Throughout my academic journey, my research has been rooted in strengthening the theoretical foundations of classical information theory, particularly in the areas of data compression and channel coding. Even though these core aspects of information theory are often considered as having reached maturity a long time ago, my research has challenged long-standing assumptions and introduced novel frameworks, breathing new life into classical information theory by addressing critical gaps that have persisted for decades. New frameworks and reformulation of foundational problems in both data compression and channel coding increases the practical utility of information theory for modern applications.

In my doctoral research, I focused on a key performance metric in data compression called the rate redundancy, which quantifies how quickly the performance of practical compression algorithms approaches the theoretical limits as the number n of symbols encoded increases. For decades, from the 1980s through the early 2000s, the $\Theta\left(\frac{\log n}{n}\right)$ rate redundancy result was well-established in both lossless and lossy compression, including in the universal setting where the source distribution is unknown; "the question was essentially settled," Kontoyiannis [1].

However, the *minimax* universal setting is the gold standard because it accounts for the worst-case performance of universal compression schemes over all source distributions, thus giving uniform convergence guarantees under potentially 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 in [2] shattered the long-standing $\Theta\left(\frac{\log n}{n}\right)$ paradigm by proving that the optimal rate redundancy for lossy compression under the universal *minimax* framework is actually $\tilde{\Theta}\left(\frac{1}{\sqrt{n}}\right)$. This result stands in stark contrast to 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 $\tilde{\Theta}\left(\frac{1}{\sqrt{n}}\right)$ rate redundancy persists even in the non-universal lossy setting. I gave a detailed study in [2] on how regularity conditions imposed in prior works led to a faster $\Theta\left(\frac{\log n}{n}\right)$ convergence by not accounting 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 into refining and extending the result to more general source models, which I am highly motivated to do.

In a related work [3], I pioneered a novel variant of lossy compression called *universal distortion* in which the distortion measure - traditionally fixed - is now an input to the encoder along with the source data to be compressed. One of the primary motivations behind my work on lossy coding with *universal distortion* was to introduce greater versatility in compression systems to meet the needs of a variety of end-users who may have discordant notions of distortion. Another critical motivation stemmed from modern compression algorithms, which apply nonlinear transforms to the source data before compression. Optimal compression then corresponds to adjusting distortion in the transform domain in such a way that the source-reconstruction distortion in the original domain is minimized. Due to nonlinear transforms, optimal distortion "minimization" in the transform domain has to be sensitive to the input source data. The *universal distortion* framework provides just the right model to study this problem because the distortion measure is a real-time input and is not known until the input source data is also known. The universal distortion framework is particularly well-suited for emerging applications in areas such as machine learning, image processing and signal compression, where data transformations and varying distortions are common.

I proved rate redundancy results under the combined framework of *minimax* and *universal distortion*, providing uniform convergence guarantees over both 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.

Beyond source coding, my work in channel coding in [5] and [6] has introduced two significant innovations.

First, I developed a more practical cost model that superseded existing ones. My new cost formulation constrains the cost (or power) of the transmission in expectation as well as in its variance. This *mean and variance (m.v.) cost constraint* provides an alternative to the two standard cost constraints in the literature: the strict peak-power constraint and the weaker expected cost constraint. The m.v. cost formulation with a variance parameter V generalizes these existing frameworks in the sense that letting $V \to 0$ recovers the peak-power constraint and letting $V \to \infty$ recovers the expected cost constraint. Beyond generalization, I showed that the *m.v. cost constraint* for $0 < V < \infty$ offers practical benefits over both prior cost models. Unlike the peak-power constraint, it allows for an 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 an ergodic and controlled use of transmission power, which is desirable for several practical reasons detailed in the introduction section of [6]. The benefits of the new cost constraint come about by merely 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 findings highlight the important 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 giving a broader understanding of how and when feedback improves the coding performance. Looking forward, I am dedicated to extending these results for a Gaussian channel, which is the canonical channel model for real-world communication, especially wireless communication.

In general, I believe that information theory will remain pivotal in advancing a range of fields, including machine learning, control systems, statistics, optimal transport theory and quantum computing. By reformulating problems in these areas using information theoretic quantities, we gain deeper insights and operational interpretations, grounding empirical results with robust theoretical foundations. Strengthening the theoretical foundations of classical information theory is thus essential for establishing the rigorous structures that enable innovation across domains. I am eager to pursue this line of inquiry, advancing core theory for increased utility for modern applications in communication and storage systems.

To this end, I am also actively expanding my expertise, particularly in quantum information theory and optimal transport theory, which offer rich frameworks for future exploration. With a deep commitment to both fundamental theory and interdisciplinary exploration, I aim to make significant contributions by working in academia.

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] ——, “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] ——, “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.