

Large-scale Wasserstein gradient flows with applications for computing diffusions

Evgeny **Burnaev**

Skolkovo Institute of Science and Technology

Wasserstein gradient flows provide a powerful means of understanding and solving many diffusion equations. Specifically, Fokker-Planck equations, which model the diffusion of probability measures, can be understood as gradient descent over entropy functionals in Wasserstein space. This equivalence, introduced by Jordan, Kinderlehrer and Otto, inspired the so-called JKO scheme to approximate these diffusion processes via an implicit discretization of the gradient flow in Wasserstein space. Solving the optimization problem associated to each JKO step, however, presents serious computational challenges. We introduce a scalable method to approximate Wasserstein gradient flows, targeted to machine learning applications. Our approach relies on input-convex neural networks (ICNNs) to discretize the JKO steps, which can be optimized by stochastic gradient descent. Unlike previous work, our method does not require domain discretization or particle simulation. As a result, we can sample from the measure at each time step of the diffusion and compute its probability density. We demonstrate our algorithm's performance by computing diffusions following the Fokker-Planck equation and apply it to unnormalized density sampling as well as nonlinear filtering.