IMPERIAL

SMC-Guided Diffusion for General Optimisation

Brendan Dowling, Dr O. Deniz Akyildiz

September 6, 2024

Imperial College London

Table of contents

Introduction

First Frame

Hello, world! Akyildiz et al. (2020), Anderson (1982), Del Moral (2011), Dou and Song (2023), and Ho et al. (2020)

Background

Annealing

Text

Inverse Problems

Text

Diffusion Models

Text

Sequential Monte Carlo

Text

SMCDiffOpt

Heuritic Motivation

Text

Algorithm

Text

Geometric Interpretation

Text

Appyling to Inverse Problems

Text

Experimental Results

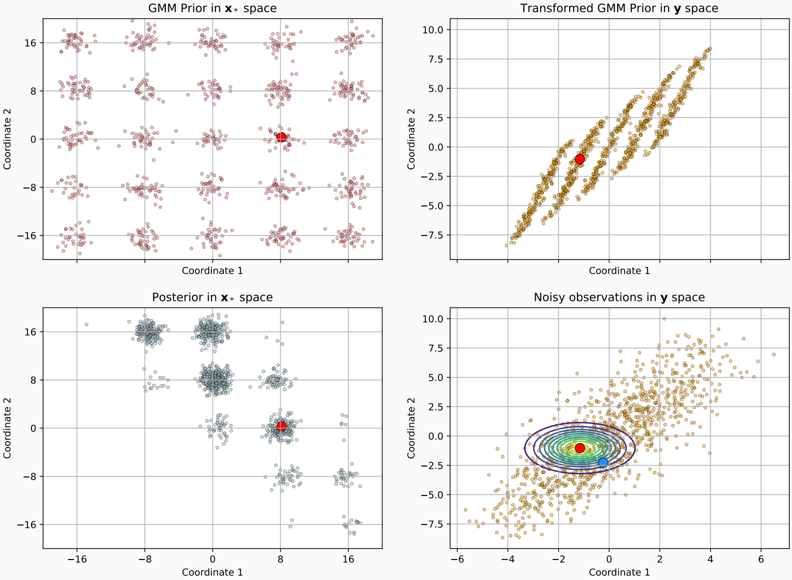
GMM Experiment — Setup

Consider a Gaussian mixture model (GMM), p_{data}, with 25 equally weighted ($\omega_{i,j} \propto 1$) d_x-dimensional Gaussian random variables with means $\mu_{i,j} := (8i, 8j, \ldots, 8i, 8j) \in \mathbb{R}^{d_x}$ for $(i,j) \in \{-2, \ldots, 2\}^2$ and unit variance. We generate some d_y-dimensional measurement, $\mathbf{y}$, according to the following process:

- Draw $\tilde{A} \sim \mathcal{N}(0,1)^{d_y \times d_x} \in \mathbb{R}^{d_y \times d_x}$ and compute its SVD decomposition, $USV^\top$.
- Sample $s_{i,j} \sim \mathcal{U}[0,1]$ for $(i,j) \in \{-2, \ldots, 2\}^2$.
- Set $A := U\, diag(\{s_{i,j}\})\, V^\top$.
- Draw $\mathbf{x}_* \sim p_{\text{data}}$ and set $\mathbf{y} := A\mathbf{x}_* + \sigma_y \epsilon,\ \epsilon \sim \mathcal{N}(\mathbf{0}_{d_y}, \mathrm{I}_{d_y})$.

The goal of the inverse problem is to take $\mathbf{y}$ (with A and σ_y known), and use it to infer $\mathbf{x}_*$. We consider $d_y < d_x$, making the problem ill-posed.

GMM Experiment — Setup Visualised



GMM Experiment — Prior Model and TMPD Guidance

GMM Experiment — `SMCDiffOpt`

GMM Experiment — Performance Comparison

d_x	d_y	SMCDiffOpt	DPS	ΠIGD	TMPD
8	1	**1.35 ± 1.0**	8.18 ± 7.5	3.16 ± 2.72	3.31 ± 2.86
	2	**0.55 ± 0.43**	2.72 ± 2.62	0.94 ± 0.89	1.34 ± 1.25
	4	0.19 ± 0.09	0.95 ± 0.9	**0.09 ± 0.03**	0.37 ± 0.33
80	1	**1.65 ± 1.41**	5.03 ± 4.63	3.08 ± 2.71	2.42 ± 1.71
	2	**1.19 ± 1.07**	3.14 ± 3.02	1.68 ± 1.62	1.35 ± 1.22
	4	1.11 ± 0.93	1.32 ± 1.21	**0.84 ± 0.79**	1.14 ± 0.9

Table 1: Sliced-Wasserstein distances of samples from particle samples.

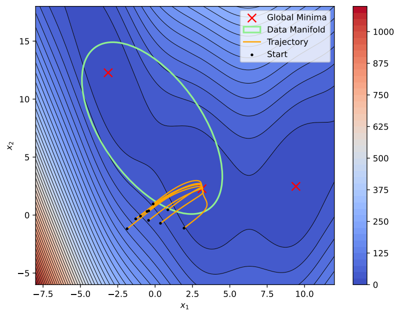
Branin Experiment — Setup

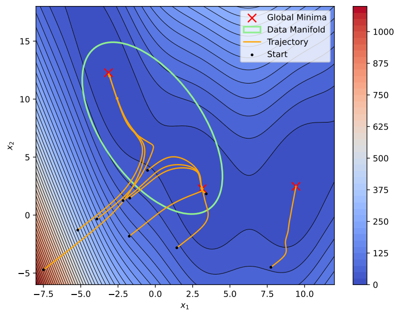
Consider optimising the Branin function as an objective over some constrained region (the data manifold) as detailed in Kong et al. (2024). The Branin function is defined by

$$f(x_1, x_2) = a(x_2 - bx_1^2 + cx_1 - r)^2 + s(1 - t)\cos(x_1) + s \tag{1}$$

where $a = 1$, $b = \frac{5.1}{4\pi^2}$, $c = \frac{5}{\pi}$, $r = 6$, $s = 10$, $t = \frac{1}{8\pi}$. It has three global minima located at $(-\pi, 12.275)$, $(\pi, 2.275)$ and $(9.42478, 2.475)$, with a value of 0.397887. We consider some uniform prior sampling distribution, p_{data}, over an ellipse region centred at $(-0.2, 7.5)$ with semi-axis lengths $(3.6, 8)$, tilted 25°. This region covers two of the global minima.

Branin Experiment — Pitfalls of Gradient Optimisation





Branin Experiment — SMCDiffOpt

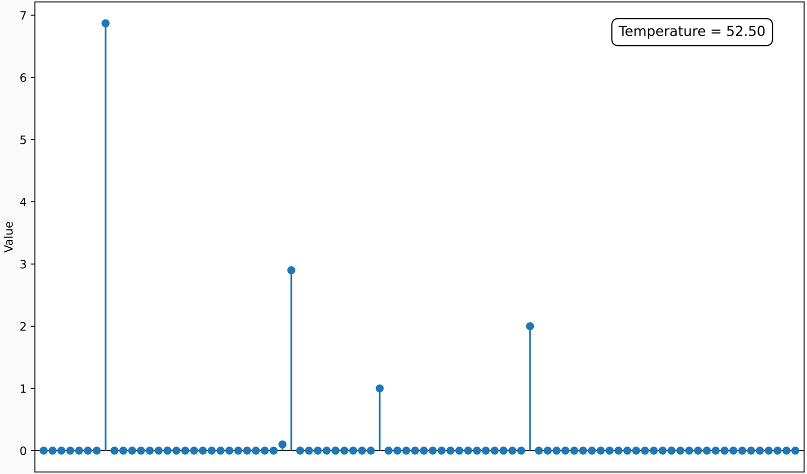
Black-Box Experiment — Motivation

Text

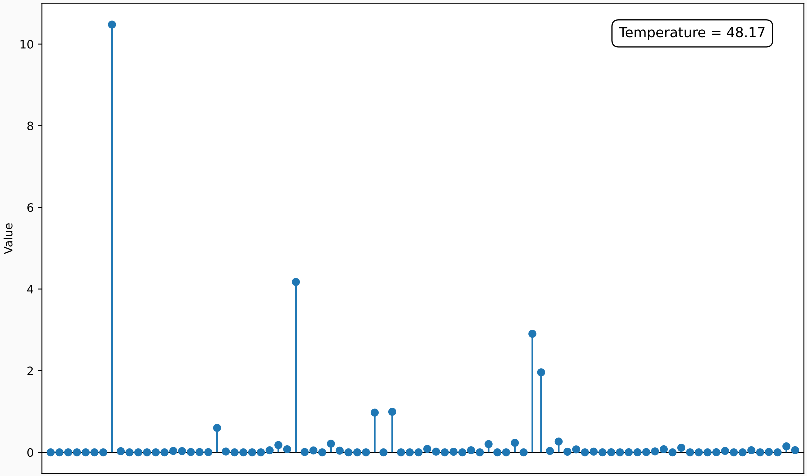
Black-Box Experiment — Setup

Text

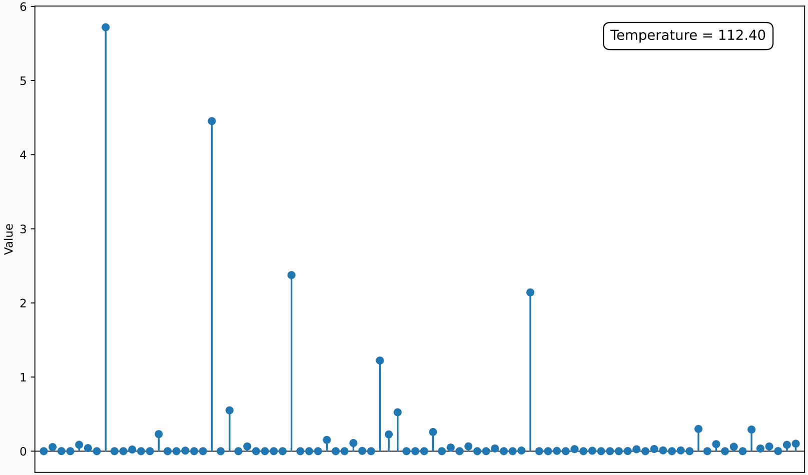
Black-Box Experiment — Real Sample



Black-Box Experiment — Generated Sample



Black-Box Experiment — Optimised Sample



Black-Box Experiment — Performance Comparison

CbAS	CMA-ES	Gradient Ascent	MINs	REINFORCE	DDOM	DiffOpt	**SMCDiffOpt**
0.433 ± 0.027	0.474 ± 0.021	0.510 ± 0.009	0.473 ± 0.003	0.483 ± 0.015	0.560 ± 0.044	0.614 ± 0.029	**0.644 ± 0.024**

Table 2: Normalised scores for SuperConductor experiment.

CbAS	CMA-ES	Gradient Ascent	MINs	REINFORCE	DDOM	DiffOpt	**SMCDiffOpt**
80.082 ± 4.933	87.774 ± 3.965	94.414 ± 1.719	87.483 ± 0.481	89.351 ± 2.708	103.600 ± 8.139	113.545 ± 5.322	**119.059 ± 4.497**

Table 3: Raw scores for SuperConductor experiment.

Conclusion

Future Work

Text

Final Remarks

Text

Thank you!

Akyildiz, Ö. D., Crisan, D., & Míguez, J. (2020). Parallel sequential Monte Carlo for stochastic gradient-free nonconvex optimization. *Statistics and Computing*, *30*(6), 1645–1663. https://doi.org/10.1007/s11222-020-09964-4

Anderson, B. D. (1982). Reverse-time diffusion equation models. *Stochastic Processes and their Applications*, *12*(3), 313–326. https://doi.org/10.1016/0304-4149(82)90051-5

Del Moral, P. (2011). Central Limit Theorems. In *Feynman-Kac formulae: Genealogical and interacting particle systems with applications* (pp. 291–330). Springer OCLC: 1063493341.

Dou, Z., & Song, Y. (2023). Diffusion Posterior Sampling for Linear Inverse Problem Solving: A Filtering Perspective. Retrieved June 8, 2024, from https://openreview.net/forum?id=tplXNcHZs1

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. https://doi.org/10.48550/ARXIV.2006.11239

Kong, L., Du, Y., Mu, W., Neklyudov, K., De Bortoli, V., Wang, H., Wu, D., Ferber, A., Ma, Y.-A., Gomes, C. P., & Zhang, C. (2024, April 29). *Diffusion Models as Constrained Samplers for Optimization with Unknown Constraints*. arXiv: 2402.18012 [cs]. Retrieved July 26, 2024, from http://arxiv.org/abs/2402.18012