Algorithms, vol.19, no.7, pp.581, July, 2026.

Kalman-Annealing: Calibrated Uncertainty for Simulated Annealing via a Probabilistic-Numerics Filter, with an Application to Reinforcement-Learning Hyperparameter Tuning

E.C. Garrido Merchán

Abstract-

Noisy, expensive, gradient-free optimisers-simulated annealing chief among them-almost never report how confident one should be in the configuration they return, and reinforcement-learning hyperparameter tuning, where the noise is large and the budget tight, is the setting where this silence hurts most. The contribution of this paper is a mechanism for uncertainty quantification, not a faster optimiser: we equip simulated annealing with a calibrated credible interval over the value of the recovered configuration, and we are explicit that this comes at an optimisation cost that only some landscapes repay. We introduce Kalman-Annealing (KA), a minimal modification of simulated annealing in which a one-dimensional Kalman filter-the canonical probabilistic numerical method is interleaved with the Metropolis acceptance step. The filter denoises each return before acceptance, and a short terminal refinement of the best visited state converts the run into a calibrated credible interval over the value of the recovered hyperparameter. A single analytical identity, Qt = c T2t, couples the filter process noise to the cooling schedule and absorbs the only free parameter of the filter into one already present in the metaheuristic. Under standard cooling assumptions the credible intervals are calibrated and the posterior variance contracts at a rate compatible with simulated-annealing convergence. On synthetic benchmarks (a noisy five-dimensional quadratic and the noisy Branin function, 200 seeds each) and on hyperparameter tuning of REINFORCE on three classic-control tasks (10 seeds each), the empirical 90% coverage of KA’s credible intervals lies within sampling error of the nominal level-a property none of the baselines provides-and the optimiser overhead is close to four orders of magnitude below that of Gaussian-process Bayesian optimisation. The interval cannot be extracted for free from an unmodified SA run: an interval built from the trailing evaluations of the vanilla trajectory fails to calibrate in every reading we test, and the repair that does calibrate is exactly KA's terminal-refinement phase grafted onto the unfiltered chain, at the same cost in diverted evaluations. Honest scoreboard: On simple regret, KA is at best on par with vanilla simulated annealing on the unimodal synthetic (the nominal advantage does not survive correction for multiple comparisons) and loses to the SA family, to CMA-ES and to Gaussian-process Bayesian optimisation on the multi-modal and ill-conditioned synthetics and on the informative reinforcement-learning tasks, under both REINFORCE and PPO. We trace this gap quantitatively to the filter acting as a low-pass, with a mean Kalman gain near one half, on the favourable-tail observations that drive SA’s basin escape, and we delineate the operating regime in which the calibrated-uncertainty contribution of KA is worth its optimisation cost.

Index Terms- simulated annealing; Kalman filter; probabilistic numerics; reinforcement learning; hyperparameter optimisation; uncertainty quantification

Citation:
Garrido-Merchán, E.C. "Kalman-Annealing: Calibrated Uncertainty for Simulated Annealing via a Probabilistic-Numerics Filter, with an Application to Reinforcement-Learning Hyperparameter Tuning", Algorithms, vol.19, no.7, pp.581, July, 2026.