Research Statement

Hyungjin Chung

My research focuses on developing modular and adaptable approaches to inference in complex, high-stakes problems by (1) constructing priors using **generative AI** and (2) designing practical approximate **posterior sampling algorithms** grounded in optimization and probability. Many challenges in science and engineering can be formulated as posterior sampling tasks. We collect measurements from various sensors and aim to maximize fidelity (likelihood) while constructing beliefs that best match the world's prior knowledge. Standard approaches such as supervised training and amortized inference lack flexibility and modularity, failing to generalize beyond the training data. Encompassing all possible scenarios during training is infeasible.

Prior work

Diffusion models for inverse problems

To address these challenges, my research aims to create a flexible composition of search space, solutions, and uncertainties through Bayesian inference. Per this goal, my research employs **generative models as priors** for solving **posterior sampling** problems. Specifically, I was among the first to advocate for using diffusion models in posterior sampling to solve general inverse problems in computer vision and imaging [1, 2, 3], demonstrating that both linear inverse problems (e.g., inpainting, super-resolution, CT/MRI reconstruction) and nonlinear problems (e.g., phase retrieval) can be addressed within a unified framework using a single generative diffusion model trained for image *synthesis*. At the time, supervised training was the standard approach for inverse problems in imaging. Through this work, I established that leveraging unconditional diffusion models with carefully designed inference algorithms could achieve state-of-the-art results in perceptual quality while having a much broader generalization capacity and flexibility.

In high-stakes applications such as medical imaging, not only should one strive for better performance, but also aim for provable guarantees for safe deployment. In this regard, my focus extends beyond designing practical algorithms to establishing theoretically grounded methods. I explored the geometry of diffusion models and their behavior under guided sampling [2], accelerated inference through stochastic contraction [1], and bounded approximation errors [3]. These theoretical insights have provided a valuable foundation for further improvements in diffusion model-based inverse problem solvers and are considered standard theoretical tools today.

Building upon my earlier works, I progressively tackled more challenging problem settings. In BlindDPS [4], I demonstrated that **composing** two different priors, each responsible for the marginals, enables solving blind inverse problems, where both the operator and the image to be reconstructed must be estimated simultaneously. In a similar spirit, for 3D reconstruction of MRI/CT scans, I proposed **factorizing** the complex 3D prior into a composition of two simpler 2D priors, leading to state-of-the-art reconstruction results [4, 5]. To make the algorithms scalable to real-world applications, I further accelerated them through the geometric lens of diffusion models, proposing an algorithm that excels in large-scale settings [6]. Additionally, I introduced a unified framework of direct diffusion bridges [7] for cases where faster inference is absolutely necessary, allowing for a

trade-off between perceptual quality and fidelity at inference time.
Finally, I tackled the case where there exists no gold standard data to train a good generative prior, a situation that is often encountered in applications with data scarcity. In such case, the best one can do is to train a generative prior with an out-of-distribution (OOD) phantom, and hope for generalization. For this, I proposed deep diffusion image prior (DDIP) [8], an online OOD adaptation algorithm for solving inverse problems and adapting the generative prior under prior mismatch. I showed that even using a single diffusion model trained with completely irrelevant phantoms that could be generated on the fly, DDIP can be adapted to diverse modalities such as MRI and CT.

Biomedical imaging applications

Theory is most compelling when it finds meaningful applications. With a background in biomedical engineering, I have worked on practical and robust algorithms for applying generative AI to diverse imaging applications, including MRI [9], optical diffraction tomography (ODT) [10], and STEM-EDX tomography [11, 12], among others. After dedicating myself to diffusion models and developing modular inference algorithms, I proposed Score-MRI [13], one of the first compressed-sensing MRI methods demonstrating exceptional generalization capacity—not only being agnostic to the forward model but also to different anatomies and contrasts. Due to the stochastic nature, Score-MRI could perform uncertainty quantification, crucial for high-stakes tasks. Moreover, while being a single generalist model never trained specifically for MRI reconstruction, it outperformed the best-in-class *specialist* supervised models trained specifically for a pre-specified forward model.
Seeing the possibility and the intriguing aspects of leveraging diffusion priors, I started working closely with radiologists and ophthalmologists to develop methods that can actually be used in practice. I developed an MRI denoising and super-resolution method [14] by training a foundational diffusion prior on a comprehensive corpus of MRI data. The software developed from this research has been commercialized and is currently operational in more than 20 hospitals across South Korea. By extending and tailoring direct diffusion bridge [7] for fundus photo enhancement, I developed FD3 [15], a fast and robust algorithm that enables ophthalmologists to plan and anticipate outcomes before conducting cataract surgery. These endeavors have also led to my contribution to MONAI Generative Models [16], aiming to democratize the use of generative models-especially diffusion models-in medical imaging.

Future directions

Multimodal generative priors for inverse problems

Information is rarely captured in a unimodal fashion. However, in inverse problems, existing works often consider only the information captured by the sensor, despite the abundance of metadata that provides additional prior knowledge and context for the signal to be recovered. For instance, in MRI, alongside capturing the k-space signal, we have access to patient information, MR imaging parameters, and medical history. Unfortunately, this additional information is often discarded during inference—a suboptimal practice according to the data processing inequality.
By training a multimodal generative model that approximates either the joint or conditional distribution of multimodal inputs, we can fully utilize previously unused information, achieving better reconstruction performance and enhanced controllability. As a proof of concept, I proposed the first works leveraging additional text conditions as input, demonstrating that it can enhance the performance of inverse problem solving [17] and serve as a control signal for selecting the target mode [18]. Focusing on text as a multimodal control knob for solving inverse problems, I posed text-to-image generation as a special case of inverse problems and proposed CFG++ [19], illustrating

significant improvements in performance with desirable properties such as smooth trajectory and better invertibility—all achieved by constructing a better inference strategy.
In future research, I aim to extend this approach to practical, real-world tasks in biomedical imaging, recognizing that as measurement signals become sparser, the role of metadata and contextual information becomes increasingly vital. In medical imaging, sparse measurements can be supplemented by detailed metadata like patient history, anatomical focus, or imaging parameters, aiding in producing sharper reconstructions and enhancing diagnostic utility. Similarly, in Cryo-EM reconstruction, metadata on factors such as sample concentration, freezing conditions, and tilt angles can guide algorithms to compensate for sparse or noisy data, ultimately improving structural fidelity and resolution. By integrating such metadata across different imaging modalities, we can push the boundaries of signal recovery and interpretation in biomedical applications where sparse data acquisition is often necessary.

Modular and adaptive priors for autonomous AI Agents

An intelligent agent should be embodied, as the perception of a multimodal world is crucial. Through interactions, the agent samples incomplete signals and must infer the best possible actions from deficient measurements. Hence, the agent naturally has to solve multimodal inverse problems in real-time, imputing missing data from known values and effectively fusing multimodal measurements. Constructing a single generative model for such agents is unrealistic due to the difficulty of collecting long-horizon, multimodal data in real-world settings. One of my research goals is to compose generative priors for each marginal—where it is feasible to collect data and construct a good generative prior—and combine them during inference so that the agent can adapt to different settings. While I previously considered the composition of generative priors in controlled, static settings [4, 20, 5], my goal is to extend this to dynamic, adaptable settings where the agent itself decides how best to leverage these different priors. In this complex scenario, it is crucial for the agent to adapt and update its priors based on the environment. Here, online adaptation of the generative prior, which I have previously explored in the inverse problem setting [8], will be essential. Autonomous, intelligent AI agents should be able to explore the search space rather than being statically guided. Reinforcement learning (RL) is pivotal in this regard, where reward functions can be used as a sparse equivalent of a dense likelihood (i.e., data fidelity) function [21]. While my previous studies have focused on modular sampling with dense guidance, I am interested in expanding this to enable exploratory behavior with sparse guidance through RL.

Towards next generation GenAI

Autoregressive models dominate language modeling, while diffusion models dominate vision, each with its own strengths and weaknesses. The next generation of generative models must preserve the strengths of both model classes: (1) the ability to generate sequences of arbitrary length, unlike current diffusion models limited to short windows, and (2) the capacity for iterative refinement, unlike autoregressive models that cannot correct previous errors. Additionally, the new model class should incorporate a memory module to tackle long-range tasks. I have explored combining the best of both worlds by integrating autoregressive models and diffusion models at test time and implementing the memory module through large language models [22]. In future research, I plan to further explore the design space of this new generative model, enabling it to autonomously allocate the required computational resources depending on the difficulty of the task. Equipped with the capabilities of arbitrary-length generation from autoregressive models and the iterative refinement from diffusion models, the new model will be able to generalize to arbitrarily long-range tasks through iterative reasoning, fully leveraging inference-time computation [23].

References

[1] **Hyungjin Chung**, Byeongsu Sim, and Jong Chul Ye. Come-Closer-Diffuse-Faster: Accelerating Conditional Diffusion Models for Inverse Problems through Stochastic Contraction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.

[2] **Hyungjin Chung***, Byeongsu Sim*, Dohoon Ryu, and Jong Chul Ye. Improving diffusion models for inverse problems using manifold constraints. *Advances in Neural Information Processing Systems*, 35:25683–25696, 2022.

[3] **Hyungjin Chung***, Jeongsol Kim*, Michael Thompson Mccann, Marc Louis Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *The Eleventh International Conference on Learning Representations*, 2023.

[4] **Hyungjin Chung***, Jeongsol Kim*, Sehui Kim, and Jong Chul Ye. Parallel diffusion models of operator and image for blind inverse problems. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023.

[5] Suhyeon Lee*, **Hyungjin Chung***, Minyoung Park, Jonghyuk Park, Wi-Sun Ryu, and Jong Chul Ye. Improving 3D imaging with pre-trained perpendicular 2D diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10710–10720, 2023.

[6] **Hyungjin Chung**, Suhyeon Lee, and Jong Chul Ye. Decomposed diffusion sampler for accelerating large-scale inverse problems. In *The Twelfth International Conference on Learning Representations*, 2024.

[7] **Hyungjin Chung**, Jeongsol Kim, and Jong Chul Ye. Direct diffusion bridge using data consistency for inverse problems. *Advances in Neural Information Processing Systems*, 36, 2024.

[8] **Hyungjin Chung** and Jong Chul Ye. Deep diffusion image prior for efficient ood adaptation in 3d inverse problems. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2024.

[9] **Hyungjin Chung**, Eunju Cha, Leonard Sunwoo, and Jong Chul Ye. Two-stage deep learning for accelerated 3D time-of-flight MRA without matched training data. *Medical Image Analysis*, 71:102047, 2021.

[10] **Hyungjin Chung***, Jaeyoung Huh*, Geon Kim, Yong Keun Park, and Jong Chul Ye. Missing Cone Artifact Removal in ODT Using Unsupervised Deep Learning in the Projection Domain. *IEEE Transactions on Computational Imaging*, 7:747–758, 2021.

[11] Yoseob Han*, Jaeduck Jang*, Eunju Cha*, Junho Lee*, **Hyungjin Chung***, Myoungho Jeong, Tae-Gon Kim, Byeong Gyu Chae, Hee Goo Kim, Shinae Jun, et al. Deep learning STEM-EDX tomography of nanocrystals. *Nature Machine Intelligence*, 3(3):267–274, 2021.

[12] Eunju Cha*, **Hyungjin Chung***, Jaeduck Jang, Junho Lee, Eunha Lee, and Jong Chul Ye. Low-Dose Sparse-View HAADF-STEM-EDX Tomography of Nanocrystals Using Unsupervised Deep Learning. *ACS nano*, 16(7):10314–10326, 2022.

[13] **Hyungjin Chung** and Jong Chul Ye. Score-based diffusion models for accelerated MRI. *Medical Image Analysis*, page 102479, 2022.

[14] **Hyungjin Chung**, Eun Sun Lee, and Jong Chul Ye. MR image denoising and super-resolution using regularized reverse diffusion. *IEEE Transactions on Medical Imaging*, 42(4):922–934, 2022.

[15] Sehui Kim*, **Hyungjin Chung***, Se Hie Park, Eui-Sang Chung, Kayoung Yi, and Jong Chul Ye. Fundus image enhancement through direct diffusion bridges. *IEEE Journal of Biomedical and Health Informatics*, 2024.

[16] Walter HL Pinaya et al. Generative ai for medical imaging: extending the monai framework. *arXiv preprint arXiv:2307.15208*, 2023.

[17] **Hyungjin Chung**, Jong Chul Ye, Peyman Milanfar, and Mauricio Delbracio. Prompt-tuning latent diffusion models for inverse problems. In *Forty-first International Conference on Machine Learning*, 2024.

[18] Jeongsol Kim*, Geon Yeong Park*, **Hyungjin Chung**, and Jong Chul Ye. Regularization by texts for latent diffusion inverse solvers. *arXiv preprint arXiv:2311.15658*, 2023.

[19] **Hyungjin Chung***, Jeongsol Kim*, Geon Yeong Park*, Hyelin Nam*, and Jong Chul Ye. Cfg++: Manifold-constrained classifier free guidance for diffusion models. *arXiv preprint arXiv:2406.08070*, 2024.

[20] **Hyungjin Chung***, Dohoon Ryu*, Michael T McCann, Marc L Klasky, and Jong Chul Ye. Solving 3d inverse problems using pre-trained 2d diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22542–22551, 2023.

[21] Kaifeng Zhao, Gen Li, and Siyu Tang. Dart: A diffusion-based autoregressive motion model for real-time text-driven motion control. *arXiv preprint arXiv:2410.05260*, 2024.

[22] **Hyungjin Chung***, Dohun Lee*, and Jong Chul Ye. Acdc: Autoregressive coherent multimodal generation using diffusion correction. *arXiv preprint arXiv:2410.04721*, 2024.

[23] OpenAI. OpenAI o1. `https://openai.com/o1/`, 2024.