Upper performance bound for blocked sparse MVM

Vuduc et al., Supercomputing 2002

Idea: Determine amount of data that needs to be transferred from memory.
Upper bound from memory bandwidth.

Amount of data for $r \times c$ BCSR:

A: $rc K_{r,c}$ doubles
$K_{r,c} + \frac{m}{r} + 1$ ints

x: n doubles

y: m doubles

$\Rightarrow 8(rc K_{r,c} + m + n) + 4(K_{r,c} + \frac{m}{r} + 1)$ bytes

Core 2: Memory bandwidth 2 B/cycle

$\Rightarrow$ runtime $\geq 4(rc K_{r,c} + m + n) + 2(K_{r,c} + \frac{m}{r} + 1)$ cycl

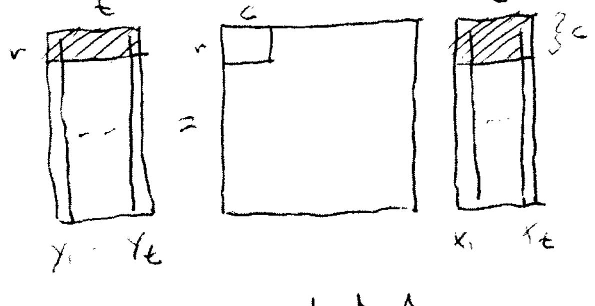
Bound based on computation:
Every $r \times c$ block incurs $2rc$ ops, total: $2rc K_{r,c}$ ops
Core 2 (no SSE): 2ops/cycle

$\Rightarrow$ runtime $\geq rc K_{r,c}$ cycles   looser bound

Multiple vectors in SMVM

$y_i = y_i + A x_i$, $i = 1..t$ Now reuse in A!

To exploit perform SMVMs interleaved and blocked



Similar to MMM