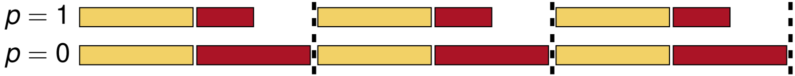
A Stochastic Performance Model for Pipelined Krylov Methods

Hannah Morgan[1], Matthew G. Knepley[2],
Patrick Sanan[3], L. Ridgway Scott[1]

[1] *University of Chicago*, [2] *Rice University*, [3] *Università della Svizzera italiana*

In a simple deterministic setting, during each iteration k of a Krylov method, each process takes a certain time c_p for local computation (yellow below) and w_p for waiting (red) so that the total running time is $T = \sum_k \max_p(c_p + w_p)$.



If we remove the synchronizations, the total running time is given by $T' = \max_p \sum_k (c_p + w_p)$ so that no speedup is achieved.

$$\text{speedup} = \frac{T}{T'} = \frac{\sum_k \max_p (c_p + w_p)}{\max_p \sum_k (c_p + w_p)} = 1$$

Instead, let T_p^k be a stochastic process time for process p and step k drawn from some distribution. The total time of a Krylov algorithm with synchronizations is given by $T = \sum_k \max_p T_p^k$ and the total time for the pipelined version is $T' = \max_p \sum_k T_p^k$.

$$\text{speedup} = \frac{E[T]}{E[T']} = \frac{\sum_k E[\max_p T_p^k]}{E[\max_p \sum_k T_p^k]} \approx \frac{E[\max_p T_p^k]}{\mu}$$

$\frac{E[T]}{E[T']}$ has an integral representation and we can compute it for common distributions.

uniform distribution: $\frac{E[T]}{E[T']} = \frac{2P}{P+1} < 2$ for all P

exponential distribution: $\frac{E[T]}{E[T']} > 2$ for all $P \geq 4$

To show that our stochastic model is a reasonable, we fit observed PGMRES (left) and PIPECG (right) runtimes to some underlying distributions.