When I sit at the piano, my music flows most naturally as my imagination leads my hands, carried by a technical foundation built from years of practice. My pursuit of a Bachelor's in Piano Performance has shaped how I now approach CS research: I start from an intuitive picture, then work through the technical grind until it becomes a precise, provably sound argument.

I have broad interests in theoretical computer science, and so far I have approached these questions through **Algorithmic Game Theory** and the **foundations of modern Machine Learning** (ML). In both, I aim to understand the limits and capabilities of complex computational systems:

- (**Foundations of modern ML.**) *On algorithmic tasks, when and why do Transformers prioritize learning statistical heuristics over generalizable algorithms, and how can we systematically induce robust reasoning?*
- (**Algorithmic Game Theory**.) *In collective decision-making, simple tournament-based voting rules are theoretically incapable of achieving optimal social welfare on their own. How can we bridge this optimality gap with minimal added overhead?*

Foundations of modern ML. While Transformer models demonstrate remarkable capabilities, they often prioritize learning statistical shortcuts over generalizable algorithms. Advised by Professors Robin Jia and Vatsal Sharan, I investigated this failure using graph connectivity and proved that the model's failure to generalize is not inherent to the architecture, but rather a deterministic consequence of the training distribution relative to the model's capacity. Crucially, this implies that by aligning the training distribution with the model's theoretical capacity, we can steer models toward more robust out-of-distribution generalization.

I developed the theoretical foundations of this project. I first proved a tight, non-asymptotic capacity bound that an L-layer model has the exact circuit complexity to solve connectivity for graphs with diameter up to 3^L. This theoretical threshold allowed us to model the learning process as a competition between two latent channels: a robust "algorithmic" channel that implements matrix powering and a shortcut "heuristic" channel that relies on a degree-counting shortcut. I then identified a sharp phase transition that depends solely on the data composition: "within-capacity" graphs drive the gradient descent trajectory toward the algorithmic solution, while "beyond-capacity" graphs promote the heuristic. This finding offers a prescriptive strategy for out-of-distribution generalization: restricting training data to graphs within the model's capacity paradoxically encourages the model to learn the generalizable algorithm by suppressing the heuristic channel, a prediction my collaborators empirically validated.

Algorithmic Game Theory. Under the metric distortion framework, the design of voting systems involves a fundamental tension between simplicity and efficiency. Tournament-based voting rules are practically appealing because they minimize cognitive load, requiring voters to only compare two options at a time. However, this simplicity comes at a theoretical cost: standard tournament rules are provably suboptimal compared to the general deterministic optimum.

Advised by Professor Kamesh Munagala, I sought to bridge this gap while preserving the simplicity of tournament rules. We proposed "Deliberation via Matching," a protocol where voters with opposing views may engage in pairwise discussions. We proved that by augmenting tournament rules with minimal pairwise deliberations, we can break the aforementioned barrier. Conceptually, this establishes that the cognitive simplicity of tournament rules does not require sacrificing social welfare: with slight overhead, they are just as powerful as general social choice rules.

In more technical details: I developed the theoretical foundations to resolve the analytical intractability that had limited prior work. The distortion objective for deliberative processes is inherently non-linear and non-convex, forcing previous studies to essentially rely on black-box numerical optimizers. I identified that the objective is *supermodular* and *convex* with respect to the metric variables. These were the crucial keys to tractability. Supermodularity implied that worst-case instances must follow a specific monotonic "structure," while convexity allowed me to apply Jensen's inequality locally to collapse a continuum of voters into a discrete set. This reformulation converted the intractable program into a bilinear relaxation, from which the optimal solutions easily followed from vertex enumerations and linear programming. While the proof is terse in its algebra, each step closely followed an intuitive story, and this precisely echos my belief that elegant research can start and be driven by clean intuitions.

At Stanford, I am eager to contribute to the Theory Group, particularly at the intersection of approximation algorithms and algorithmic game theory. I hope to work with Professors **Moses Charikar** and **Ashish Goel** (courtesy), whose foundational work on metric distortion and deliberative social choice directly influenced my recent paper on deliberation via matching. I am excited by how the Theory Group's broader strengths in algorithms, complexity, and algorithmic game theory would let me push these ideas toward more realistic, dynamic models of collective decision-making.