Introduction: The field of information theory is both mathematically profound and fundamentally practical. With deep connections to probability, statistics, and ergodic theory, it provides a unifying framework for understanding the quantification, storage, and communication of information. The two pillars of information theory—data compression and channel coding—are often considered to have reached maturity. However, my research has brought fresh perspectives to both these areas by challenging long-standing assumptions and introducing new frameworks. A key insight driving my research is that many foundational models and assumptions in information theory, though historically powerful, fail to fully capture the complexities of real-world scenarios. As new applications emerge, they expose gaps between theory and practice. My work bridges these gaps by reformulating foundational problems, offering more robust, flexible, and realistic models for contemporary applications. By deriving theoretical performance limits under these new models, my work provides more precise benchmarks for modern applications.

Data Compression: The Shannon entropy and rate-distortion function characterize the theoretical performance limits in lossless and lossy data compression, respectively. In my doctoral research, I focused on a key performance metric in data compression called the rate redundancy, which measures how quickly practical compression algorithms converge to these limits as the number n of encoded symbols increases. Since the 1980s and 1990s, the optimal rate redundancy has been firmly established as $\Theta\left(\frac{\log n}{n}\right)$ in both lossless and lossy compression, specifically in the *universal* setting where the source distribution is unknown. As Kontoyiannis [1] stated, "the question was essentially settled."

However, the *minimax* universal setting, which accounts for the worst-case performance of universal compression schemes over all source distributions, is the gold standard because it gives uniform convergence guarantees, even under adversarial conditions. In the minimax universal setting, the same $\Theta\left(\frac{\log n}{n}\right)$ rate redundancy result has been proven for lossless codes since 1981. However, minimax universal results for lossy compression remained elusive until 2023, when my work [2] broke new ground. My work shattered the long-standing $\Theta\left(\frac{\log n}{n}\right)$ paradigm by proving that the optimal rate redundancy for lossy compression under the minimax universal framework is actually $\tilde{\Theta}\left(\frac{1}{\sqrt{n}}\right)$. This contrasts sharply with the prior $\Theta\left(\frac{\log n}{n}\right)$ results in lossy compression, which only gave pointwise convergence guarantees. Even more importantly, I showed that the $\tilde{\Theta}\left(\frac{1}{\sqrt{n}}\right)$ rate redundancy result holds even when the source distribution is known (the non-universal setting). I provided a detailed study in [2] showing how regularity conditions imposed in prior works led to the faster $\Theta\left(\frac{\log n}{n}\right)$ convergence because they did not account for all i.i.d. sources and distortion measures. This fundamentally redefines the landscape of lossy compression theory, upending the $\Theta\left(\frac{\log n}{n}\right)$ standard across the board. This breakthrough merits renewed efforts in refining and extending the result to more general source models, which I am highly motivated to undertake in the future.

In another work [3], I introduced a novel paradigm in lossy compression called *universal distortion*, in which the distortion measure—traditionally fixed—is now a runtime input only to the encoder, along with the source data to be compressed. The universal distortion framework affords greater adaptability and is an especially useful model for modern compression algorithms based on nonlinear transforms, where optimal distortion in the transform domain must dynamically adjust to the source data characteristics. The universal distortion framework provides just the right model to study this problem because the distortion measure is a runtime input and is not known until the source data is also known. I proved rate redundancy results under the combined framework of *minimax* and *universal distortion*, providing uniform convergence guarantees over all i.i.d. sources and all distortion measures. I consider the extension of these results to source distributions with infinite alphabets a promising avenue for future research, especially after the recent work by Silva and Piantanida [4] addressing infinite source alphabets in the traditional lossy compression setting.

Channel Coding: In my work on channel coding [5], [6], I introduced two significant innovations. First, I developed an advanced cost model that supersedes the two cost constraints that have dominated the literature: the strict peak-power constraint and the weaker expected cost constraint. This novel cost formulation constrains the cost (or power) of the transmission both in expectation and variance. With a variance parameter V, the *mean and variance (m.v.) cost constraint* generalizes the existing frameworks, with $V \to 0$ recovering the (first- and second-order) coding performance of the peak-power constraint and $V \to \infty$ recovering the expected cost constraint. Beyond generalization, I showed that the m.v. cost constraint for $0 < V < \infty$ has practical advantages over both prior cost models. Unlike the peak-power constraint, it enables improved coding performance with feedback; even without feedback, the coding performance under the m.v. cost constraint is superior. Unlike the expected cost constraint, it enforces a controlled, ergodic use of transmission power. This is essential for several practical reasons, such as operating circuitry in the linear regime, minimizing power consumption, and reducing interference with other terminals [6]. The new cost constraint achieves its benefits by allowing the cost to fluctuate above the threshold in a manner consistent with a noise process, thus making it a more realistic and natural cost model in practice than the restrictive peak-power constraint.

My second innovation was in feedback communication, where I unveiled new ways in which feedback can improve communication performance. For *any* $V > 0$, I showed that feedback improvement is possible for a significantly larger class of channels than in the prior study on unconstrained channels [7]. Additionally, I proved that for a broad class of channels, feedback improvement is possible *if and only if* $V > 0$. These results reveal the critical role of cost variability V in enabling feedback mechanisms to improve coding performance. These are also the first results to establish second-order feedback improvement for discrete memoryless channels *with* cost constraints, thus providing a broader understanding of how and when feedback improves coding performance. Looking forward, I am dedicated to extending these results to the Gaussian channel, which is the canonical channel model for real-world communication, especially wireless communication.

Looking Ahead: The relevance of information theory continues to grow, especially as it intersects with fields such as machine learning, quantum computing, control systems, and optimal transport theory. Expanding the theoretical boundaries of information theory is thus essential for establishing the rigorous structures that enable innovation across disciplines. By reformulating problems in these areas using information-theoretic quantities, we gain deeper insights and operational interpretations, grounding empirical results with robust theoretical foundations. I am eager to pursue this line of inquiry, advancing core theory for modern applications as well as exploring interdisciplinary applications, particularly in quantum information theory and optimal transport theory.

Working in academia provides the ideal environment for me to pursue my research vision. Universities offer the intellectual freedom to pursue both basic and applied research, in addition to opportunities for collaboration and mentoring the next generation of scholars. For me, this combination makes academia a fulfilling career.

References

[1] I. Kontoyiannis, "Pointwise redundancy in lossy data compression and universal lossy data compression," *IEEE Transactions on Information Theory*, vol. 46, no. 1, pp. 136–152, 2000.

[2] A. Mahmood and A. B. Wagner, "Minimax rate-distortion," *IEEE Transactions on Information Theory*,

vol. 69, no. 12, pp. 7712–7737, 2023.

[3] A. Mahmood and A. B. Wagner, "Lossy compression with universal distortion," *IEEE Transactions on Information Theory*, vol. 69, no. 6, pp. 3552–3573, 2023.

[4] J. F. Silva and P. Piantanida, "On universal d-semifaithful coding for memoryless sources with infinite alphabets," *IEEE Transactions on Information Theory*, vol. 68, no. 4, pp. 2782–2800, 2022.

[5] A. Mahmood and A. B. Wagner, "Channel coding with mean and variance cost constraints," in *2024 IEEE International Symposium on Information Theory (ISIT)*, 2024, pp. 510–515.

[6] A. Mahmood and A. B. Wagner, "Improved channel coding performance through cost variability," 2024. [Online]. Available: https://arxiv.org/abs/2407.05260

[7] A. B. Wagner, N. V. Shende, and Y. Altuğ, "A new method for employing feedback to improve coding performance," *IEEE Transactions on Information Theory*, vol. 66, no. 11, pp. 6660–6681, 2020.