A Mixture-based Framework for Guiding Diffusion Models

Alain Oliviero-Durmus

Inverse problems—such as reconstructing images from partial or noisy measurements, or separating individual sources from mixed signals—are inherently challenging due to their ill-posed nature. In such settings, Bayesian inference, when combined with generative modelling, provides a systematic and principled approach. By using generative models trained on representative data distributions, these methods incorporate meaningful prior knowledge, which can then be integrated with the likelihood function describing the observed data. This leads to a posterior distribution, whose samples represent plausible solutions that harmonize both the observed data and prior assumptions. In recent developments, diffusion models have emerged as state-of-the-art generative models, demonstrating exceptional capabilities in image and audio generation tasks. Diffusion models function by first progressively adding noise to data samples through a forward diffusion process, ultimately converting them into pure Gaussian noise. The generative model is then trained to reverse this noising process, effectively learning to reconstruct original data from noise. While diffusion models provide powerful priors, directly using them for inverse problems typically requires constructing a posterior denoiser that blends this learned prior with the gradient of the log-likelihood function derived from the observations. However, existing posterior sampling methods for diffusion models often rely on crude approximations of the likelihood gradient and require significant heuristic tuning and adjustments specific to each task. In this talk, I will introduce a novel principled approach specifically designed to overcome these limitations. The core contribution of this approach is the construction of a mixture approximation of intermediate posterior distributions defined by the diffusion model. The sampling is carried out sequentially via Gibbs sampling, a Markov Chain Monte Carlo method, using a careful data augmentation scheme. Gibbs sampling is employed here due to its simplicity and theoretical guarantees, allowing for exact conditional updates at each iteration, thus ensuring stability and efficiency. One key advantage of the presented algorithm is its flexibility: it adapts to varying levels of computational resources by adjusting the number of Gibbs iterations. Consequently, substantial performance gains can be achieved by increasing inference-time computational effort. I will present extensive experimental results demonstrating strong empirical performance across ten diverse image restoration tasks, involving both pixel-space and latentspace diffusion models, and showcase its successful application in musicalsource separation.