Milestone 2

Nested Sampling with Peers

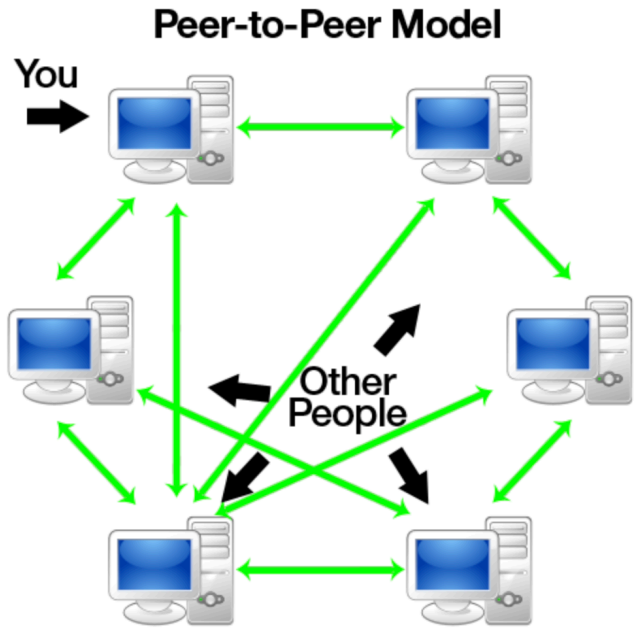
Student: Feishuang Wang
School of Data Science

Supervisor: Dr Brendon James Brewer
School of Computer Science

Project description

Nested Sampling is a Monte Carlo method (not necessarily MCMC) that was introduced by John Skilling in 2004.

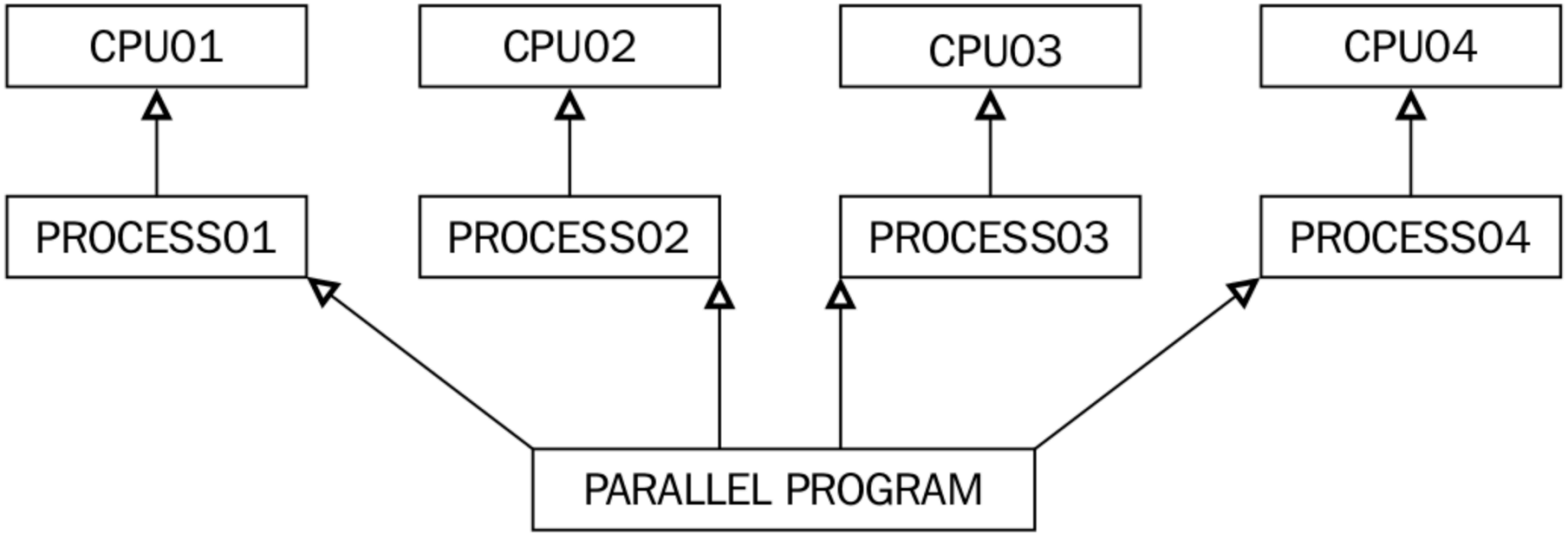
We introduce a general Monte Carlo method based on Nested Sampling, which name is Nested Sampling with peers.



Reasearch Questions

For computer code running only serially, the nested sampling method would considerably increase the wall clock time necessary to reach convergence.

How can we use the parallel programming to speed up the nested sampling and update its accuracy?



Literature review

Bayesian Analysis and nested sampling

1. Let X be a continuous random variable and g be any function. The expectation of g(X) can be obtained as follows (Russel, 2017; Dybowski et al., 2013):

$$E[g(X)] = \int_X g(X) f_X(x) \qquad (1.7)$$

2. The enclosed prior mass is usually highly right skewed with most of its mass concentrated at around X = 0 because in pratice, the maximum likelihood is located in a small region of the prior space (Russel, 2017).

3. The evidence, Z is greatly influenced by the size of the region of the prior mass where the posterior mass is most concentrated. H quantifies where this region is to a single value. According to Skilling (2006), this region is e^{-H} the fraction of the prior mass, which relates to the estimation of X discussed in the previous section.

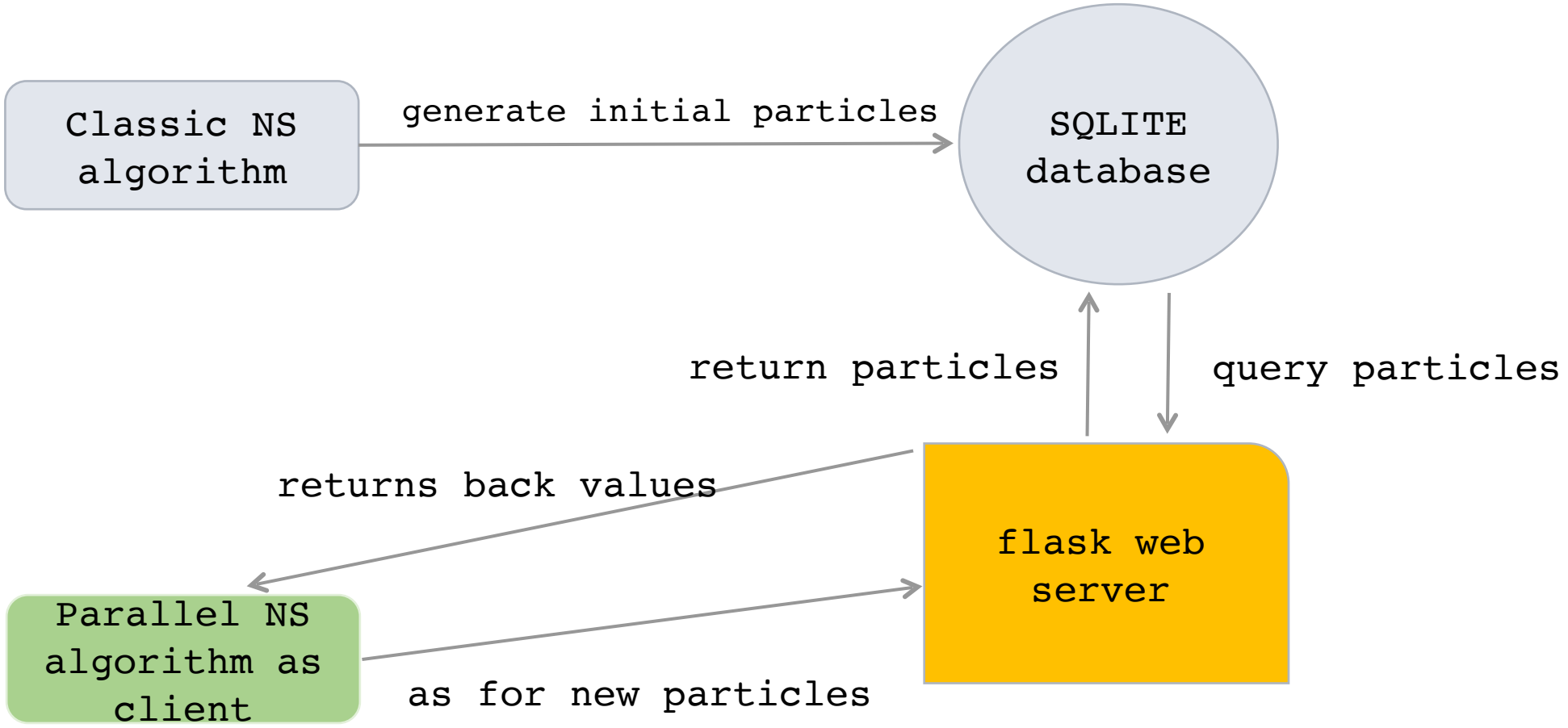
Literature review

Parallel Nested Sampling

1. While the usual approach with nested sampling is to discard and replace only one sample at each iteration ($r = 1$), larger values of r can be used. This idea was proposed previously by Burkoff, et al.

2. Nested sampling requires the r live samples with the lowest likelihood to be discarded and replaced. When $r = 1$, this can be accomplished by simply finding the minimum of the likelihood values in the live sample population. However, when $r > 1$, a more sophisticated minimization algorithm is required.

Methodology


Classic NS algorithm
generate initial particles
SQLITE database
return particles
query particles
returns back values
flask web server
Parallel NS algorithm as client
as for new particles

Milestones

- proprosed timeline to finish the dissertation
  - ✓ 2022.1.1 Building the basic programming environment:
    - ✓ Operating system: Ubuntu 16.04, macOS 10.15.4
    - ✓ Compiler: gcc 9.3.0
    - ✓ Language: C++17, Python 3.8
    - ✓ Editor: Visual Studio Code
  - ✓ 2022.1.10 Implement the basic nested sampling algorithm
  - ✓ 2022.2.10 Calculating MSE function
  - ✓ 2022.2.20 Implement the nested sampling with peers algorithm
    - ✓ 2022.3.10 Implement the Flask web server
    - ✓ 2022.4.6 Implement the updating particles from peers
    - ✓ 2022.4.18 Adding MSE function to the new model
  - ✓ 2022.5.4 Re-test the model by updating the database every iteration
  - ✓ 2022.5.15 Comparing models
  - ✓ 2022.5.22 Summarise the results
  - ✓ 2022.5.27 finish the milestones 2
  - ☐ 2022.6.27 writing the dissertation and submit
  - ☐ 2022.7.3 submit the poster file

References

[1] Farge Marie, *Wavelet Transforms and Their Applications to Turbulence*, Ann. Rev. Fluid Mech. volume 24, pages 395-457, 1992.

[2] Salmond Jennifer, *Vertical Mixing of Ozone in the Very Stable Nocturnal Boundary Layer*, PhD Thesis, University of British Columbia, 2001.

[3] Stull B. Ronald, *Introduction to Boundary Layer Meteorology*, Dordrecht; Boston: Kulwer Academic Publishers, 1988.

[4] Torrence Christopher, Compo Gilbert P., *A Practical Guide to Wavelet Analysis*, Bulletin of the American Meteorological Society volume 79, pages 61-78, 1998.