NESTED SLICE SAMPLING

VECTORIZED NESTED SAMPLING

DAVID YALLUP, NAMU KROUPA, WILL HANDLEY

Find out more

handley-lab.co.uk/nested-sampling-book

yallup@github.io

david.yallup@gmail.com

yallup

NESTED SAMPLING

Nested Sampling[1,2] estimates the normalising constant associated with $P(x)$,

$$P(x) = \frac{e^{-\beta E(x)}\Pi(x)}{Z},$$

$$Z(\beta) = \int e^{-\beta E(x)}\Pi(x)dx\,.$$

A particle Monte Carlo method[8], construct series of interpolating reference distribution $\Pi(x)$ to target $P(x)$.

Annealing a common path, NS constructs unique path of constrained regions of prior density. Constraint term arises from increasing energy (likelihood) level[7].

Unique challenges formed with respect to other popular approaches, How do we draw samples from constrained region of a target density[5,6]? How do we use gradient information in such a setting[4]?

Algorithm:

1. Population of initial particles $\{x_i\}_0^m$
2. Sort $\{x_i\}_0^m$ by $E(x)$ removing lowest k particles (defines E_{min})
3. Boolean **weight**, $w_i \leftarrow E(\{x_i\}_0^{m-k}) > E_{min}$
4. Duplicate k particles **resampled** with w_i
5. **Propagate** with short chains until $\{x_i'\}_0^k$ decorrelated new points achieved
6. Return $\{\{x_i\}_0^{m-k}, \{x_i'\}_0^k\}, E_{\min}$

Action of this outer Kernel yields a set of increasing ordinates in energy,

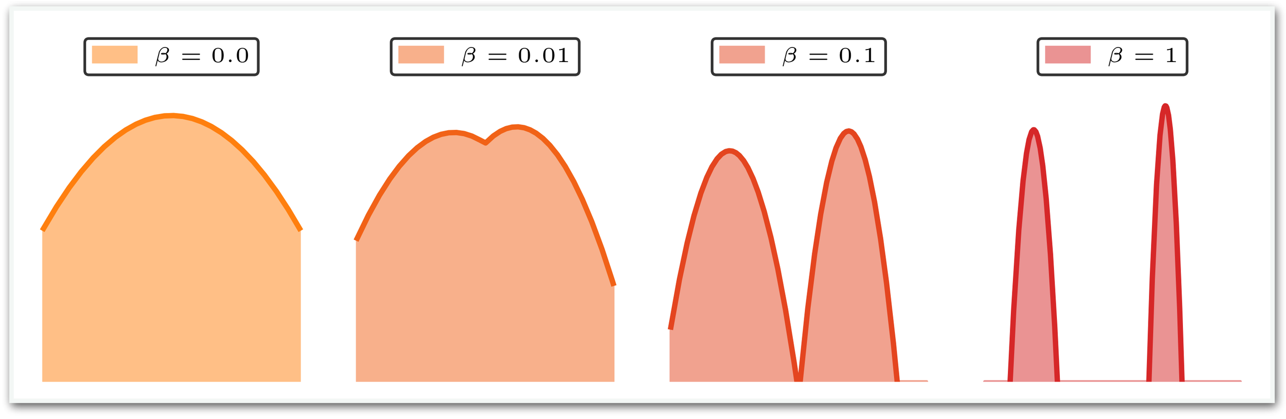
$$-E_1 < -E_2 < \ldots < -E_n\,,$$

And particles that give a geometric estimate of volume decrease associated with transition,

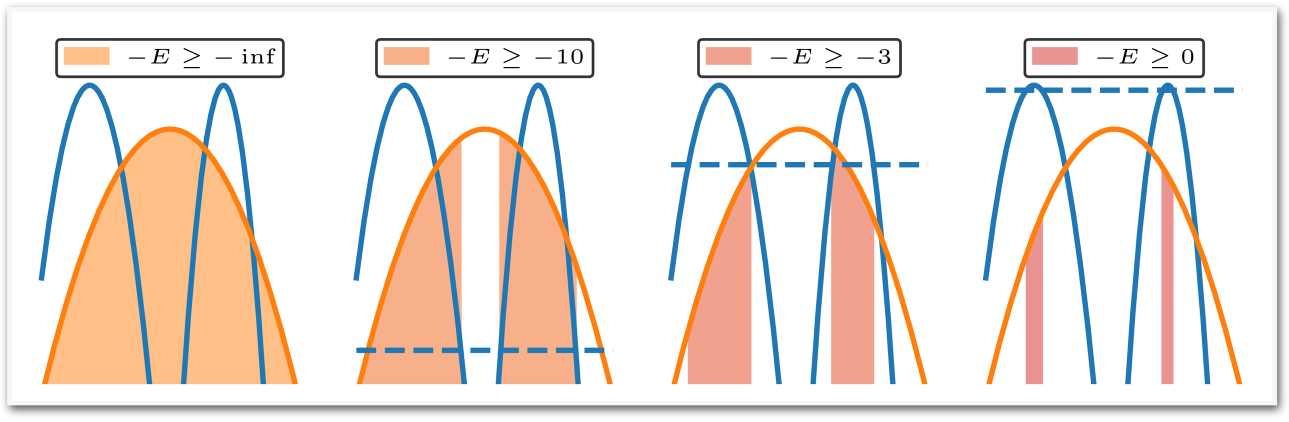
$$\Delta \ln X_i = -k/m\,,$$

These sum to form a Lebesgue integral,

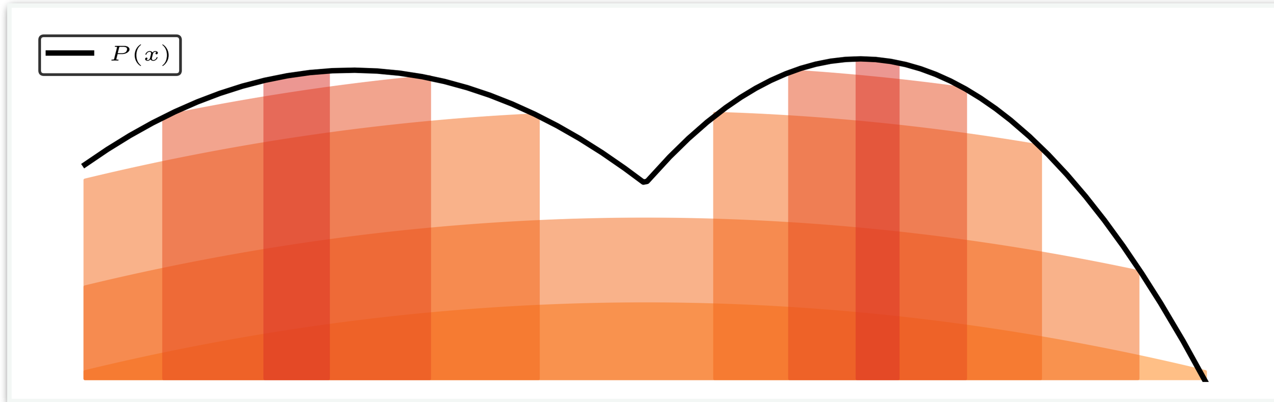
$$Z = \sum_{i=1}^{n} \exp(-\beta E_i)\Delta X_i\,.$$



Annealing path of distributions

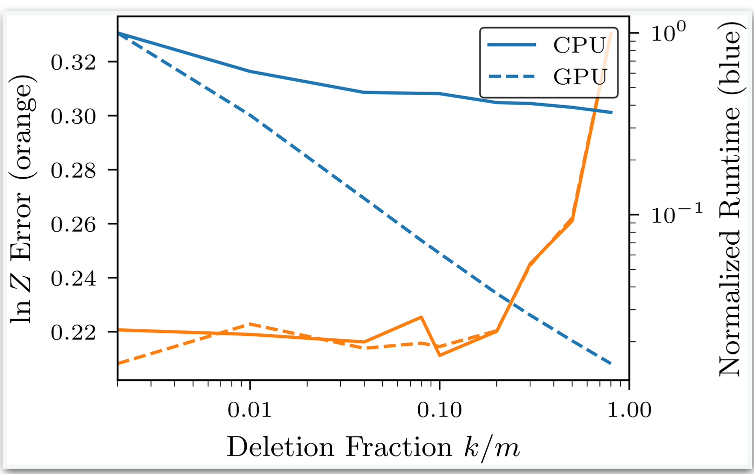


Nested Sampling algorithm running Left to Right



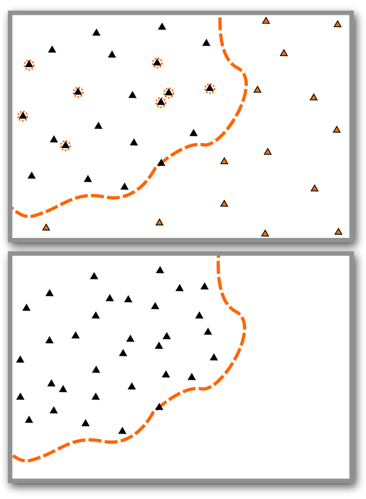
Nested sampling reconstruction of Posterior and Volume estimate

VECTORIZATION



Vectorization speedup as percentage of population deleted

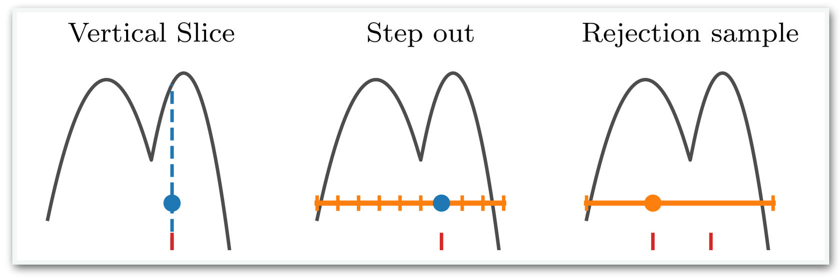
Achieves high GPU utilization for parallel particle updates, but the variable cost of constrained sampling introduces inefficiencies and limits perfect scaling.



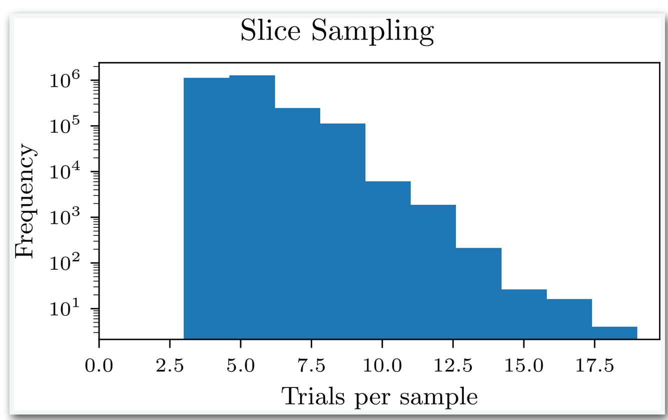
Reweight, Resample, Propagate

Batch updates (k/m) trade evidence accuracy for GPU Speed: Updating particle subsets yields significant walltime speedups on GPUs, incurring a manageable cost in evidence estimate accuracy.

SLICE SAMPLING



Action of Slice Sampling kernel to generate new samples



Suppressed tail in walk length c.f. Random walks for constrained sampling

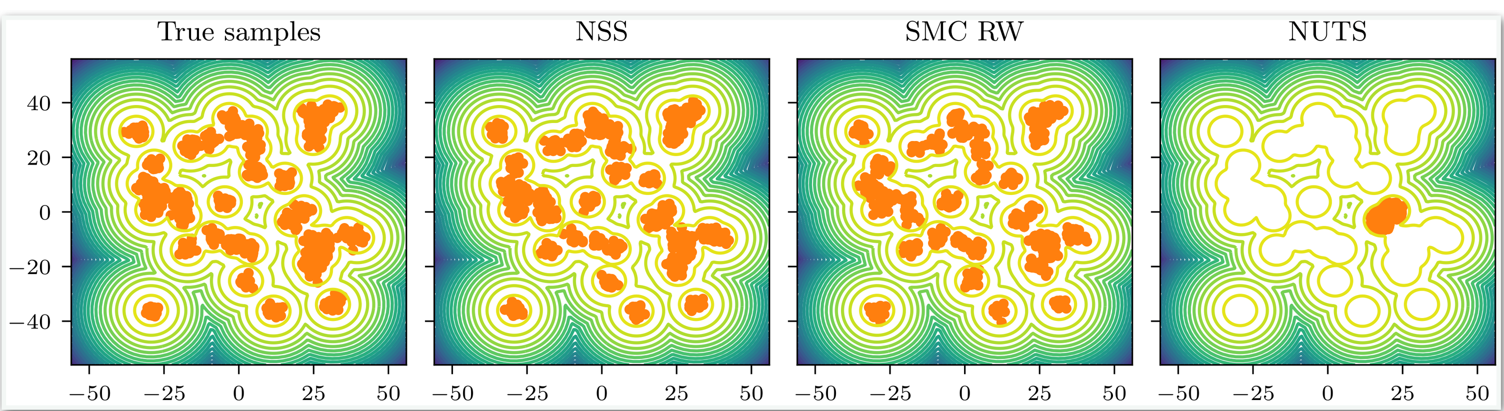
Slice Sampling [3] enables Constrained NS Step: Efficiently samples points within the required likelihood constraint essential for Nested Sampling.

Adaptive Slicing is Key: Performance hinges on adapting slice proposals (currently population-based) to match the target likelihood level.

Neural Slice Proposals: Explore ML models to learn highly adaptive slice proposals for improved NS

RESULTS

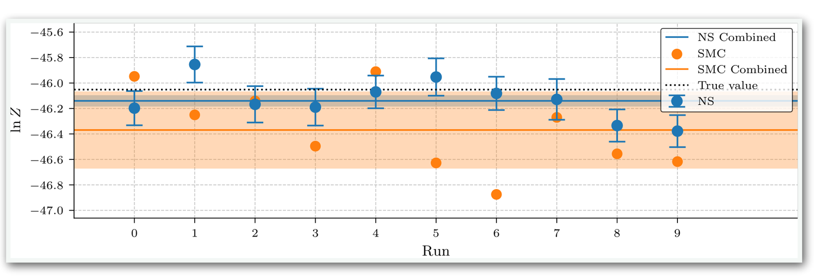
handley-lab/blackjax



Pedagogical Mixture of Bivariate Gaussians

NS compares favourably with SMC for State of the Art performance on certain classes of challenging inference tasks.

Probabilistic volume estimates produced by NSS give unique advantages with respect to SMC[9]. Evidence estimation with error out of the box.

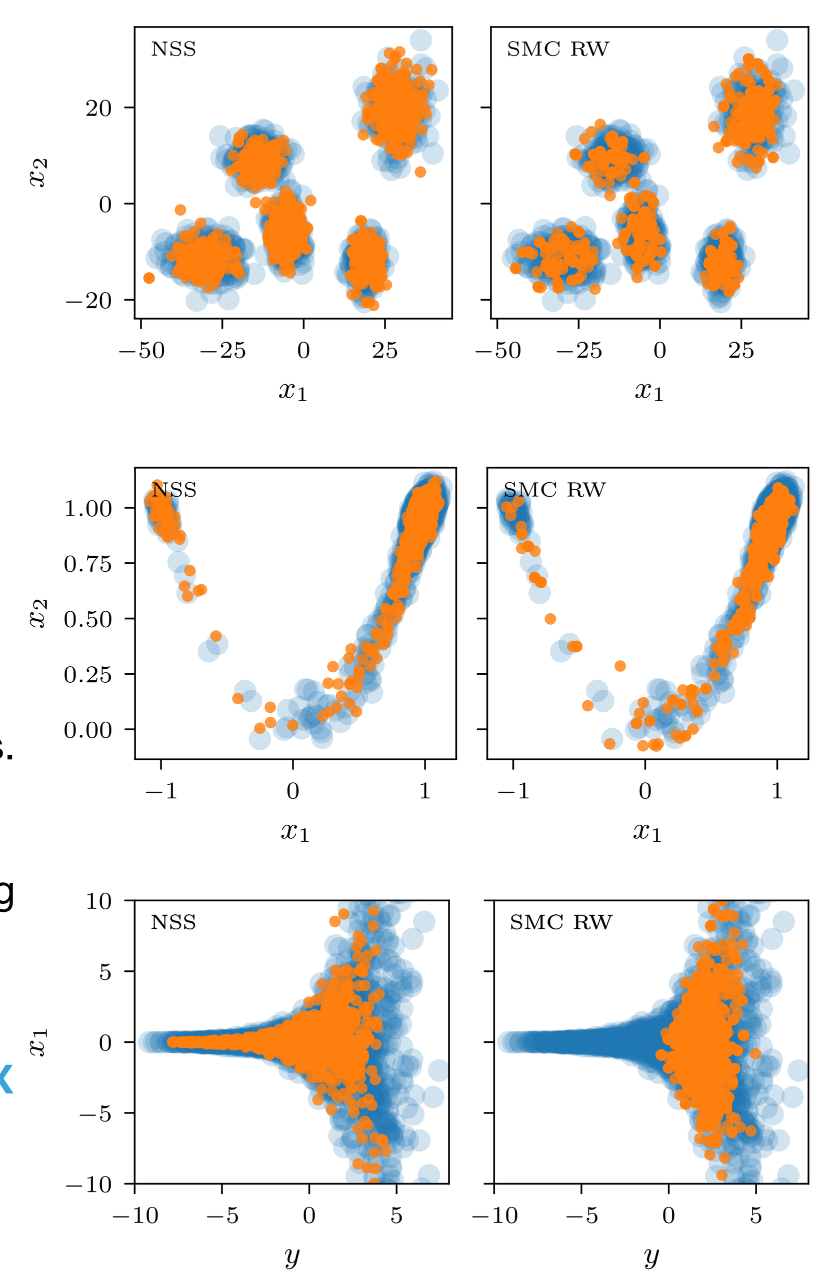


Probabilistic volume estimation

Robust Bayesian Evidence Estimation: Excels on complex geometries, scaling efficiently demonstrated to $\mathcal{O}(100)$ dimensions.

Gradient-Free State-of-the-Art Sampling: Achieves scalable, leading performance where gradients are unavailable or intractable.

Open Source & Reusable BlackJAX Kernel: Public code integrated with BlackJAX [10,11], featuring a modular core for broader use.



10D Synthetic challenging posterior problems

KEY TAKEAWAYS

1. High-Performance Nested Sampling in JAX/blackJAX
2. Massively Parallel Execution: Achieves significant speedups on GPUs/TPUs via vectorization.
3. Runtime Competitive with State-of-the-Art SMC: Matches or exceeds performance of leading Population MC methods on benchmark tasks.
4. New Frontiers for ML-Enhanced Sampling: Opens doors for integrating learned proposals, surrogates; Ideas welcome!

References:

1. Nested sampling for general Bayesian computation, Skilling
2. Nested sampling for physical scientists, Ashton et al.
3. Slice sampling, Neal
4. Improving Gradient-guided Nested Sampling, Lemos et al.
5. Multimodal nested sampling, Feroz et al.,
6. polychord: next-generation nested sampling, Handley et al.
7. Unbiased and Consistent Nested Sampling, Salomone et al.
8. Sequential Monte Carlo samplers, Del Moral et al.
9. An Adaptive Sequential Monte Carlo Sampler, Fearnhead et al.
10. BlackJAX: Composable Bayesian inference in JAX, Cabezas et al.
11. JAX: composable transformations of Python+NumPy programs, Bradbury et al

UNIVERSITY OF CAMBRIDGE