Fast nested sampling with deep neural network model emulators

Johannes Buchner, MPI for extraterrestrial Physics, Giessenbachstr, 85748 Garching, Germany
Jack O'Brien, Dept. of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA
Wolfgang Kerzendorf, Dept. of Computational Mathematics, Science, and Engineering, ibid

MAX PLANCK INSTITUTE
FOR EXTRATERRESTRIAL PHYSICS

MOTIVATION

Bayesian model comparison in astrophysics is popular with nested sampling

Various algorithms for sampling replacement points at each iteration:

- MultiNest heuristic ellipsoid clustering
- Slice sampling: default in PolyChord, for high-dimensions
- MLFriends: highly robust, default in UltraNest

Need to evaluate a likelihood function for a set of proposed points

Paradigm shift:

GPUs can process deep learning models equally fast for 1 input as for 1000 inputs

Novel algorithms that take advantage of this:

UltraNest's new vectorized MLFriends is capable of >100,000 model evaluations per second on a laptop

WHY DEEP LEARNING EMULATORS?

Elaborate simulations of physical systems can be approximated by deep learning model emulators, aka surrogate models with training data generated from the full model

model emulators are becoming more and more common in astronomy

LITERATURE EXAMPLE: DALEK

Deep neural network emulator presented in Kerzendorf et al. (2020), overcoming limitations of previous Gaussian process approach.

with radiative transfer model (TARDIS):

600s per evaluation

Deep learning emulator (Dalek)

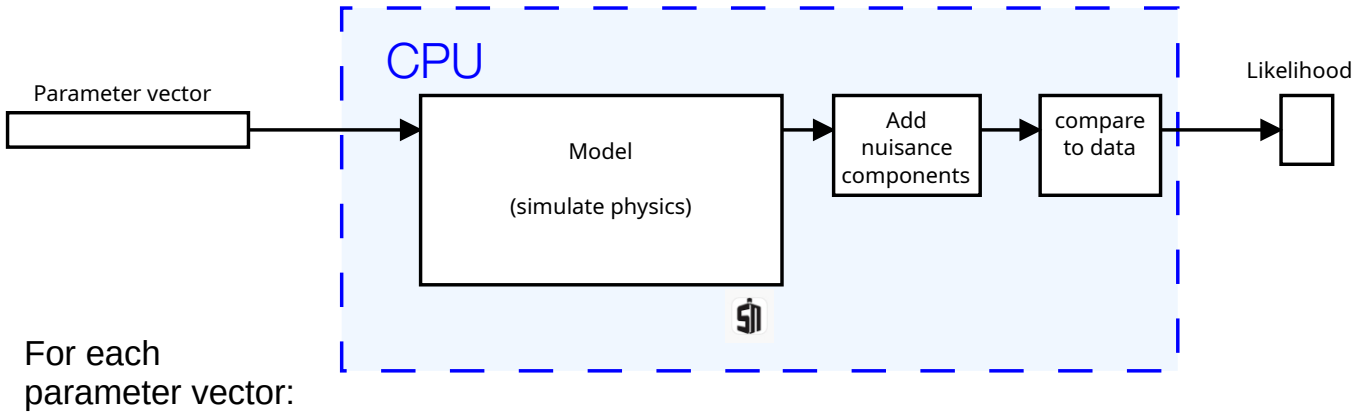
training: 70 CPU-days

0.01s per 1000 evaluations

Kerzendorf+22 presented an emulator predicting a spectrum with uncertainties (probabilistic Dalek)

COMPUTING MODELS

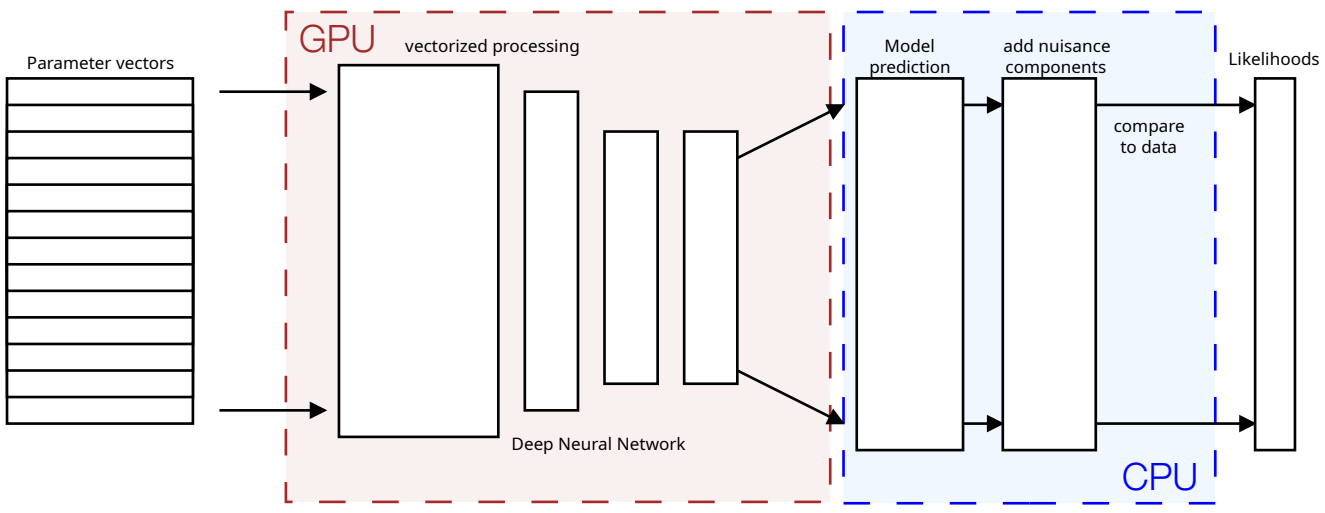
Old computing model:



For each parameter vector:

1. Set up simulation for given parameters
2. Predict model, for example a spectrum
3. Compute likelihood of observed data

New computing model:

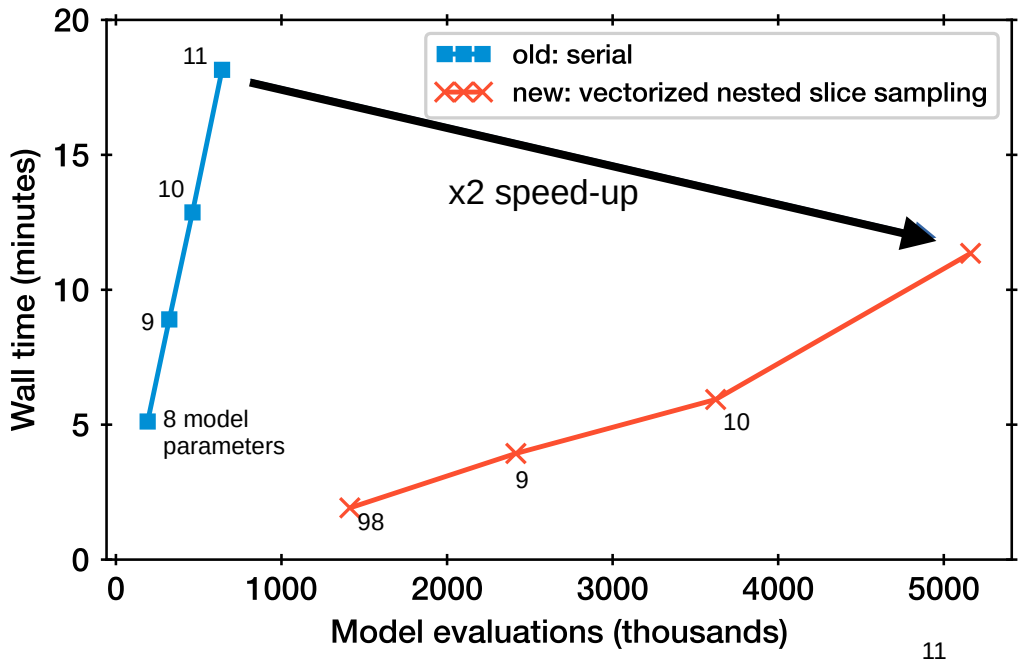


many model instances evaluated at once (vectorization), without much additional cost

Within UltraNest, we implemented vectorized MLFriends, vectorized Gaussian Random Walk Metropolis, and a **vectorized slice sampler**. The vectorized slice nested sampling maintains an ensemble of live points undergoing slice sampling. Care is needed to maintain detailed balance.

RESULTS

with radiative transfer model TARDIS:	420 years	per spectrum
nested slice sampling + Probabilistic Dalek:	18 minutes	per spectrum
vectorized nested slice sampling:	**11** minutes	per spectrum



Despite nominally making more model evaluations it is **twice as fast**. (the number of nested sampling iterations is the same)