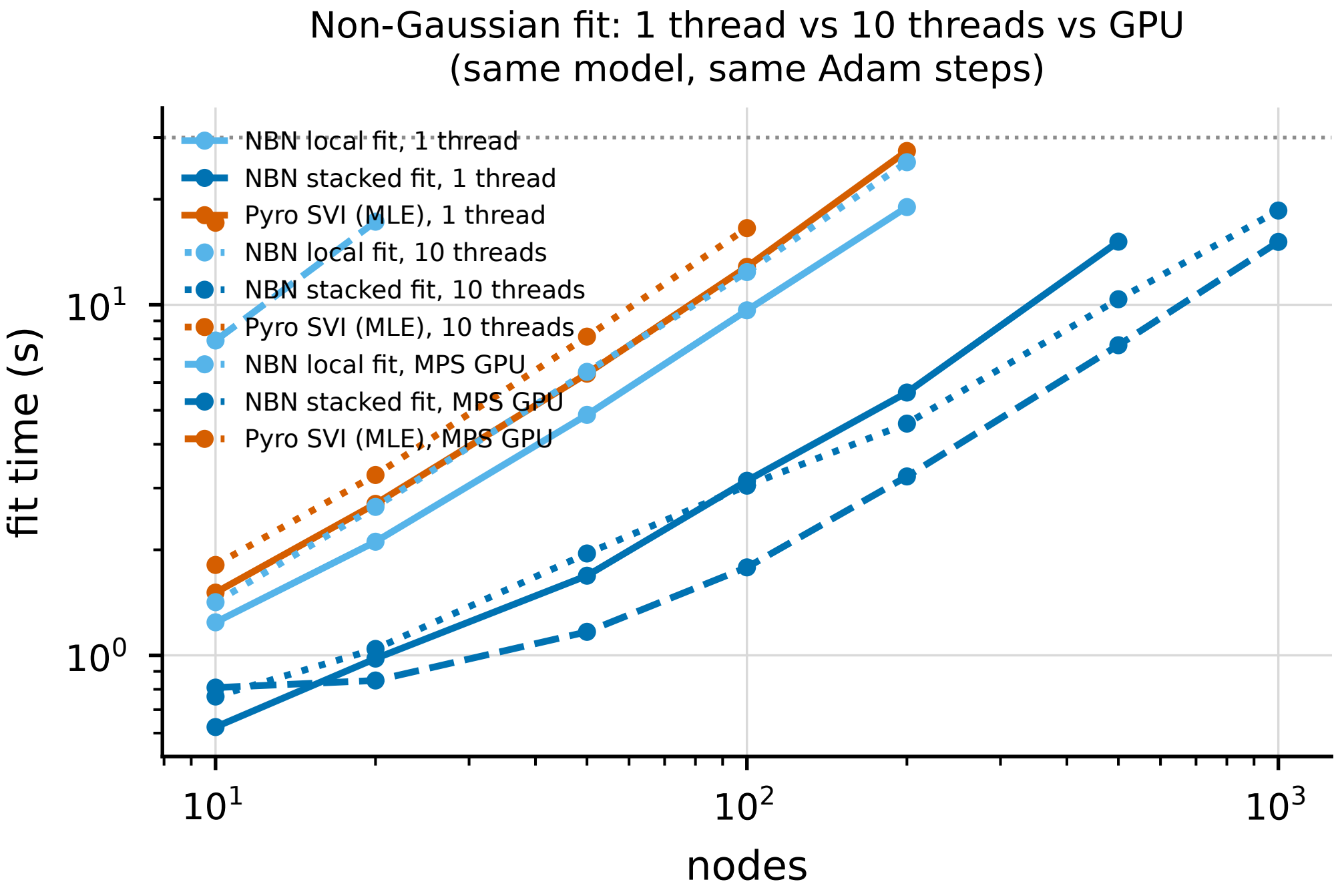

Non-Gaussian fit: 1 thread vs 10 threads vs GPU
(same model, same Adam steps)
NBN local fit, 1 thread
NBN stacked fit, 1 thread
Pyro SVI (MLE), 1 thread
NBN local fit, 10 threads
NBN stacked fit, 10 threads
Pyro SVI (MLE), 10 threads
NBN local fit, MPS GPU
NBN stacked fit, MPS GPU
Pyro SVI (MLE), MPS GPU
fit time (s)
$10^1$
$10^0$
$10^1$
$10^2$
$10^3$
nodes