# Scalable Fast Multipole Methods for Vortex Element Methods

Qi Hu[1], Nail A. Gumerov[1], Rio Yokota[2], Lorena Barba[3], Ramani Duraiswami[1]

Institute for Advanced Computer Studies, University of Maryland College Park[1]

Strategic Initiative for Extreme Computing King Abdullah University of Science and Technology[2]

Mechanical Engineering Department Boston University[3]

## • Problem Formulation:

1. Compute the induced local velocity and stretching at M locations $\mathbf{y}_j$ given N vortex elements $\mathbf{x}_i$

$$\mathbf{v}_j = \mathbf{v}(\mathbf{y}_j) = \sum_{i=1}^{N} \frac{\boldsymbol{\omega}_i \times (\mathbf{y}_j - \mathbf{x}_i)}{|\mathbf{y}_j - \mathbf{x}_i|^3} = \nabla \times \frac{\boldsymbol{\omega}_i}{|\mathbf{y}_j - \mathbf{x}_i|}$$

$$\frac{d\boldsymbol{\omega}_j}{dt} = \boldsymbol{\omega}_j \cdot \nabla \mathbf{v}_j$$

2. For the viscous flows, let $r_{ij} = |\mathbf{y}_j - \mathbf{x}_i|$.

I. The singular Biot-Savart kernel is smoothed by Gaussian basis function

$$K_\sigma(\mathbf{y}_j, \mathbf{x}_i) = \mathrm{erf}(\sqrt{\frac{r_{ij}^2}{2\sigma_i^2}}) - \sqrt{\frac{4}{\pi}}\sqrt{\frac{r_{ij}^2}{2\sigma_i^2}}\exp(-\frac{r_{ij}^2}{2\sigma_i^2})$$

II. The stretching term is smoothed by

$$G(\mathbf{y}_j, \mathbf{x}_i) = 3K(\mathbf{y}_j, \mathbf{x}_i) - 2\frac{r_{ij}^2}{2\sigma_i^2}\sqrt{\frac{2r_{ij}^2}{\pi\sigma_i^2}}\exp\left(-\frac{r_{ij}^2}{2\sigma_i^2}\right)$$

## • Lamb-Helmholtz Decomposition:

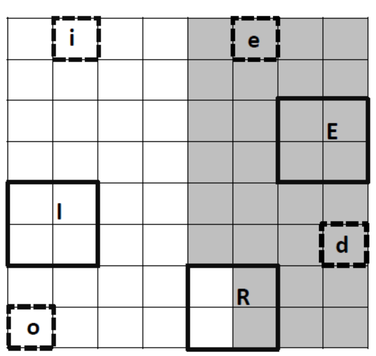
1. Given the divergence constraints, the 3D velocity field can be described by only two harmonic scalar potentials:

$$\mathbf{v}(\mathbf{r}) = \nabla\phi(\mathbf{r}) + \nabla \times (\mathbf{r}\chi(\mathbf{r}))$$

2. This decomposition [1] can perform the FMM translation for "velocity+stretching" calculation at the cost of **two** Laplace potential kernels instead of six [2].

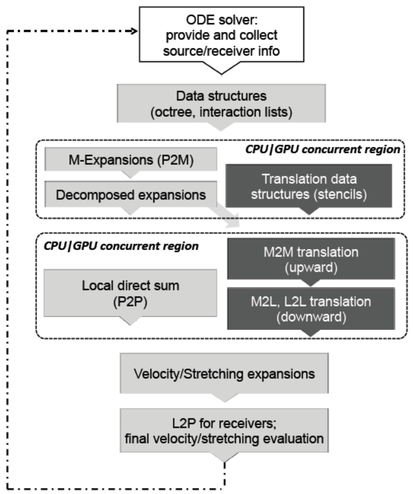
3. Algorithmic speeds-up the overall vortex element method

## • Data Structures to Manage Communications:



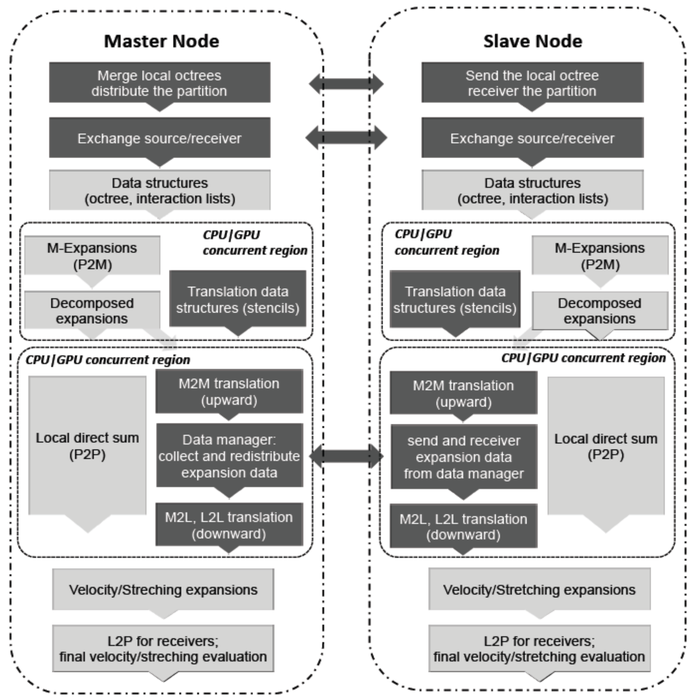

1. Classify all the spatial boxes as five categories [3]
2. New data structures build on local essential tree (LET)
3. Use master-slave model to manage communications
4. Fine parallel granularity constructions on GPUs

## • Single-Node FMM Algorithm




1. Heterogeneous [4]
2. Particle related processes are on GPUs:
   - Data structures
   - M-expansions
   - P2P computations
   - Velocity/stretching expansions
   - Final consolidations
3. Box related processes are on CPU. i.e., all M2M, M2L, L2L translations

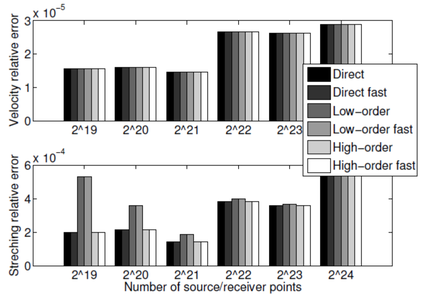
## • Multiple-node FMM Algorithm



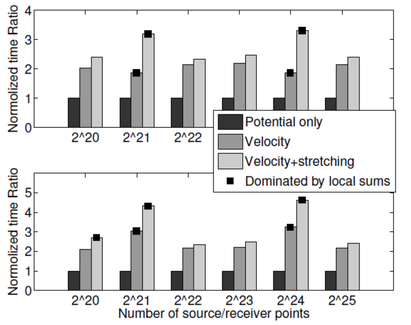

1. Master collects receiver boxes and distributes work regions (to achieve work balance)
2. Assigns particle data according to assigned work regions
3. Translations and P2P sums are fully distributed and are executed concurrently on CPUs and GPUs
4. One-time communication between master and slaves to exchange M-data during the upward translation stage

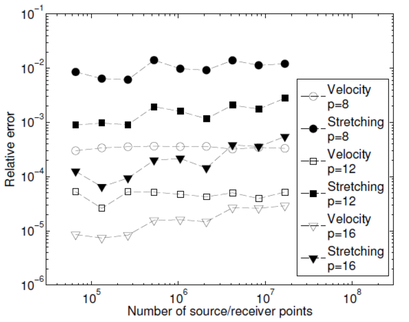
## • Vortex Core Evaluations

1. Fast approximations of transcendental functions
2. Accuracy is guaranteed
3. Fast implementations on GPU (can be applied to other problems)

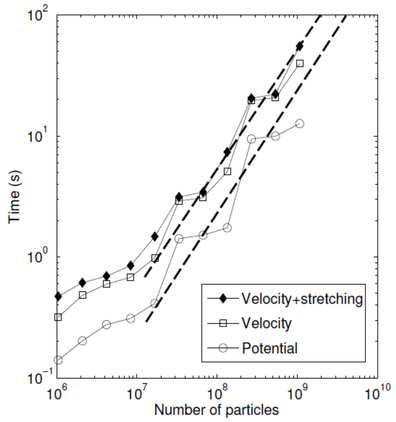



## • Performance




- Time comparisons between potential, velocity only and velocity+stretching computations
- Difference between top and bottom figures is without or with the smoothing kernels
- Accuracy tests are performed by varying truncation numbers (smoothing kernels used in all these cases)




- The overall runtime on a 32-node cluster: Chimera
- Only best timing of 1 or 2 GPUs used is reported
- Calculate velocity+stretching 1 billion source and receiver (different) in 55.9 seconds
- 49.12 Tflop/s on 32 nodes with 64 GPUs (933 Gflop/s peak performance for each GPU reported by NVIDIA)

References:

[1]. N. A. Gumerov and R. Duraiswami, "Efficient FMM accelerated vortex methods in three dimensions via the Lamb-Helmholtz decompositions," ArXiv e-prints, January 2012

[2]. R. Yokota, T. Narumi, R. Sakamaki, S. Kameoka, S. Obi, and K. Yasuoka, "Fast multipole methods on a cluster of GPUs for the meshless simulation of turbulence," Computer Physics Communications, vol. 180, no. 11, pp. 2066–2078, 2009

[3]. Q. Hu, N. A. Gumerov, and R. Duraiswami, "Scalable distributed fast multipole methods," in Proceedings of the 14th International Conference on High Performance Computing and Communications (HPCC-2012), ser. HPCC '12. ACM, 2012

[4]. Q. Hu, N. A. Gumerov, and R. Duraiswami, "Scalable fast multipole methods on distributed heterogeneous architectures," in Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis, ser. SC '11. ACM, 2011, pp. 36:1–36:12.