Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

NESTED DENOISING DIFFUSION SAMPLING FOR GLOBAL OPTIMIZATION

Yuhao Wang[1]

[1]School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT

We propose Nested Denoising Diffusion Sampling (NDDS), a novel method for global optimization of expensive black-box functions. NDDS leverages conditional denoising diffusion probabilistic models to approximate the evolving solution distribution, eliminating the need for extensive additional function evaluations. Unlike prior diffusion-based optimization methods that rely on heuristically chosen conditioning variables, NDDS systematically generates them using a statistically principled mechanism. Furthermore, we introduce a likelihood ratio–based data reweighting strategy to correct the mismatch between the empirical training distribution and the current target distribution. Numerical experiments on benchmark problems demonstrate that NDDS consistently outperforms the Extended Cross-Entropy method under the same evaluation budget, with notable efficiency gains in high-dimensional settings.

1 INTRODUCTION

Deterministic global optimization problems with black-box, expensive-to-evaluate objectives arise in fields such as engineering design, hyperparameter tuning, and simulation optimization. Many sampling-based algorithms iteratively update a probability distribution over candidate solutions, but their efficiency is hampered by the cost of function evaluations.

To address this challenge, we propose *Nested Denoising Diffusion Sampling (NDDS)*, which integrates generative modeling—specifically, conditional diffusion models—into the optimization process. NDDS leverages diffusion-based sampling to improve efficiency by generating high-quality candidate solutions without requiring a large number of additional expensive evaluations.

2 NESTED DENOISING DIFFUSION SAMPLING (NDDS)

Let $p_k(x)$ denote the estimated distribution of the optimal solution at iteration k. Classical distribution-based methods update

$$p_{k+1}(x) \;\propto\; S\big(H(x)\big)\, p_k(x), \tag{1}$$

where $S(\cdot)$ is an increasing and positive transformation of the objective $y = S\big(H(x)\big)$.

NDDS trains a *conditional* denoising diffusion model $q_{X|Y}(x \mid y)$ on evaluated solution–value pairs $\{(x_i, y_i)\}$ with $y_i = S(H(x_i))$. At each iteration, it generates new candidates via the following two-stage procedure:

1. **Condition sampling.** Draw a label y from a learned distribution with density proportional to $y\, q_Y(y)$.
2. **Conditional generation.** Generate $x \sim q_{X|Y}(\cdot \mid y)$ via the denoising process of the diffusion model to obtain candidate solutions conditioned on the sampled label y.

Here, q_X and q_Y denote the marginals for x and y, respectively, and $q_{X|Y}$ is the conditional distribution of x given y. Theoretically, if q is perfectly trained, the generated samples follow

$$f(x) \;\propto\; S\big(H(x)\big)\, q_X(x) \;\approx\; p_{k+1}(x),$$

i.e., NDDS produces samples proportional to the updated target distribution.

3 ITERATIVE OPTIMIZATION FRAMEWORK

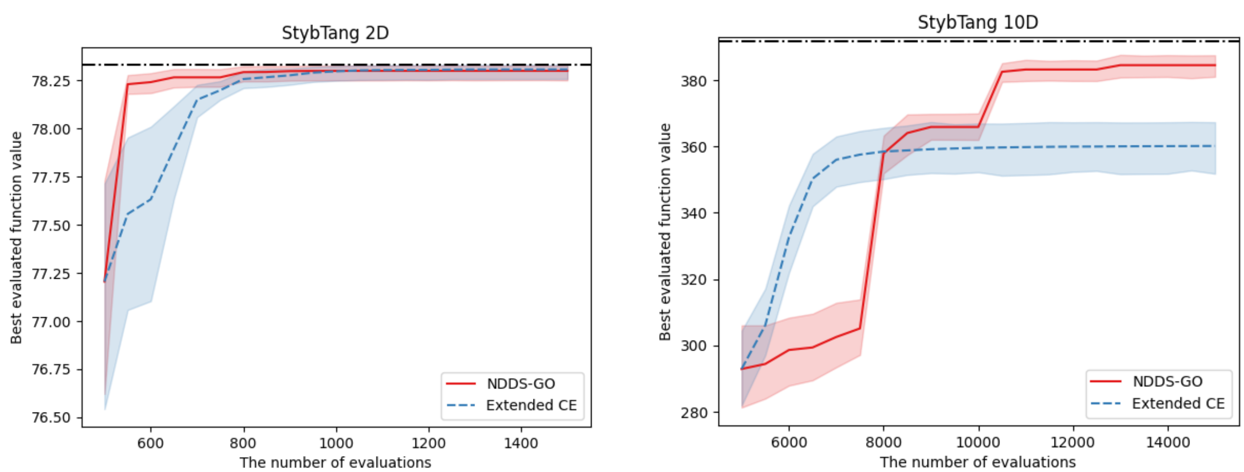
Our NDDS-based global optimization algorithm (NDDS-GO) alternates between the following steps:

1. **Evaluation.** Sampling and evaluating a small batch of solutions using a mixture of the current estimate and an exploration distribution.
2. **Model training.** Retraining the conditional diffusion model on reweighted data, where we apply a likelihood ratio–based reweighting of historical data to address the mismatch between the dataset distribution and the target distribution.
3. **Sampling.** Generating a large batch of synthetic solutions via NDDS and fitting a parametric distribution for the next iteration.

This framework reduces the number of expensive evaluations while maintaining a large effective sample size for updating the distribution estimate.

4 EXPERIMENTAL RESULTS

We evaluated NDDS-GO against the Extended Cross-Entropy method on 2D and 10D Styblinski–Tang functions. As shown in Figure 1, NDDS-GO converged more quickly to the global optimum in 2D and avoided local maxima in 10D, achieving higher final objective values with the same evaluation budget.



(a) Styblinski–Tang 2D

(b) Styblinski–Tang 10D

Figure 1: Best evaluated function value against the number of evaluations for NDDS-GO and Extended CE over 10 macro-replications.

5 CONCLUSION

By combining conditional diffusion models with a principled conditioning mechanism and likelihood ratio–based data reweighting, NDDS enables more sample-efficient global optimization. Experimental results demonstrate substantial improvements over traditional distribution-based methods, particularly in high-dimensional problems.

ACKNOWLEDGEMENT

The author thanks Haowei Wang, Enlu Zhou, and Szu Hui Ng for their valuable collaboration and insights that contributed to this work.