Performance comparison of CFD-DEM solver MFiX-Exa, on GPUs and CPUs

Shandong Lao\textsuperscript{a}, Aaron Holt\textsuperscript{a}, Deepthi Vaidhynathan\textsuperscript{b}, Hariswaran Sitaraman\textsuperscript{b}, Christine M. Hrenya\textsuperscript{c}, Thomas Hauser\textsuperscript{a}

\textsuperscript{a}Research Computing, University of Colorado, Boulder
\textsuperscript{b}Computational Science Center, National Renewable Energy Laboratory
\textsuperscript{c}Chemical and Biological Engineering, University of Colorado, Boulder

Abstract

We present computational performance comparisons of gas-solid simulations performed on current CPU and GPU architectures using MFiX Exa, a CFD-DEM solver that leverages hybrid CPU+GPU parallelism. A representative fluidized bed simulation with varying particle numbers from 2 to 67 million is used to compare serial and parallel performance. A single GPU was observed to be about 10 times faster compared to a single CPU core. The use of 3 GPUs on a single compute node was observed to be 4x faster than using all 64 CPU cores. We also observed that using an error controlled adaptive time stepping scheme for particle advance provided a consistent 4x speed-up on both CPUs and GPUs. Weak scaling results indicate superior parallel efficiencies when using GPUs compared to CPUs for the problem sizes studied in this work.

Keywords: Computational Fluid Dynamics, Discrete Element Method, High Performance Computing

1. Introduction

Gas-solid multiphase systems are ubiquitous in several applications such as pharmaceutical, mining and petrochemical industry. The Computational fluid dynamics (CFD)-Discrete element method (DEM) is a well-established theoretical approach for modeling granular phenomenon in these systems.

Email address: shandong.lao@colorado.edu (Shandong Lao)
The motion of individual solid particles are resolved in the CFD-DEM approach along with a continuum description for the background gaseous phase. There are minimal modeling assumptions compared to two-fluid [1,2] or mixture models [3,4] in this approach which are mainly associated with particle drag and inter-particle collisions. CFD-DEM is therefore a rigorous method which is amenable to small-scale systems with ~ few million particles while its application to large-scale industrial systems (~ few billion particles) [5,6,7] is constrained by high computational requirements. Recent advances in high performance computing (HPC) architectures with multiple Central Processing Units (CPU) cores and Graphics Processing Units (GPU) acceleration provide a viable pathway to perform large-scale CFD-DEM simulations. Several CFD-DEM solvers have been ported to parallel computing frameworks [8,9,10,11] using the message passing interface (MPI) and exhibit improved performance. CFD-DEM coupling though can impose conflicting configurations [12,10] for parallel domain decomposition for which dual grids or unified load-balancing algorithms may be necessary. DEM has received considerable attention with regard to use of GPU acceleration due to the simplified nature of particle advance [13,6,14], while CFD using traditional finite-volume methods most often require linear-solvers that impose algorithmic constraints on extracting GPU performance. Therefore, CFD solvers that are efficient on CPUs have been coupled with accelerator-enhanced DEM solvers [15,16]. Linear solvers that use scalable multi-grid [17] and Krylov Subspace methods [18] have now been demonstrated to show superior performance on GPU accelerators. MFiX-Exa, the software employed in this work, is an emerging CFD-DEM solver that utilizes GPU acceleration for both continuum and particle solves within a hybrid MPI+GPU parallelism framework. The performance implications for such a novel framework requires careful analysis of computational performance with respect to scalability, memory utilization and arithmetic intensity over hybrid HPC architectures, which is the main contribution of this work.

This paper is organized as follows. We first provide a brief description of the numerical methods and the parallelization algorithms within MFiX-Exa. We then present comparison of compute and memory performance on CPU and GPU architectures for a physically relevant fluidized bed simulation. Finally we show parallel performance in weak-scaled settings that utilize the hybrid MPI+GPU paradigm.
2. Numerical methods and algorithms

CFD-DEM solver, MFiX-Exa [19, 20], that leverages software library, AMReX [21], is used in this work. AMReX is a performance portable library specifically tailored towards massively parallel, block-structured adaptive-mesh-refinement (AMR) based applications and has been actively employed in several recent scalable frameworks in combustion [22], wind [23] and astrophysics [24]. The equations for gas-phase mass and momentum are solved in the incompressible form for the fluid update while the particle position, translational and angular velocities are updated using Newton’s laws of motion. Readers are directed to references [25, 26, 27, 19] for details on formulation and numerical methods. A brief description of the fluid-particle update algorithm in a typical CFD-DEM time-step where particle collision time-scales are smaller than fluid time-scales is described below:

1. The mass and momentum equations for the fluid phase are advanced using an incompressible projection scheme for a time-step of $\Delta t_f$
2. Particles are advanced using multiple sub-iterations with time-step $\Delta t_p$ for time $\Delta t_f$.
   (a) Particle drag is computed using interpolated fluid variables.
   (b) Collisional forces are computed using particle neighbor-lists.
   (c) Particle velocity and position are advanced using an explicit scheme.
3. Solids volume fraction and momentum transfer terms are computed from particle data on the fluid grid using volume filtering.

A hybrid distributed (MPI) and shared (GPU) memory paradigm is used in the solver for parallelization. The domain decomposition approach is as shown in Fig. 1. The computational domain that uses an uniform Cartesian grid is divided into sub-domains or *boxes* with co-located particles, as shown in Fig. 1(b). A space-filling curve [28] algorithm is used to distribute the *boxes* among processors. The number of *boxes* typically is equal or greater than the number of MPI ranks, as indicated by an example decomposition in Fig. 1(d). Shared memory parallelism is achieved by using GPU threads that operate on one or more particles (for DEM) or fluid-cells (for CFD) within each *box*, as shown in Fig. 1(e) and (f). An important aspect of this approach is the need to store *ghost* particles at the boundary of each *box*. This is shown in Fig. 1(c) for an example case of two adjacent *boxes*. The *ghost* particles are stored within a single layer of cells and are updated every particle time-step among all *boxes* and processors. Particles that move outside of a box
Figure 1: (a) shows an example distribution of particles in a Cartesian domain, (b) shows domain decomposition into 16 boxes with 4 boxes along each direction, (c) shows two adjacent boxes with ghost-layers that are used for communication of particles between MPI ranks, (d) shows an example distribution of 16 boxes among 4 MPI ranks, each rank owns 4 boxes, (e) shows the assignment of the 4 boxes local to an MPI rank to multiple GPU streams and (f) shows the allocation of particles in a stream to GPU threads.
are redistributed to MPI ranks that own adjacent boxes. Particle collision forces are computed from interactions within a sphere of influence that is six times the particle radius. Potential collision partners for a given particle are stored in a neighbor-list data structure that contain partner identifiers. This data-structure is dynamically updated during particle advance after every few time-steps.

3. Results

3.1. Problem description

A rectangular fluidized bed simulation setup is used to analyze the performance of MFiX-Exa on both CPU and GPU in this study. The geometry and particle transients at statistically stationary state are illustrated in Fig. 2 for a representative run with 40,000 particles. A series of simulations of different domain sizes and particle numbers are carried out to analyze computational performance on CPUs and GPUs. The range of domain sizes and particle numbers are listed in Table 1. Tables 2 and 3 define the fluid and solid properties used in our simulations, respectively.
| Parameter                  | Range          |
|---------------------------|----------------|
| H (m)                     | 0.0100 - 0.6400|
| W (m)                     | 0.0032 - 0.1024|
| D (m)                     | 0.0032 - 0.0256|
| Solid Volume fraction (%)  | 20.0 - 50.0    |
| Number of Particles       | 40,000 - 67,000,000|

Table 1: Ranges of domain size and particle numbers for the fluidized bed cases.

| Parameter                  | Value          |
|---------------------------|----------------|
| Viscosity (kg·m⁻¹·s⁻¹)    | 2.0 × 10⁻⁵     |
| Density (kg/m³)           | 1.000          |
| Gas inlet Velocity (m/s)  | 0.015          |
| Drag model                | BVK[29]        |

Table 2: Fluid properties used in the fluidized bed simulations.

| Parameter                  | Value          |
|---------------------------|----------------|
| Diameter (m)              | 1.0 × 10⁻⁴     |
| Density (kg/m³)           | 1000.00        |
| Spring Constant (N/m)     | 10.00          |
| Sliding Friction Coefficient | 0.00        |
| Particle-Particle Restitution Coefficient | 0.80 |
| Particle-Wall Restitution Coefficient | 1.00 |
| Tangential-to-Normal Spring Constant Factor | 0.28 |
| Tangential-to-Normal Damping Coefficient Factor | 0.50 |

Table 3: Particle properties used in the fluidized bed simulations.
Hardware Specification

| Hardware       | Specification                  |
|----------------|--------------------------------|
| CPU            | 2x Dual AMD EPYC 7452          |
| Total Cores    | 64                             |
| Total Threads  | 128                            |
| L3 Cache Per Core | 128 MB                       |
| Memory         | 264 GB                         |
| GPU            | 3x NVIDIA A100 PCIe            |
| CUDA Cores     | 6912                           |
| GPU Memory     | 40 GB                          |

Table 4: Blanca CPU+GPU node hardware

| Hardware       | Specification                  |
|----------------|--------------------------------|
| CPU            | Dual Intel Xeon Gold Skylake 6154 |
| Total Cores    | 36                             |
| Total Threads  | 36                             |
| L3 Cache Per Core | 75 MB                       |
| Memory         | 96 GB                          |

Table 5: Eagle CPU only node hardware

3.2. Compute architectures

Serial and single node benchmarking and profiling was completed on University of Colorado Research Computing’s Blanca CPU+GPU nodes. Each Blanca CPU+GPU node contains two AMD EPYC 7452 processors and three Nvidia A100 PCIe cards each with 40 GB of GPU device memory. The hardware on these nodes are more recent and have more resources per node allowing for larger problem sizes. Additionally, hardware counters were enabled on all of these nodes allowing for more in-depth profiling such as roofline analysis. A summary of Blanca’s hardware is as shown below in Table 4. MFiX-Exa was compiled on the Blanca using GCC 8.2.0, OPENMPI 4.0.0, and CUDA 11.2 compilers. However, The Blanca CPU+GPU nodes only have a 10 GB ethernet connecting them, which may reduce the performance of multi-node computational jobs. Multi-node performance benchmarking was completed on the National Renewable Energy Laboratory (NREL)’s supercomputer, Eagle. Both a CPU only node and a CPU+GPU node were used for benchmarking runs. Both compute node types use two Intel Xeon
Gold Skylake 6154 processors and are connected by a 100 GB/s EDR Infini-band network. The CPU only nodes have 96 GB of memory whereas the CPU+GPU nodes have 768 GB. The CPU+GPU nodes have two NVIDIA Tesla V100 PCIe cards with 16 GB of device memory for each one of them. Tables 5 and 6 give a hardware summary of the Eagle compute nodes used in this work. MFiX-Exa was compiled using GCC 8.4.0, OpenMPI 4.0.4 and CUDA 10.2 compilers on Eagle.

3.3. Serial and Single Node Performance

3.3.1. Serial performance variation over time steps

A single V100 graphics card on Eagle ran a 160,000 particle simulation on a $32 \times 50 \times 32$ grid for 10 ms in 270 seconds which is $\sim 10x$ faster than a single Intel Skylake CPU that took 2400 seconds to completion as shown in Fig. 3(a) and (b). These figures also demonstrate the step performance during the early transients and at statistical steady state. Both the CPU and GPU runs follow a similar pattern in terms of fluid update time-step size over the course of the simulation. The main difference being the GPU case climbed to a time per step of 2.4 seconds at simulation time of 0.004 before stabilizing whereas the CPU case hit a time per step of 24 seconds, dropped to 19 seconds and then stabilized. Overall, the single GPU case was expected to be faster than a single CPU core case because a multi-threaded V100 GPU has higher peak performance compared to a single CPU core.

3.3.2. Serial performance with problem size variation

We compare serial performance on Blanca with varying problem sizes in this section. For this purpose, a series of simulations with varying domain

| Hardware     | Specification            |
|--------------|--------------------------|
| CPU          | Dual Intel Xeon Gold Skylake 6154 |
| Total Cores  | 36                       |
| Total Threads| 36                       |
| L3 Cache Per Core | 75 MB                   |
| Memory       | 768 GB                   |
| GPU          | 2x NVIDIA Tesla V100 PCIe |
| CUDA Cores   | 5120                     |
| GPU Memory   | 16 GB                    |

Table 6: Eagle CPU+GPU node hardware
Figure 3: Time history of time-step size and costs per time-step for a 160,000 particle fluidized bed simulation performed on (a) single CPU core and (b) a single GPU, respectively.
| Problem size factor | Domain $W \times H \times D$ ($m^3$) | Particle no: ($\times 10^6$) |
|---------------------|--------------------------------------|-------------------------------|
| 1x                  | $0.0128 \times 0.02 \times 0.0128$   | 2.10                          |
| 2x                  | $0.0128 \times 0.04 \times 0.0128$   | 4.20                          |
| 4x                  | $0.0128 \times 0.08 \times 0.0128$   | 8.40                          |
| 8x                  | $0.0128 \times 0.16 \times 0.0128$   | 16.76                         |
| 16x                 | $0.0128 \times 0.32 \times 0.0128$   | 33.52                         |
| 32x                 | $0.0128 \times 0.64 \times 0.0128$   | 67.03                         |

Table 7: Domain size scaling and associate particle numbers for performance studies. The Cartesian domain for the 1x case is a $64 \times 100 \times 64$ grid, and is decomposed into $8 \times 100 \times 8$ boxes. The number of cells along the axial direction in the domain and boxes are proportionally increased with increasing problem size.

Figure 4: Serial performance on a single CPU core and a GPU with varying problem sizes.
sizes and particle numbers were carried out. The domain sizes and particle numbers for these experiments are listed in Table 7. The serial performance is shown in Figure 4. The computational time increases with larger problem sizes in both CPU and GPU cases. The wall-clock times are higher than the ideal scaling for the CPU mainly due to the increased memory requirements for larger problem sizes. On the other hand, the GPU wall clock times tend to be lower than the ideal scaling (25% lower wall clock time at 16x problem size compared to ideal) indicating efficient vectorization and memory utilization among multiple threads on the GPU for larger problem sizes. The GPU runs are about 10x faster than the serial CPU run and is consistent across varying problem sizes.

3.3.3. Single node performance

There are 64 CPU cores and 3 GPUs in a single node of Blanca (Table 4). Each CPU core is enabled to run 2 threads simultaneously. Wall-clock times on one node (64 and 128 MPI tasks on 64 CPU cores, and 3 MPI tasks on 3 GPUs) were compared and is as shown in Fig. 5. The performance of using 3 GPUs is higher than using all CPU cores. The GPU speedup factor with respect to CPU runs is higher at larger problem sizes and is about 4x faster.
at 32 times the baseline problem size (see Table 7) with 67 million particles. We also observe that enabling multiple threads on a single CPU core does not provide appreciable performance improvements.

### 3.3.4. Performance improvements with adaptive time-stepping

We utilize an error controlled adaptive time-stepping method (ATS) [27] for particle advance where the particle time-steps are dynamically varied based on velocity errors. This allows for increased time-steps when particle dynamics are less stiff thus reducing the total number of particle time-steps per fluid time step. ATS significantly improves the simulation performance for all sizes investigated on both CPU and GPU. The detailed results are plotted in Figure 6 which indicates approximately 4x improvement with using ATS compared to the default constant time step method on both CPUs and GPUs.

### 3.4. Parallel performance on multiple nodes

Figure 7 illustrates the weak scaling results on both CPU and GPU with and without ATS. These simulations were performed on Eagle supercomputer which has 36 CPU cores and 2 GPUs per node. The wall-clock times for the
Figure 7: Weak Scaling results on CPU and GPU with and without using ATS. Upper half is the CPU performance while lower half is for GPU cases. Processor here means a single CPU core for the CPU case while it is a single GPU for the GPU case. The number of processors correspond to the problem size factor listed in Table 7. It is also the same as the number of MPI ranks where each MPI task runs on one CPU core or a GPU.
CPU case tends to increase with increasing number of MPI tasks mainly due to greater memory requirement on a single node and MPI communication costs. On the other hand, the GPU runs tend to stabilize at larger number of MPI tasks (8-32) indicating effective utilization of GPU threads and memory at larger problem sizes. ATS tends to give a consistent performance improvement of approximately 4x across varying number of MPI tasks and problem sizes on both GPUs and CPUs.

4. Conclusions

In this work, we presented computational performance comparisons of CFD-DEM simulations performed on current CPU and GPU architectures using emerging multiphase gas-solid solver, MFiX-Exa. A representative fluidized bed case with varying problem sizes (2 to 67 million particles) were simulated and CPU vs GPU performance in serial and parallel paradigms were compared. A single GPU was observed to be about 10 times faster compared to a single CPU core. The use of 3 GPUs on a single compute node was observed to be 4x faster than using all 64 CPU cores. We also observed that using an error controlled adaptive time stepping scheme for particle advance provided a consistent 4x speed-up on both CPUs and GPUs. Weak scaling results indicate superior parallel efficiencies when using GPUs compared to CPUs. The GPUs show near perfect parallel efficiencies at large problem sizes while the CPU cases show reduction in efficiencies due to increased memory utilization and communication costs at larger problem sizes.

Acknowledgments

This research was supported by the Department of Energy for project titled “MFiX-DEM enhancement for industry relevant flows” under Grant No. DE-FE0026298. The authors acknowledge helpful discussions with their collaborator Peiyuan Liu at University of Colorado, Boulder, Jordan Musser and William D. Fullmer from National Energy Technology Laboratory and Ann Almgren from Lawrence Berkeley National Laboratory A portion of this work utilized resources from the University of Colorado Boulder Research Computing Group, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University. A portion of this research was performed
using computational resources sponsored by the Department of Energy’s Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.
References

[1] A. Passalacqua, R. O. Fox, Implementation of an iterative solution procedure for multi-fluid gas–particle flow models on unstructured grids, Powder technology 213 (1-3) (2011) 174–187.

[2] M. J. Rahimi, H. Sitaraman, D. Humbird, J. J. Stickel, Computational fluid dynamics study of full-scale aerobic bioreactors: Evaluation of gas–liquid mass transfer, oxygen uptake, and dynamic oxygen distribution, Chemical Engineering Research and Design 139 (2018) 283–295.

[3] H. Sitaraman, E. M. Kuhn, A. Nag, M. A. Sprague, M. P. Tucker, J. J. Stickel, Multiphysics modeling and simulation of high-solids dilute-acid pretreatment of corn stover in a steam-explosion reactor, Chemical Engineering Journal 268 (2015) 47–59.

[4] H. Sitaraman, N. Danes, J. J. Lischeske, J. J. Stickel, M. A. Sprague, Coupled cfd and chemical-kinetics simulations of cellulosic-biomass enzymatic hydrolysis: Mathematical-model development and validation, Chemical Engineering Science 206 (2019) 348–360.

[5] P. Liu, C. M. Hrenya, Challenges of dem: I. competing bottlenecks in parallelization of gas–solid flows, Powder technology 264 (2014) 620–626.

[6] J. Gan, Z. Zhou, A. Yu, A gpu-based dem approach for modelling of particulate systems, Powder Technology 301 (2016) 1172–1182.

[7] R. Cocco, W. D. Fullmer, P. Liu, C. M. Hrenya, Cfd-dem: Modeling the small to understand the large, Chemical Engineering Progress 113 (9) (2017) 38–45.

[8] P. Gopalakrishnan, D. Tafti, Development of parallel dem for the open source code mfix, Powder technology 235 (2013) 33–41.

[9] D. Kafui, S. Johnson, C. Thornton, J. Seville, Parallelization of a lagrangian–eulerian dem/cfd code for application to fluidized beds, Powder Technology 207 (1-3) (2011) 270–278.

[10] G. Pozzetti, X. Besseron, A. Rouset, B. Peters, A co-located partitions strategy for parallel cfd–dem couplings, Advanced Powder Technology 29 (12) (2018) 3220–3232.
[11] Y. Dufresne, V. Moureau, G. Lartigue, O. Simonin, A massively parallel cfd/dem approach for reactive gas-solid flows in complex geometries using unstructured meshes, Computers & Fluids 198 (2020) 104402.

[12] H. Sitaraman, R. Grout, Balancing conflicting requirements for grid and particle decomposition in continuum-lagrangian solvers, Parallel Computing 52 (2016) 1–21.

[13] J. Qi, K.-C. Li, H. Jiang, Q. Zhou, L. Yang, Gpu-accelerated dem implementation with cuda, International Journal of Computational Science and Engineering 11 (3) (2015) 330–337.

[14] Y. Tian, S. Zhang, P. Lin, Q. Yang, G. Yang, L. Yang, Implementing discrete element method for large-scale simulation of particles on multiple gpus, Computers & Chemical Engineering 104 (2017) 231–240.

[15] C. Kloss, C. Goniva, A. Hager, S. Amberger, S. Pirker, Models, algorithms and validation for opensource dem and cfd–dem, Progress in Computational Fluid Dynamics, an International Journal 12 (2-3) (2012) 140–152.

[16] Y. He, F. Muller, A. Hassanpour, A. E. Bayly, A cpu-gpu cross-platform coupled cfd-dem approach for complex particle-fluid flows, Chemical Engineering Science 223 (2020) 115712.

[17] P. Basu, S. Williams, B. Van Straalen, L. Oliker, P. Colella, M. Hall, Compiler-based code generation and autotuning for geometric multigrid on gpu-accelerated supercomputers, Parallel Computing 64 (2017) 50–64.

[18] K. Matsumoto, Y. Idomura, T. Ina, A. Mayumi, S. Yamada, Implementation and performance evaluation of a communication-avoiding gmres method for stencil-based code on gpu cluster, The Journal of Supercomputing 75 (12) (2019) 8115–8146.

[19] J. Musser, A. S. Almgren, W. D. Fullmer, O. Antepara, J. B. Bell, J. Blaschke, K. Gott, A. Myers, R. Porcu, D. Rangarajan, et al., Mfix-exa: A path toward exascale cfd-dem simulations, The International Journal of High Performance Computing Applications (2021) 109434202110009293.

17
[20] M. Syamlal, W. Rogers, T. J. O’Brien, Mfix documentation: Theory guide, National Energy Technology Laboratory, Department of Energy, Technical Note DOE/METC-95/1013 and NTIS/DE95000031.

[21] W. Zhang, A. Almgren, V. Beckner, J. Bell, J. Blaschke, C. Chan, M. Day, B. Friesen, K. Gott, D. Graves, et al., Amrex: a framework for block-structured adaptive mesh refinement, Journal of Open Source Software 3 (37) (2019) 1370.

[22] H. Sitaraman, S. Yellapantula, M. T. H. de Frahan, B. Perry, J. Rood, R. Grout, M. Day, Adaptive mesh based combustion simulations of direct fuel injection effects in a supersonic cavity flame-holder, Combustion and Flame 232 (2021) 111531.

[23] Amr-wind: a massively parallel, block-structured adaptive-mesh, incompressible flow solver for wind turbine and wind farm simulations., https://github.com/Exawind/amr-wind.

[24] M. Zingale, A. Almgren, M. B. Sazo, V. Beckner, J. Bell, B. Friesen, A. Jacobs, M. Katz, C. Malone, A. Nonaka, et al., Meeting the challenges of modeling astrophysical thermonuclear explosions: Castro, maestro, and the amrex astrophysics suite, in: Journal of Physics: Conference Series, Vol. 1031, IOP Publishing, 2018, p. 012024.

[25] W. D. Fullmer, A. S. Almgren, M. Rosso, J. Blaschke, J. Musser, Benchmarking of a preliminary mfix-exa code (2019). arXiv:1909.02067.

[26] Mfix-exa documentation., https://amrex-codes.github.io/MFIX-Exa/docs_html/Introduction.html

[27] H. Sitaraman, D. Vaidhynathan, R. Grout, T. Hauser, C. M. Hrenya, J. Musser, An error-controlled adaptive time-stepping method for particle advancement in coupled cfd-dem simulations, Powder Technology 379 (2021) 203–216.

[28] M. N. Farooqi, T. Nguyen, W. Zhang, A. S. Almgren, J. Shalf, D. Unat, Phase asynchronous amr execution for productive and performant astrophysical flows, in: SC18: International Conference for High Performance Computing, Networking, Storage and Analysis, IEEE, 2018, pp. 880–893.
[29] R. Beetstra, M. A. van der Hoef, J. Kuipers, Drag force of intermediate reynolds number flow past mono-and bidisperse arrays of spheres, AIChE journal 53 (2) (2007) 489–501.