Sparse-Matrix Belief Propagation
Supplemental Material

Reid Bixler
Department of Computer Science
Virginia Tech
Blacksburg, VA 24061

Bert Huang
Department of Computer Science
Virginia Tech
Blacksburg, VA 24061

A ADDITIONAL EXPERIMENTS

We present here additional experiments completed after initial submission of the main paper.

A.1 SEQUENTIAL BELIEF PROPAGATION

The power of sparse-matrix belief propagation comes from its ability to flexibly implement *parallel* belief propagation on various backends. However, in many settings, parallel belief propagation is not as desirable as sequential belief propagation. Sequential updates can often converge faster when hardware parallelization is limited.

We run a variation of our main experiments with with an additional comparison method: a sequential belief propagation (Sequential-CPU), which updates the belief for each variable then immediately updates all outgoing messages from that variable. Each variable and its outgoing messages is updated in sequence, reducing redundant computation within each belief propagation iteration.

We again run trials on multiple sized grids with randomized potential functions and different variable cardinalities. For cardinalities $c \in \{8, 16, 32\}$, sequential BP consistently faster than Python loop-based parallel BP but slower than either PyTorch or SciPy sparse parallel BP. For $c = 64$, sequential BP is the fastest of all CPU-based methods.

On these same grids, the GPU remains orders of magnitude faster than CPU-based methods, but the fact that sequential BP can be faster than sparse-matrix BP in some settings suggests that there is a tradeoff between massive parallelism and the benefits of sequential inference. A direction of future work may be to find abstractions that enable sequential updates, though a key challenge is that the macro-operations of bulk message and belief updates is what enables the sparse-matrix abstraction to work well.

A.2 ADDITIONAL BENCHMARKS

We will include experiments on established benchmark graphical models in a later revision of this supplemental document.

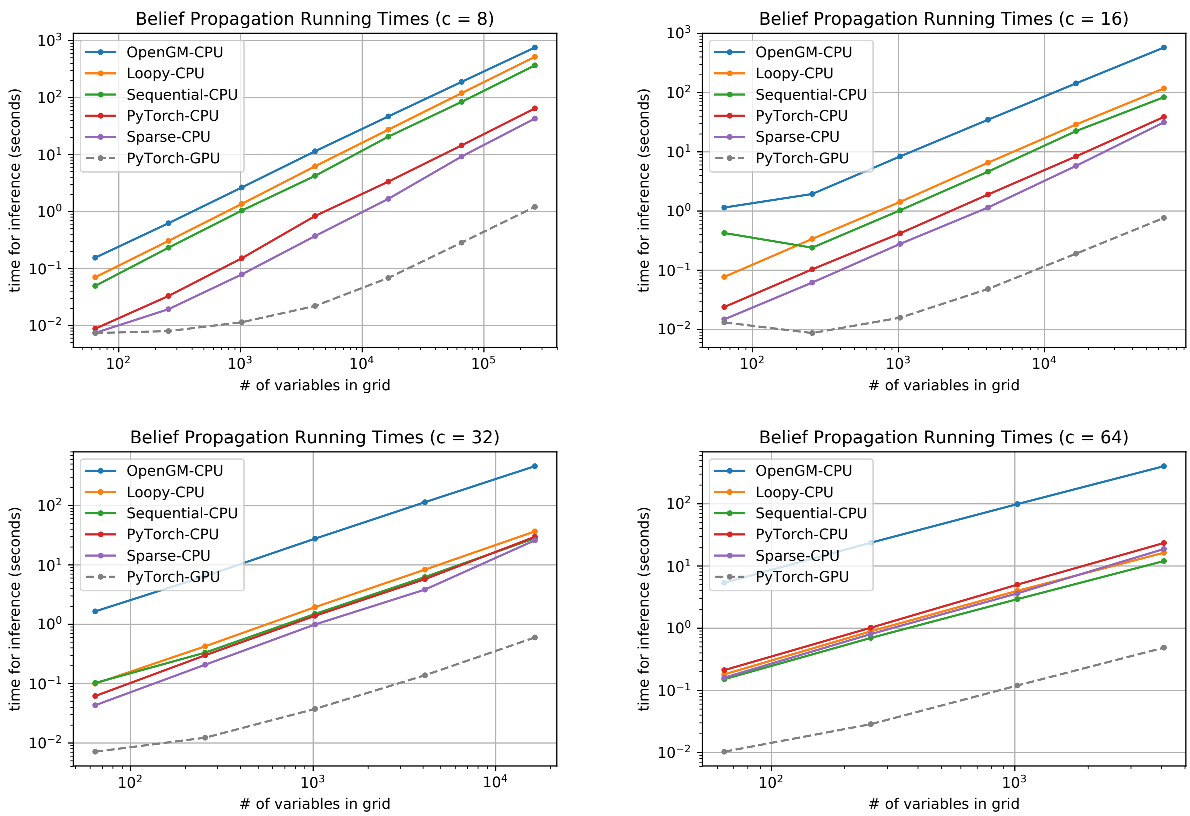

Figure 1: Running times for variants of belief propagation on a 4 Ghz i7 CPU and an Nvidia GTX 1080 GPU. In these experiments, we compare against CPU-based sequential belief propagation, which was not implemented in our main experiments.