference samples both the discrete lattice sites and the continuous values of hyperfine parameters that are close to the DFT-computed hyperfine values simultaneously. This hybrid scheme preserves the ability to tie hyperfine couplings to physical lattice locations while introducing the flexibility needed to match experimental spectra.

Third, the higher spin density in SiC means nuclear–nuclear interactions cannot be neglected. Because the central spin coherence must be modeled with CCE2 rather than CCE1, these interactions enter the Hamiltonian explicitly [131]. We will compute their strengths using the same procedure as for hyperfine couplings: starting from dipolar interactions between nuclear pairs within a given radius, then modeling them as Gaussians centered on the dipolar values to account for other effects (e.g., J-coupling). Including these terms may increase convergence time for the posterior distribution, but they should not be computationally prohibitive. Together, these developments will yield a scalable, isotope-aware, and experimentally anchored inference pipeline for SiC defects, expanding the range of quantum systems we can characterize in an automated, high-throughput fashion.

5.2.2 Adaptive and online experimental design

We are currently developing a high-throughput computational workflow with real-time data assimilation for automated characterization of spin-defects in semiconductors and insulators, such as nitrogen-vacancies (NVs) in diamond and divacancies in silicon carbide. The workflow serves as a bridge between large-scale spin-defect synthesis and application-specific

quantum defect selection by enabling automated inference of key properties such as coherence times and nuclear spin environments. Coherence times and the surrounding nuclear spin environment (specifically, the positions of nuclear spins in the lattice and their coupling strengths to the spin-defect) directly impact spin-defect device performance. However, different applications demand different properties. For quantum sensing, strongly coupled nuclear spins are typically undesirable, as they introduce magnetic noise that can obscure the target signal. In contrast, for quantum computing and quantum communication, such couplings can be beneficial, enabling entanglement with nearby nuclear spins that serve as local memory units. The computational pipeline we are developing integrates experimental data with simulation through an iterative, Bayesian feedback loop between experiments and simulations. Standard experimental measurements are initially performed on a spin-defect of interest. The experimental data is then fed into a trans-dimensional Bayesian inference algorithm presented in Ch. 3 to generate a set of "digital twins", which are simulated realizations of the most likely local nuclear spin environments. The inferences are then used to optimize the next set of experiments for maximal information gain. This loop repeats until a desired characterization confidence is reached.

We face several HPC-related challenges in implementing this adaptive experimental scheme:

1. I/O and Data Movement: A major challenge is the real-time integration of experimental data into our Bayesian inference pipeline, as well as between our Bayesian inference pipeline (on CPUs) and digital twin simulations (on GPUs). These workflows require timely transfer of data across facilities (e.g.,from experimental setups at Argonne or UChicago to HPC resources). Since each experimental measurement takes on the order of a few microseconds, we need to find the optimal real-time data batching strategy.

2. Workflow Variability and Scalability: The number of digital twins to simulate varies dynamically depending on the posterior uncertainty and the defect under study. This

makes it difficult to predefine static job sizes or resource requirements,complicating queueing and resource allocation. Moreover, digital twin simulations are computationally expensive and require scalable multi-node GPU resources for nuclear spin-bath simulations, while Bayesian inference is CPU-bound. Thus, we will require flexible and dynamic resource orchestration across both CPU and GPU architectures.

3. Software Interoperability and Portability: Our workflow combines multiple open-source codes:

    - PyCCE: GPU-accelerated software to compute coherence times using the cluster-correlation expansion(CCE) method [130]
    - A custom Python Bayesian inference engine [45, 46]
    - ATOMATE2 and FIREWORKS: Workflow management tools to schedule and launch CPU and GPU jobs [61]
    - Quantum ESPRESSO: GPU-accelerated software to perform density functional theory (DFT)calculations used to inform Bayesian priors
    - WEST: massively-parallel, GPU-accelerated software to perform time-dependent DFT, many-body perturbation theory, and quantum embedding theory calculations used to computationally characterize the spin-defects and inform Bayesian priors [146–149]

    These are heterogeneous in their hardware demands, language stacks, and I/O formats. Ensuring seamless communication across them is a continual challenge, especially across experimental sites and computational architectures.

4. Porting and Profiling: While PyCCE supports GPU execution, integration with other Python-based CPU-only libraries for inference and orchestration remains complex. We lack fine-grained tools to profile the multi-step performance of the entire workflow end-

to-end, making it difficult to identify and mitigate bottlenecks in either data movement, inference runtime, or digital twin simulations.

5. Productivity Challenges: Developing and maintaining a flexible, modular software stack across different HPC systems, experimental sites, and spin-defect targets requires substantial effort. Minor changes in experimental protocols or target materials often demand reconfiguration of the pipeline. Debugging errors that occur across this stack, especially in automated multi-step workflows, is time-consuming and often requires deep familiarity with all integrated tools.

It currently takes a few days to characterize a single spin-defect manually, and no fully automated workflows currently exist for this kind of characterization of spin-defects. This level of high-throughput characterization would be transformative for the practical deployment of spin-defects into quantum sensing and computing technologies. In the short term, we will validate our automated workflow by manually confirming the predicted coherence times, nuclear spin locations, and hyperfine couplings for a subset of characterized defects. Demonstrating agreement between predicted and measured properties would confirm the utility of our approach for both basic quantum science and applied quantum engineering tasks. In the longer term, this platform would enable a new mode of operation: rapid, automated characterization of spin-defects that directly feed into the design and selection of defects for specific applications, whether for high-coherence quantum sensors or multi-qubit registers for quantum computing. It would also allow us to bridge the gap between computational screening of new defect candidates and experimental validation. When new candidate spin-defects are proposed through first-principles modeling or high-throughput computational screening, this workflow could be used to experimentally confirm their properties in a streamlined, automated fashion.

5.2.3 Adaptive pulse sequence optimization

A central outcome of this thesis is the ability to reconstruct and characterize the microscopic spin environments surrounding individual defects. Having established methods to infer the structure of spin baths from experimental data, a next step is to move from characterization to control: how can detailed knowledge of the local environment be leveraged to actively mitigate magnetic noise and prolong quantum coherence? Rather than treating decoherence as an unavoidable limitation, the reconstructed spin bath provides a quantitative starting point for designing control strategies tailored to a specific defect realization.

In conventional experiments, dynamical decoupling (DD) protocols suppress environmental noise. Experimental pulse sequences consisting of evenly spaced π-pulses—such as Carr–Purcell–Meiboom–Gill (CPMG)–type protocols—are widely considered to be nearly optimal in situations where little is known about the surrounding spin bath [128]. Their optimality follows from assumptions about the noise spectrum: magnetic fluctuations are typically modeled as stationary, Gaussian-distributed noise with a smooth spectral density dominated by low-frequency components [150]. Under these assumptions, evenly-spaced π-pulses act as spectral filters that suppress low-frequency noise while extending coherence times in an average-case sense. However, these protocols implicitly assume ensemble-averaged environments and stochastic noise models, whereas the experiments considered here probe single defects embedded in specific, well-characterized nuclear spin baths. In such systems, decoherence arises not from featureless noise but from coherent interactions with a finite number of spins possessing well-defined couplings and dynamics. As a result, evenly spaced pulse sequences are optimal only in the absence of microscopic knowledge; once the bath is characterized, they no longer represent the fundamental performance limit.

With detailed reconstructions of individual spin baths in hand, an important direction for future work is the design of defect-specific control sequences that maximize coherence for a given environment rather than for an assumed noise distribution. This suggests reframing

pulse-sequence design as an inference and optimization problem over a discrete–continuous parameter space. Hybrid Markov chain Monte Carlo (MCMC) approaches developed in this thesis naturally extend to this setting. For example, random-walk Metropolis–Hastings updates can explore continuous parameters such as interpulse spacings and pulse rotation angles, while reversible-jump MCMC enables trans-dimensional moves that vary the number of pulses in the sequence itself. Such a framework allows the complexity of the control protocol to emerge adaptively from the data, balancing sequence length against achievable coherence enhancement.

Finally, combining pulse-sequence optimization with the adaptive, online experimental design strategies discussed in the previous chapter points toward a unified closed-loop framework. In this framework, experiments iteratively characterize the spin bath, update probabilistic models of the environment, and propose new control sequences predicted to maximize coherence. Measurements performed under these optimized sequences then refine the inferred bath parameters, enabling simultaneous learning and noise mitigation. This integration of inference, control, and experiment establishes a pathway toward autonomous quantum sensing and defect engineering, where computational models and laboratory measurements co-evolve to both understand and suppress decoherence at the level of individual quantum systems.

5.3 Concluding remarks

The methods and frameworks presented in this thesis provide a foundation for bridging fundamental materials science with the realization of low-energy, high-performance, and quantum-enabled computing technologies, underscoring the central principle that the future of computation depends as much on materials as on algorithms. By developing scalable and physics-informed computational methods for automated polarization evaluation, high-throughput materials screening, and probabilistic reconstruction of spin environments, this thesis presents computational tools that advance our ability to screen, design, and char-

acterize materials with specific functional properties. These methods enable systematic exploration of compositional and structural design spaces, support iterative feedback between computation and experiment, and generate reliable data for machine-learning models, accelerating the discovery and optimization of functional materials for application in next-generation computing architectures.

BIBLIOGRAPHY

[1] G. E. Moore, "Cramming more components onto integrated circuits", Electronics **38**, 1–4 (1965).

[2] R. Dennard, F. Gaensslen, H.-N. Yu, V. Rideout, E. Bassous, and A. LeBlanc, "Design of ion-implanted mosfet's with very small physical dimensions", IEEE Journal of Solid-State Circuits **9**, 256–268 (1974).

[3] D. Frank, R. Dennard, E. Nowak, P. Solomon, Y. Taur, and H.-S. P. Wong, "Device scaling limits of si mosfets and their application dependencies", Proceedings of the IEEE **89**, 259–288 (2001).

[4] M. Bohr, "A 30 year retrospective on dennard's mosfet scaling paper", IEEE Solid-State Circuits Society Newsletter **12**, 11–13 (2007).

[5] M. Horowitz, "1.1 computing's energy problem (and what we can do about it)", in 2014 ieee international solid-state circuits conference digest of technical papers (isscc) (IEEE, 2014), pp. 10–14.

[6] C. E. Leiserson, N. C. Thompson, J. S. Emer, B. C. Kuszmaul, B. W. Lampson, D. Sanchez, and T. B. Schardl, "There's plenty of room at the top: what will drive computer performance after moore's law?", Science **368**, eaam9744 (2020).

[7] R. Ramesh, S. Salahuddin, S. Datta, C. H. Diaz, D. E. Nikonov, I. A. Young, D. Ham, M.-F. Chang, W.-S. Khwa, A. S. Lele, et al., "Roadmap on low-power electronics", APL Materials **12**, 099201 (2024).

[8] A. Shehabi, A. Newkirk, S. J. Smith, A. Hubbard, N. Lei, M. A. B. Siddik, B. Holecek, J. Koomey, E. Masanet, and D. Sartor, *2024 united states data center energy useage report*, tech. rep. (Lawrence Berkeley National Laboratory, 2024).

[9] S. S. Gill, A. Kumar, H. Singh, M. Singh, K. Kaur, M. Usman, and R. Buyya, "Quantum computing: a taxonomy, systematic review and future directions", Software: Practice and Experience **52**, 66–114 (2022).

[10] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer", SIAM review **41**, 303–332 (1999).

[11] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, "Quantum computational chemistry", Reviews of Modern Physics **92**, 015003 (2020).

[12] P. Sureshkumar and M. Narayan, "Quantum computing in material science", in *From bits to qubits: the quantum transformation of computing: the power of quantum computing* (Springer, 2026), pp. 393–406.

[13] N. P. De Leon, K. M. Itoh, D. Kim, K. K. Mehta, T. E. Northup, H. Paik, B. Palmer, N. Samarth, S. Sangtawesin, and D. W. Steuerman, "Materials challenges and opportunities for quantum computing hardware", Science **372**, eabb2823 (2021).

[14] Z. Guo, J. Yin, Y. Bai, D. Zhu, K. Shi, G. Wang, K. Cao, and W. Zhao, "Spintronics for energy-efficient computing: an overview and outlook", Proceedings of the IEEE **109**, 1398–1417 (2021).

[15] M. M. Vopson, "Fundamentals of multiferroic materials and their possible applications", Critical Reviews in Solid State and Materials Sciences **40**, 223–250 (2015).

[16] M. Trassin, "Low energy consumption spintronics using multiferroic heterostructures", Journal of Physics: Condensed Matter **28**, 033001 (2016).

[17] J. A. C. Incorvia, T. P. Xiao, N. Zogbi, A. Naeemi, C. Adelmann, F. Catthoor, M. Tahoori, F. Casanova, M. Becherer, G. Prenat, et al., "Spintronics for achieving system-level energy-efficient logic", Nature Reviews Electrical Engineering **1**, 700–713 (2024).

[18] P. Barla, V. K. Joshi, and S. Bhat, "Spintronic devices: a promising alternative to cmos devices", Journal of Computational Electronics **20**, 805–837 (2021).

[19] G. Finocchio, M. Di Ventra, K. Y. Camsari, K. Everschor-Sitte, P. K. Amiri, and Z. Zeng, "The promise of spintronics for unconventional computing", Journal of Magnetism and Magnetic Materials **521**, 167506 (2021).

[20] K. Rallis, I. Liliopoulos, E. Tsipas, G. D. Varsamis, N. Melissourgos, I. G. Karafyllidis, G. C. Sirakoulis, and P. Dimitrakis, "Hardware-level interfaces for hybrid quantum-classical computing systems", in 2025 6th international conference in electronic engineering & information technology (eeite) (IEEE, 2025), pp. 1–6.

[21] W. Jeon, G. Ko, J. Lee, H. Lee, D. Ha, and W. W. Ro, "Deep learning with gpus", in *Advances in computers*, Vol. 122 (Elsevier, 2021), pp. 167–215.

[22] B. Bauer, S. Bravyi, M. Motta, and G. K.-L. Chan, "Quantum algorithms for quantum chemistry and quantum materials science", Chemical Reviews **120**, 12685–12717 (2020).

[23] D. Xu, Q. Zhang, X. Huo, Y. Wang, and M. Yang, "Advances in data-assisted high-throughput computations for material design", Materials Genome Engineering Advances **1**, e11 (2023).

[24] G. Tom, S. P. Schmid, S. G. Baird, Y. Cao, K. Darvish, H. Hao, S. Lo, S. Pablo-García, E. M. Rajaonson, M. Skreta, et al., "Self-driving laboratories for chemistry and materials science", Chemical Reviews **124**, 9633–9732 (2024).

[25] F. Naccarato, F. Ricci, J. Suntivich, G. Hautier, L. Wirtz, and G.-M. Rignanese, "Searching for materials with high refractive index and wide band gap: a first-principles high-throughput study", Physical Review Materials **3**, 044602 (2019).

[26] K. Choudhary, I. Kalish, R. Beams, and F. Tavazza, "High-throughput identification and characterization of two-dimensional materials using density functional theory", Scientific Reports **7**, 5179 (2017).

[27] M. Y. Toriyama, J. Zhan, S. Kanai, and G. Galli, "Strategies to search for two-dimensional materials with long spin qubit coherence time", npj 2D Materials and Applications **9**, 108 (2025).

[28] S. Kanai, F. J. Heremans, H. Seo, G. Wolfowicz, C. P. Anderson, S. E. Sullivan, M. Onizhuk, G. Galli, D. D. Awschalom, and H. Ohno, "Generalized scaling of spin qubit coherence in over 12,000 host materials", Proceedings of the National Academy of Sciences **119**, e2121808119 (2022).

[29] D. Vanderbilt, *Berry phases in electronic structure theory: electric polarization, orbital magnetization and topological insulators* (Cambridge University Press, 2018).

[30] T. E. Smidt, S. A. Mack, S. E. Reyes-Lillo, A. Jain, and J. B. Neaton, "An automatically curated first-principles database of ferroelectrics", Scientific Data **7**, 72 (2020).

[31] F. Ricci, S. E. Reyes-Lillo, S. A. Mack, and J. B. Neaton, "Candidate ferroelectrics via ab initio high-throughput screening of polar materials", npj Computational Materials **10**, 15 (2024).

[32] M. Abolhasani and E. Kumacheva, "The rise of self-driving labs in chemical and materials sciences", Nature Synthesis **2**, 483–492 (2023).

[33] L. Childress and R. Hanson, "Diamond nv centers for quantum computing and quantum networks", MRS Bulletin **38**, 134–138 (2013).

[34] T. Taminiau, J. Wagenaar, T. Van der Sar, F. Jelezko, V. V. Dobrovitski, and R. Hanson, "Detection and control of individual nuclear spins using a weakly coupled electron spin", Physical Review Letters **109**, 137602 (2012).

[35] M. Abobeih, J. Randall, C. Bradley, H. Bartling, M. Bakker, M. Degen, M. Markham, D. Twitchen, and T. Taminiau, "Atomic-scale imaging of a 27-nuclear-spin cluster using a quantum sensor", Nature **576**, 411–415 (2019).

[36] N. Zhao, S.-W. Ho, and R.-B. Liu, "Decoherence and dynamical decoupling control of nitrogen vacancy center electron spins in nuclear spin baths", Physical Review B **85**, 115303 (2012).

[37] H. Park, J. Lee, S. Han, S. Oh, and H. Seo, "Decoherence of nitrogen-vacancy spin ensembles in a nitrogen electron-nuclear spin bath in diamond", npj Quantum Information **8**, 95 (2022).

[38] G. Galli, A. Castillo, S. Chattaraj, S. Chen, M. Govoni, Y. Jin, J. Nagura, A. N. Poteshman, V. Somjit, M. Y. Toriyama, et al., "Strategies to predict and design spin defects for quantum technologies", `10.26434/chemrxiv.15000706/v1` (2026).

[39] J. Damewood, J. Karaguesian, J. R. Lunger, A. R. Tan, M. Xie, J. Peng, and R. Gómez-Bombarelli, "Representations of materials for machine learning", Annual Review of Materials Research **53**, 399–426 (2023).

[40] E. C.-Y. Yuan, Y. Liu, J. Chen, P. Zhong, S. Raja, T. Kreiman, S. Vargas, W. Xu, M. Head-Gordon, C. Yang, et al., "Foundation models for atomistic simulation of chemistry and materials", Nature Reviews Chemistry, 1–19 (2026).

[41] Y. Liu, C. Niu, Z. Wang, Y. Gan, Y. Zhu, S. Sun, and T. Shen, "Machine learning in materials genome initiative: a review", Journal of Materials Science & Technology **57**, 113–122 (2020).

[42] K. Choudhary, B. DeCost, C. Chen, A. Jain, F. Tavazza, R. Cohn, C. W. Park, A. Choudhary, A. Agrawal, S. J. Billinge, et al., "Recent advances and applications of deep learning methods in materials science", npj Computational Materials **8**, 59 (2022).

[43] R. Kondor, "The principles behind equivariant neural networks for physics and chemistry", Proceedings of the National Academy of Sciences **122**, e2415656122 (2025).

[44] A. N. Poteshman, F. Ricci, and J. B. Neaton, "High-throughput computation of electric polarization in solids via berry flux diagonalization", npj Computational Materials **12**, 87 (2026).

[45] A. N. Poteshman, J. Yun, T. H. Taminiau, and G. Galli, "Trans-dimensional hamiltonian model selection and parameter estimation from sparse, noisy data", Quantum **10**, 2055 (2026).

[46] A. N. Poteshman, M. Onizhuk, C. Egerstrom, D. P. Mark, D. D. Awschalom, F. J. Heremans, and G. Galli, "High-throughput spin-bath characterization of spin defects in semiconductors", Physical Review Applied **24**, 054048 (2025).

[47] J. Scott, "Applications of modern ferroelectrics", Science **315**, 954–959 (2007).

[48] Y. Shin and G. Galli, "Tunable ferroelectricity in oxygen-deficient perovskites with grenier structure", npj Computational Materials **9**, 218 (2023).

[49] N. A. Spaldin and R. Ramesh, "Advances in magnetoelectric multiferroics", Nature Materials **18**, 203–212 (2019).

[50] R. Resta and D. Vanderbilt, "Theory of polarization: a modern approach", in *Physics of ferroelectrics: a modern perspective*, edited by J.-M. T. K. Rabe C. H. Ahn (Springer, 2007), pp. 31–68.

[51] J. B. Neaton, C. Ederer, U. V. Waghmare, N. A. Spaldin, and K. Rabe, "First-principles study of spontaneous polarization in multiferroic bifeo$_3$", Physical Review B **71**, 014113 (2005).

[52] R. Resta, "Macroscopic polarization in crystalline dielectrics: the geometric phase approach", Reviews of Modern Physics **66**, 899 (1994).

[53] R. King-Smith and D. Vanderbilt, "Theory of polarization of crystalline solids", Physical Review B **47**, 1651 (1993).

[54] J. M. Munro, V. S. Liu, V. Gopalan, and I. Dabo, "Implementation of distortion symmetry for the nudged elastic band method with dispy", npj Computational Materials **5**, 52 (2019).

[55] J. Bonini, D. Vanderbilt, and K. M. Rabe, "Berry flux diagonalization: application to electric polarization", Physical Review B **102**, 045141 (2020).

[56] N. A. Spaldin, "A beginner's guide to the modern theory of polarization", Journal of Solid State Chemistry **195**, 2–10 (2012).

[57] D. Vanderbilt and R. King-Smith, "Electronic polarization in the ultrasoft pseudopotential formalism", arXiv preprint cond-mat/9801177, `10.48550/arXiv.cond-mat/9801177Focustolearnmore` (1998).

[58] D. Vanderbilt, "Berry-phase theory of proper piezoelectric response", Journal of Physics and Chemistry of Solids **61**, 147–151 (2000).

[59] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple", Physical Review Letters **77**, 3865 (1996).

[60] G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method", Physical Review B **59**, 1758 (1999).

[61] A. M. Ganose, H. Sahasrabuddhe, M. Asta, K. Beck, T. Biswas, A. Bonkowski, J. Bustamante, X. Chen, Y. Chiang, D. C. Chrzan, et al., "Atomate2: modular workflows for materials science", Digital Discovery **4**, 1944–1973 (2025).

[62] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, et al., "Commentary: the materials project: a materials genome approach to accelerating materials innovation", APL materials **1**, 011002 (2013).

[63] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, et al., "Advanced capabilities for materials modelling with quantum espresso", Journal of Physics: Condensed matter **29**, 465901 (2017).

[64] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, "Python materials genomics (pymatgen): a robust, open-source python library for materials analysis", Computational Materials Science **68**, 314–319 (2013).

[65] D. Brunato, P. Delugas, G. Borghi, and A. Fonari, *Qeschema*, `https://github.com/QEF/qeschema`.

[66] K. Bystrom, D. Broberg, S. Dwaraknath, K. A. Persson, and M. Asta, "Pawpyseed: perturbation-extrapolation band shifting corrections for point defect calculations", arXiv preprint arXiv:1904.11572, `10.48550/arXiv.1904.11572Focustolearnmore` (2019).

[67] S. Castelletto and A. Boretti, "Silicon carbide color centers for quantum applications", Journal of Physics: Photonics **2**, 022001 (2020).

[68] L. V. Rodgers, L. B. Hughes, M. Xie, P. C. Maurer, S. Kolkowitz, A. C. Bleszynski Jayich, and N. P. de Leon, "Materials challenges for quantum technologies based on color centers in diamond", MRS Bulletin **46**, 623–633 (2021).

[69] C. P. Anderson, E. O. Glen, C. Zeledon, A. Bourassa, Y. Jin, Y. Zhu, C. Vorwerk, A. L. Crook, H. Abe, J. Ul-Hassan, et al., "Five-second coherence of a single spin with single-shot readout in silicon carbide", Science Advances **8**, eabm5912 (2022).

[70] A. Bourassa, C. P. Anderson, K. C. Miao, M. Onizhuk, H. Ma, A. L. Crook, H. Abe, J. Ul-Hassan, T. Ohshima, N. T. Son, G. Galli, and D. D. Awschalom, "Entanglement

and control of single nuclear spins in isotopically engineered silicon carbide", Nature Materials **19**, 1319–1325 (2020).

[71] C. Bradley, S. de Bone, P. Möller, S. Baier, M. Degen, S. Loenen, H. Bartling, M. Markham, D. Twitchen, R. Hanson, D. Elkouss, and T. TH, "Robust quantum-network memory based on spin qubits in isotopically engineered diamond", npj Quantum Information **8**, 122 (2022).

[72] A. Waeber, G. Gillard, G. Ragunathan, M. Hopkinson, P. Spencer, D. Ritchie, M. Skolnick, and E. Chekhovich, "Pulse control protocols for preserving coherence in dipolar-coupled nuclear spin baths", Nature Communications **10**, 3157 (2019).

[73] W. Dong, F. Calderon-Vargas, and S. E. Economou, "Precise high-fidelity electron–nuclear spin entangling gates in nv centers via hybrid dynamical decoupling sequences", New Journal of Physics **22**, 073059 (2020).

[74] J. Cramer, N. Kalb, M. A. Rol, B. Hensen, M. S. Blok, M. Markham, D. J. Twitchen, R. Hanson, and T. H. Taminiau, "Repeated quantum error correction on a continuously encoded qubit by real-time feedback", Nature Communications **7**, 11526 (2016).

[75] T. H. Taminiau, J. Cramer, T. van der Sar, V. V. Dobrovitski, and R. Hanson, "Universal control and error correction in multi-qubit spin registers in diamond", Nature Nanotechnology **9**, 171–176 (2014).

[76] J. C. Marcks, M. Onizhuk, N. Delegan, Y.-X. Wang, M. Fukami, M. Watts, A. A. Clerk, F. J. Heremans, G. Galli, and D. D. Awschalom, "Guiding diamond spin qubit growth with computational methods", Physical Review Materials **8**, 026204 (2024).

[77] C. P. Horn, C. Wicker, A. Wellisz, C. Zeledon, P. V. K. Nittala, F. J. Heremans, D. D. Awschalom, and S. Guha, "Controlled spalling of 4h silicon carbide with investigated spin coherence for quantum engineering integration", ACS Nano **18**, 31381–31389 (2024).

[78] R. Budakian, A. Finkler, A. Eichler, M. Poggio, C. L. Degen, S. Tabatabaei, I. Lee, P. C. Hammel, S. P. Eugene, T. H. Taminiau, et al., "Roadmap on nanoscale magnetic resonance imaging", Nanotechnology **35**, 412001 (2024).

[79] A. Laraoui, F. Dolde, C. Burk, F. Reinhard, J. Wrachtrup, and C. A. Meriles, "High-resolution correlation spectroscopy of 13c spins near a nitrogen-vacancy centre in diamond", Nature Communications **4**, 1651 (2013).

[80] K. Jung, M. Abobeih, J. Yun, G. Kim, H. Oh, A. Henry, T. Taminiau, and D. Kim, "Deep learning enhanced individual nuclear-spin detection", npj Quantum Information **7**, 41 (2021).

[81] B. Varona-Uriarte, C. Munuera-Javaloy, E. Terradillos, Y. Ban, A. Alvarez-Gila, E. Garrote, and J. Casanova, "Automatic detection of nuclear spins at arbitrary magnetic fields via signal-to-image ai model", Physical Review Letters **132**, 150801 (2024).

[82] N. Kura and M. Ueda, "Finite-error metrological bounds on multiparameter hamiltonian estimation", Physical Review A **97**, 012101 (2018).

[83] W. Yu, J. Sun, Z. Han, and X. Yuan, "Robust and efficient hamiltonian learning", Quantum **7**, 1045 (2023).

[84] H.-Y. Huang, Y. Tong, D. Fang, and Y. Su, "Learning many-body hamiltonians with heisenberg-limited scaling", Physical Review Letters **130**, 200403 (2023).

[85] A. Anshu, S. Arunachalam, T. Kuwahara, and M. Soleimanifar, "Sample-efficient learning of interacting quantum systems", Nature Physics **17**, 931–935 (2021).

[86] P. J. Green, "Reversible jump markov chain monte carlo computation and bayesian model determination", Biometrika **82**, 711–732 (1995).

[87] W. Vousden, W. M. Farr, and I. Mandel, "Dynamic temperature selection for parallel tempering in markov chain monte carlo simulations", Monthly Notices of the Royal Astronomical Society **455**, 1919–1937 (2016).

[88] D. G. Denison, C. C. Holmes, B. K. Mallick, and A. F. Smith, *Bayesian methods for nonlinear classification and regression*, Vol. 386 (John Wiley & Sons, 2002).

[89] D. Sanz-Alonso and O. Al-Ghattas, "A first course in monte carlo methods", arXiv preprint arXiv:2405.16359, 10.48550/arXiv.2405.16359 (2024).

[90] P. J. Green and D. I. Hastie, *Reversible jump mcmc*, tech. rep. (University of Bristol, 2009).

[91] A. Toubiana, M. L. Katz, and J. R. Gair, "Is there an excess of black holes around 20 m⊙? optimizing the complexity of population models with the use of reversible jump mcmc.", Monthly Notices of the Royal Astronomical Society **524**, 5844–5853 (2023).

[92] M. Zevin, C. Pankow, C. L. Rodriguez, L. Sampson, E. Chase, V. Kalogera, and F. A. Rasio, "Constraining formation models of binary black holes with gravitational-wave observations", The Astrophysical Journal **846**, 82 (2017).

[93] D. Zhu and R. Gibson, "Seismic inversion and uncertainty quantification using transdimensional markov chain monte carlo method", Geophysics **83**, R321–R334 (2018).

[94] Y. Cho, R. L. Gibson Jr, and D. Zhu, "Quasi 3d transdimensional markov-chain monte carlo for seismic impedance inversion and uncertainty analysis", Interpretation **6**, T613–T624 (2018).

[95] J. R. Oaks, P. L. Wood Jr, C. D. Siler, and R. M. Brown, "Generalizing bayesian phylogenetics to infer shared evolutionary events", Proceedings of the National Academy of Sciences **119**, e2121036119 (2022).

[96] M. Pagel and A. Meade, "Modelling heterotachy in phylogenetic inference by reversible-jump markov chain monte carlo", Philosophical Transactions of the Royal Society B: Biological Sciences **363**, 3955–3964 (2008).

[97] I. Takács and V. Ivády, "Accurate hyperfine tensors for solid state quantum applications: case of the nv center in diamond", Communications Physics **7**, 178 (2024).

[98] M. D. Hoffman, A. Gelman, et al., "The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo.", Journal of Machine Learning Research **15**, 1593–1623 (2014).

[99] G. Van de Stolpe, D. Kwiatkowski, C. Bradley, J. Randall, M. Abobeih, S. Breitweiser, L. Bassett, M. Markham, D. Twitchen, and T. Taminiau, "Mapping a 50-spin-qubit network through correlated sensing", Nature Communications **15**, 2006 (2024).

[100] B. Engquist, B. D. Froese, and Y. Yang, "Optimal transport for seismic full waveform inversion", arXiv preprint arXiv:1602.01540, `10.48550/arXiv.1602.01540` (2016).

[101] A. Dréau, J.-R. Maze, M. Lesik, J.-F. Roch, and V. Jacques, "High-resolution spectroscopy of single nv defects coupled with nearby 13 c nuclear spins in diamond", Physical Review B **85**, 134107 (2012).

[102] H. Oh, J. Yun, M. Abobeih, K.-H. Jung, K. Kim, T. Taminiau, and D. Kim, "Algorithmic decomposition for efficient multiple nuclear spin detection in diamond", Scientific Reports **10**, 14884 (2020).

[103] D. M. Walker, F. J. Pérez-Barbería, and G. Marion, "Stochastic modelling of ecological processes using hybrid gibbs samplers", Ecological Modelling **198**, 40–52 (2006).

[104] B. Mathew, A. Bauer, P. Koistinen, T. Reetz, J. Léon, and M. Sillanpää, "Bayesian adaptive markov chain monte carlo estimation of genetic parameters", Heredity **109**, 235–245 (2012).

[105] J. Zhang, J. Li, X. Chen, and Y. Li, "Geological structure-guided hybrid mcmc and bayesian linearized inversion methodology", Journal of Petroleum Science and Engineering **199**, 108296 (2021).

[106] S. Reuschen, F. Jobst, and W. Nowak, "Efficient discretization-independent bayesian inversion of high-dimensional multi-gaussian priors using a hybrid mcmc", Water Resources Research **57**, e2021WR030051 (2021).

[107] B. D. Nguyen, P. Potapenko, A. Demirci, K. Govind, S. Bompas, and S. Sandfeld, "Efficient surrogate models for materials science simulations: machine learning-based prediction of microstructure properties", Machine Learning with Applications **16**, 100544 (2024).

[108] A. Babbar, S. Ragunathan, D. Mitra, A. Dutta, and T. K. Patra, "Explainability and extrapolation of machine learning models for predicting the glass transition temperature of polymers", Journal of Polymer Science **62**, 1175–1186 (2024).

[109] C. Nyshadham, M. Rupp, B. Bekker, A. V. Shapeev, T. Mueller, C. W. Rosenbrock, G. Csányi, D. W. Wingate, and G. L. Hart, "Machine-learned multi-system surrogate models for materials prediction", npj Computational Materials **5**, 51 (2019).

[110] M. Verriere, N. Schunck, I. Kim, P. Marević, K. Quinlan, M. N. Ngo, D. Regnier, and R. D. Lasseri, "Building surrogate models of nuclear density functional theory with gaussian processes and autoencoders", Frontiers in Physics **10**, 1028370 (2022).

[111] T. Roongcharoen, G. Conter, L. Sementa, G. Melani, and A. Fortunelli, "Machine-learning-accelerated dft conformal sampling of catalytic processes", Journal of Chemical Theory and Computation **20**, 9580–9591 (2024).

[112] A. Chandrasekaran, D. Kamal, R. Batra, C. Kim, L. Chen, and R. Ramprasad, "Solving the electronic structure problem with machine learning", npj Computational Materials **5**, 22 (2019).

[113] M. A. Clyde, J. Ghosh, and M. L. Littman, "Bayesian adaptive sampling for variable selection and model averaging", Journal of Computational and Graphical Statistics **20**, 80–101 (2011).

[114] G. Seni and J. Elder, *Ensemble methods in data mining: improving accuracy through combining predictions* (Morgan & Claypool Publishers, 2010).

[115] P. Bühlmann, "Bagging, boosting and ensemble methods", in *Handbook of computational statistics: concepts and methods* (Springer, 2011), pp. 985–1022.

[116] M. Adrian, J. A. Soloff, and R. Willett, "Stabilizing black-box model selection with the inflated argmax", arXiv preprint arXiv:2410.18268, 10.48550/arXiv.2410.18268 (2024).

[117] Y. Li, Y. Wang, and L. Yan, "Surrogate modeling for bayesian inverse problems based on physics-informed neural networks", Journal of Computational Physics **475**, 111841 (2023).

[118] C. Lieberman, K. Willcox, and O. Ghattas, "Parameter and state model reduction for large-scale statistical inverse problems", SIAM Journal on Scientific Computing **32**, 2523–2542 (2010).

[119] A. Morello, J. J. Pla, P. Bertet, and D. N. Jamieson, "Donor spins in silicon for quantum technologies", Advanced Quantum Technologies **3**, 2000005 (2020).

[120] M. Sutula, I. Christen, E. Bersin, M. P. Walsh, K. C. Chen, J. Mallek, A. Melville, M. Titze, E. S. Bielejec, S. Hamilton, D. Braje, P. B. Dixon, and D. R. Englund, "Large-scale optical characterization of solid-state quantum emitters", Nature Materials **22**, 1338–1344 (2023).

[121] V. Vorobyov, J. Javadzade, M. Joliffe, F. Kaiser, and J. Wrachtrup, "Addressing single nuclear spins quantum memories by a central electron spin", Applied Magnetic Resonance **53**, 1317–1330 (2022).

[122] A. A. Neath and J. E. Cavanaugh, "The bayesian information criterion: background, derivation, and applications", Wiley Interdisciplinary Reviews: Computational Statistics **4**, 199–203 (2012).

[123] D. I. Hastie and P. J. Green, "Model choice using reversible jump markov chain monte carlo", Statistica Neerlandica **66**, 309–338 (2012).

[124] L. Childress, M. Gurudev Dutt, J. Taylor, A. Zibrov, F. Jelezko, J. Wrachtrup, P. Hemmer, and M. Lukin, "Coherent dynamics of coupled electron and nuclear spin qubits in diamond", Science **314**, 281–285 (2006).

[125] B. Rose, C. Weis, A. Tyryshkin, T. Schenkel, and S. A. Lyon, "Spin coherence and 14n eseem effects of nitrogen-vacancy centers in diamond with x-band pulsed esr", Diamond and Related Materials **72**, 32–40 (2017).

[126] N. Zhao, J. Wrachtrup, and R.-B. Liu, "Dynamical decoupling design for identifying weakly coupled nuclear spins in a bath", Physical Review A **90**, 032319 (2014).

[127] S. Kolkowitz, Q. P. Unterreithmeier, S. D. Bennett, and M. D. Lukin, "Sensing distant nuclear spins with a single electron spin", Physical Review Letters **109**, 137601 (2012).

[128] G. A. Álvarez and D. Suter, "Measuring the spectrum of colored noise by dynamical decoupling", Physical Review Letters **107**, 230501 (2011).

[129] W. Yang and R.-B. Liu, "Quantum many-body theory of qubit decoherence in a finite-size spin bath", Physical Review B **78**, 085315 (2008).

[130] M. Onizhuk and G. Galli, "Pycce: a python package for cluster correlation expansion simulations of spin qubit dynamics", Advanced Theory and Simulations **4**, 2100254 (2021).

[131] H. Seo, A. L. Falk, P. V. Klimov, K. C. Miao, G. Galli, and D. D. Awschalom, "Quantum decoherence dynamics of divacancy spins in silicon carbide", Nature Communications **7**, 12935 (2016).

[132] M. Biercuk, A. Doherty, and H. Uys, "Dynamical decoupling sequence construction as a filter-design problem", Journal of Physics B: Atomic, Molecular and Optical Physics **44**, 154002 (2011).

[133] L. Scorzato, "Reliability and interpretability in science and deep learning", Minds and Machines **34**, 27 (2024).

[134] L. Lüer, I. M. Peters, A. S. Smith, E. Dorschky, B. M. Eskofier, F. Liers, J. Franke, M. Sjarov, M. Brossog, D. M. Guldi, et al., "A digital twin to overcome long-time challenges in photovoltaics", Joule **8**, 295–311 (2024).

[135] R. W. Epps, M. S. Bowen, A. A. Volk, K. Abdel-Latif, S. Han, K. G. Reyes, A. Amassian, and M. Abolhasani, "Artificial chemist: an autonomous quantum dot synthesis bot", Advanced Materials **32**, 2001626 (2020).

[136] J. Park, J.-J. Zhou, Y. Luo, and M. Bernardi, "Predicting phonon-induced spin decoherence from first principles: colossal spin renormalization in condensed matter", Physical Review Letters **129**, 197201 (2022).

[137] G.-L. Zhang, W.-L. Ma, and R.-B. Liu, "Cluster correlation expansion for studying decoherence of clock transitions in spin baths", Physical Review B **102**, 245303 (2020).

[138] J. C. Marcks, M. Onizhuk, Y.-X. Wang, Y. Zhu, Y. Jin, B. S. Soloway, M. Fukami, N. Delegan, F. J. Heremans, A. A. Clerk, G. Galli, and D. D. Awschalom, "Quantum spin probe of single charge dynamics", Physical Review Letters **133**, 130802 (2024).

[139] C. Zeledon, B. Pingault, J. C. Marcks, M. Onizhuk, Y. Tsaturyan, Y.-x. Wang, B. S. Soloway, H. Abe, M. Ghezellou, J. Ul-Hassan, T. Ohshima, N. T. Son, F. J. Heremans, G. Galli, C. P. Anderson, and D. D. Awschalom, "Minute-long quantum coherence enabled by electrical depletion of magnetic noise", arXiv preprint arXiv:2504.13164, `10.48550/arXiv.2504.13164` (2025).

[140] J. McCallum, B. Johnson, and T. Botzem, "Donor-based qubits for quantum computing in silicon", Applied Physics Reviews **8**, 031314 (2021).

[141] D. Maile and J. Ankerhold, "Performance of quantum registers in diamond in the presence of spin impurities", Physical Review B **110**, 035407 (2024).

[142] T. Li, S. Deng, H. Liu, and J. Chen, "Insights into strain engineering: from ferroelectrics to related functional materials and beyond", Chemical Reviews **124**, 7045–7105 (2024).

[143] P. Zhong, D. Kim, D. S. King, and B. Cheng, "Machine learning interatomic potential can infer electrical response", npj Computational Materials **11**, 384 (2025).

[144] S. Falletta, A. Cepellotti, A. Johansson, C. W. Tan, M. L. Descoteaux, A. Musaelian, C. J. Owen, and B. Kozinsky, "Unified differentiable learning of electric response", Nature Communications **16**, 4031 (2025).

[145] B. A. Martin, A. M. Ganose, V. Kapil, T. Li, and K. T. Butler, "General learning of the electric response of inorganic materials", arXiv preprint arXiv:2508.17870, `10.48550/arXiv.2508.17870` (2025).

[146] M. Govoni and G. Galli, "Large scale gw calculations", Journal of Chemical Theory and Computation **11**, 2680–2696 (2015).

[147] V. W.-z. Yu and M. Govoni, "Gpu acceleration of large-scale full-frequency gw calculations", Journal of Chemical Theory and Computation **18**, 4690–4707 (2022).

[148] Y. Jin, V. W.-z. Yu, M. Govoni, A. C. Xu, and G. Galli, "Excited state properties of point defects in semiconductors and insulators investigated with time-dependent density functional theory", Journal of Chemical Theory and Computation **19**, 8689–8705 (2023).

[149] V. W.-z. Yu, Y. Jin, G. Galli, and M. Govoni, "Gpu-accelerated solution of the bethe–salpeter equation for large and heterogeneous systems", Journal of Chemical Theory and Computation **20**, 10899–10911 (2024).

[150] D. Suter and G. A. Álvarez, "Colloquium: protecting quantum information against environmental noise", Reviews of Modern Physics **88**, 041001 (2016).