Challenges of efficient Fast Fourier Transform

Mahendra Verma
IIT Kanpur

http://turbulencehub.org

mkv@iitk.ac.in

Definition of Fourier Transform

x → k

$$f(x) = \sum_{k} f(k)\exp(ikx)$$

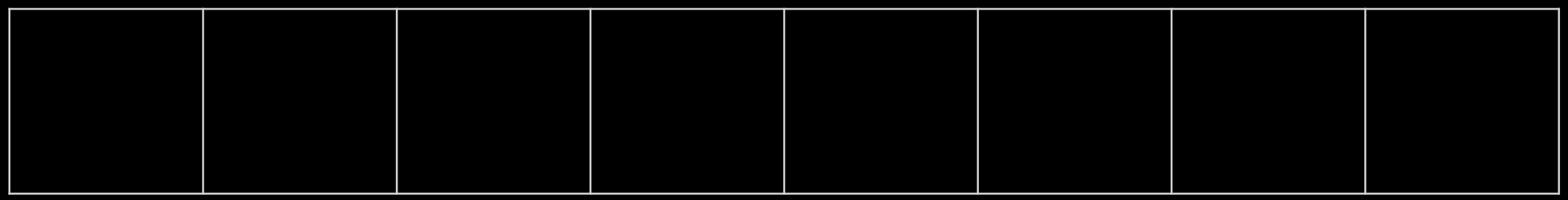
k → x

$$f(k) = \sum_{x} f(x)\exp(-ikx)$$

Time complexity: $O(N^2)$

Fast Fourier transform

Cooley and Tukey, 1965



0 1 2 3 4 5 6 7

FT(f) split into FT of even and odd indexed data.

Continue this operation at all levels.

Divide and conquer

Time complexity: O(N log N)

Why is it important?

Multiscale problems

Image processing

Structures at different scales…

Interactions across scales

Accurate derivative solver

$$f(x) = \sum_k f(k)\exp(ikx)$$

$$\frac{d}{dx} f(x) = \sum_k [ikf(k)]\exp(ikx)$$

Magnetic field reversal

Geomagnetism

Glatzmaier & Roberts
Nature, 1995

Polarity reversals after random time intervals (tens of millions of years to 50K years).

Last reversal took place around 780,000 years ago.

Complexity

Atmospheric simulation

1 km grid on Earth x 10 m along vertical

$40000 \times 40000 \times 10000 = 64 \times 10^{12}$

Does not fit in a single machine.

PARALLEL programming

Parallel FFT

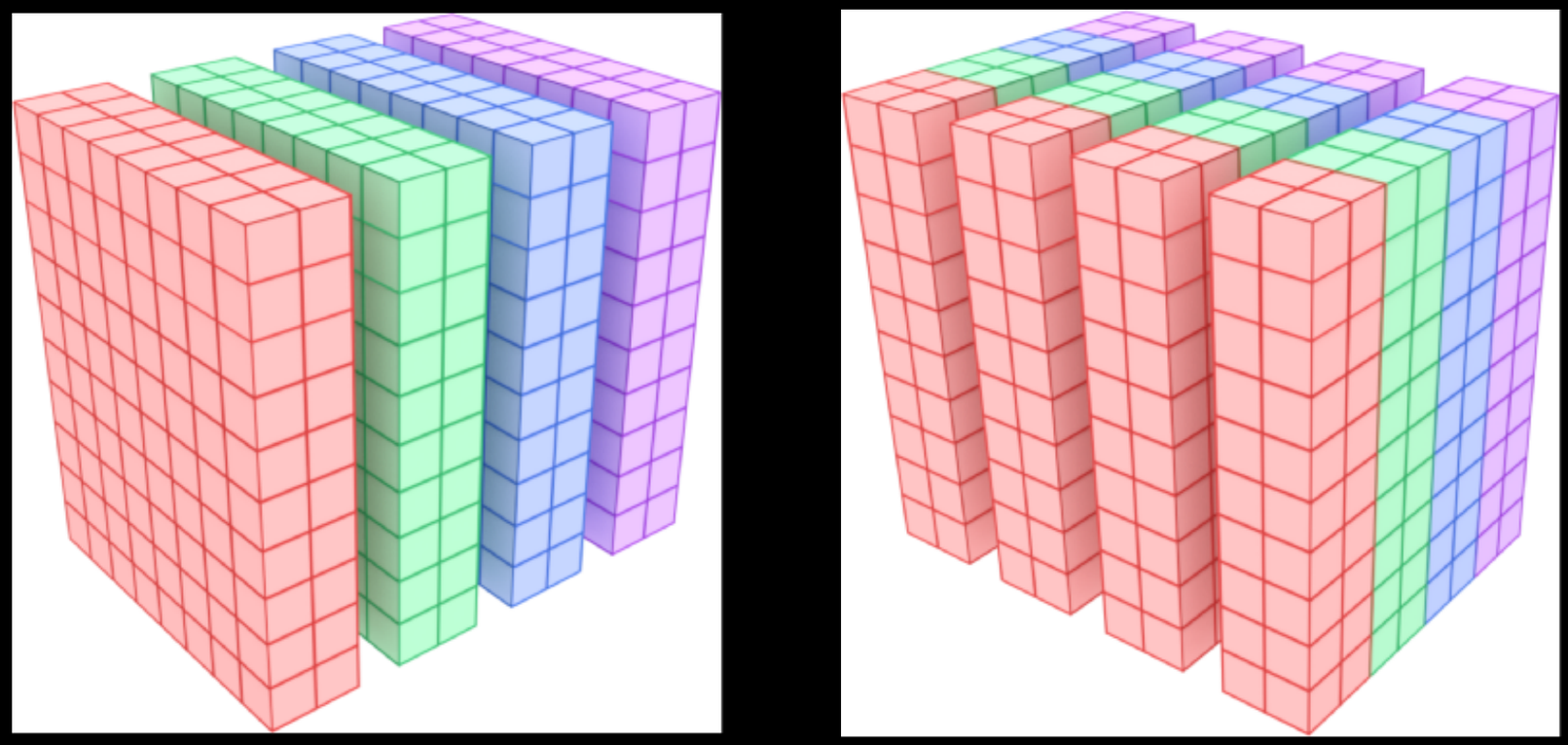
Parallel libraries

- FFTW
- P3DFFT
- FFTK
- PFFT
- Hybrid FFT
- GPU FFT

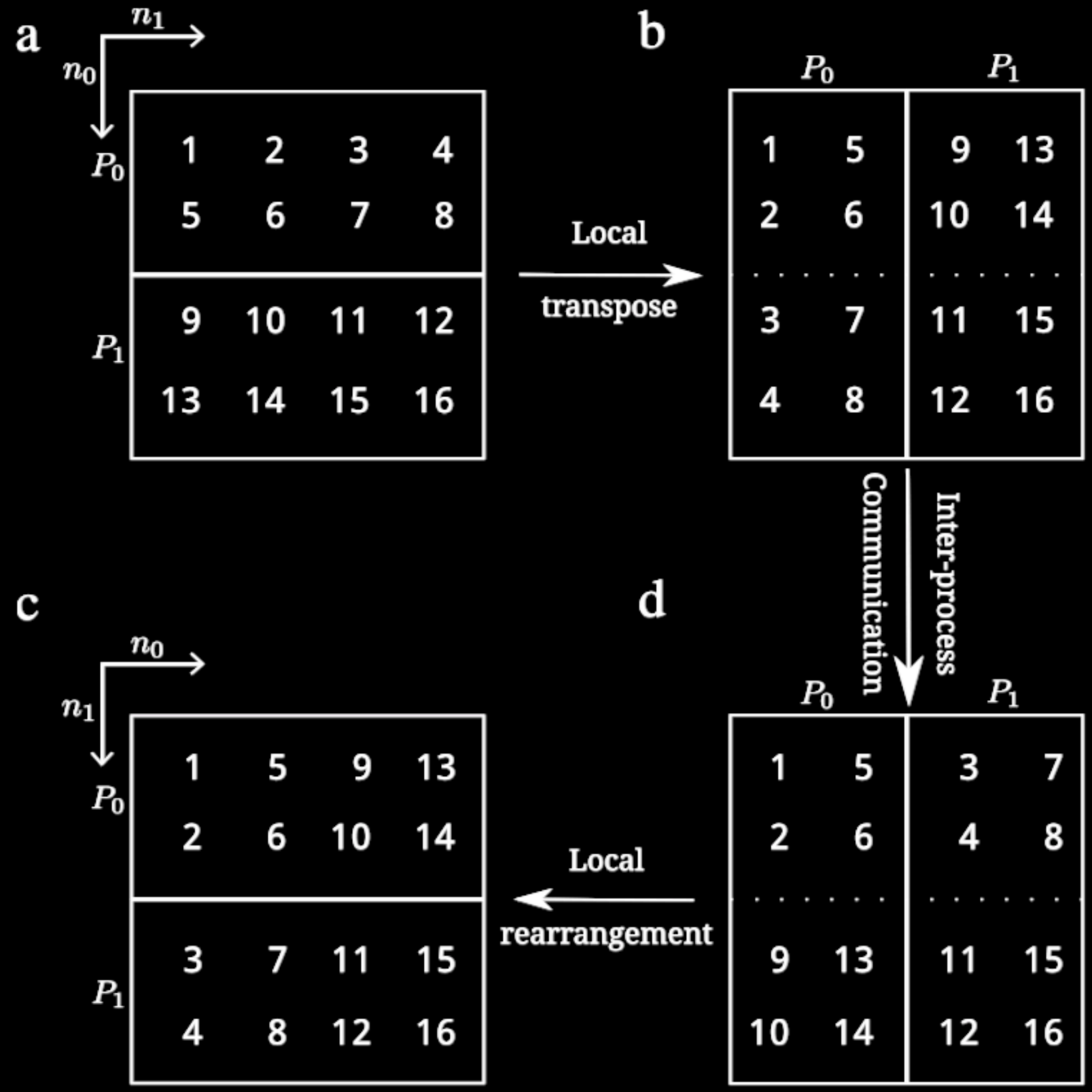
Slab decomposition

$$f(x,y,z)=\sum_{k_x}\sum_{k_y}\sum_{k_z}\hat{f}(k_x,k_y,k_z)\exp[i(k_x x+k_y y+k_z z)]$$

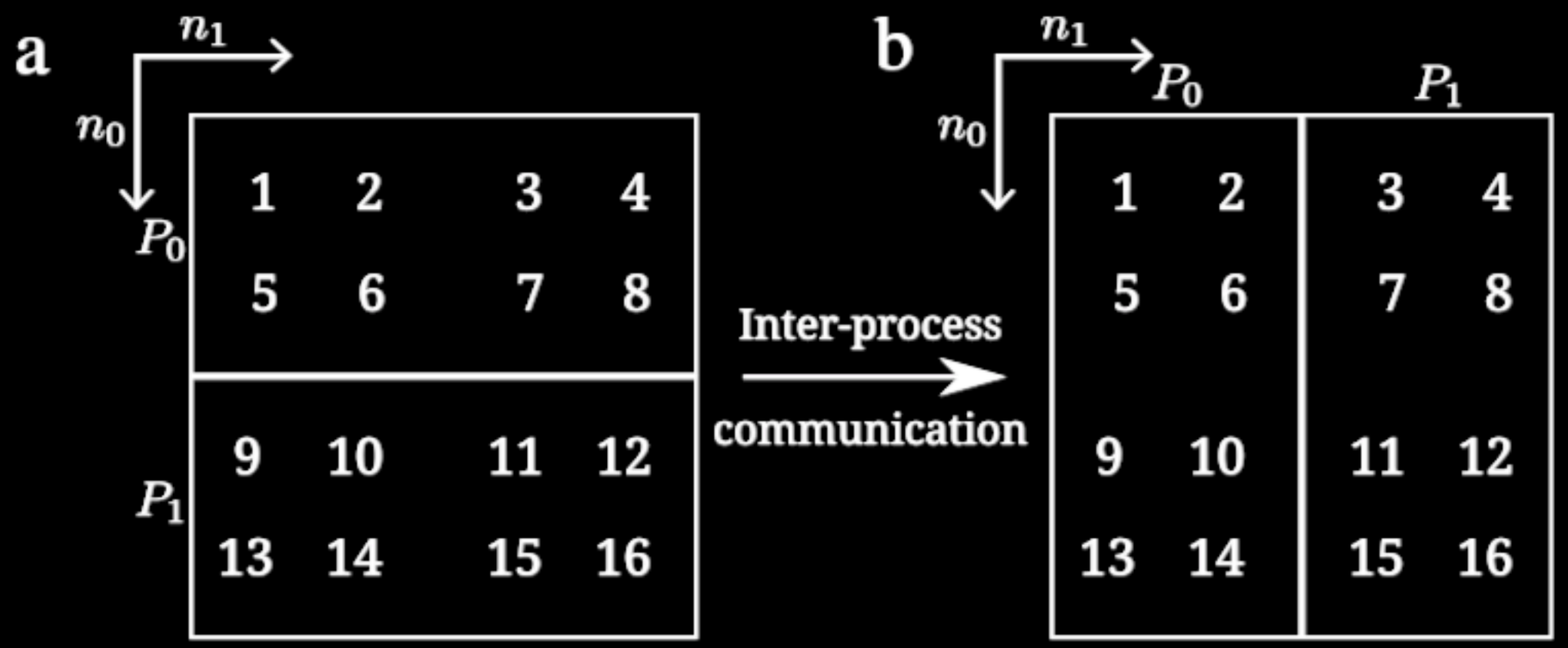
Divide the data among 4 procs



All_to_all communication


a
n_1
n_0
P_0
1 2 3 4
5 6 7 8
P_1
9 10 11 12
13 14 15 16
Local
transpose
b
P_0
P_1
1 5 9 13
2 6 10 14
3 7 11 15
4 8 12 16
Communication
Inter-process
c
n_0
n_1
P_0
1 5 9 13
2 6 10 14
P_1
3 7 11 15
4 8 12 16
Local
rearrangement
d
P_0
P_1
1 5 3 7
2 6 4 8
9 13 11 15
10 14 12 16

Transpose-free communication

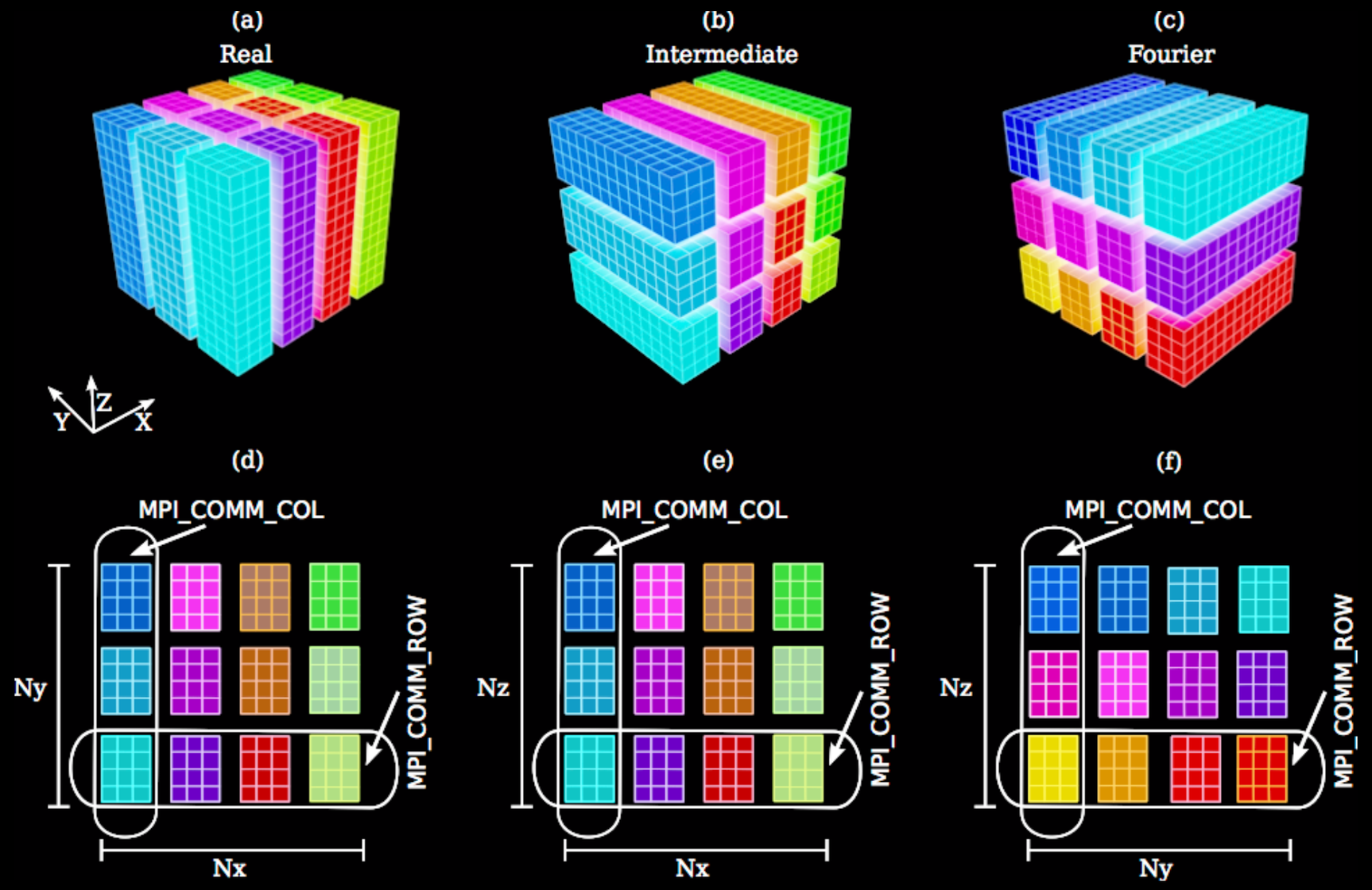


12-15% faster compared to FFTW

Limitations

- For N^3 grid, maximum number of processor = N

Pencil decomposition



Two sets of communications

But only among a set of processors

FFTW

- Slab decomposition
- Frigo and Johnson, MIT, ~1999
- C language
- Optimises for a given hardware

P3DFFT

- First pencil decomposition
- Dmitry Pekurovsk, San Diego Supercomputer Center (SDSC) ~1999
- Communication time ~ $p^{-2/3}$
- 8192^3 grid, 65536 cores (SIAM, J. Sci Comput. 2012)
- Used for spectral codes of PK Yeung with 262114 cores.

PFFT

- Pencil
- Pippig
- 1024^3 grid, 262144 cores (SIAM 2013)

GPU FFT

- DiGPUFFT, Czechowski et al. (ACM conference 2012)
- Takahashi, Japan

Hybrid FFT

- OpenMP+MPI
- Fortran
- Mininni et al. (Parallel Comput., 2011)
- 3072^3 grid, 20000 cores with 6/12 threads.

FFTK

- FFT Kanpur
- Pencil-based
- Chatterjee, Verma, and group members of Kanpur
- Scaled up to 196608 cores of Shaheen II of KAUST for 3072^3
- Tested up 6144^3 grid.
- Fluid solver TARANG uses it.

Scaling of FFT pencil

Computation $$T_{\text{comp}} = \frac{1}{c_1} p^{-\gamma_1},$$

Communication $$T_{\text{comm}} = \frac{1}{c_1} n^{-\gamma_2},$$

Total $$T^{-1} = \frac{1}{C} p^{\gamma},$$

$$T = c_1 D\left(\frac{1}{p^{\gamma_1}}\right) + c_2 D\left(\frac{1}{n^{\gamma_2}}\right) = C\left(\frac{1}{p^{\gamma}}\right)$$

On Shaheen 2 at KAUST
with Anando Chatterjee, Abhishek Kumar,
Ravi Samtaney, Bilel Hadri, Rooh Khurram

Cray XC40
ranked ~20th in top500

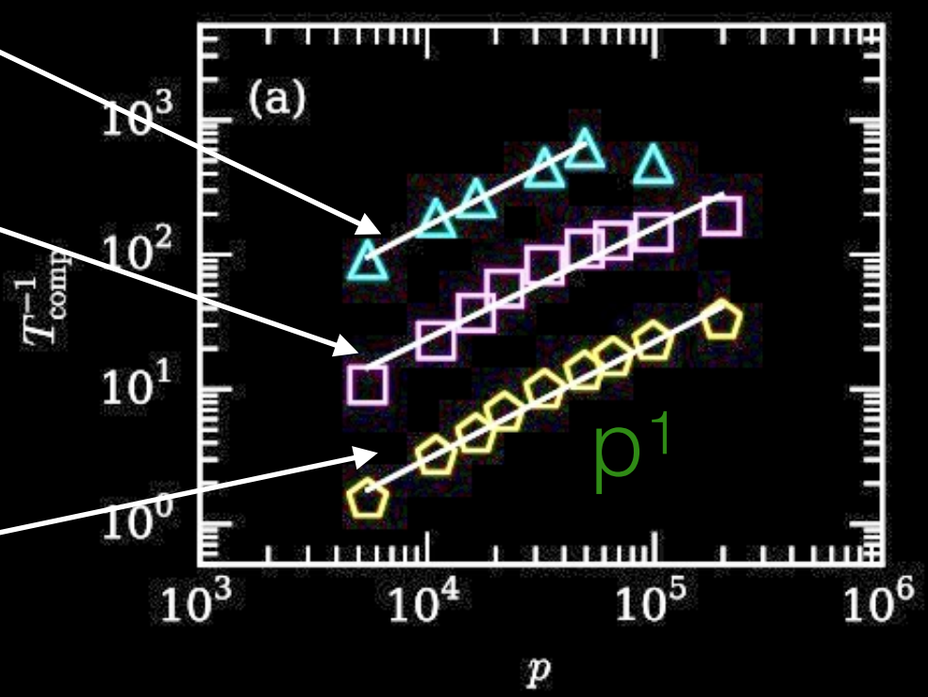
Chatterjee et al., JPDC 2018

Comp

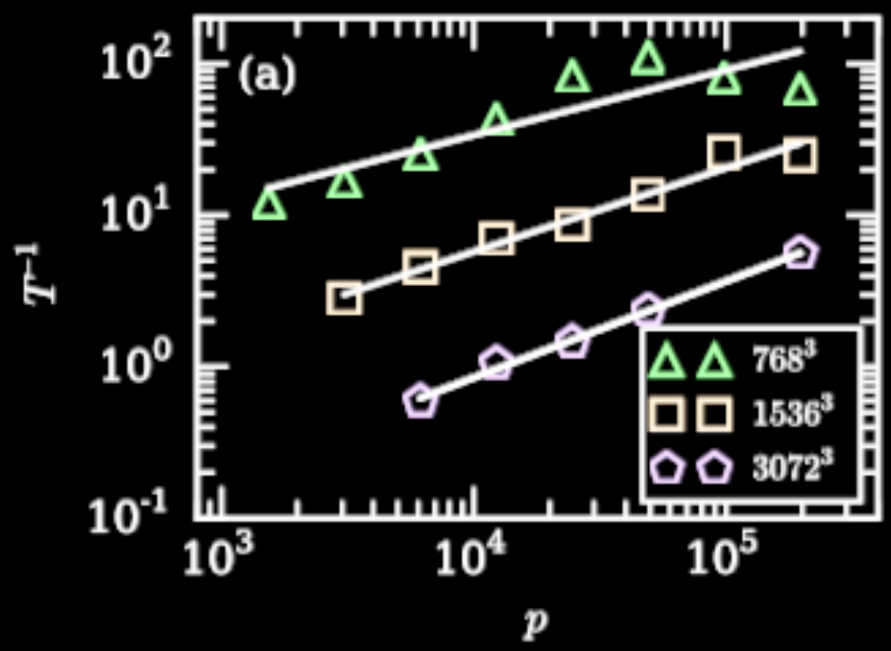
768^3

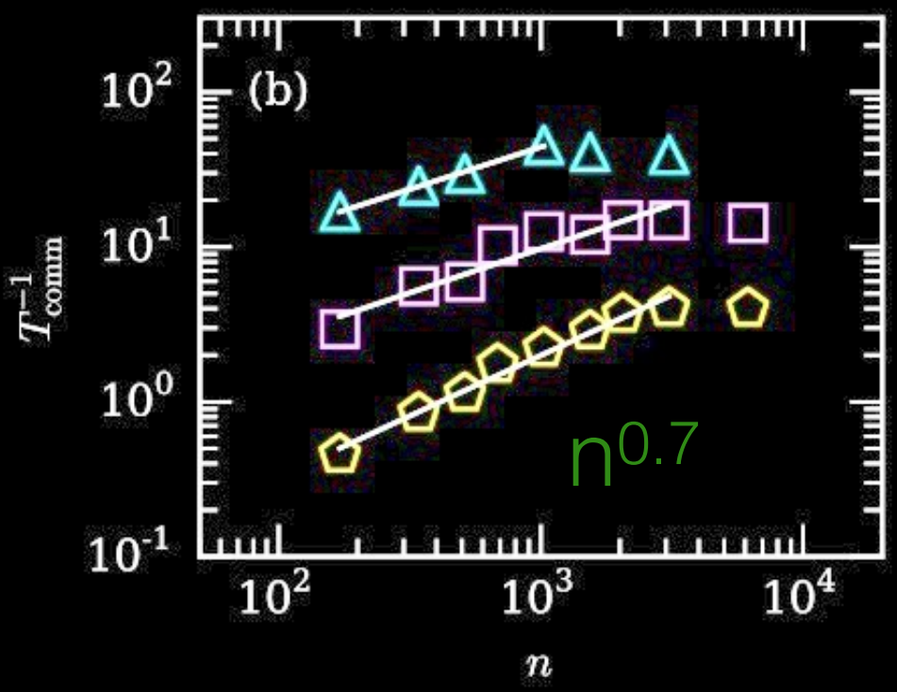
1536^3

3072^3

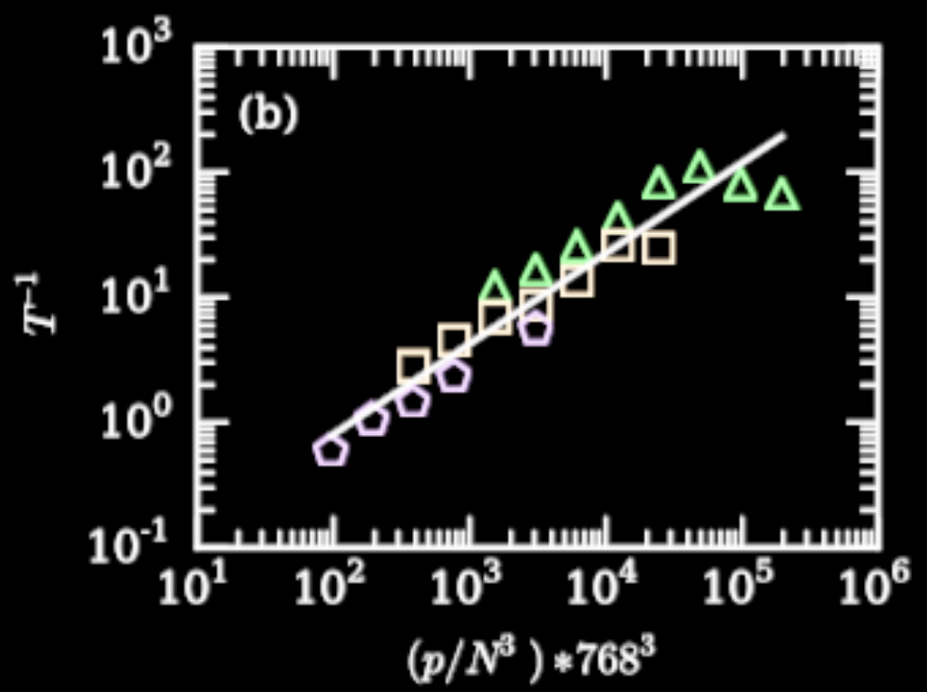


Strong scaling





Comm



Weak scaling

Table 4
FFTK scaling on Cray XC40 for the FFF and SFF basis: The exponents γ_1 for the computation time (T_{comp}), γ_2 for the communication time (T_{comm}), and γ for the total time (T) [refer to Eq. (17) for definition]. Maximum cores used: 196608.

Grid	γ_1	γ_2	γ
FFF			
768^3	0.79 ± 0.14	0.43 ± 0.09	0.43 ± 0.09
1536^3	0.93 ± 0.08	0.52 ± 0.04	0.55 ± 0.04
3072^3	1.08 ± 0.03	0.60 ± 0.02	0.64 ± 0.02
SFF			
768^3	0.82 ± 0.13	0.44 ± 0.03	0.46 ± 0.04
1536^3	0.97 ± 0.07	0.63 ± 0.02	0.66 ± 0.01
3072^3	0.99 ± 0.04	0.70 ± 0.05	0.73 ± 0.05

Table 5
FFTK on Cray XC40: Effective FLOP rating in Giga FLOP/s of Cray XC40 cores for various grid sizes and ppn. The efficiency E is the ratio of the effective per-core FLOP rating and the peak FLOP rating of each core (approximately 36 G FLOP/s).

Grid size	768^3	1536^3	3072^3
GFlop/s	0.45	0.53	0.64
E	0.013	0.015	0.018

Blue Gene/P

ning on Blue Gene/P. The exponents γ_1 for the computation time (T_{comp}), communication time (T_{comm}), and γ for the total time (T) [refer to Eq. (1) nition]. The maximum nodes used is 16 384 with 1ppn, 2ppn, and 4ppn.

ppn	2048^3	4096^3
1	1.00 ± 0.01	0.97 ± 0.01
2	1.00 ± 0.02	0.96 ± 0.01
4	1.00 ± 0.03	0.95 ± 0.03
1	0.7 ± 0.1	0.9 ± 0.1
2	0.7 ± 0.1	0.8 ± 0.2
4	0.7 ± 0.1	0.8 ± 0.2
1	0.87 ± 0.05	0.94 ± 0.05
2	0.81 ± 0.05	0.96 ± 0.09
4	0.76 ± 0.07	0.9 ± 0.1

Table 3
FFTK on Blue Gene/P: Effective FLOP rating in Giga FLOP/s of Blue Gene/P cores for various grid sizes and ppn. The efficiency E is the ratio of the effective per-core FLOP rating and the peak FLOP rating of each core (approximately 3.4 G FLOP/s).

Grid	ppn	Giga FLOP/s	E
2048^3	1	0.38	0.11
	2	0.28	0.082
	4	0.17	0.050
4096^3	1	0.36	0.11
	2	0.25	0.073
	4	0.14	0.041
8192^3	1	0.36	0.11
	2	0.26	0.076
	4	0.15	0.044

Effective flop rating/core

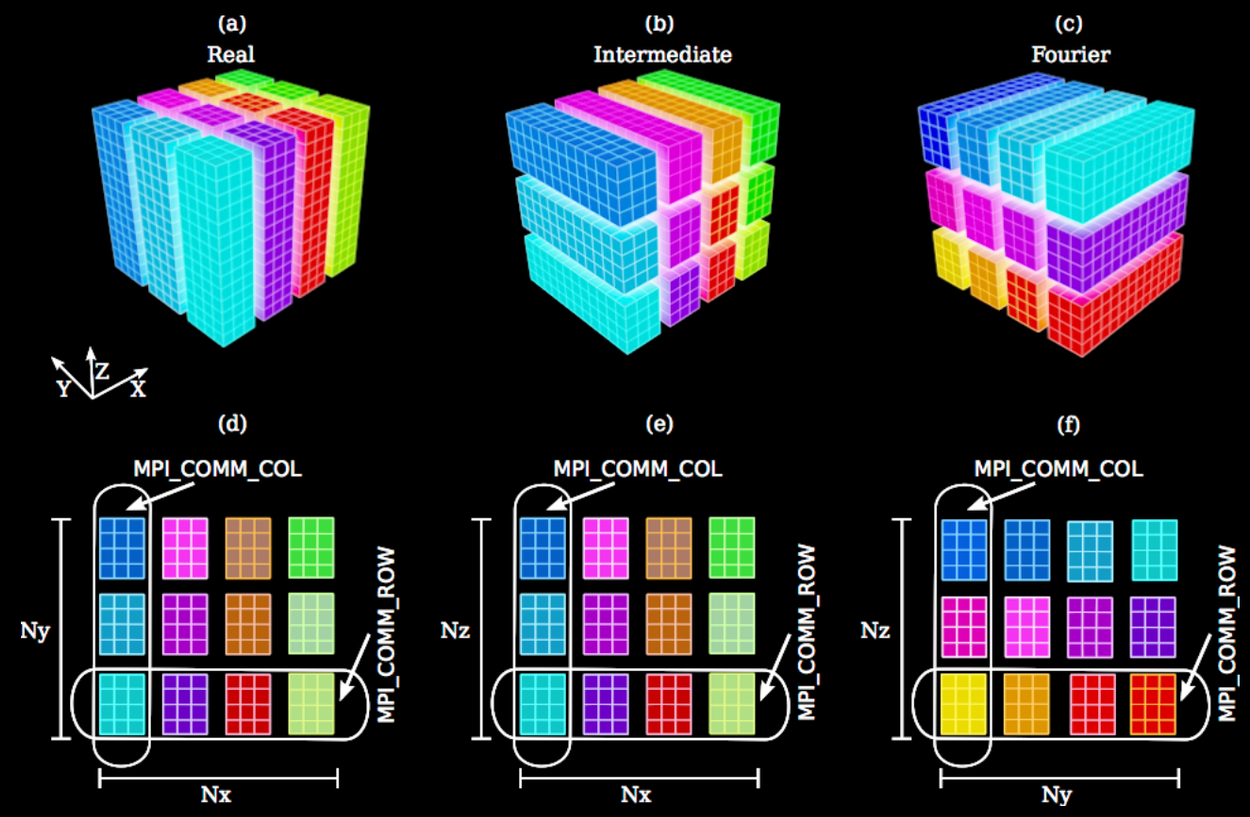
Cray XC40 (~1.5 %)

BlueGene/P (~10%)

Speed up

- Overlap communication and computation ?
- GPU ?
- Xeon Phi ?
- Optimise communication

Node configuration



Adjacent nodes in comm: both along x & y

Anando Chatterjee, Samar Aseeri, David Keys

Spectral method

velocity field

Pressure

Ext. Force

$$\partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla p + \nu \nabla^2 \mathbf{u} + \mathbf{F}$$

kinematic viscosity

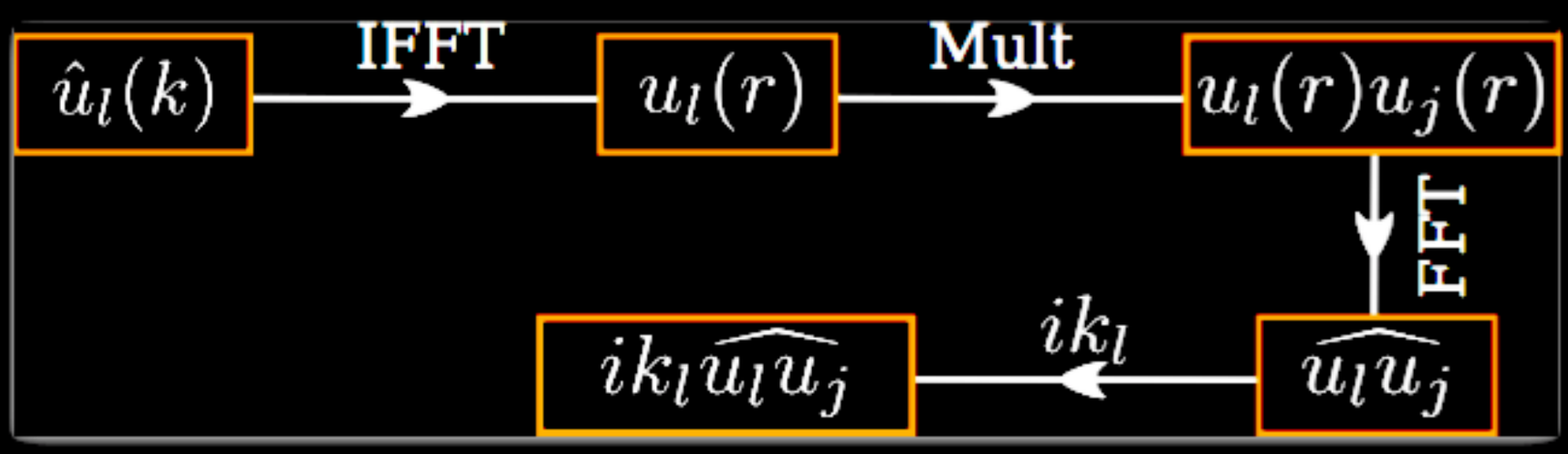
$$\nabla \cdot \mathbf{u} = 0$$

Incompressibility

Reynolds no = $\frac{UL}{\nu}$

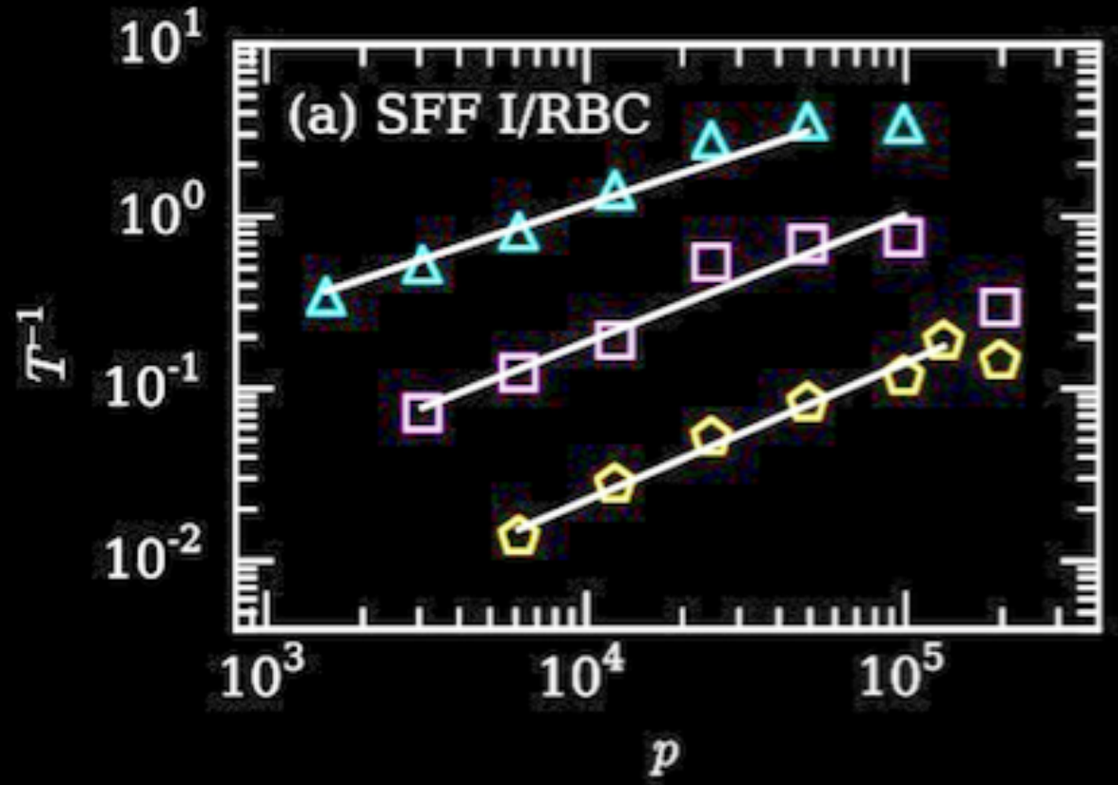
Set of ODEs

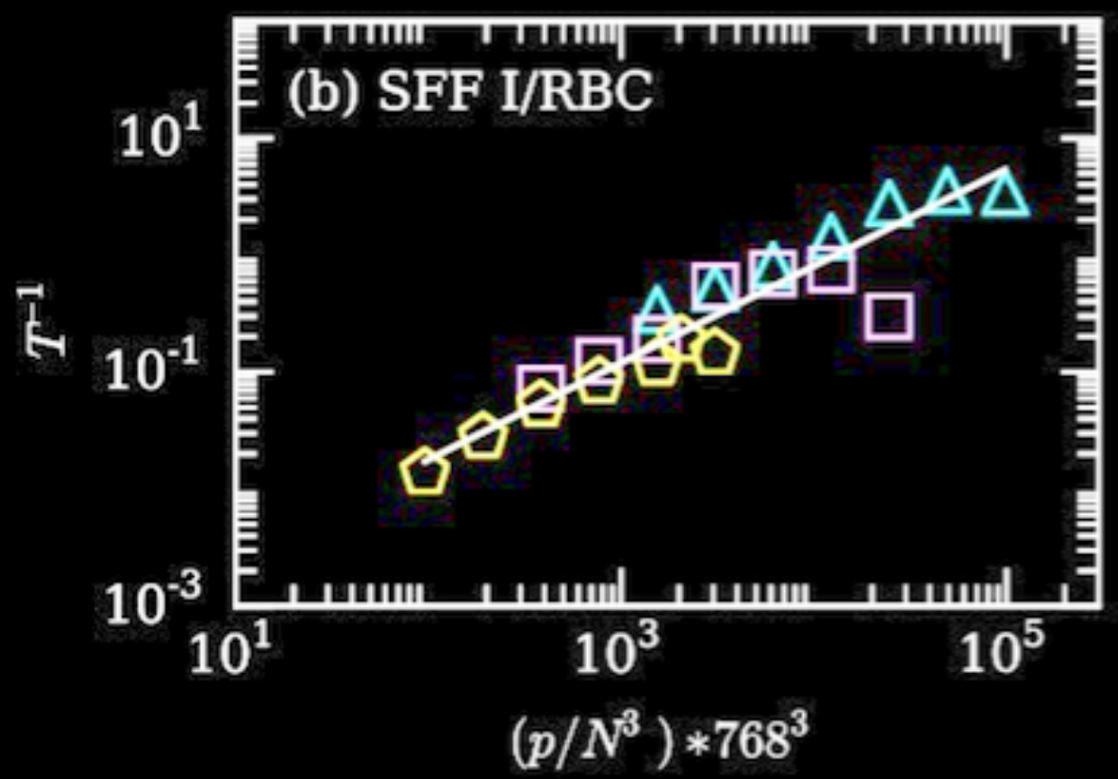
$$\frac{du_i(\mathbf{k})}{dt} = -jk_m \widehat{u_m(\mathbf{r})u_i(\mathbf{r})} - jk_i p(\mathbf{k}) - \nu k^2 u_i(\mathbf{k})$$



Tarang scaling

On Shaheen at KAUST


(a) SFF I/RBC
10^1
10^0
10^{-1}
10^{-2}
T^{-1}
10^3
10^4
10^5
p


(b) SFF I/RBC
10^1
10^{-1}
10^{-3}
T^{-1}
10^1
10^3
10^5
$(p/N^3) * 768^3$

Acknowledgements

Students:

Anando Chatterjee

Abhishek Kumar

Roshan Samuel

Sandeep Reddy

Mani Chandra

Sumit Kumar & Vijay

Faculty:

Ravi Samtaney

Fahad Anwer

Ported to:

PARAM, CDAC

Shaheen, KAUST

HPC system IITK

Funding

Dept of Science and Tech., India

Dept of Atomic Energy, India

KAUST (computer time)

Thank you!