Unfreezing Topology With Nested Sampling

Dan Hoying

(Collaboration between Urs Wenger and myself)

August 1, 2023

Intro/Background – Frozen Topology

- At low temperatures (meaning a small change in energy generates a large change in accessible microstates) some theories have observables which freeze:
- the observable (e.g. topology) does not change much in value if we let our system evolve
- Substantial free energy barriers make it unlikely to find our system in different topological region.
- System eventually tunnels, but not efficiently (takes a long time of running monte-carlo evolution).

Intro/Background – (Quenched) Schwinger Model

- QED in $1+1$ D; $U(1)$ Gauge theory
- exhibits frozen topology
- A usual toy model for QCD

Intro/Background – Solutions for Topological Freezing

- Ignore the problem (some observables maybe don't care about topology/topology invariant). Not usually an option (finite volume errors).
- Open boundary conditions: different topologies come in from outside the lattice (but then one must avoid the boundaries)
- Other ideas: Metadynamics[1]
- Nested Sampling[2]

Nested Sampling (NS)

- Originally, Nested Sampling (NS) is a way to do Bayesian inference:
- Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$Z \equiv \sum_A P(B|A)P(A)$$

Nested sampling integral:

$$Z = \int L(\vec{\theta})\pi(\vec{\theta})d\vec{\theta}$$

$$= \int L(X)dX$$

$$P(A) \leftrightarrow \pi(\vec{\theta})d\vec{\theta}, dX; \; P(B|A) \leftrightarrow L; \; Z \leftrightarrow P(B)$$

Picture of Nested Sampling
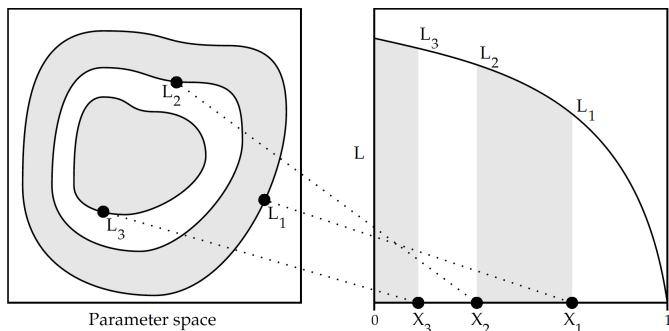

Figure 3: Nested likelihood contours are sorted to enclosed prior mass X.

Figure: From Skilling's original paper[3]

Nested Sampling for Temperature/Energy Dependent Observables

- Boltzmann factor as likelihood: $P(T|E) = exp(-E/T)$
- prior: $P(E) = \rho(E)dE$ with $\rho(E)$ the density of configurations at energy E
- Posterior (what we infer): $P(E|T)$
- Think we want to calculate: Observable O at temperature T: $O(T) = \sum_E O(E)P(E|T)$
- What we need: a list of energies at constant densities, and how big our step sizes are (step size: $\rho(E)dE$)

Nested Sampling – Key Idea: Uniform Sampling Implies (Average) Step Size is Known

- We sample (Nlive samples) uniformly from the (energy) space of all possible lattice configurations
- The expected volume of lattices with energies greater than the sample with the highest energy is equivalent to the expected max value of Nlive random real numbers drawn from the interval $[0, 1]$: $\frac{Nlive}{Nlive+1}$.
- Define shrinkage factor t, such that $dX = (1 - t)X$; $X_{\text{new}} = tX_{\text{old}}$
- $\Rightarrow \langle t \rangle = \frac{Nlive}{Nlive+1}$
- Minimize systematic error from $\rho(E)$ not being constant over interval dE by increasing Nlive.

Nested Sampling – Procedure

1. Setup: Uniformly sample an Nlive number of configurations (points).

2. Loop: Find the point with the max energy. This is our first dead point. Save the energy of this iteration i, E_i.

3. Replace the dead point with one of our live points. Update it enough to de-correlate it from the copied point.

4. Find the point with the new max energy among our live points. /Loop.

5. (One possible) stopping condition: stop if $\Delta Z / Z < tol$ (user specified)

6. After the NS loop, do the final integral via the trapezoidal rule:

$$O(T) = \frac{1}{Z(T)} \sum_i \frac{1}{2} exp(-E_i/T) (X_{i-1} - X_{i+1}) O(E_i)$$

Caveats 1/3

- The volumes X_i we use are estimates. We only obtain the true $O(T)$ after many NS runs (with some statistical error bar).
- The volume factor $(X_{i-1} - X_{i+1})$ goes to zero as Nlive $\to \infty$ (step size $\to 0$); As Nlive $\to 0$, the total volume X becomes large for a small number of steps. The volume occupied by energies relevant to T may become a negligible (numerically zero) portion of that volume.
- Too many live points sends our step size to zero and our compute time to infinity.
- Too few live points implies:
  1. Systematic error from coarse density resolution (as mentioned)
  2. Volume of energies relevant to the temperature T we are interested in may be negligible ($\Rightarrow$ volume factor goes to zero)
  3. Our observable freezes: the live points either become correlated or under-sample from all the relevant regions.

Caveats 2/3

- Density $\rho(E)$ goes up for bigger lattices (E is extensive), meaning for a fixed number of live points, our average energy step size dE goes down.
- This implies we must run NS longer/need more compute for bigger lattices, or else we must (substantially) decrease the number of live points (giving systematic error).
- If Nlive is too small, we under-sample our observable regions, causing our observable to be frozen again.

Caveats 3/3

- Stopping condition: without being able to compute the ground state of our system, there is no stopping condition which can guarantee that we did not stop the NS run too early. There could always be a spike of likelihood we miss at some energy we did not run NS long enough to find.
- Missing modes: for finite Nlive, we can not guarantee that we did not miss some density spike of configurations. This is also a problem for Markov Chain Monte Carlo (MCMC).

Some Possible Solutions

- problem: too few live points, sending volume factor to zero. solution: Don't count the parts of the volume where the likelihood (Boltzmann factor) is zero (equivalent to re-scaling the density of the remaining energies by a constant amount). We only save dead point energies below a thermalization threshold set by the temperature we are interested in.

- problem: frozen observable (topology) as $V \to \infty$. possible solution: Cool the individual points by applying updates constrained not to raise the energy, until we get below our target energy. We apply updates constrained by our target energy until the points become uniformly distributed. We then run NS.

- problem: too slow of thermalization. solution: potentially the same as for MCMC: a more efficient update strategy.

Past Lattice Results (1/2)

- NS has been used already on the Potts Model[4], a type of simple $U(1)$ lattice model with a finite number of possible phases allowed at each site.

Past Lattice Results (2/2)
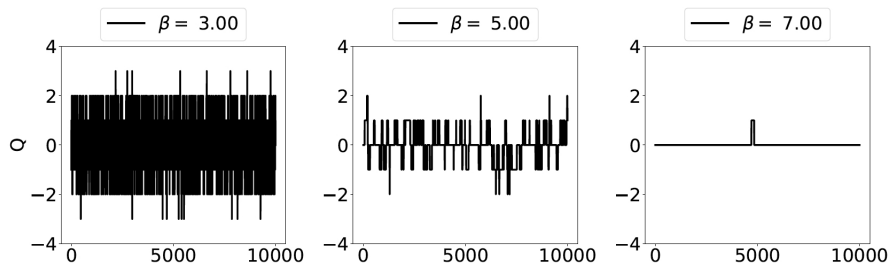
Figure 4.1: Monte Carlo history for topological charge Q. Quenched simulations with $L = L_t = 8$, were used here. We made $10'000$ thermalization steps and $10'000$ measurements afterwards. Between each measurement we skipped 100 metropolis update steps.

Figure: from thesis of Bühlmann, Patrick

Our (Very Preliminary) Result

8x8 Schwinger Model, Nested Sampling, Nlive=2000, E>0.53
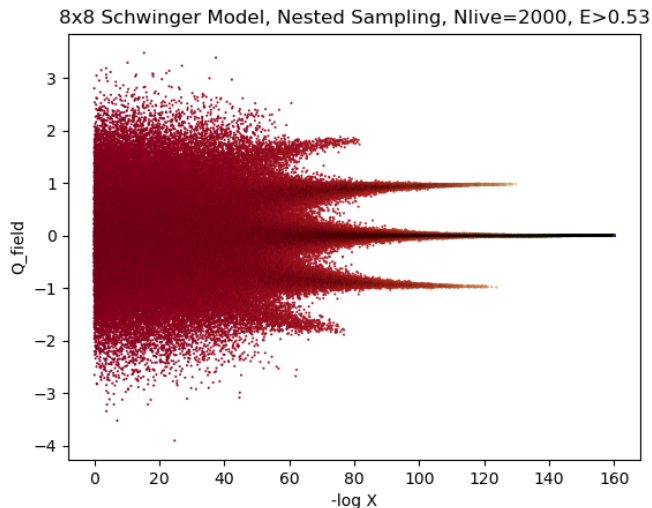
Conclusion(s)

- Nested Sampling (NS) may be useful for solving the problem of frozen observables/topological freezing, in some limited cases
- Potentially very useful for finite T/high T observables.
- In terms of topological freezing, NS is likely most useful as a conceptual tool, another way to think about/formulate the problem.
- Beyond the lattice: NS is useful as a general purpose tool for evaluating multi-dimensional integrals, and for use in Bayesian inference.

Thanks

Thanks!

References

[1] Philip Rouenhoff et al. "Metadynamics Surfing on Topology Barriers in the Schwinger Model". In: *PoS* LATTICE2022 (2022), p. 253. DOI: 10.22323/1.430.0253. arXiv: 2212.11665 [hep-lat].

[2] Greg Ashton et al. "Nested sampling for physical scientists". In: *Nature* 2 (2022). DOI: 10.1038/s43586-022-00121-x. arXiv: 2205.15570 [stat.CO].

[3] John Skilling. "Nested Sampling". In: *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. Ed. by Rainer Fischer, Roland Preuss, and Udo Von Toussaint. Vol. 735. American Institute of Physics Conference Series. Nov. 2004, pp. 395–405. DOI: 10.1063/1.1835238.

[4] Manuel J. Pfeifenberger, Michael Rumetshofer, and Wolfgang von der Linden. *Nested sampling, statistical physics and the Potts model*. 2017. arXiv: 1603.02516 [physics.comp-ph].