As a military child and the oldest of six children, I profoundly identify with the values of hard work, service, and sacrifice. Snapshots from my childhood—my father missing my sister's birth while serving on submarine duty, tutoring mathematics while waiting hours to be picked up from school, missing sports competitions to take care of my siblings, doing grocery runs as soon as I could drive, and my parents' eventual messy divorce—influence who I am today. I have learned firsthand that success is determined by perseverance and collaboration rather than natural talent. I care deeply about giving back to my family, my community, and my country. I hope to embody these values throughout my career, leading by example as a researcher focused on building an inclusive and interdisciplinary scientific community.

After moving constantly as a kid, I didn't truly find my home until I was recruited to the University of Southern California. I was awarded both the merit-based, **full-tuition** Trustee Scholarship and the Viterbi Fellowship, a grant worth **$24,000 in research funding** across four years. Despite my high school's lack of resources for advanced opportunities, I immediately began to flourish in the unstructured collegiate environment. I became a leader in a variety of exceptional research projects and service organizations, resulting in a number of submitted publications as well as measurable impact on my community.

I have had the privilege to work in **interdisciplinary research laboratories across academia, industry, and government**, where I learned that the collaboration of diverse researchers from varied facets of the scientific community is a critical ingredient for innovative research. These impactful and productive research experiences motivated me to pursue a **Ph.D. in theoretical machine learning**. My goal in graduate school is to continue this cross-domain collaboration, reconciling theoretical understanding with empirical phenomena in deep learning by performing mathematical study and experimental analysis in tandem.

I am particularly interested in **optimization and generalization in deep learning**, wherein theoretical advancements have enormous potential to increase neural networks' efficiency, accuracy, and robustness. I am also passionate about **fairness in machine learning**, where I can combine my theoretical and practical expertise to ensure that deep models remain fair and equitable when deployed at scale.

<u>Intellectual Merit:</u> My undergraduate research is characterized by both breadth and depth. I have collaborated with professors, industry researchers, and government scientists in Bayesian learning, theoretical machine learning, computer vision, and convex optimization. I have also taken four Ph.D. level computer science courses to prepare me for rigorous interdisciplinary research. Throughout my undergrad, I have focused on simultaneously advancing the technical state-of-the-art in my field and engineering measurable impact for users and scientists in other domains. My research has resulted in **two submitted publications** with a third in preparation; additionally, I have twice transitioned my code to production as an integral component of impactful projects in materials science and computational agriculture.

<u>Broader Impacts:</u> My teaching and outreach initiatives are centered around (1) increasing the diversity and inclusiveness of STEM opportunities in K-12 education, and (2) involving practitioners to promote machine learning fairness. The first was inspired by my sister, who was discouraged by teachers and peers from pursuing a STEM career but persevered to join me at USC as a Biomedical Engineering major; the second, by my hackathon project and essay which won Top 20 Ethical Hacking Finalist out of 4500 teams. My outreach has resulted in the creation and funding of robotics programs at five schools in underprivileged neighborhoods of L.A., **creating STEM opportunities for over 200 minority students**. And, my fairness projects have mobilized a dozen USC students in analyzing and rectifying biases in deep learning models, resulting in two presentations and an open-source product in development.

<u>Research Background:</u> My first major research experience was my machine learning research internship at Sandia National Laboratories. I designed Bayesian convolutional neural networks (BCNN) for uncertainty quantification of CT scan segmentations of materials such as thermal protection systems for spacecraft reentry. ***Intellectual Merit:*** While previous work considered Bayesian neural networks computationally infeasible, I independently developed a novel BCNN architecture which scales to billion-voxel 3D volumes. My BCNN outperforms current methods by up to 40% in recent uncertainty metrics and generates statistically credible geometric uncertainty maps which can be smoothly probed during simulation. My internship culminated in a poster presentation at the Sandia Math and Analytics Research Symposium as one of three undergraduates among 25 Ph.D. students, where I was awarded the SIMLR Award for Outstanding Intern. My **first-author publication** on the novel technical merit of my BCNN is under submission to WACV 2021, and my **second-author journal publication** on my BCNN's interdisciplinary impact is under submission to Nature Communications. My first-author publication has already been cited multiple times, by follow-up work as well as researchers working on Graph CNNs for organ segmentation. I collaborated with a team of materials scientists, statisticians, and software engineers to integrate my BCNN with Sandia's supercomputers, then open-sourced my codebase on GitHub where it is now the 6th most starred Sandia repository (out of 69). Sandia hired an additional intern and a full-time researcher to continue my work in 2020-21, accelerating the interdisciplinary application of my research to segmentations of graphite electrodes, laser-welded metal joints, and, most recently, medical CT imaging. ***Broader Impacts:*** My research enables materials scientists to ascertain bounds on physical properties (*e.g.*, porosity and conductance) of materials during computational simulations. This innovation is a critical component of the automobile and aerospace industries' development of next-generation vehicles, increasing America's economic competitiveness and making our transportation network safer and cheaper.

My theory interests prompted me to become an undergraduate researcher with Prof. Jason D. Lee, developing provable methods for deep learning. Funded by my Viterbi Fellowship, I investigated generalization in overparameterized neural networks. ***Intellectual Merit:*** I independently developed a GPU-parallelized testbed and evaluated various complexity measures—used to bound generalization—on computer vision tasks, characterizing each measure's robustness to regularization. Then, I empirically verified the prominence of the accuracy gap between ResNets and their kernel approximations, suggesting a shortcoming in the theory of infinite-width models. I worked with Prof. Lee on theoretical reasons for this gap, including implicit advantages of overparameterization not captured by kernels. ***Broader Impacts:*** The unpredictable behavior of overparameterized neural networks contradicts classical theory and leads to unexpected results in practice. By contributing to the fundamental understanding of deep learning, my research could increase effectiveness, credibility, and robustness of deep learning in its myriad applications. While Prof. Lee's transfer to Princeton prematurely ended our collaboration, the importance of this research ignited my desire to pursue a Ph.D., and my research proposal is based on our work.

As a machine learning research intern at Google X on the computational agriculture team, I developed a deep learning architecture for temporal identity preservation in multiple object tracking. ***Intellectual Merit:*** Time-series data is critical for predicting crop health and yield, but our current models were unable to associate crops over time due to their drastic physical transitions. I independently designed a novel CNN-LSTM architecture for understanding object evolution, generalizing previous work in multiple object tracking and enabling recovery of biological lifecycles from images. I collaborated with crop scientists, biologists, and engineers to identify a high-potential raspberry strain and initiate a time-lapse experiment in breeding fields, building a dataset with over 20,000 annotated images. My model can simultaneously track over 100 raspberries' growth from bud to flower to fruit, for the first time enabling extraction of temporal features such as flowering rate and ripening speed. My internship culminated in a keynote presentation to Google executives, who approved the hire of a full-time researcher to continue my work in 2021. My **first-author publication** is currently in preparation as we are targeting an ICCV 2021 submission, and my object tracking code has been transitioned to the official TensorFlow GitHub repository. I also took initiative to join a Google Brain group studying generalization in deep learning, where fruitful discussions with inspiring researchers cemented my Ph.D. topic. ***Broader Impacts:*** With a projected world population of ten billion by 2050, computational agriculture is necessary for ensuring sufficient food production. By automatically analyzing phenotype development over time, my research contributes to breeding high-yield, disease-resistant plants, helping farms grow more food for less money. Additionally, we are collaborating with the life sciences team to apply my work to microscopy imagery of cancer cell development, which could improve treatments for early-stage cancer.

Currently, I am an undergraduate researcher with Prof. Shaddin Dughmi working on my senior thesis. Funded by my Viterbi Fellowship, I am combining ideas from convex optimization and approximation algorithms to investigate alternative oracles for solving the linear feasibility problem, critical in the study of linear programming. ***Intellectual Merit:*** Typically, one requires a separation oracle in the use of the Ellipsoid Method, but in cases such as packing problems, solving the separation problem can be NP-hard or otherwise infeasible. Using ideas from convex geometry, I am developing an efficient randomized algorithm to solve the feasibility problem using only an oracle which reports the distance to the polytope. I am also investigating approximate versions of separation and distance oracles. My ultimate goal in this research is to show a lower bound on the information necessary to solve linear programs, which would both deepen our understanding of optimization problems and have important ramifications in algorithms and game theory. I aim to publish my results by the time I graduate, likely in a theory conference such as ICALP. ***Broader Impacts:*** Linear programming is used daily in engineering applications, but real-life scenarios are often too complex for a simple separation oracle to arise. My research would enable the use of linear programming in low-feedback settings such as satellite communication and energy grid optimization, lowering costs and assisting in the development of America's infrastructure.

Teaching and Outreach: In senior year of high school, I led my team to qualify for the VEX Robotics World Championship but couldn't attend due to lack of funding. I was driven to ensure that local students would not experience the same disappointment. Luckily, the vice principal Rick Monteilh of the local K-8 St. Odilia School had won a grant to start a VEX robotics program but lacked an experienced mentor. I saw my opportunity and quickly contacted Rick to volunteer. I coordinated with administration and grant patrons to purchase a $1,000 robotics kit, and for the next year **I mentored four female Hispanic students in St. Odilia's first middle-school robotics program.** I led my students through constructing their first robot out of steel parts and using the C language to program their robot to maneuver around the field and pick up foam balls. I'll never forget their faces when they drove their robot for the first time. In 2019, my mentees earned first place in the L.A. Clippers SciFest SoCal robotics competition, taking home $2,000 in prize money as well as free HP laptops, and we were interviewed for a local news article.

My desire to scale my volunteer experience led me to join the USC Viterbi School of Engineering K-12 STEM Center as the associate director of robotics outreach. I spearheaded the development of the Building Opportunities with Teachers in Schools (BOTS) initiative, whose goal is to **foster digital equity in East L.A. elementary schools** via a community of practice of local teachers. I created a curriculum using Sphero robots to teach STEM concepts to 1st and 2nd graders, coordinated a $66,000 budget to purchase and distribute Spheros to three schools, and led five USC volunteers in introducing Spheros to over 30 teachers. Impactful results of my contributions include the inclusion of an additional partner school for a total of 225 students, a bilingual program called *El Círculo Familiar* engaging 35 families of Hispanic students in East L.A. with their child's STEM experience through BOTS, and three USC news articles.

I then joined the USC Center for AI in Society (CAIS), an undergraduate organization focused on applications of machine learning for social good, where **I specialized in bias and fairness in machine learning.** My experience as an undergraduate teaching assistant for discrete mathematics—where I held triple the required amount of office hours—stoked my love of teaching and inspired me to join CAIS leadership as a curriculum director. I contributed lessons on CNNs, GANs, and VAEs to our open-source curriculum and mentored five undergraduates (two of whom now hold leadership roles) in deep learning topics. This culminated in a project and presentation in sentiment analysis of cyberbullying on Reddit. I then became project director in machine learning fairness, where I led four undergraduates in implementing the Word2Vec embedding algorithm from scratch and training it on 50,000 news articles, then analyzing the model's biases via the Euclidean distances between target adjectives and gendered nouns. We open-sourced our work on GitHub and presented our results, which inspired a continuation of the project focused on developing a machine learning-based resume screener to reduce gender and cultural biases in hiring.

Future Goals: With my experience and insight into the interplay between theory and practice in machine learning, I am uniquely positioned to contribute to the mathematical foundations of learning techniques while ensuring their fair and equitable application. To this end, I am pursuing a **Ph.D. in theoretical machine learning.** I would most like to join Prof. Tengyu Ma at Stanford University to analyze the generalization of practical neural networks. After my Ph.D., I plan to **pursue a career in academia** where I can combine my drive for cutting-edge research with my passion for mentoring the next generation of scientists. Then, I will **enter university leadership as a dean** to scale research partnerships between academia, industry, and government organizations. The NSF Graduate Research Fellowship would have an enormous effect on my career by enabling me to pursue my most ambitious, impactful research ideas and accelerate my efforts to foster a diverse, inclusive, and equitable scientific community.