Research Statement

Xuanzhao Gao

OVERVIEW

My expertise encompasses applied and computational mathematics, scientific computing, and computational physics. I am particularly focused on developing efficient numerical algorithms for modeling and simulating complex systems, with an emphasis on high-performance implementation. Specifically, I am working on **fast summation algorithms** designed for long-range interactions in doubly periodic systems, particularly in their application to molecular dynamics simulations. Additionally, I have a keen interest in **tensor network based algorithms** and their application in solving combinatorial optimization problems. In these projects, I focused on designing new algorithms and their high-performance implementation across various platforms, including both CPU and GPU architectures, and accelerating them through modern parallel and distributed computing techniques.

In the following parts, I present an overview of my past and current work, emphasizing my research contributions and achievements, while also outlining my future research plans that aim to integrate applied mathematics with scientific exploration.

RESEARCH ACCOMPLISHMENTS

Fast summation algorithms

In my previous works, I focused on the fast summation algorithms for quasi-2D systems, which are three-dimensional systems exhibiting periodicity in only two dimensions as shown in Figure 1 (a). They are ubiquitous in nature, e.g. charged colloidal suspensions, liquid crystals, and biological systems. However, simulating these systems is challenging due to the long range nature of the electrostatic interactions and the doubly-periodic boundary conditions, thus the fast summation algorithms for such systems are highly desirable. Throughout my Ph.D. studies, I developed a range of innovative methods to address these systems, which will be detailed in the subsequent sections.

The fast sum-of-Exponential Ewald2D method. Building on the success of the fast Gaussian transform via the sum-of-Exponential (SOE) approximation [10], I integrated the SOE approximation with the Ewald splitting technique to derive an accurate quasi-2D lattice summation formula with a complexity of $O(N^{1.4})$ called the SOEwald2D method [2]. This method is numerically
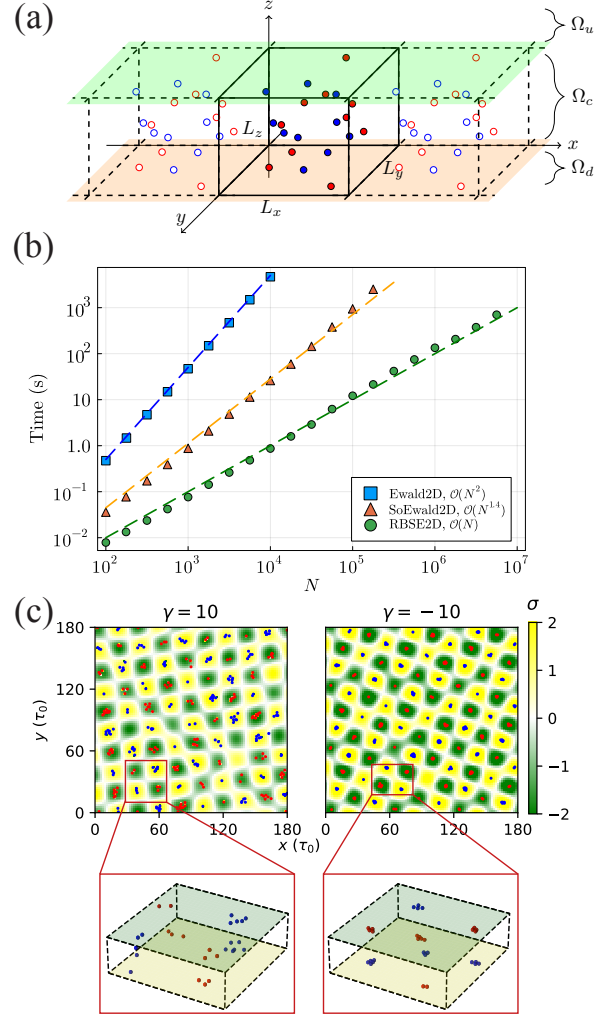

FIG. 1. (a) Illustration of a doubly periodic charged system. (b) The scaling of the SOEwald2D method [2], which shows that with the random batch sampling technique, the method (RBSE2D) can achieve linear complexity. (c) The SSB phenomenon in the system with strong negative dielectric confinement [4], showing the checkerboard pattern formation of the surface charge density and the ion distribution.

stable and maintains its accuracy regardless of the system's thickness. Furthermore, the method is enhanced by employing a random batch sampling technique, which achieves linear complexity, as shown in Figure 1 (b). It has been demonstrated that a speedup of $2 \sim 3$ orders of magnitude comparing with Ewald2D method, enabling molecular dynamics (MD) simulations with up to 10^6 particles on a single CPU core. My contributions include the derivation of the quasi-2D lattice summation formula,

the implementation of the method (see `SoEwald2D.jl`), and the execution of related numerical tests.

The fast sum-of-Gaussian method. To further enhance both the efficiency and accuracy of the fast summation methods for quasi-2D charged systems, my collaborators and I developed a spectral method that leverages the sum-of-Gaussian (SOG) expansion in conjunction with fast Fourier transforms. The fast sum-of-Gaussian method [6] is a spectral method based on the sum-of-Gaussian (SOG) expansion and fast Fourier transform, where we integrate SOG expansion with a novel Fourier-Chebyshev solver to develop a fast summation method that achieves a complexity of $\mathcal{O}(N \log N)$ per time step. Utilizing the SOG approach allows for a significant reduction in the required mesh resolution, down to approximately 60% of that needed by the Ewald-based methods, without necessitating upsampling. This method is also highly effective for systems with large aspect ratios, accommodating ratios of up to $L_{x/y}/L_z = 10^{3.5}$ without incurring additional computational costs. My contributions primarily involved the implementation of this method, which includes a highly efficient particle-mesh solver `ChebParticleMesh.jl` as a backend (similar to the `FINUFFT` package), along with the Fourier-Chebyshev solver `FastSpecSoG.jl`. The resulting implementation is both efficient and scalable, capable of calculating long-range electrostatic interactions in quasi-2D systems with 10^6 particles on a single CPU core in less than 10 seconds. Additionally, I conducted related numerical tests and collaborated on the algorithm design.

The spontaneous symmetry breaking in doubly periodic charged systems. We developed stochastic methods to investigate doubly periodic charged systems with dielectric mismatches [3, 5]. By employing appropriate singular subtraction techniques, this method can be effectively applied to systems exhibiting negative dielectric confinement. The findings indicate that SSB occurs in systems with strongly polarizable dielectric confinement [4], leading to the formation of checkerboard patterns as ions aggregate due to the strong polarization at the dielectric boundary as shown in Figure 1 (c). This work includes a prediction of the spontaneous symmetry breaking (SSB) phenomenon, detailing the conditions for lattice formation and the *wavelength* of the checkerboard patterns, which have been validated through molecular dynamics simulations. My contributions include the derivation and the implementation of the method (see `QuasiEwald.jl`), and the study of the SSB phenomenon via MD simulations.

Tensor network based algorithms

Tensor network is a powerful computational model developed in the last few decades and has been widely
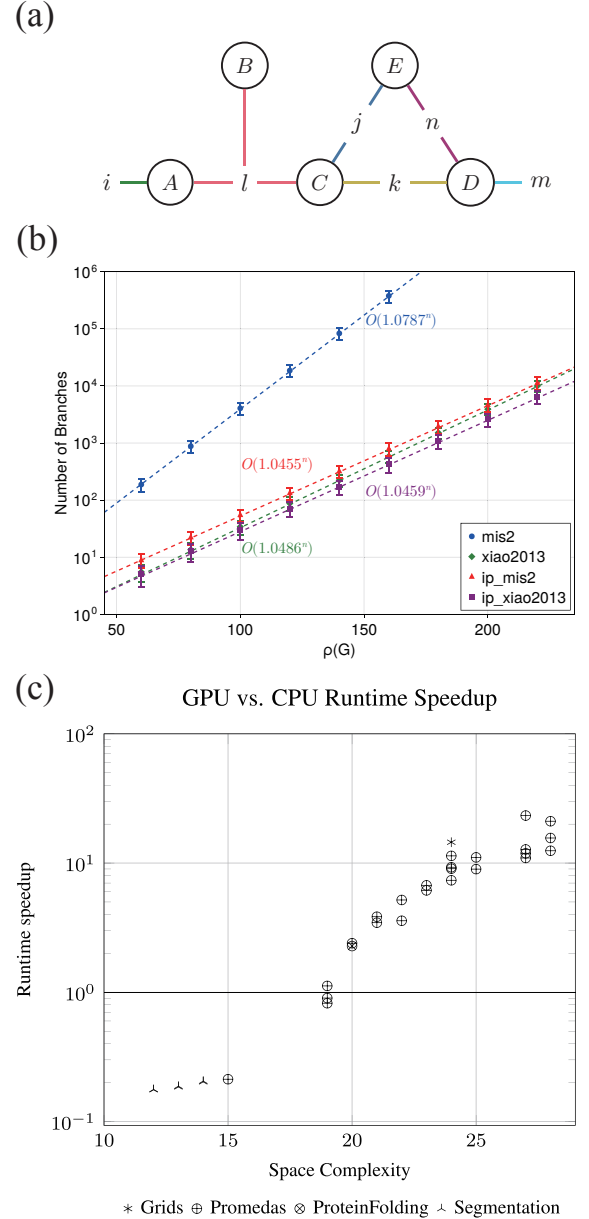

FIG. 2. (a) Illustration of a simple tensor network, where the white circles represent the tensor nodes and the lines represent the indices. (b) Number of branches of the optimal branching algorithm (`ip_mis2` and `ip_xiao2013`) compared with the traditional branching algorithms (`mis2` and `xiao2013`), $\rho(G)$ represents the number of vertices here. (c) Speedup of the GPU accelerated inference algorithm compared to the CPU version.

used in mathematics, condensed matter physics and quantum computing, an illustration is shown in Figure 2 (a). In my work, I focused on developing and implementing tensor network based algorithms and applied them to solve the combinatorial optimization problem [7, 8] and the inference problems in probabilistic graphical models [16]. I also developed softwares to

Project	Description
`ExTinyMD.jl`	A basic framework for simulating the particle systems.
`EwaldSummations.jl`	Implementations of various Ewald summation based methods.
`ChebParticleMesh.jl`	A particle-mesh solver for Poisson's equations for arbitrary dimensions and periodicity.
`TropicalNumbers.jl`	Implementation of the tropical semiring.
`CuTropicalGEMM.jl`	A GPU implementation of the general matrix multiplication of tropical semiring.
`TreeWidthSolver.jl`	A solver for the exact tree width and corresponding tree decomposition of graphs.

TABLE I. List of open source projects developed.

speed up the tensor network contraction, including implementing the general matrix multiplication of the tropical semiring on GPUs (`CuTropicalGEMM.jl`), and optimizing the contraction order by the tree width algorithm (`TreeWidthSolver.jl`).

Optimal branching via generic tensor networks. I and collaborators integrated tensor network techniques with the branch-and-bound algorithm to identify the optimal branching rule that minimizes the number of leaves in the searching tree [8]. We employ tensor network techniques to effectively capture the local structure of the satisfiability problem, followed by the application of integer programming to determine the optimal branching rule. Our approach has been successfully implemented for the maximum independent set problem. Our method addresses the bottleneck case of the state-of-the-art algorithm [18], as illustrated in Figure 2 (b). For 3-regular graphs, we compared our method with the traditional `mis2` and the state-of-the-art algorithm [18] (`xiao2013`). The results show that our method (`ip_xiao2013`) outperforms the previous methods both in scaling and absolute number of branches. It can practically decrease the branching complexity in practice to $\mathcal{O}(1.0455^n)$, where n is the number of vertices in the graph. My contributions include the design and implementation of the algorithm (see `OptimalBranching.jl`), as well as the execution of numerical tests.

Probabilistic inference via tensor networks and automatic differentiation. We convert the probabilistic inference problem into the tensor network contraction problem [16]. Using the state-of-the-art tensor network contraction order optimization algorithms and the automatic differentiation technique [14], we show that an exponential speedup can be achieved compared with the traditional methods. In this work, I developed a GPU implementation of the general matrix multiplication for tropical semiring (see `CuTropicalGEMM.jl`), achieving a remarkable speedup of over 4000 times compared to its CPU version. This advancement leads to a 10 times acceleration in large-scale inference problems, as illustrated in Figure 2 (c). My contributions encompass the implementation of the GPU kernel, optimization of the tensor network contraction order, and execution of numerical tests.

Open source software development

Throughout my Ph.D. studies, I actively contributed to open source software development within the scientific computing community, primarily utilizing the Julia programming language and C-CUDA. I participated in two notable open source projects, including:

- Open Source Promotion Plan 2023, TropicalGEMM on GPU,

- Google Summer of Code 2024, Tensor network contraction order optimization and visualization.

I have developed a diverse array of tools and libraries related to my research, covering a wide range of topics, including molecular dynamics simulation tools, tensor network based algorithms and graph theory tools. Some of the them are listed in Table I. I have also contributed to some popular open source projects, such as the `OMEinsumContractionOrders.jl`, which has been used as a backend of various popular tensor network libraries, including `OMEinsum.jl` and `ITensorNetworks.jl`. For more details, please refer to my Github page at https://github.com/ArrogantGao.

FUTURE RESEARCH PLANS

My future research will primarily concentrate on developing rapid and precise algorithms for simulating complex scientific processes, as well as applying these methods to tackle open scientific questions and generate new predictions. Additionally, I will focus on creating efficient implementations of these methods across various platforms, including GPUs and supercomputers. The areas of my interest are outlined as follows:

Integral equation methods for solving boundary value problems. I am interested in developing efficient algorithms for addressing systems with complex boundary conditions through the integral equation method. Currently, I am working on a boundary element solver for Poisson's equation involving multiple polarizable spheres. In the future, I aim to explore this field further and create more versatile solvers applicable to a wider range of boundary value problems and various types of partial dif-

ferential equations (PDEs) by making use of the state-of-the-art methods such as the RCIP method [11].

Fast summation methods for long-range interactions. I am keen on developing efficient summation methods for long-range interacting systems, such as those described by Poisson's equation and the Helmholtz equation. For example, I am interested in extending the recently introduced dual-space multilevel kernel-splitting (DMK) [9, 12] framework to a variety of long-range interactions and systems with periodic/partially periodic boundary conditions. These algorithms will also facilitate the solution of integral equations that arise in systems with complex geometries.

Tensor network based methods for quantum many-body systems. Tensor networks with specialized structures have long served as ansatz for quantum many-body systems [1], including matrix product states (MPS) and projected entangled-pair states (PEPS). With the advanced tensor network contraction techniques [16], I am interested in exploring more flexible tensor network structures as ansatz, which may yield more accurate representations of quantum many-body states. Additionally, I aim to develop efficient algorithms for evolving these states by integrating them with the Dirac–Frenkel/McLachlan variational principle [15], the automatic differentiation techniques [14] and state-of-the-art algorithms for the Schrödinger equation [13, 17].

---

[1] Matthew Fishman, Steven R. White, and E. Miles Stoudenmire. The ITensor software library for tensor network calculations. *SciPost Phys. Codebases*, page 4, 2022.

[2] Zecheng Gan, Xuanzhao Gao, Jiuyang Liang, and Zhenli Xu. Fast algorithm for quasi-2D Coulomb systems. *arXiv preprint arXiv:2403.01521*, 2024. Under the 2nd round of revision at the Journal of Computational Physics.

[3] Zecheng Gan, Xuanzhao Gao, Jiuyang Liang, and Zhenli Xu. Random batch Ewald method for dielectrically confined Coulomb systems. *arXiv preprint arXiv:2405.06333*, 2024. Under the 1st round of revision at the SIAM Journal on Scientific Computing.

[4] Xuanzhao Gao and Zecheng Gan. Broken symmetries in quasi-2D charged systems via negative dielectric confinement. *The Journal of Chemical Physics*, 161(1), 2024.

[5] Xuanzhao Gao and Zecheng Gan. Efficient particle-based simulations of Coulomb systems under dielectric nanoconfinement. 2024. Preprint available upon request.

[6] Xuanzhao Gao, Shidong Jiang, Jiuyang Liang, Zhenli Xu, and Qi Zhou. A fast spectral sum-of-Gaussians method for electrostatic summation in quasi-2D systems. *arXiv preprint arXiv:2412.04595*, 2024.

[7] Xuanzhao Gao, Xiaofeng Li, and Jinguo Liu. Programming guide for solving constraint satisfaction problems with tensor networks. *arXiv preprint arXiv:2501.00227*, 2024.

[8] Xuanzhao Gao, Yi-Jia Wang, Pan Zhang, and Jin-Guo Liu. Automated discovery of branching rules with optimal complexity for the maximum independent set problem. *arXiv preprint arXiv:2412.07685*, 2024.

[9] Leslie F. Greengard, Shidong Jiang, Manas Rachh, and Jun Wang. A new version of the adaptive fast Gauss transform for discrete and continuous sources. *SIAM Review*, 66(2):287–315, 2024.

[10] Shidong Jiang & Leslie Greengard. Approximating the Gaussian as a sum of exponentials and its applications to the fast Gauss transform. *Communications in Computational Physics*, 2022.

[11] Johan Helsing. Solving integral equations on piecewise smooth boundaries using the RCIP method: A tutorial. *Abstract and Applied Analysis*, 2013(1):938167.

[12] Shidong Jiang and Leslie Greengard. A dual-space multilevel kernel-splitting framework for discrete and continuous convolution, 2023.

[13] Jason Kaye and Hugo UR Strand. A fast time domain solver for the equilibrium Dyson equation. *Advances in Computational Mathematics*, 49(4):63, 2023.

[14] Hai-Jun Liao, Jin-Guo Liu, Lei Wang, and Tao Xiang. Differentiable programming tensor networks. *Phys. Rev. X*, 9:031041, Sep 2019.

[15] A. Raab. On the dirac–frenkel/mclachlan variational principle. *Chemical Physics Letters*, 319(5):674–678, 2000.

[16] Martin Roa-Villescas, Xuanzhao Gao, Sander Stuijk, Henk Corporaal, and Jin-Guo Liu. Probabilistic inference in the era of tensor networks and differential programming. *Physical Review Research*, 6(3):033261, 2024.

[17] Mathias Schwendt and Walter Pötz. Transparent boundary conditions for higher-order finite-difference schemes of the Schrödinger equation in (1+1)D. *Computer Physics Communications*, 250:107048, 2020.

[18] Mingyu Xiao and Hiroshi Nagamochi. Confining sets and avoiding bottleneck cases: A simple maximum independent set algorithm in degree-3 graphs. *Theoretical Computer Science*, 469:92–104, 2013.