

Gavin S. Hartnett

THEORETICAL PHYSICIST & MACHINE LEARNING RESEARCHER

 email  Google Scholar  Github  LinkedIn

Professional Experience

Q-CTRL

LEAD SCIENTIST

Santa Monica, CA

June 2022 - present

- Lead of the Applications and Algorithms team
- Co-lead a project aimed at experimentally demonstrating large-scale GHZ state generation
- Developed a machine learning approach to rank circuit performance
- Supported the development of a linear optics circuit synthesis code base
- Supported the development of a Quantum Approximate Optimization pipeline

The RAND Corporation

INFORMATION SCIENTIST

Santa Monica, CA

Aug. 2017 - May 2022

- AI/ML Lablet Lead at the [Tech and Narrative Lab](#)
- Organizer of company-wide AI seminar series and study group
- Co-PI for project on generative modeling for networks
- Investigated the vulnerability of autonomous agents to adversarial examples
- Investigated domain adaptation for object detection using synthetic data sets
- Tech lead for a project investigating how COVID-19 spreads across real-world contact networks

Large Language Model Red-Teaming

INDEPENDENT CONSULTING

Santa Monica, CA

Sept. 2022 - Oct. 2023

- Worked with Meta to red-team their language models
- Worked with OpenAI to red-team [GPT-4](#)

University of Southampton

POSTDOCTORAL RESEARCH FELLOW

Southampton, United Kingdom

Sept. 2015 - Aug. 2017

- Studied theoretical properties of black holes and quantum field theories
- Worked on multiple projects as part of an international collaboration
- Co-organized 3 seminar series
- Traveled extensively to present research and facilitate collaborations

Education

University of California, Santa Barbara

PHD AND MA IN PHYSICS

Santa Barbara, CA

Aug. 2009 - May 2015

- Adviser: Prof. Gary Horowitz
- Dissertation: [Aspects of Black Holes in Higher Dimensions](#)

Syracuse University

BSC IN PHYSICS AND MATHEMATICS

Syracuse, NY

Sept. 2005 - May 2009

- Summa Cum Laude
- Honors Thesis: [Spiral Patterns in Liquid Crystals](#)

Technical Publications

• Learning to rank quantum circuits for hardware-optimized performance enhancement

G. S. Hartnett, A. Barbosa, P. S. Mundada, M. Hush, M. J. Biercuk, Y. Baum.

[arXiv:2404.06535 \[quant.ph\]](#)

Preprint

• Twisty-puzzle-inspired approach to Clifford synthesis

N. Bao, G. S. Hartnett.

[arXiv:2307.08684 \[quant.ph\]](#)

Phys. Rev. A 109, issue 3, 032409 (2024)

- **The Karzas-Latter-Seiler Model of a High-Altitude Electromagnetic Pulse: A New Numerical Code for an Old Model**
G. S. Hartnett.
[arXiv:2402.14864\[physics.plasm-ph\]](#)
Preprint
- **The hierarchical parity model**
G. S. Hartnett.
[arXiv:2208.13316 \[cond-mat.dis-nn\]](#)
Physica A: Statistical Mechanics and its Applications 617 (2023): 128679.
- **Modeling the Impact of Social Distancing and Targeted Vaccination on the Spread of COVID-19 through a Real City-Scale Contact Network**
G. S. Hartnett, E. Parker, T. R. Gulden, R. Vardavas, D. Kravitz.
[arXiv:2107.06213 \[physics.soc-ph\]](#)
Journal of Complex Networks 9.6 (2021): cnab042.
- **Protecting the Most Vulnerable by Vaccinating the Most Active**
T. R. Gulden, G. S. Hartnett, R. Vardavas, D. Kravitz.
RAND Perspective PE-A1068-1
- **Deep Generative Modeling in Network Science with Applications to Public Policy Research**
G. S. Hartnett, R. Vardavas, L. Baker, M. Chaykowsky, C. B. Gibson, F. Giroi, D. P. Kenedy, O. A. Osoba.
[arXiv:2010.07870 \[cs.LG\]](#)
RAND Working Paper WRA843-1
- **Self-Supervised Learning of Generative Spin-Glasses with Normalizing Flows**
G. S. Hartnett, M. Mohseni.
[arXiv:2001.00585 \[cs.LG\]](#)
Preprint
- **A Probability Density Theory for Spin-Glass Systems**
G. S. Hartnett, M. Mohseni.
[arXiv:2001.00927 \[cond-mat.dis-nn\]](#)
Preprint
- **Operationally Relevant Artificial Training for Machine Learning: Improving the Performance of Automated Target Recognition Systems**
G. S. Hartnett, L. Menche, J. Léveillé, D. Baveye, L. Zhang, D. Gold, J. Hagen, J. Xu.
RAND Report RRA683-1 (2020)
- **Covariant Noether charges for type IIB and 11-dimensional supergravities**
O. J. C. Dias, G. S. Hartnett, J. E. Santos.
[arXiv:1912.01030 \[hep-th\]](#)
Class. Quant. Grav. 31, no. 1, 015003 (2021)
- **Adversarial Examples for Cost-Sensitive Classifiers**
G. S. Hartnett, A. J. Lohn, A. P. Sedlack.
[arXiv:1910.02095 \[stat-ML\]](#)
Workshop on Safety and Robustness in Decision Making, NeurIPS 2019
- **Holographic dual of hot Polchinski-Strassler quark-gluon plasma**
I. Bena, O. J. C. Dias, G. S. Hartnett, Benjamin. E. Niehoff, J. E. Santos.
[arXiv:1805.06463 \[hep-th\]](#)
JHEP 9, 33 2019
- **Replica Symmetry Breaking in Bipartite Spin Glasses and Neural Networks**
G. S. Hartnett, E. Parker, E. Geist.
[arXiv:1803.06442 \[cond-mat.dis-nn; cs.LG\]](#)
Phys. Rev. E 98, issue 2, 022116 (2018)
- **Constraining the mass of dark photons and axion-like particles through black-hole superradiance**
V. Cardoso, O. J. C. Dias, G. S. Hartnett, M. Middleton, P. Pani, J. E. Santos.
[arXiv:1801.01420 \[gr-qc\]](#)
JCAP 1803, no.03, 043 (2018)

- **Mass-deformed M2 branes in Stenzel space**
O. J. C. Dias, G. S. Hartnett, B. E. Niehoff, J. E. Santos
[arXiv:1704.02323 \[hep-th\]](#)
JHEP 1711, 105 (2017)
- **Localised Anti-Branes in Flux Backgrounds**
G. S. Hartnett.
[arXiv:1501.06568 \[hep-th\]](#)
JHEP 1506, 007 (2015)
- **A No Black Hole Theorem**
G. S. Hartnett, G. T. Horowitz and K. Maeda.
[arXiv:1410.1875 \[hep-th\]](#)
Class. Quant. Grav. 32, no. 5, 055011 (2015)
- **Quasinormal modes of asymptotically flat rotating black holes**
O. J. C. Dias, G. S. Hartnett and J. E. Santos.
[arXiv:1402.7047 \[hep-th\]](#)
Class. Quant. Grav. 31, no. 24, 245011 (2014)
- **Holographic thermalization, quasinormal modes and superradiance in Kerr-AdS**
V. Cardoso, O. J. C. Dias, G. S. Hartnett, L. Lehner and J. E. Santos.
[arXiv:1312.5323 \[hep-th\]](#)
JHEP 1404, 183 (2014)
- **Non-Axisymmetric Instability of Rotating Black Holes in Higher Dimensions**
G. S. Hartnett and J. E. Santos.
[arXiv:1306.4318 \[gr-qc\]](#)
Phys. Rev. D 88, 041505 (2013)
- **Geons and Spin-2 Condensates in the AdS Soliton**
G. S. Hartnett and G. T. Horowitz
[arXiv:1210.1606 \[hep-th\]](#)
JHEP 1301, 010 (2013)

Policy Publications

- **Cybersecurity and Supply Chain Risk Management Are Not Simply Additive**
V. A. Greenfield, J. W. Welburn, K. Schwindt, D. Ish, A. J. Lohn, G. S. Hartnett
RAND Report RR-A532-1 (2023)
- **Operational Feasibility of Adversarial Attacks Against Artificial Intelligence**
L. A. Zhang, G. S. Hartnett, J. Aguirre, A. J. Lohn, I. Khan, M. Herron, and C. O'Connell
RAND Report RR-A866-1 (2022)
- **Empirical Evaluation of Physical Adversarial Patch Attacks Against Overhead Object Detection Models**
G. S. Hartnett, L. Zhang, C. O'Connell, A. J. Lohn, J. Aguirre
[arXiv:2206.12725](#)
- **Airline Security Through Artificial Intelligence**
S. McKay, G. S. Hartnett, B. Held
RAND Report PEA731-1
- **Maintaining the Competitive Advantage in Artificial Intelligence and Machine Learning**
R. Waltzman, L. Ablon, C. Curriden, G. Hartnett, M. Holliday, L. Ma, B. Nichiporuk, A. Scobell, D. Tarraf
RAND Report RRA200

Teaching

Pardee RAND Graduate School

CORE FACULTY MEMBER/PROFESSOR

- Introduction to Modern AI
- Introduction to Blockchain Technology

Santa Monica, CA

2018-2022

University of Southampton

LECTURER

- MATH1052 Differential Equations
- MATH1008 Mathematical Methods
- MATH3071 Light and Waves

Southampton, UK

Sept. 2015 - May 2015

University of California, Santa Barbara

HEAD TEACHING ASSISTANT

- Managed team of 40+ TA's for the entire Physics Department
- Worked with faculty and staff to assign TA's to courses

Santa Barbara, CA

Aug. 2010 - Aug. 2012

TEACHING ASSISTANT

- PHYS6L Introductory Physics (3 quarters)
- PHYS21 General Physics
- PHYS105 Classical Mechanics
- PHYS115 Quantum Mechanics (2 quarters)
- PHYS219 Statistical Mechanics (graduate level)

Sept. 2009 - May 2015

Professional Activities

FOUNDER AND ORGANIZER OF AI SEMINAR SERIES AT THE RAND CORPORATION

2018 - 2022

ORGANIZER OF GRADUATE STUDENT HIGH-ENERGY JOURNAL CLUB

2012 - 2014

REFEREE FOR

- *Ethics Reviewer for NeurIPS 2021*
- *ACM Conference on Fairness, Accountability, and Transparency (FAccT) 2020*
- *NeurIPS 2019 Workshop: Machine Learning and the Physical Sciences*
- *Journal of High Energy Physics (JHEP)*
- *Physical Letters B*
- *Classical and Quantum Gravity*
- *General Relativity and Gravitation*

Awards

2021	RAND Spotlight Award , awarded for a study assessing how AI could be used to improve the TSA baggage screening process	Santa Monica, CA
2020	RAND Bronze Medal Award , company-wide annual award, awarded for "vision, integrity, and leadership" in the course of a project on adversarial machine learning for cyber defense systems.	Santa Monica, CA
2019	RAND Spotlight Award , awarded for "developing a new game theoretic approach with Machine Learning techniques to assess cyber defense capabilities."	Santa Monica, CA
2019	RAND Project Air Force Team Innovation Award , awarded for our team's "high-risk/high-reward approach to solving a complex technical problem – understanding how machine learning-based algorithms might be vulnerable to cyber attack"	Santa Monica, CA
2014	Dean's Fellowship , Competitive University-wide fellowship	Santa Barbara, CA
2013	James Hartle Award , Best graduate student talk	Warsaw, Poland
2011	Chairs Certificate of Appreciation , Outstanding service as Head TA	Santa Barbara, CA
2009	Syracuse University Scholar , Highest undergraduate academic honor	Syracuse, NY
2008	Barry Goldwater Scholarship , Most prestigious undergraduate national science award	Syracuse, NY