PARTICLE--MALA AND PARTICLE--MGRAD: GRADIENT--BASED MCMC METHODS FOR HIGH--DIMENSIONAL STATE-SPACE MODELS

AXEL FINKE

State--of--the--art methods for Bayesian inference in state--space models are (a) conditional sequential Monte Carlo (CSMC) algorithms; (b) sophisticated `classical' MCMC algorithms like MALA, or mGRAD from Titsias and Papaspiliopoulos (2018). The former propose $N$ particles at each time step to exploit the model's `decorrelation--over--time' property and thus scale favourably with the time horizon, T, but break down if the dimension of the latent states, D, is large. The latter leverage gradient--/prior--informed local proposals to scale favourably with D but exhibit sub-optimal scalability with T due to a lack of model--structure exploitation. We introduce methods which combine the strengths of both approaches. The first, Particle--MALA, spreads N particles locally around the current state using gradient information, thus extending MALA to $T>1$ time steps and $N>1$ proposals. The second, Particle--mGRAD, additionally incorporates (conditionally) Gaussian prior dynamics into the proposal, thus extending the mGRAD algorithm to $T>1$ time steps and $N>1$ proposals. We prove that Particle--mGRAD interpolates between CSMC and Particle--MALA, resolving the `tuning problem' of choosing between CSMC (superior for highly informative prior dynamics) and Particle--MALA (superior for weakly informative prior dynamics). We similarly extend other `classical' MCMC approaches like auxiliary MALA, aGRAD, and preconditioned Crank--Nicolson--Langevin (PCNL) to $T>1$ time steps and $N>1$ proposals. In experiments, for both highly and weakly informative prior dynamics, our methods substantially improve upon both CSMC and sophisticated `classical' MCMC approaches.