AFRL-AFOSR-VA-TR-2023-0204

A Data-driven Approach to Correlated Quantum Many-Body Problems

Booth, George
KING'S COLLEGE LONDON
THE STRAND
LONDON, , WC2R 2LS
GBR

11/25/2022
Final Technical Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE 20221125	**2. REPORT TYPE** Final	**3. DATES COVERED** **START DATE** 20180930	**END DATE** 20210929

4. TITLE AND SUBTITLE
A Data-driven Approach to Correlated Quantum Many-Body Problems

5a. CONTRACT NUMBER	**5b. GRANT NUMBER** FA9550-18-1-0515	**5c. PROGRAM ELEMENT NUMBER** 61102F
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
George Booth

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) KING'S COLLEGE LONDON THE STRAND LONDON WC2R 2LS GBR	**8. PERFORMING ORGANIZATION REPORT NUMBER**

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/AFOSR RTA2	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)** AFRL-AFOSR-VA-TR-2023-0204

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The final report of grant FA9550-18-1-0515 details the developments and successes made in the computational challenge posed by quantum many-body problems throughout chemistry, materials science and condensed matter fields of research, as part of the AFOSR computational mathematics programme. The work focused on the development of the Gaussian Process State as a novel, data-driven approach to describing quantum many-body states, their optimization and physical understanding. It has brought together the fields of machine-learning, electronic structure and function optimization in a novel approach to enable beyond state-of-the-art calculations on a number of key correlated quantum systems.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	5

19a. NAME OF RESPONSIBLE PERSON FARIBA FAHROO	**19b. PHONE NUMBER** *(Include area code)* 426-8429

Final Performance Report (30th Sept 2018-29th April 2022)

A Data-driven Approach to Correlated Quantum Many-body Problems

Abstract

The final report of grant FA9550-18-1-0515 details the developments and successes made in the computational challenge posed by quantum many-body problems throughout chemistry, materials science and condensed matter fields of research, as part of the AFOSR computational mathematics programme. The work focused on the development of the Gaussian Process State as a novel, data-driven approach to describing quantum many-body states, their optimization and physical understanding. It has brought together the fields of machine-learning, electronic structure and function optimization in a novel approach to enable beyond state-of-the-art calculations on a number of key correlated quantum systems.

Scientific Summary and accomplishments

The broad ambition of this 3-year project was to establish a formal link between *quantum many-body physics* and *machine learning* domains. This approach was built around the Bayesian inference of a quantum state in an explicitly data-driven approach, bringing together the fields of Gaussian process regression and strongly entangled quantum states. In this endeavour, we believe we have made a step-change in the field, establishing a new paradigm in the description of quantum states as viewed through this lens. The culmination of this work has been the derivation and numerical demonstration of the flexibility and expressive power of an entirely new description of quantum states in this project, which we have termed the ***Gaussian Process State***. The novel feature of this state is that it can rigorously connect the domains of kernel learning to quantum states, and therefore has opened the door to the development of new techniques for e.g. a Bayesian interpretation for the learning of quantum systems, which we have shown has a number of benefits for both the formal expressibility of the state, and for practical learning algorithms, exposing the entanglement power of these models. We believe there is an extremely broad scope for further developments within this framework, which holds the potential for revolutionary advances in our understanding and manipulation of compact descriptions of quantum states which have previously been stymied by the exponential complexity of their information content.

This new paradigm for defining and manipulating quantum information was first presented in a publication in *Physical Review X* (impact factor 14.5) in the second year of the grant period, along with four other publications concerning the refinement of these states over the duration of the grant, with more in preparation. The approach has also caught the attention of the wider field, with a large number of invited presentations related to the work, including an invited talk at the large ‘*Psi-k*’ conference with over 1250 delegates in 2022. Furthermore, through links established within the Computational Mathematics portfolio, a productive collaboration emerged during the course of the project, which resulted in joint publications with another AFOSR-funded PI (Denys Bondar, Tulane University), combining our expertise in description of quantum states, with their AFOSR-funded work into dynamics. This led to two joint publications during this grant period, including one in *Physical Review Letters* (impact factor 9.1), which received additional popular press articles which both PIs contributed to.

Output, publications and accomplishments

This 3-year PDRA grant has seen research output of **5 published papers** (all available via open-access journals or via publicly-accessible repositories) related to this grant;

[1] *Gaussian Process States: A data-driven representation of quantum many-body physics*, Glielmo, Rath, Csanyi, De Vita, **Booth**, *Physical Review X*, **10**, 041026 (2020).

This paper established the motivation and formal description of the Gaussian process state, linking the description of a quantum system to a data-driven Bayesian framework. Furthermore, it established a numerical procedure based on machine learning principles (maximal marginal likelihood) for the 'learning' of this state which resulted in state-of-the-art descriptions of fermionic lattice models.

[2] *Driven Impostors: Controlling expectations in many-body systems*, McCaul, Orthodoxou, Jacobs, **Booth**, Bondar, *Physical Review Letters*, 124, 183201 (2020)
[3] *Controlling arbitrary observables in correlated many-body systems*, McCaul, Orthodoxou, Jacobs, **Booth**, Bondar, *Physical Review A*, 101, 053408 (2020)

This series of two papers arose from the collaboration of the PI with another PI (Denys Bondar) via AFOSR computational mathematics portfolio, where we aimed to combine the expertise in the representation of quantum states with the dynamical quantum control aspects of the work by Bondar (FA9550-16-1-0254). This work was very well received, being awarded an 'Editors Choice' in PRL, with the editor also commissioning an external 'viewpoint' companion article on the significance of the work. This further spawned a number of write-ups of the work in popular science magazines including 'Physics World', 'Nature Materials', and 'Quanta';

- **Physics**: Making Materials Mimic Each Other
- **PhysicsWorld**: Quantum control using laser light could turn insulators into conductors and vice-versa
- **US Army**: Research shows how to make lead act like gold, enabling optical computing
- Nature Materials: Masters of Disguise
- https://www.quantamagazine.org/alchemy-arrives-in-a-burst-of-laser-light-20200930/

This collaboration is continuing, with the aim to further combine this dynamical modelling with efficient representations, in order to control the high-harmonic generation of 2D correlated materials in a more applied direction, with a publication in the process of preparation.

[4] *A Bayesian inference framework for compression and prediction of quantum states*, Rath, Glielmo, **Booth**, *Journal of Chemical Physics*, **153**, 124108 (2020).

In this work, we extended the probabilistic Bayesian interpretation of quantum states, showing how *any* quantum state could be efficiently 'compressed' into our new representation of a 'Gaussian Process State'. Furthermore, the effective use of a maximal marginal likelihood estimator was described, for the selection of the data by which the state is defined (the support points of the model), inspired by probabilistic modelling in the machine-learning community. We showed how this data exhibited features of the correlated physics which emerged from the system naturally, demonstrating that not only could the state qualitatively describe the physics, but that new qualitative insights could be obtained from the trained model. This paper was an invited article towards an edition on machine learning in quantum physics.

[5] *Quantum Gaussian Process State: A kernel-inspired state with quantum support data*, Rath, **Booth**, *Physical Review Research*, **4**, 023126 (2022).

In this work, we fundamentally generalize the Gaussian Process State (GPS) construction, in order to allow the 'data' which explicitly defines the state to be unentangled product states, rather than individual 'classical' configurations. This formally enlarges the class of states which can be efficiently described, while maintaining the connection of these states to their Bayesian genesis. These states are now entirely defined via continuous valued parameters, rather than the mixed continuous-discrete optimization required for the 'classical' GPS. This hugely simplifies the optimization of these states, allowing any state to be continuously deformed into any other. Furthermore, these fundamental developments have exposed further connections of the GPS to other parametric quantum states such as neural quantum states and tensor networks, with constructive algorithms devised for transforming between these representations. This allows for improved analytic studies of the connections between these expressive forms and a future study of the entanglement scaling of these states, and their relation to the formal class of quantum states which can be efficiently represented in this form. The further connections and opportunities afforded by these connections to other quantum states established in the wider literature are still being explored.

This generalized form of the GPS was further entwined with machine learning principles, by the demonstration of a Bayesian supervised learning scheme for this 'quantum' GPS, which showed that previous difficulties with the (variational) learning of highly flexible quantum states via Monte Carlo sampling could be thought of through the lens of a 'generalization' error in the optimization. This is where a significant advantage of the 'quantum' GPS was demonstrated, that from its probabilistic interpretation and Bayesian approach, a more principled and rigorous approach to the fitting of quantum states (specifically, highly entangled quantum spin systems) could be achieved. We believe that this holds the promise for not just a formal definition of these flexible states, but robust numerical schemes for their optimization. This is an avenue we are currently still exploring.

Further to these published research outputs from this project, we are also in preparation of additional manuscripts, detailing an auto-regressive adaptation of these states for direct sampling, and a generalized linear model for the posterior distribution in the fitting procedures. It is clear that this establishment of this new framework has many opportunities for impact across multiple disciplines in both quantum physics and machine learning, and we are actively planning for its further exploitation in conjunction with AFOSR.

While the COVID-19 pandemic limited the number of conferences over the course of this project, there were still opportunities for further dissemination of the research of this project via a number of conferences and invited institutional seminars (12 which directly related to the work of this project over its funded duration). These included high-profile speaking opportunities as an invited presented at the *Psi-k* conference (August 2022), the free-electron laser institute at Hamburg (October 2021), the conference series '*Mathematical and Scientific Machine Learning*' (August 2021), and the conference '*Machine Learning in Quantum Many-Body Systems*' at the CCQ Flatiron institute in New York, amongst others.

Impacts

The work of this project clearly cuts across traditional scientific disciplines, and has therefore had a broad impact on a number of fields. Its principal discipline is in the efficient computational description of quantum many-body states from a very fundamental perspective. This impact on the field of quantum chemistry, materials science and nuclear matter. We anticipate future works looking to consider the use of the GPS as a 'solver' within larger multi-scale simulations in these fields across larger lengthscales, or as a description of the system in its own right for frustrated or strongly entangled systems. The work also impacts on the field of machine learning. From the beginning of the project where the GPS was considered as a simple kernel model, the evolution of the model at its core has impacted upon how traditional kernel models can be viewed for a host of other supervised learning or generative modelling applications, including image recognition. As the exponential of a CP factorization, the final evolution of the model in this project brought connections to a new form of low-rank tensor factorization too from the perspective of low-rank linear algebra and optimization theory, combining a Bayesian approach with alternative least squares method of optimization.

Changes to staffing and COVID-19 disruption

The original broad scientific ambitions of the project outlined in the proposal were fully met or surpassed with the developments in the Gaussian Process State and its exploitation as outlined above. However, there has been disruption to the planned staffing and implementation of the project from a management perspective compared to the one originally envisaged, primarily due to the COVID-19 pandemic. These include disruption to recruitment, staffing, access to required HPC as well as opportunities for dissemination of the research. Deviations from planned staffing have been discussed and approved at the point they were made with the programme officer, and the best approach to minimize disruption to the scientific ambitions were found. These are also discussed individually in the annual reporting of the grant. Most pertinently, the project sought (and was granted) a 7-month no-cost extension to the project duration, which enabled the original ambitions of the project to be achieved. Overall, we believe that an effective and justified approach was taken in the project management to ensure success of the project objectives. No significant deviations from the original project direction and ambitions were changed in this timescale and staffing adjustments.