My research takes two main directions: 1) I research the *theoretical foundations of computation* and what it means for computational problems to be inherently *complex* and how that complexity can be utilized. My research connects disparate areas of Complexity Theory to each other and explores when computational difficulty can be leveraged to attain Cryptography guarantees. 2) I research the *social impact of computation and its research* has in the real-world and which communities stand to be most affected and how to include their voices in the field. My research here is inherently interdisciplinary and focuses on building theoretical infrastructure and language to serve as interfaces for disparate disciplines to meet at more holistic understandings of Privacy and the young field of Algorithmic Fairness.

Foundations: Complexity Theory and Cryptography

Understanding the foundations of Cryptography is understanding the foundations of Hardness. Complexity Theory studies the nature of computational hardness – i.e. lower bounds on the time required in solving computational problems – not only to understand what is difficult, but to understand the utility hardness offers. By hiding secrets within 'structured hardness,' Complexity Theory, only as recently as the 1970s, transformed the ancient ad-hoc field of Cryptography into a science with rigorous theoretical foundations. But *which* computational hardness can we feel comfortable basing Cryptography on?

My research studies a question foundational to Complexity Theory and Cryptography:

Does Cryptography's existence follow from $\mathsf{P} \neq \mathsf{NP}$?

P is the class of problems computable in polynomial time and is qualitatively considered the set of 'easy' problems, while NP is the class of problems whose solutions, when obtained, can be *checked* as correct in polynomial time. NP thus captures the computational problems that could be considered of practical interest, lest we wouldn't even know a correct solution when saw it. Since it seems that it should sometimes be harder to *find* solutions than to *verify* them, it is widely believed that $\mathsf{P} \neq \mathsf{NP}$ and resolving this is the largest open problem in Computer Science and one of the seven Clay Millennium "million-dollar" problems in mathematics. Further, $\mathsf{P} \neq \mathsf{NP}$ is the weakest possible assumption on which Cryptography might follow from, intuitively since you must be able to *check* that you have correctly decoded an encrypted message.

While most of Cryptography is based on the assumptions of the hardness of *specific* problems, basing Cryptography on $\mathsf{P} \neq \mathsf{NP}$ is no longer cherry-picked but instead achieves a structural theorem relating the the existence of Cryptography to the hardness of a natural *class* of problems. This would show that NP's problems lack structure that can be exploited for attaining fast algorithms yet retain enough structure to accomplish Cryptography, giving us a window into the nature and flexibility of the type of hardness within one of the most natural and studied complexity classes. My work aims to make progress towards, define variants of, and show barriers to basing Cryptography on $\mathsf{P} \neq \mathsf{NP}$.

Past Work. My work uses the nascent field of Fine-Grained Complexity Theory to open new directions on this long-studied foundational question. By studying "Hardness within P" (see [31, 32]) and the connections of problems computable in, say, n^2 time versus n^3 time, Fine-Grained Complexity addresses the *practical* efficiency of problems. However, while this more *quantitative* approach addresses practical hardness problem-by-problem, we lose connection to key *qualitative* claims of classical Complexity Theory such as the general ability to hide secrets (Cryptography), the ability to show that a world where we have access to randomness is no more powerful computationally than a deterministic one (Derandomization), and the ability to transform a problem that is hard in the worst-case scenario into one that is hard almost always (Hardness Amplification).

My work shows that these connections can be recovered and that the problem-specific claims of Fine-Grained Complexity Theory cannot exist in a vacuum without ramifications to structural Complexity Theory.

• **Hardness Amplification:** My work [3] shows how to take core Fine-Grained problems believed to be hard to compute in the worst-case and, by interpolating the problem's input-output relation with a low-degree multilinear extension over a finite field, create a new function that is now moderately hard *on all but a small fraction* of the problem's instances. This sort of 'average-case hardness' is *necessary* for Cryptography since Cryptographic objects should be secure almost always and not only if the adversary is unlucky and receives a worst-case instance. Our low-degree multilinear framework has spawned many works in the average-case hardness of fine-grained problems and in delegating computation – e.g. [1, 13, 14, 15, 16]. By achieving average-case hardness within the Fine-Grained world, we use this as a stepping stone to achieve both Cryptographic primitives and Derandomization.

• **Cryptography:** Introduced in [11] to combat spam and now serving as the heart of many cryptocurrencies, Proofs of Work (PoWs) ask a fundamental question about hardness: Can you *prove* that you expended a

certain amount of computational work? We obtain the first PoWs from worst-case complexity assumptions, thus finally shifting them from a heuristic idea to a mathematically grounded Complexity Theoretic object. [4]. Most importantly, we show that these PoWs follow from strengthenings of $\mathsf{P} \neq \mathsf{NP}$, including from the Strong Exponential Time Hypothesis (SETH): Whereas $\mathsf{P} \neq \mathsf{NP}$ is know to be equivalent to a specific problem in NP requiring more than polynomial time [10, 24], SETH states that it requires *exponential* time. Thus, our work not only gives the first PoWs from worst-case assumptions, we give the first Cryptographic objects that follow from a worst-case assumption that is a natural strengthening of $\mathsf{P} \neq \mathsf{NP}$.

• **Derandomization:** We achieve complexity-theoretic Pseudorandom Generators (PRGs) from the core assumptions of Fine-Grained Complexity [9]. These objects are sufficient to *derandomize* algorithms [26], thus showing how a deterministic world may be just as computationally powerful as one where we have free access to true randomness. Our results not only connect the resource-centric coarse-grained study of derandomization to the problem-centric hardness of Fine-Grained Complexity, but improves over all previous derandomizations shown from algorithmic hardness assumptions [7, 21, 30]. Further, complexity-theoretic PRGs are weakened versions of Cryptographic PRGs, which are *equivalent* to attaining secret-key Cryptography. Thus, this can be viewed as showing that from Fine-Grained assumptions including from the natural strengthening of $\mathsf{P} \neq \mathsf{NP}$, SETH, we get progress towards secret-key Cryptography.

Continuing and Future Work. My continuing work in Complexity progresses towards, defines variants of, and shows barriers to and consequences of basing Cryptography on $\mathsf{P} \neq \mathsf{NP}$.

• **Fine-Grained Cryptography:** Beyond the PoWs we achieve in [4], common Cryptographic abilities such as encrypting messages may also be desired. Towards this end, my work [3] introduced the concept of Fine-Grained Cryptography which should be *moderately* secure (hard for n^k time adversaries for some k instead of being super-polynomially hard) and gives definitions of the core objects needed to begin realizing it. This has already inspired the works of [23] and [8] which make progress towards Fine-Grained public-key Cryptography and Fine-Grained secure computation, respectively. Continuing this line of progress is promising since Fine-Grained Cryptography may be a form of Cryptography that *is* achievable from Fine-Grained assumptions and SETH even if traditional Cryptography isn't!

• **Barriers:** One of the trademark qualities of Complexity Theory is its abundance of impossibility and barrier results, proving that certain proof techniques are impossible to achieve a desired theorem or would have to first prove results we either don't believe or seem far beyond our current reach. In particular, there is a long line of research showing barriers to achieving Cryptography from $\mathsf{P} \neq \mathsf{NP}$ – e.g. [2, 6, 12].

The 'flavor' of these barriers, however, are all of the form that a *specific standard set of proof techinques* achieving the goal would yield consequences that we *don't* believe are true. I am in ongoing work to instead show that *any* proof of "$\mathsf{P} \neq \mathsf{NP} \Rightarrow$ Cryptographic PRGs" would yield breakthrough circuit lower bounds that have been conjectured *true* for decades, thus making our barriers closer in 'flavor' to those in [22].

Social Impact: Algorithmic Fairness and Privacy

As algorithmic decision-making becomes increasingly ubiquitous, structuring society and daily life, it has grown unavoidably clear that it can perpetuate harms at new and terrifying speeds. Academia has responded to this: The nascent field Algorithmic Fairness has quickly built an array of powerful mathematical tools with insightful proofs to yield less biased algorithms that are quickly being adopted and deployed by companies. However, there is growing concern, especially across disciplines in this inherently interdisciplinary field, that even algorithms attempting "Fairness" codify and legitimize the systems using them and that mathematical stamps of approval of "Fair" in one real-world context may yield very inequitable results in another. My work aims to address the areas of Theoretical CS that are inherently interdisciplinary and provide theoretical infrastructure and language to serve as interfaces for disparate disciplines to engage with in ways that can address cross-disciplinary critiques, both in the young field of Algorithmic Fairness and in Privacy.

Past Work. My entrance into Algorithmic Fairness was *both* as a Computer Scientist and as a member of the Queer, Trans, and Person of Color (POC) community. From this intersection, I had reservations that an abstract notion of a "Fair" algorithm would not be portable to multiple social contexts and runs the risk of harming my communities while still allowing companies a "Fairness" stamp of approval. I entered the field to make sure my community and their concerns and values were reflected.

After participating heavily in the Simons Institute at UC Berkeley's 2019 Summer program on Algorithmic Fairness, I found that many of my reservations were directly reflected in numerous recent critiques – e.g.

[5, 18, 19, 20, 29]. I thus collaborated with an interdisciplinary group of graduate students from Computer Science, Medicine, and Philosophy across multiple universities to create and organize a workshop for FAccT* 2020 [17]. The workshop's audience of CS Fairness researchers, troubled by the emerging critiques but unsure how to begin addressing them within their discipline, were guided through discussions and activities towards incorporating praxis-centered methods and principles from the area of Community Organizing.

Since this workshop I have engaged deeply with the CS community and across disciplines on these issues and the social impact of CS research more generally: I co-organized the Resistance AI workshop at NeurIPS 2020, focusing on discussing how AI shifts power in the world and centering a large line-up of Black and Indigenous activists, researchers, and organizers to present and discuss how to shift power back to marginalized communities. I served as discussant with keynote speaker Stephanie Dinkins at the 2020 Mechanism Design for Social Good workshop. And I was a panelist with Kade Crockford and Alex Hanna on AI policy and privacy in the Queer in AI workshop at ICML 2020.

In mid-2020 I joined a postdoc at the ERC-funded COHUBICOL project working with lawyers and legal philosophers to construct theoretical foundations for how Machine Learning (ML) and smart contracts were altering both how Law is practiced and its nature and definitions. I remotely gave talks and joined discussions in workshops with this group and helped create a cross-disciplinary vocabulary between ML and Law for six months before COVID-19 made continuing this Netherlands position untenable.

In 2021 I gave a talk at Aalborg University's SECURE workshop that gathered researchers across disciplines to explore the real-world impact of crypto-systems and find shared vocabulary. At CVPR 2021 I served on the program committee of Emily Denton and Timnit Gebru's *Beyond Fairness* workshop.

Continuing and Future Work.

• **DLA Framework:** Using the insights gained from co-organizing our FAccT* 2020 workshop, we are writing up the Discriminatory and Liberatory Algorithms (DLA) framework to address concerns within Algorithmic Fairness. This framework is a reframing and contextualizing of Fairness questions by both delineating different objectives within the field as well as attaching methodologies and domain boundaries to its categories. This allows the DLA framework to serve as an interface for disparate disciplines to engage with different aspects of Fairness questions. Thus, our work is spiritually similar to Nissenbaum's defining of Contextual Integrity [27] as a way of separating out technical from contextual aspects of Privacy so that multiple disciplines could interface with Privacy questions cohesively.

The DLA framework most directly addresses the concerns in [29], which outlines a series of socially harmful "traps" the field of Fairness seems to routinely fall into. We show that our framework doesn't just categorize Fairness work but instead, by design, avoids these traps by giving an edifice for different disciplines to engage with "Fairness" where their field's expertise and notions rigor are needed and relevant. Thus, our framework aims at the same goal of Fairness in producing a more just world with respect to algorithm usage while being much less prone to these traps. I have widely given well-received talks on this framework to Fairness groups at Boston University, MIT, the MD4SG community, UC San Diego, and others.

Crypto-Anthropology. After being invited to speak at social anthropologist Adrienne Mannov's cross-disciplinary SECURE workshop on how Privacy is approached across disciplines, I am now working with Mannov to explore how Theoretical Cryptography in done in the real-world with respect to how other disciplines address the same questions. Mannov has previously called for a Crypto-Anthropology [25] which we are now pursuing, thus extending the work of [28] into an Anthropology lens. By studying the ways in which core Cryptographic primitives yield hierarchies and values akin to [33], we can hope to bridge language and concerns across disciplines and work towards a more holistic understanding of Privacy.

Conclusion

In Complexity Theory, my longterm goal is to understand the foundations of Cryptography and its relationship to NP. While ambitious, I have already begun novel progress on this front and the emergence of Fine-Grained Complexity seems to be shedding new light onto this old and studied question.

My goals in Fairness and Privacy might be yet more ambitious: I want to build interdisciplinary bridges and interfaces that allow disparate disciplines to meet at a more holistic study of these fields and for researcher demographics to begin to reflect the communities that stand to be most affected by the algorithms in question. My workshop organizing across disciplines and communities along with the cross-disciplinary DLA framework take important steps in this direction.

References

[1] Enric Boix Adserà, Matthew Brennan, and Guy Bresler. The average-case complexity of counting cliques in erdős-rényi hypergraphs. *CoRR*, abs/1903.08247, 2019.

[2] Adi Akavia, Oded Goldreich, Shafi Goldwasser, and Dana Moshkovitz. On basing one-way functions on np-hardness. In Jon M. Kleinberg, editor, *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*, pages 701–710. ACM, 2006.

[3] Marshall Ball, Alon Rosen, Manuel Sabin, and Prashant Nalini Vasudevan. Average-case fine-grained hardness. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 483–496. ACM, 2017.

[4] Marshall Ball, Alon Rosen, Manuel Sabin, and Prashant Nalini Vasudevan. Proofs of work from worst-case assumptions. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology - CRYPTO 2018 - 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2018, Proceedings, Part I*, volume 10991 of *Lecture Notes in Computer Science*, pages 789–819. Springer, 2018.

[5] Cynthia L. Bennett and Os Keyes. What is the point of fairness? disability, ai and the complexity of justice, 2019.

[6] Andrej Bogdanov and Luca Trevisan. On worst-case to average-case reductions for NP problems. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings*, pages 308–317. IEEE Computer Society, 2003.

[7] Jin-yi Cai, Ajay Nerurkar, and D. Sivakumar. Hardness and hierarchy theorems for probabilistic quasi-polynomial time. In Jeffrey Scott Vitter, Lawrence L. Larmore, and Frank Thomson Leighton, editors, *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing, May 1-4, 1999, Atlanta, Georgia, USA*, pages 726–735. ACM, 1999.

[8] Matteo Campanelli and Rosario Gennaro. Fine-grained secure computation. In Amos Beimel and Stefan Dziembowski, editors, *Theory of Cryptography - 16th International Conference, TCC 2018, Panaji, India, November 11-14, 2018, Proceedings, Part II*, volume 11240 of *Lecture Notes in Computer Science*, pages 66–97. Springer, 2018.

[9] Marco L. Carmosino, Russell Impagliazzo, and Manuel Sabin. Fine-grained derandomization: From problem-centric to resource-centric complexity. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPIcs*, pages 27:1–27:16. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018.

[10] Stephen A. Cook. The complexity of theorem-proving procedures. In Michael A. Harrison, Ranan B. Banerji, and Jeffrey D. Ullman, editors, *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing, May 3-5, 1971, Shaker Heights, Ohio, USA*, pages 151–158. ACM, 1971.

[11] Cynthia Dwork and Moni Naor. Pricing via processing or combatting junk mail. In Ernest F. Brickell, editor, *Advances in Cryptology - CRYPTO '92, 12th Annual International Cryptology Conference, Santa Barbara, California, USA, August 16-20, 1992, Proceedings*, volume 740 of *Lecture Notes in Computer Science*, pages 139–147. Springer, 1992.

[12] Joan Feigenbaum and Lance Fortnow. On the random-self-reducibility of complete sets. In *Proceedings of the Sixth Annual Structure in Complexity Theory Conference, Chicago, Illinois, USA, June 30 - July 3, 1991*, pages 124–132. IEEE Computer Society, 1991.

[13] Oded Goldreich. On doubly-efficient interactive proof systems. *Foundations and Trends in Theoretical Computer Science*, 13(3):158–246, 2018.

[14] Oded Goldreich and Guy N. Rothblum. Worst-case to average-case reductions for subclasses of P. *Electronic Colloquium on Computational Complexity (ECCC)*, 24:130, 2017.

[15] Oded Goldreich and Guy N. Rothblum. Counting t-cliques: Worst-case to average-case reductions and direct interactive proof systems. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:46, 2018.

[16] Oded Goldreich and Guy N. Rothblum. Counting t-cliques: Worst-case to average-case reductions and direct interactive proof systems. In Mikkel Thorup, editor, *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 77–88. IEEE Computer Society, 2018.

[17] Ezra Goss, Lily Hu, Manuel Sabin, and Stephanie Teeple. Manifesting the sociotechnical: Experimenting with methods for social context and social justice. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20, page 693, New York, NY, USA, 2020. Association for Computing Machinery.

[18] Ben Green. Data science as political action: Grounding data science in a politics of justice, 2019.

[19] Ben Green and Lily Hu. The myth in the methodology: Towards a recontextualization of fairness in machine learning. In *Machine Learning: The Debates workshop at the 35th International Conference on Machine Learning (ICML)*, Stockholm, Sweden, 2018.

[20] Anna Lauren Hoffmann. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, 22(7):900–915, 2019.

[21] Russell Impagliazzo and Avi Wigderson. Randomness vs. time: De-randomization under a uniform assumption. In *39th Annual Symposium on Foundations of Computer Science, FOCS '98, November 8-11, 1998, Palo Alto, California, USA*, pages 734–743. IEEE Computer Society, 1998.

[22] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. In Lawrence L. Larmore and Michel X. Goemans, editors, *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, pages 355–364. ACM, 2003.

[23] Rio LaVigne, Andrea Lincoln, and Virginia Vassilevska Williams. Public-key cryptography in the fine-grained setting. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology - CRYPTO 2019 - 39th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2019, Proceedings, Part III*, volume 11694 of *Lecture Notes in Computer Science*, pages 605–635. Springer, 2019.

[24] L. A. Levin. Universal sequential search problems. *Problems of Information Transmission*, 9(3):265–266, 1973.

[25] Adrienne Mannov. A call for crypto-anthropology. MIT Cryptography and Information Security (CIS) Seminar 2019, October 2019. https://www.csail.mit.edu/event/adrienne-mannov-call-crypto-anthropology.

[26] Noam Nisan and Avi Wigderson. Hardness vs. randomness (extended abstract). In *29th Annual Symposium on Foundations of Computer Science, White Plains, New York, USA, 24-26 October 1988*, pages 2–11. IEEE Computer Society, 1988.

[27] Helen Nissenbaum. *Privacy in Context - Technology, Policy, and the Integrity of Social Life*. Stanford University Press, 2010.

[28] Phillip Rogaway. The moral character of cryptographic work. *IACR Cryptology ePrint Archive*, 2015:1162, 2015.

[29] Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. Fairness and abstraction in sociotechnical systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency, FAT* 2019, Atlanta, GA, USA, January 29-31, 2019*, pages 59–68. ACM, 2019.

[30] Luca Trevisan and Salil P. Vadhan. Pseudorandomness and average-case complexity via uniform reductions. In *Proceedings of the 17th Annual IEEE Conference on Computational Complexity, Montréal, Québec, Canada, May 21-24, 2002*, pages 129–138. IEEE Computer Society, 2002.

[31] Virginia Vassilevska Williams. Hardness of easy problems: Basing hardness on popular conjectures such as the strong exponential time hypothesis (invited talk). In Thore Husfeldt and Iyad A. Kanj, editors, *10th International Symposium on Parameterized and Exact Computation, IPEC 2015, September 16-18, 2015, Patras, Greece*, volume 43 of *LIPIcs*, pages 17–29. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2015.

[32] Virginia Vassilevska Williams. On some fine-grained questions in algorithms and complexity. In *Proceedings of the ICM*, 2018.

[33] Langdon Winner. Do artifacts have politics? *Daedalus*, 109(1):121–136, 1980.