



Development of an equation-based parallelization method for multiphase particle-in-cell simulations

Mino Woo^{1,2} · Terry Jordan¹ · Tarak Nandi¹ · Jean François Dietiker³ · Christopher Guenther¹ · Dirk Van Essendelft¹

Received: 18 May 2022 / Accepted: 21 November 2022 / Published online: 22 December 2022
This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2022

Abstract

Manufacturers have been developing new graphics processing unit (GPU) nodes with large capacity, high bandwidth memory and very high bandwidth intra-node interconnects. This enables moving large amounts of data between GPUs on the same node at low cost. However, small packet bandwidths and latencies have not decreased, which makes global dot products expensive. These characteristics favor a new kind of problem decomposition called “equation decomposition” rather than traditional domain decomposition. In this approach, each GPU is assigned one equation set to solve in parallel so that the frequent and expensive dot product synchronization points in traditional distributed linear solvers are eliminated. In exchange, the method involves infrequent movement of state variables over the high bandwidth, intra-node interconnects. To test this theory, our flagship code Multiphase Flow with Interphase eXchanges (MFiX) was ported to TensorFlow. This new product is known as MFiX-AI and can produce near identical results to the original version of MFiX with significant acceleration in multiphase particle-in-cell (MP-PIC) simulations. The performance of a single node with 4 NVIDIA A100s connected over NVLINK 2.0 was shown to be competitive to 1000 CPU cores (25 nodes) on the JOULE 2.0 supercomputer, leading to an energy savings of up to 90%. This is a substantial performance benefit for small- to intermediate-sized problems. This benefit is expected to grow as GPU nodes become more powerful. Further, MFiX-AI is poised to accept native artificial intelligence/machine learning models for further acceleration and development.

Keywords High-performance computing · Computational fluid dynamics · Parallel computation · Multiphase flow · TensorFlow · GPU acceleration

1 Introduction

Recent advances in high-performance computing (HPC), including those in both multi-central processing units (CPUs) and graphics processing units (GPUs), have greatly improved the speed of numerical computations in areas like computational fluid dynamics (CFD) [1, 2], molecular dynamics [3, 4], lattice-Boltzmann methods [5], and deep learning [6, 7]. Computationally expensive CFD computations requiring high mesh resolution have significantly

benefited from these advances. However, developing parallel computer codes to utilize multicore architecture is still challenging, as it requires rewriting and optimization of existing serial codes, particularly when both CPUs and GPUs are used in a collaborative manner [8, 9], where the GPU to CPU memory transfer can be a serious bottleneck [10]. Although parallel computing has been widely used in a variety of CFD applications [11, 12], development of a computer code to fully utilize multicore HPC capabilities on heterogeneous systems is still challenging [13].

TensorFlow (TF) [14] is an open-source software library capable of utilizing both CPU and GPU architectures and can potentially help accelerate CFD computations. TF was the most popular and well-documented framework at the beginning of this project and is an effective tensor algebra library, which makes it useful for many scientific computing challenges. In TF, numerical computations are represented as a graph of connected operations [14] that can be executed on multiple devices

✉ Dirk Van Essendelft
dirk.vanessendelft@netl.doe.gov

¹ National Energy Technology Laboratory, Morgantown, WV 26505, USA

² Oak Ridge Institute for Science and Education, Oak Ridge, TN 37830, USA

³ Leidos Research Support Team, Pittsburgh, PA 15236-0940, USA

including multicore CPUs, general purpose GPUs, and custom-designed application-specific integrated circuits (ASICs). This enables TF-based applications to be run on a wide variety of hardware and to scale well from powerful GPU servers down to mobile devices [15]. In addition, TF offers multiple levels of abstraction and operates on multiple languages and operating systems [14]. These features allow TF to substitute or supplement the existing CFD frameworks to accelerate computations. At its heart, TF is an extremely powerful math library that allows for both hardware agnostic programming and hardware-specific optimization depending on the requirements. Recently, Zhao et al. [16] reported that a TF-based CFD simulation with GPU acceleration required almost ten times less computational time compared to the same simulation carried out on CPUs using a Fortran code.

The MFIX (Multiphase Flow with Interphase eXchanges) CFD code was chosen as the base development platform to investigate GPU acceleration using TF. MFIX is an open-source multiphase flow solver developed at the National Energy Technology Laboratory (NETL) and employs the two-fluid model (TFM), the discrete element model (DEM), and the multiphase particle-in-cell (MP-PIC) model for modeling a wide range of multiphase flows [17]. MFIX has been widely used for more than three decades and is well validated, particularly for gas–solid fluidized beds, by TFM [18, 19], DEM [20–22] and MP-PIC models [23–25]. In industrial-scale semi-dense multiphase flows, where the focus is more on trends than solution precision, the MP-PIC model is best suited because it uses a statistical averaging technique that enables simulations to be quickly advanced with minimal particle-level overhead [17]. In addition, using solid-phase normal stress for particle interactions rather than direct consideration of particle collisions can remarkably accelerate computation speed on CPU–GPU hybrid computing platforms [26].

The present study aims to develop an equation-based parallelization method toward efficient and effective GPU acceleration. The TF-based platform allows independent computations for each transport equation by assigning variables into separate processor devices. This enables memory- and computation-efficient, intuitive parallelization of solution procedures depending on the computation load for each equation. The developed MFIX-AI code has been rigorously verified by comparing the number of matching digits in the results, and a dramatic increase in speed was established by using a single GPU node rather than dozens of multicore parallel CPU nodes. Furthermore, MFIX-AI with multiple GPUs can natively deploy arbitrary artificial intelligence and machine learning (AI/ML) models at any point within the CFD code, which allows further optimization in performance.

2 Methodology

2.1 MP-PIC model

The MP-PIC model in MFIX consists of a Eulerian model for continuous fluid phase and a Lagrangian model to track the position and trajectory of solid particles. Instead of resolving each individual particle, MP-PIC uses the concept of “parcels,” wherein groups of identical particles are represented by parcels that are tracked in a Lagrangian manner. A parcel’s volume is defined by $V_{\text{parcel}} = \omega V_{\text{particle}}$ where ω is the statistical weight, representing the average number of particles per parcel. Several assumptions are involved in the MFIX MP-PIC model [17], for example, particles within a computational parcel are assumed to be spherical in shape. Also, instead of considering particle collisions or using any Newtonian mechanics to calculate the individual particle displacement and velocity, the MP-PIC model creates an aggregated solids stress momentum source term that directly affects local solids velocity. Furthermore, the parcels are assumed to maintain a mean density and not experience any rotation. With these assumptions, the three-dimensional gas phase continuity and momentum equations can be written as follows [23]:

$$\frac{\partial}{\partial t}(\epsilon_G \rho_G) + \nabla \cdot (\epsilon_G \rho_G \mathbf{u}_G) = 0, \quad (1)$$

$$\frac{\partial}{\partial t}(\epsilon_G \rho_G \mathbf{u}_G) + \nabla \cdot (\epsilon_G \rho_G \mathbf{u}_G \mathbf{u}_G) = -\epsilon_G \nabla p + \nabla \cdot \bar{\bar{\tau}}_G - \mathbf{F} + \epsilon_G \rho_G \mathbf{g}, \quad (2)$$

and the equations for parcel motion can be represented as

$$\frac{d\mathbf{x}_p}{dt} = \mathbf{u}_p, \quad (3)$$

$$\epsilon_p \rho_p \frac{d\mathbf{u}_p}{dt} = -\epsilon_p \nabla p + \mathbf{F}_p + \epsilon_p \rho_p \mathbf{g} - \nabla \tau_p, \quad (4)$$

where ϵ_G and ϵ_p are the volume fraction of gas and parcel, ρ_G and ρ_p are gas density and parcel density, and \mathbf{u}_G and \mathbf{u}_p are three-dimensional velocity vectors for gas and parcel, respectively. \mathbf{g} denotes the gravity vector. $\bar{\bar{\tau}}_G$ in Eq. (2) is the fluid shear stress tensor and τ_p in Eq. (4) is a frictional stress term proposed by Snider [27] as

$$\tau_p = \frac{P_s \epsilon_p^\beta}{\max[\epsilon_{CP} - \epsilon_p, \alpha(1 - \epsilon_p)]}. \quad (5)$$

Here, ϵ_{CP} is the close pack volume fraction, and α , β and P_s are scalar model parameters. The fluid–particle momentum exchange is represented by the interphase drag force terms \mathbf{F} and \mathbf{F}_p in Eqs. (2) and (4), respectively, where \mathbf{F} is the interpolated force from the parcel location to the

corresponding fluid cell, and F_p is the force at the parcel position. The Syamlal–O’Brien model [28] was employed among the various types of drag models tested to form the force terms. Further discussions on drag models, numerical scheme and discretization are beyond the scope of this study, and interested readers may refer to the MFiX Theory Guide [17] for more details.

2.2 Code structure

This section introduces the coupling strategy to establish MFiX-AI and its basic code structure. This study is based on MFiX version 19.3.1 (available at <https://mfix.netl.doe.gov/doc/mfix/19.3.1/about.html>) and TF version 1 or 2 (with TF 2.x behavior deactivated in the latter).

2.2.1 Coupling strategy

In MFiX-AI, the core solution procedure occupying the most computation time was rewritten using TF, while miscellaneous functions such as computational domain handling, initial particle population and post-processing remained in the basic Fortran of classic MFiX. Figure 1 demonstrates the workflow within MFiX-AI. The mfix.f file is placed on the top level of the code hierarchy in MFiX. This file controls major code behavior such as initialization, time marching, and calls the fluid and PIC solvers. A flag ‘USE_TF_SOLVERS’ was introduced to activate the use of the tf_solver.f

file for MFiX-AI instead of following the classic MFiX workflow. There are two calls—the first call INITIALIZE_TF is used to initialize the TF graph, and the second call SOLVE_TF is used to take a separate timestep inside the TF graph. The INITIALIZE_TF subroutine in the tf_solver.f file gathers the required model parameters and truth tables defined during MFiX pre-processing and passes them into the MFiX-AI code. The data exchange between the Fortran-based MFiX code and the Python-based MFiX-AI code is established by the additional C wrapper code tfpywrapper.c. The wrapper code is used to directly transfer memory from Fortran to TF. The Python functions are defined in the tf_loading_nn.py file that feeds the placeholders in MFiX-AI and runs the TF graph. The body of the CFD computations are handled within the mfixTF_mod.py file and associated modules. For visualization and data recall, the field variables of interest are transferred back to the Fortran code through the same call chain in the reverse direction. Aside from initialization and data recall, no memory transfers into or out of TF during the solution procedure except the timestep value. Further, the entire graph is GPU compatible and can be executed with no host/GPU data transfer except for the desired timestep at the head of every timestep.

2.2.2 Parallel computing strategy

MFiX-AI is a direct port from Fortran to the TF Python ecosystem and was designed to run on GPUs. As such, the

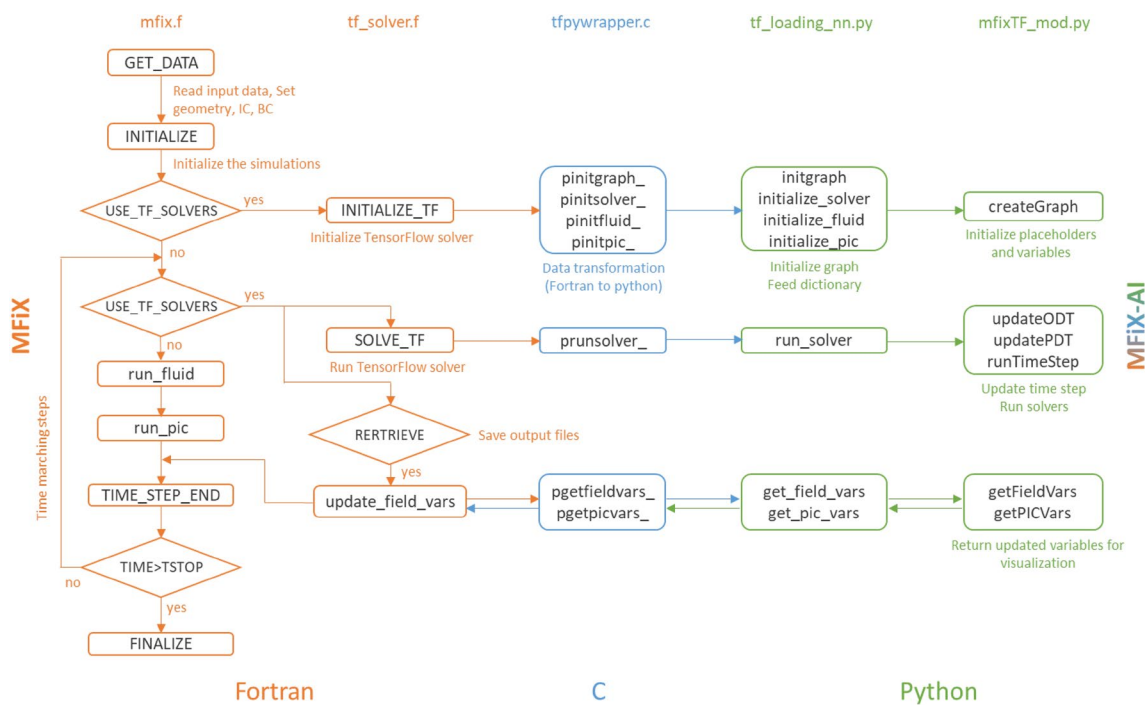


Fig. 1 Code structure of the data exchange process between MFiX and MFiX-AI

solution procedure used in MFiX was rewritten using graph representation in TF. MFiX-AI employs placeholders, tensors, and variables supported by TF. Placeholders are used for data input, tensors are immutable data structures within graphs, and variables are mutable data structures. The main part of MFiX-AI is initialized by using placeholders that store values in variables for later recall. A TF variable is designed to maintain a shared, persistent state manipulated by a program [14], so the variables in MFiX-AI are used for the data that should remain on the device between session run calls. Tensors are immutable data structures used for intermediate data that does not need to persist between session run calls. Initialization flows from placeholders through tensors and ends in variable assignment. For calculation of timesteps, the graph flows from reads of variables through tensor calculations to variable updates. Because it is only safe to update variables once per session run call, it was decided to create a graph that does a single complete timestep and to make repeated calls into TF to advance the time from the Fortran base code. This method also allows for easy matrix construction and solving without the need for repeated field variable and/or matrix data transfers between host and device.

An equation-based parallelization scheme that can utilize multiple GPUs was developed. First, required variables for each component of the momentum equations were declared independently on individual devices (referred to as Udev, Vdev and Wdev), which allowed simultaneous solutions of the momentum equations. If other transport equations such as species and energy are needed, they can also be solved

simultaneously on different devices. Next, a set of devices were utilized to solve for pressure, which is a dependent variable dictated by the momentum equation. In the current version of the code, a multi-GPU linear solver feature can be used for pressure. A PICdev was assigned to calculate MP-PIC particle motion and flow. For simplicity, the PICdev is always GPU 1 and Pdev for pressure is also GPU 1 when the multi-GPU linear solver mode is disabled. Currently, the PIC equations are solved over a single GPU, which creates a serial workload. Future versions of the code will distribute this across all available devices.

Figure 2 shows schematic diagrams of parallel computing strategies in MFiX and MFiX-AI. The parallel computing strategy for multicore CPU computations in MFiX is a typical MPI-based approach that requires domain decomposition and significant MPI communications to transfer data among the subdomains. The need for MPI communications is often a bottleneck in HPC. In particular, the latencies to communicate dot products often limit the minimum iteration time. Several dot product updates are required at every iteration of the linear solver for every SIMPLE solver iteration. Frequent small communications such as dot products are not advantageous, as many studies have shown that as much as 80% of the computation time in modern HPC codes is spent waiting for data to arrive [29–31]. This creates a situation where time to solution does not decrease as processor workload decreases past a certain point, as is seen in most strong scaling curves. The situation can be even more complicated with GPU based solvers because the device efficiency also depends on having enough workload to balance

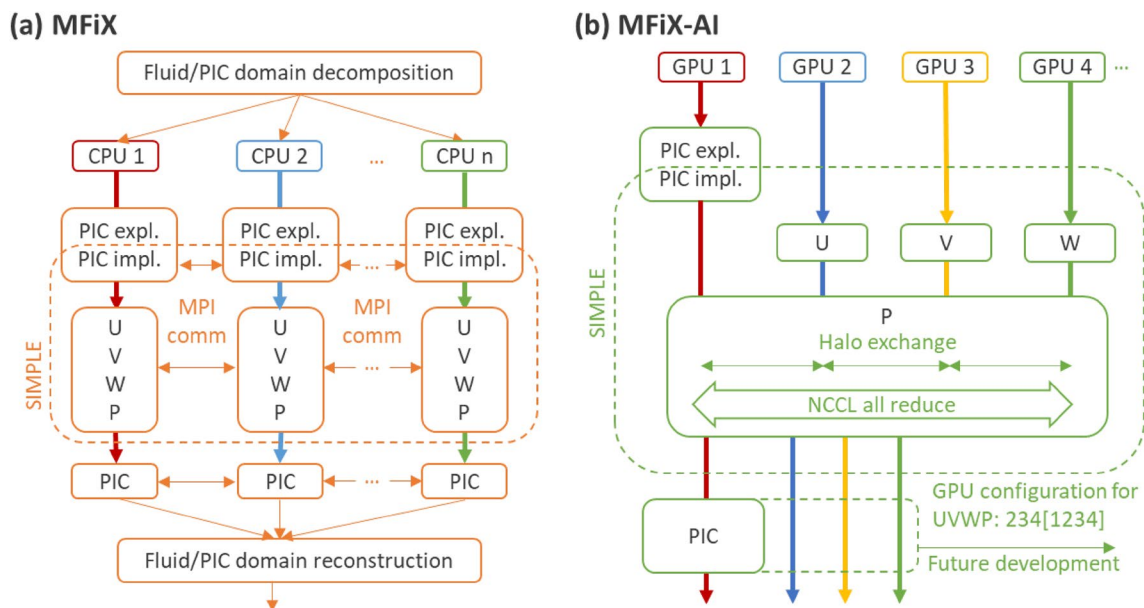


Fig. 2 Parallel computing strategy for MFiX on multicore CPUs and MFiX-AI on multi-GPUs. **a** Typical MPI process for the decomposed computational domain used in MFiX, **b** multi-GPU data communication used in MFiX-AI in the TF graph

the processing relative to the device global memory access. Even without latency limitations, performance does not scale to small workloads on GPUs, which limits minimum time to solution in parallel workloads as well.

MFiX-AI utilizes an equation-based parallel computing strategy. This kind of strategy is ideal for modern GPU nodes like the DGX A100 with the NVLink switch. This strategy requires very infrequent communication of entire field variables and no dot product synchronization points in the linear solver unless the parallel pressure solver is turned on. This software architecture trades very frequent small data transfers for infrequent large data transfers. Modern GPU nodes have exceptionally high inter-device bandwidths of up to 7.2 terabyte/s for NVLink version 3.0, which makes this software architecture tractable since the time to move large data sets between GPUs on the same node is minimal. The inter-device data transfer rates are over 144 times the fastest 400 Gb/s inter-node transfer rates. While this is a substantial increase in bandwidth, latencies have not improved much and are still on the order of 10–30 μ s. Thus, the equation-based parallelization methodology is expected to work very well on problem scales that can fit in the memory space of a single node. MFiX-AI also has an advantage because formation and solve performed on device and host/device data transfers are all but eliminated. The equation-based parallelization strategy can also be more efficient than the current block partitioning in MFiX, where load imbalance can severely affect parallel efficiency. Finally, the equation-based strategy is expected to continue to pay dividends in future HPC systems where GPUs become larger with more memory, as inter-device bandwidth grows significantly (especially relative to inter-node bandwidth) and communication latencies only make marginal improvements.

The parallelization strategy for the pressure equation works well provided that the work per GPU is high and that enough iterations are needed for convergence to overcome the setup overhead. However, the synchronization latencies and workload utilization efficiencies are expected to limit how many GPUs can effectively be used at a given problem scale. This strategy could be extended to transport equations, but these equations usually converge in just a few iterations and it is difficult to make up the costs of setting up the solver for parallel work.

Depending on the available number of GPUs, users can specify GPUs for each Udev, Vdev, Wdev, and Pdev. PICdev is always on GPU 1 in the current version of the code. For example, when only a single GPU is available, all equations are assigned to one GPU, which can be symbolized as 111[1] representing UVW[P] device where PICdev is always assigned to 1. Figure 2b shows an example case of a parallel run with four GPUs (234[1234] decomposition) where three GPUs are assigned to the three velocity components and one GPU is assigned to

PICdev. The four numbers in the square bracket denote the workload for pressure distributed to the four GPUs. This configuration is currently the most memory and computationally efficient way to run the code.

For fluid–particle coupling, PICdev is used at least twice in the solution procedure of both the codes depending on whether implicit or explicit coupling is used. At the head of the SIMPLE iteration, coupling from particle to fluid is achieved where the momentum exchange terms \mathbf{F} in Eq. (2) are calculated. In explicit coupling (PIC expl. in Fig. 2), the particle to fluid terms are calculated before the first SIMPLE iteration and are not updated during subsequent SIMPLE iterations. In implicit coupling (PIC impl. in Fig. 2), those terms are updated at the beginning of every SIMPLE iteration. Thus, calculation times, which are compared in the performance analysis in Sect. 4, depend on the coupling method. The second PICdev utilization comes after the SIMPLE iteration and is used to update parcel positions and velocities. The interpolation between particle positions and Eulerian grid values is a known area for further parallel development. The computations could be spread over multiple devices using the high bandwidth available between GPUs on the same node, much like the fluid solver.

2.2.3 Use of TF graph and custom operator

MFiX-AI uses TF in graph execution mode, where an efficient computational graph is created before carrying out the actual calculations. A computational graph represents a series of prearranged TF operations that manipulate tensors or tensor-like objects. MFiX-AI employs rank-1 tensors for variables used in calculations and for indices indicating the positions of the computation cells corresponding to the variables. Instead of using element-wise looping, a combination of merge, gather, and scatter operations is used for vector arithmetic. To avoid conditional statements, truth tables are predefined in Fortran, fed into the graph as rank-1 tensors, and converted to compressed int64 representations to save memory. Although most calculations in MFiX can be done with combinations of standard TF operations, a few complex calculations containing many irregular conditional statements and performance-critical operations are handled through custom operators. Custom operators (or custom ops) in TF allow users to develop functionality that is not defined as default TF operators. The custom ops are written in C++ and whatever language is compatible with other supported devices (such as CUDA for NVIDIA devices and HIP for AMD). The use of custom ops can also be beneficial for reducing the number of kernel calls, optimizing memory usage during the operations, and optimizing memory access patterns.

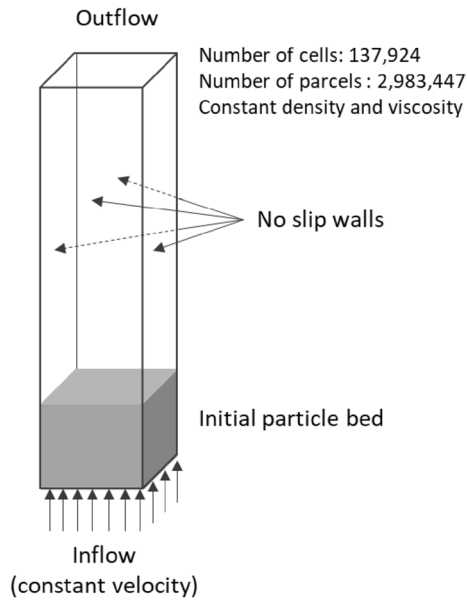


Fig. 3 Computational domain and boundary conditions for the verification case

3 Code verification

MFiX-AI was verified by conducting a code-to-code comparison against the classic MFiX code. Figure 3 shows the schematics of the verification case. The test case was an isothermal simulation of a fluidized bed in a rectangular cuboid geometry with dimensions of 12 cm, 72 cm, 12 cm in x , y , and z direction, respectively. The computational domain was discretized using a structured mesh with 27, 162 and 27 cells in x , y , z direction, respectively, such that the total number of cells is 137,924. The constant density and viscosity of the gas phase were set to 1.093 kg/m^3 and $1.9 \times 10^{-5} \text{ Pa s}$, respectively. Gravity acts in the $-y$ direction. The domain inlet was set at the bottom ($y = 0$), with gas flow velocity ranging between 0.3 and 1.0 m/s, and the outlet at the top ($y = 72 \text{ cm}$). The vertical planes were treated as no-slip walls. The particles were assumed to be of identical size with diameter of $400 \mu\text{m}$ and density of 2000 kg/m^3 . The initial bed height was set to 12 cm with the gas volume fraction set to 0.42 for the bed area. With a statistical weight of 10 (the number of particles per parcel), the total number of parcels present in the domain was 2,983,447. The initial timestep was set to 1 ms and varied depending on the convergence of SIMPLE iterations with the given tolerance and maximum number of iterations. The average Courant number according to the initial timestep is in the range of 6.75×10^{-4} to 2.25×10^{-3} depending on the gas flow velocity. No preconditioners were used for the linear solvers in either of the codes to ensure that the solver settings are identical.

3.1 Verification for a single timestep

The test case was conducted using both codes in 64-bit precision to maintain conformity between the MFiX and MFiX-AI simulations. Usually, the initial positions of the particles are set in a randomized manner that precludes an exact comparison of the time evolution of the two codes' simulations. Hence, both simulations were restarted from an MFiX calculation where the bubbling particle motion was active. Residuals of the velocity components and pressure calculated from both codes were in excellent agreement as presented in Fig. 4, and a converged solution was obtained in 33 SIMPLE iterations for both codes with the same tolerance for the SIMPLE iterations. In addition, both codes were modified to write the same field variables to file for every cell and every particle at the same point in the solution process. This allowed a comparison of the number of matching significant digits by using the following equation for the field variable ϕ from each code.

$$\text{Number of digits matching} = -\log_{10} \frac{|\phi_{\text{MFiX}} - \phi_{\text{MFiX-AI}}|}{|\phi_{\text{MFiX}}|} \quad (6)$$

This is a very restrictive comparison method as the error is relative to the order of magnitude, and not the absolute difference.

Figure 5 displays histograms for the matching number of digits for the selected variables from both codes at different steps in the solution procedure. The matching number of digits for the velocity components and pressure at the first SIMPLE iteration are shown in Fig. 5a. The velocity

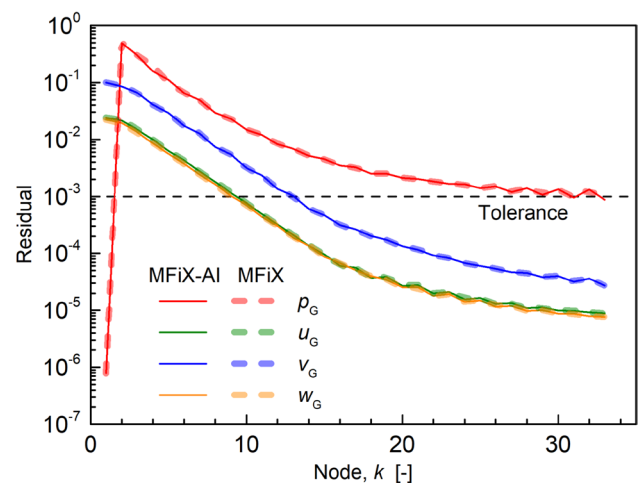


Fig. 4 Residual plots of velocity and pressure from MFiX-AI (solid lines) and MFiX (dashed lines) during SIMPLE iterations at the first timestep

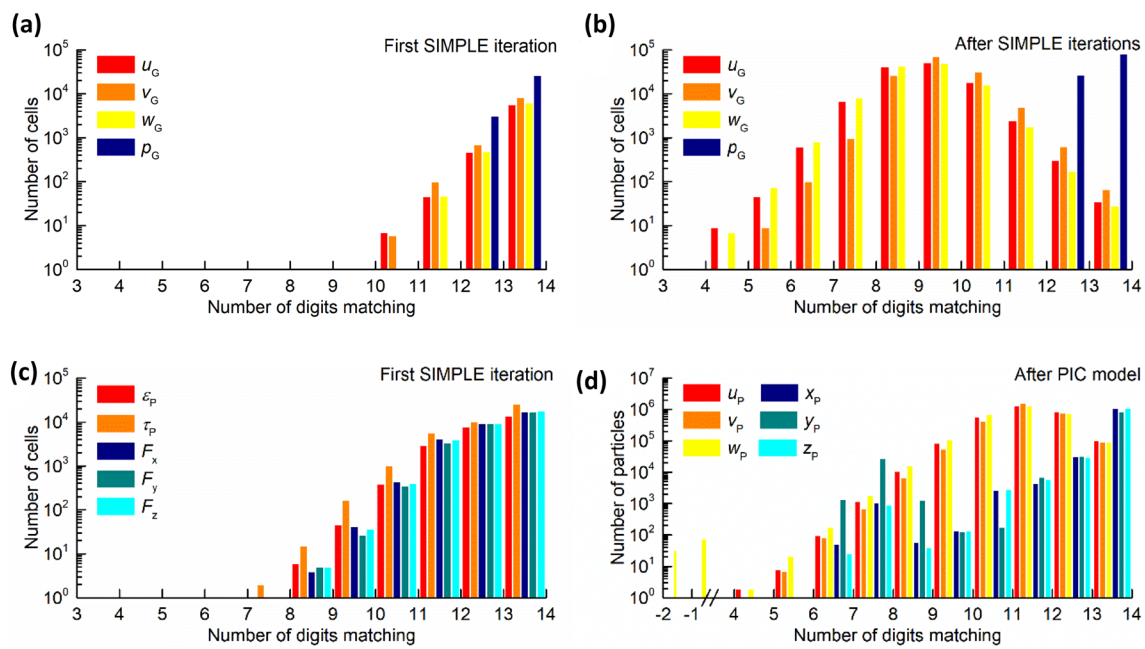


Fig. 5 Histograms for the number of matching digits for non-zero quantities obtained from MFiX-AI and MFiX. **a** Comparison of velocity and pressure at the first SIMPLE iteration, **b** at the first

timestep, **c** variables affecting fluid–particle momentum exchange at the first SIMPLE iteration, and **d** particle position and velocity after PIC calculation

components match within 10 digits and most cells have 14 matching significant digits. However, the pressure matches with fewer digits of precision as it is a quantity derived from the velocity components. In MFiX-AI, the pressure used in the SIMPLE iterations is gauge pressure while absolute pressure was used in MFiX, which allows MFiX-AI to maintain five more digits of precision than MFiX. Also, the linear equations for pressure generally take many more iterations to converge than those of the velocity components, which may contribute to error accumulation. Hence, the corrected velocity field at the end of the first SIMPLE iteration (that includes contributions from the predicted pressure field) also shows fewer matching significant digits (similar to that for pressure). That said, most cells match within nine significant digits and the worst cell matched with four. Even with the worst number of digits matching, they are still in very good agreement as it represents less than 0.01% relative difference, given that the two simulations were carried out on different hardware. Note that the pressure values used in Fig. 5a, b represent absolute pressure.

Figure 5c shows that the variables related to the fluid–particle momentum transfer can also influence error propagation. The solid volume fraction ϵ_p is calculated using the bilinear numerical interpolation scheme based on the given particle positions from MFiX, but it only matches within eight significant digits. The solids shear stress, τ_p , lost one more matching significant digit after Eq. (5) presumably due to the power function of ϵ_p in the numerator. F_x , F_y and F_z

represent the components of the particle force vector \mathbf{F} in Eq. (2). Although the particle-related variables in Fig. 5c did not significantly affect the gas velocity components as shown in Fig. 5a, those errors further propagated into the MP-PIC model’s calculation of solid motion. Another significant divergence in behavior was observed during the treatment of parcel reflections at the boundary. The problem stems from hard conditionals that determine the occurrence of parcel reflections, which are affected by very small differences in position. A bit perfect match is needed to ensure the same conditional execution in both codes, and this may not be possible because the computations were performed on different hardware. This unavoidable difference resulted in negative values in Fig. 5d. Apart from the discrepancies observed on a few particles, the velocity of most particles matched within 11 significant digits and the position of most particles matched within 14 digits, while in the worst case, they matched within 4 and 6 digits, respectively.

3.2 Validation of bed characteristics

Despite very good agreement observed at the first timestep, the accumulation of bitwise differences leads both results to quickly diverge after a few timesteps. This is not a concern as results are expected to be non-deterministic; however, making a side-by-side comparison at every timestep is impossible. Therefore, instead of instantaneous differences between the predicted variables, the physical behavior of

the fluidized bed was time averaged at a variety of gas inlet velocities. To obtain sufficiently time-independent mean fields, the simulations were run for 102 s (in physical time) starting from the initial conditions. The first two seconds were discarded to eliminate the influence of the initial conditions on the time averaged field. A fast Fourier transform of the time series of gas volume fraction at a point (located at the middle of the initial bed at $[x,y,z] = [6 \text{ cm}, 6 \text{ cm}, 6 \text{ cm}]$) was carried out to obtain a bubble frequency spectrum.

Figure 6 demonstrates the results of the fluidized bed simulations from MFiX and MFiX-AI. Figure 6a, b compare the instantaneous and mean fields of gas volume fraction and gas velocity magnitude from both codes at two gas inlet velocities: 0.5 and 1.0 m/s. As expected, the instantaneous comparisons do not match, but the mean fields from both codes are qualitatively almost identical. Figure 6c compares both mean gas volume fraction and gas velocity magnitude profiles at the center of the domain in y direction.

The maximum differences between the two code profiles are only 0.02 for gas volume fraction and 0.15 m/s for gas velocity magnitude. Moreover, Fig. 6d shows a similar dominant frequency and bubble frequency spectrum, which indicates that the bubbling phenomena obtained from the two cases are in excellent agreement. Hence, the series of evidence supports that the bubbling flow dynamics from both codes are very similar.

In addition, Fig. 7 exhibits the pressure drop obtained from both codes for inlet velocity of 0.3–1.0 m/s. The pressure drop obtained from the two codes agree very well, with the relative difference being less than 1%. Those values are also close to the bed pressure representing the total weight of the particle bed. The series of verification and validation tests conducted in this section concludes that MFiX-AI can successfully replicate the predictions from the MFiX code, which is already well validated with physical systems [23–25].

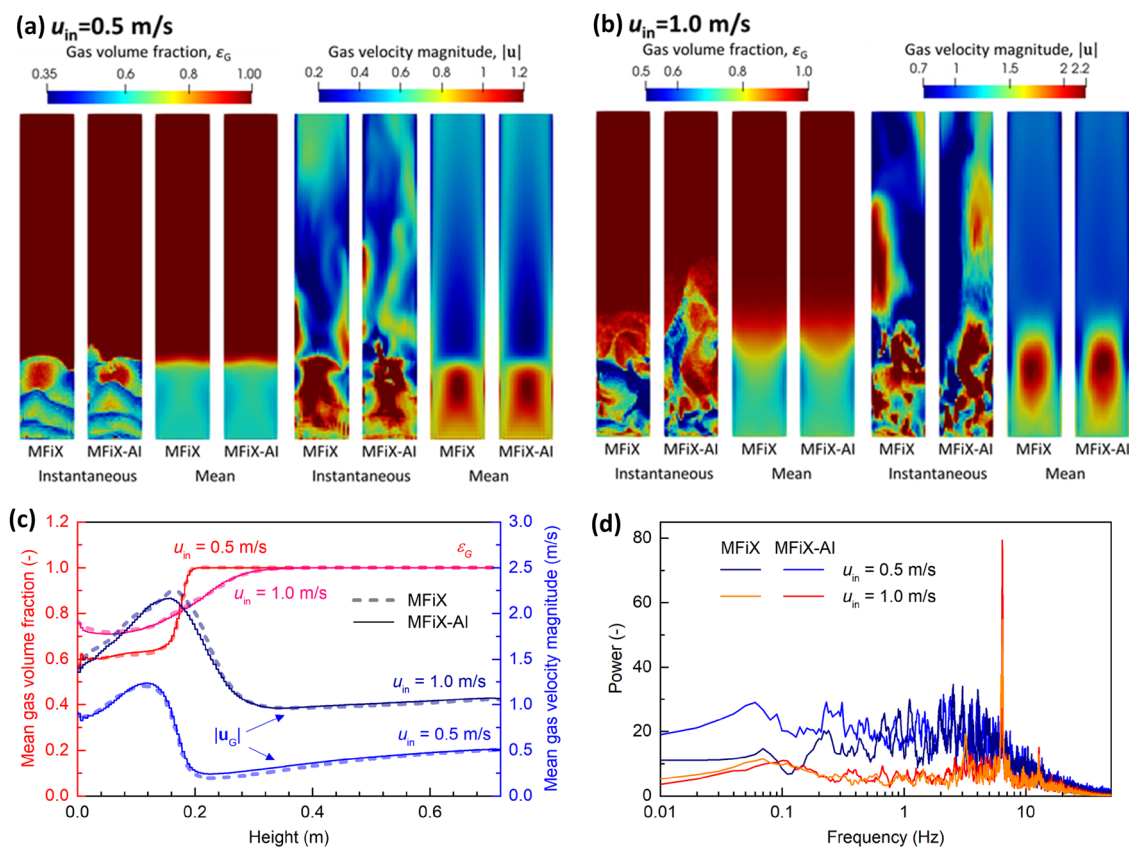


Fig. 6 Result of fluidized bed simulations from MFiX and MFiX-AI codes for inlet gas velocity of 0.5 and 1.0 m/s. Instantaneous and time mean fields of gas volume fraction and gas velocity magnitude from both codes for inlet gas velocity of **a** 0.5 m/s and **b** 1.0 m/s. **c** Comparison of mean gas volume fraction and mean gas velocity magnitude profiles at the center of the reactor. **d** Comparison of bubble frequency spectrum from both codes

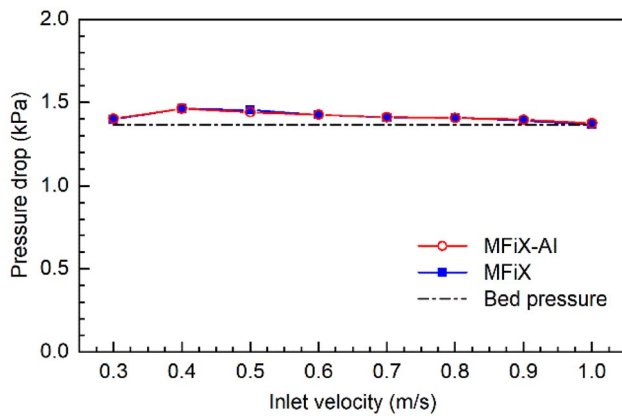


Fig. 7 Comparison of pressure drop obtained from the results of both codes running with the gas inlet velocity ranging from 0.3 to 1.0 m/s

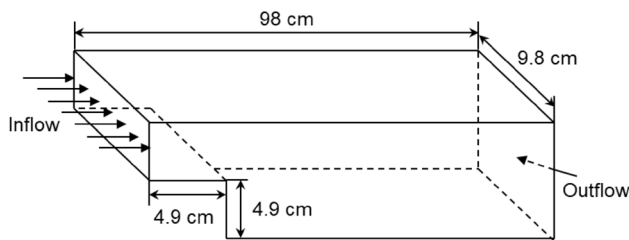


Fig. 8 Computational domain and boundary conditions for the BFS benchmark case

4 Parallel scaling performance

The performance enhancement of MFiX-AI against MFiX was evaluated by running identical example cases on Nvidia A100 GPUs (NVLink connected on a DGX A100), each with 80 GB memory for multi-GPU runs (A100-SXM-80 GB). The JOULE 2.0 supercomputer at NETL, utilizing Xeon Gold CPU (6148 20C 2.4 GHz) with Intel Omni-Path 100Gbit interconnect, was used for the MPI-based multicore CPU runs. The performance analysis consists of strong scaling analysis with fixed problem size, and weak scaling analysis with problem size fixed per CPU node or GPU. First, single-phase calculations without particles were conducted to investigate the speed gains solely for solving the steps in the SIMPLE algorithm. A backward-facing step (BFS) flow case was chosen as the test case for the single-phase calculations. Figure 8 shows the computational domain spanned 9.8 cm, 4.9 cm and 98 cm in x , y , z directions where an internal block with dimensions of 4.9 cm, 4.9 cm, and 9.8 cm was used to construct the step. Fluid viscosity and density were set to 1.8×10^{-5} Pa s and 1.0 kg/m^3 , respectively, and the inlet velocity (directed along the $+z$ direction) was set to

1 m/s. Secondly, a PIC example case was also used for performance analysis, although the domain was slightly different from that used in the validation study. The computational domain was a cuboid as for the verification case shown in Fig. 3. The dimensions are 12 cm, 12 cm, 36 cm in the x , y and z directions, and the initial bed height was set to 12 cm. For efficient parallel computation, the longest side was aligned along the z direction, and gravity acts in the $-z$ direction. The density and viscosity were the same as those in the performance analysis study. The inlet gas velocity was set to 0.15 m/s, and the particle diameter and density were specified as $200 \mu\text{m}$ and 2000 kg/m^3 , respectively. For both single-phase and PIC examples, the number of cells and particles varied depending on the scaling methods. For both cases, Norm_g, which is a factor to normalize the gas continuity equation residual in MFiX, was set to unity.

4.1 Strong scaling

Strong scaling analyses for single- and multiphase cases were conducted to compare the speed gains and scalability of both MFiX-AI and MFiX codes. Figure 9a, b exhibits the results of strong scaling for single-phase BFS calculations. While the number of cells was fixed as 10,001,880 (10 M), 23,708,160 (20 M) and 30,802,500 (30 M) in this analysis, the number of GPUs used for MFiX-AI calculations varied from 1 to 8, and the number of nodes for MFiX calculations varied from 1 to 25, where one node consists of 40 CPU cores. The timestep was initially set to 10^{-4} s to guarantee convergence from the initial condition, and gradually increased to the maximum timestep of 5×10^{-4} s. The physical time for calculations was 0.5 s. Figure 9a compares the average computation time per SIMPLE iteration of MFiX and MFiX-AI with regards to the size of test problem and the number of CPU nodes or GPUs. MFiX-AI showed a dramatic boost in performance compared to MFiX for all three cases. Figure 9b shows the speed gain of MFiX-AI against MFiX for three combinations of the number of GPUs and the number of CPU nodes. The speed gain generally increased as the computation size increased. The maximum speed gain was 12.4, which corresponds to the use of a single GPU node against a single CPU node for the 10 M case. It should be noted that the calculation times for four GPUs were even slightly faster than the computation times for 25 nodes corresponding to 1,000 MPI ranks on the CPU. Scaling for MFiX-AI stopped improving with four GPUs, as this problem involves only four equations. If more equations were used, the scaling may continue provided that the parallel solver for pressure can continue to efficiently scale.

Figure 9c, d show the results of strong scaling for a benchmark case of PIC calculations with a fixed number

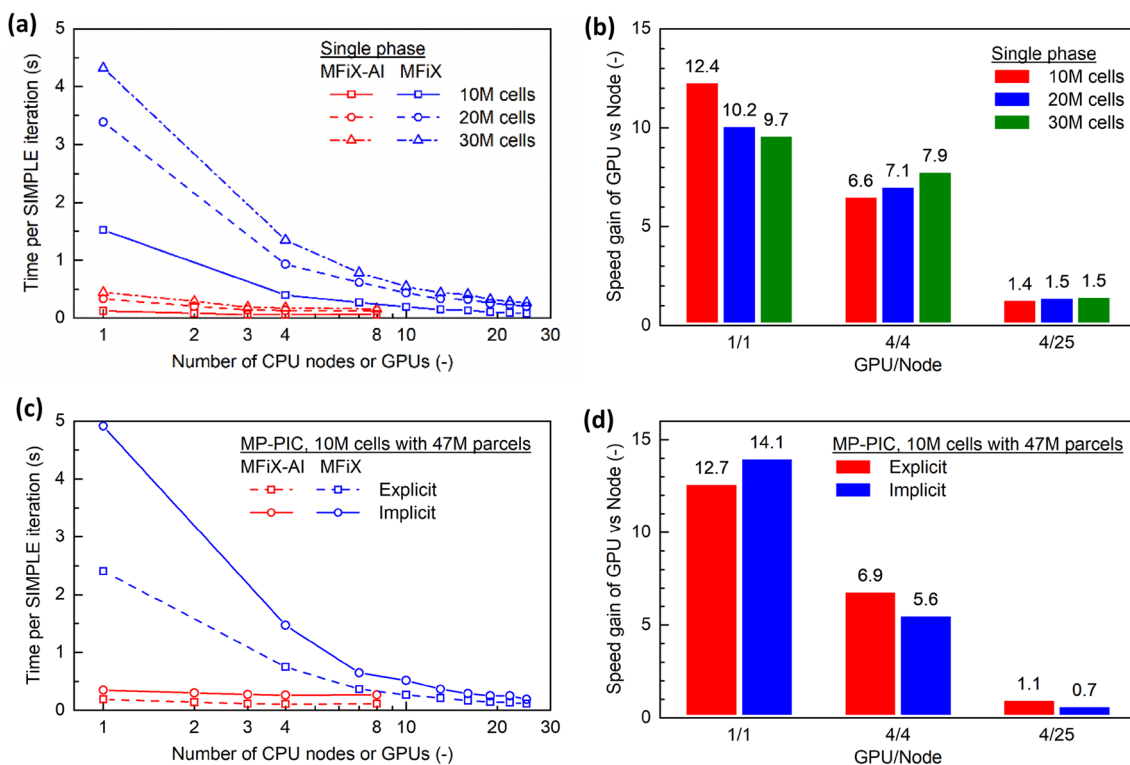


Fig. 9 Strong scaling analysis for **a**, **b** single-phase backward-facing step flow computations and **c**, **d** MP-PIC fluidized bed computations obtained by MFiX-AI with multiple A100 GPUs and MFiX with multicore CPU nodes. The average time per SIMPLE iteration of both

codes for **a** single phase in terms of cell counts, and **c** those for MP-PIC computations in terms of particle coupling method. Speed gain of MFiX-AI against MFiX for three pairs of GPU/CPU node combinations for **b** single-phase computations and **d** MP-PIC computations

of cells (10,125,000; ~ 10 M) and a fixed number of parcels (47,853,435; ~ 47 M). The initial timestep was set to 5×10^{-4} s, and the 0.5 s (physical time) long calculation was used for the strong scaling analysis. The number of GPUs and CPU nodes employed are identical to those used in the single-phase strong scaling analysis. However, instead of comparing the cases with different computation sizes, this comparison focuses on computation times depending on the coupling method. The time per SIMPLE iteration was estimated by dividing the total wall time spent in the SIMPLE routine by the total number of SIMPLE iterations. Thus, the time for PIC computations in both explicit and implicit coupling were included. The difference in calculation times for implicit and explicit coupling is only dictated by the different number of calculations for the momentum exchange terms \mathbf{F} in Eq. (2). In explicit coupling, this is calculated once before the start of the SIMPLE loop, while, in implicit coupling, the calculation is performed at the start of every SIMPLE iteration. In Fig. 9c, the average computation time per SIMPLE iteration for MFiX-AI with a single GPU is much less than MFiX with a single CPU node and is comparable to MFiX with 25 CPU nodes. This dramatic boost in the MFiX-AI computation speed is more clearly shown in Fig. 9d, which presents the speed gain of MFiX-AI against

MFiX. The speed gain of single GPU compared to a single CPU node is approximately 14. The optimal case of MFiX-AI using four GPUs shows excellent performance compared to MFiX using multiple CPU nodes. The computation time for the optimal case is very close to that for MFiX using 25 nodes (= 1000 CPU cores); for the explicit case, MFiX-AI is 10% faster than MFiX while for the implicit case MFiX-AI is only 30% slower than MFiX. The observations in both single- and multiphase computations clearly show the dramatic reduction in computation time for MP-PIC calculations using MFiX-AI compared to MFiX. It can be further improved by parallelizing the MP-PIC solver, optimizing the custom operations for hardware-specific implementations, and applying machine learning algorithms for initial guesses of SIMPLE iteration.

4.2 Weak scaling of MFiX vs. MFiX-AI with a fixed number of GPUs

A weak scaling analysis for MFiX with a fixed number of cells and/or parcels per number of CPU cores was compared to the MFiX-AI calculations with a fixed number of GPUs for single-phase BFS flow, PIC with a constant number of parcels, and PIC with a variable number of parcels. The test

conditions were identical to those used for the strong scaling analysis, with different initial timesteps and physical times for calculation, which were 10^{-3} s, 50 s for the BFS case and 5×10^{-4} s, and 1 s for both PIC cases. For MFiX, the cell per node is fixed to approximately 400,000. For PIC with constant parcels, the number of parcels is set to 3,998,062 (4 M), and for PIC with variable parcels, the number of parcels per node is set to approximately 4 M. For comparison, identical calculations were employed in MFiX-AI, but with a constant number of GPUs.

Figure 10 shows the results of the three scaling analyses. In the BFS case, as shown in Fig. 10a, the time per

SIMPLE iteration for MFiX-AI increased almost linearly as the number of cells increased, while for MFiX it showed a gentler increasing trend. The increasing trend for MFiX-AI was expected, given that the workload increased per device. However, the scaling curves for MFiX should be nearly flat in an ideal case. However, almost 2% of the volume was taken out for the step, resulting in a load imbalance, which led to less-than-ideal scaling.

If more than 8 M cells are used for MFiX-AI with two GPUs, the computation time will be larger than MFiX for the single-phase calculations (since the MFiX runs were not restricted by the maximum wall time). The MFiX-AI runs

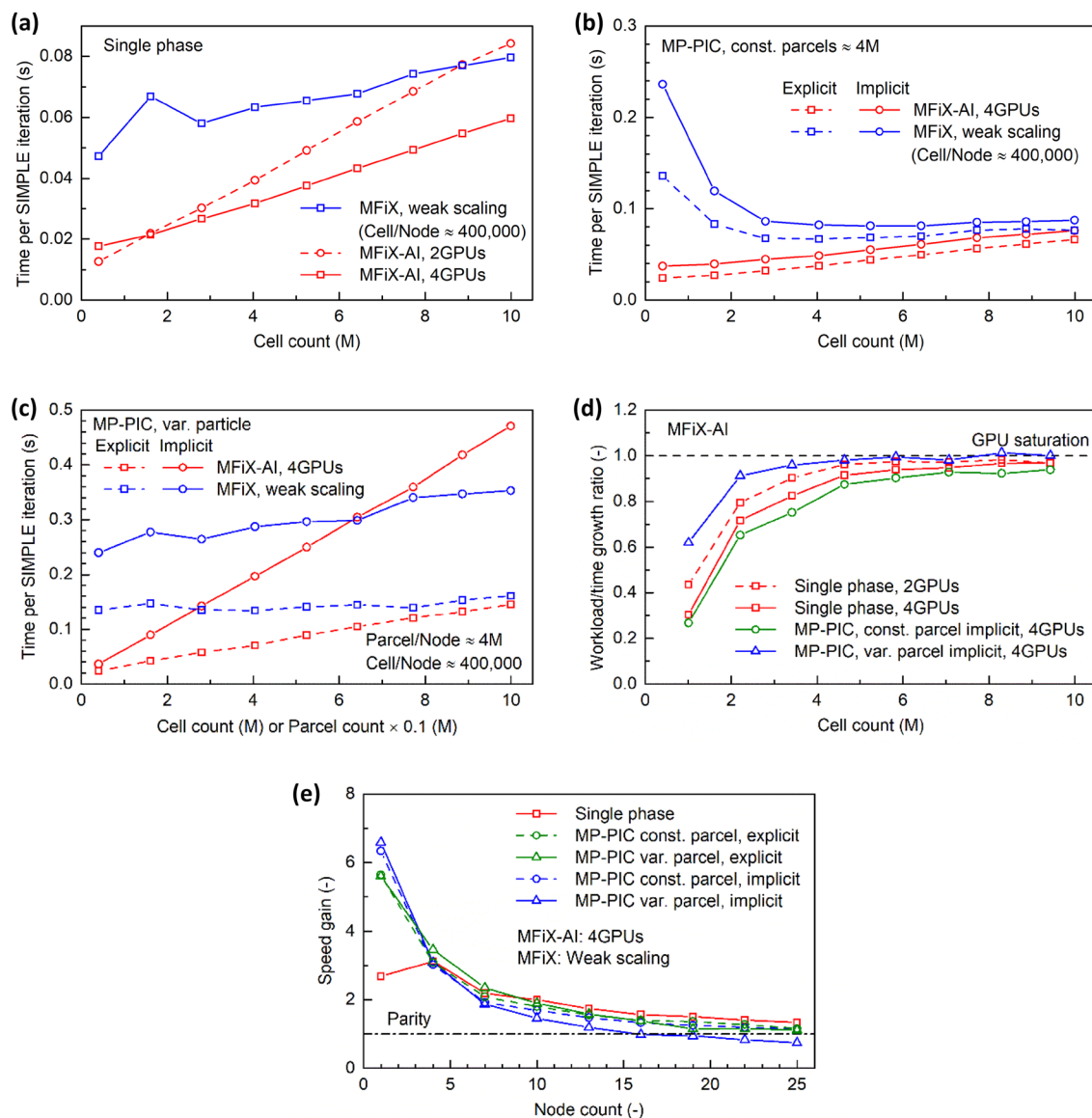


Fig. 10 Weak scaling analysis of MFiX for **a** single-phase BFS flow, **b** MP-PIC fluidized bed calculation with constant 4 million (M) parcels, and **c** variable parcels with 4 M parcels per node. Identical cases of MFiX-AI with fixed number of GPUs were compared. The number

of cells is 400,000 per node for all three cases. **d** The ratio of increments of time per SIMPLE iteration and workload for MFiX-AI computations. **e** Comparison of speed gain of MFiX-AI against MFiX for three weak scaling cases

with four GPUs were almost consistently faster than MFiX, with a difference in time per iteration of approximately 0.02 s (corresponding to total wall time of 1 h) for all test conditions in this study. It is interesting to note that MFiX-AI with four GPUs scaled almost as well through increased load as MFiX in weak scaling with a slight load imbalance.

Figure 10b displays the weak scaling results for PIC calculations with constant parcels. As the number of cells increases, the MFiX time per iteration dramatically decreases until 3 M cells due to the decreasing number of parcels per CPU. The benefit of continued mesh refinement is lost because of particle load balancing in the fluid bed. However, this behavior is not observed with MFiX-AI because the PIC calculations are carried out in a single device (PICdev) and there is no spatial domain decomposition. The computation time for MFiX-AI increases approximately linearly with the cell count. There is almost no difference between computation times of explicit and implicit coupling in MFiX. For MFiX-AI, the computation times for implicit and explicit coupling shifted by an approximately constant value, where the offset represents the additional amount of work in PICdev in SIMPLE iterations. This illustrates the need to parallelize the PIC workloads in future versions of the code. It is noted that MFiX-AI is faster than MFiX for all explicit workloads tested as the increased serial workload for explicit coupling was negligible in performance.

Figure 10c exhibits the same type of scaling analysis for PIC with variable parcels and cells. In this case, the computation times for MFiX-AI show a linear increasing trend, while those for MFiX are almost flat with a slight increase. For explicit coupling, MFiX-AI is always faster than MFiX, while, for implicit coupling, MFiX-AI is only faster than MFiX for less than 7 M cells and 70 M parcels. The rapidly increasing computation time in the implicitly coupled weak scaling shows the incredible expense involved in interpolating between the Lagrangian particles/parcels and the Eulerian grid, especially in MFiX-AI. This workload scales in proportion to both grid size and particle/parcel count. In general, implicit coupling should not be used unless there is a clear need for accuracy or advantage for it. Further this illustrates the need to parallelize this workload effectively in future versions of MFiX-AI as the current version of the code only uses one GPU for the PIC portions of the code. That said, the explicit coupling in MFiX-AI is very competitive to the CPU-based MFiX version.

Figure 10d shows the ratio between time spent for SIMPLE iterations and workload in MFiX-AI computations. The ratios for different cases vary depending on the problem size per number of GPUs employed. A ratio of unity implies GPU workload saturation. In general, workloads become saturated in the 2–5 million cells/parcel range, which is consistent with other reported saturation limits for stencil problems on GPUs [32]. Figure 10e compares the speed gain

of MFiX-AI against MFiX for the three weak scaling cases. For the single-phase BFS case, the maximum speed gain is approximately three times and it almost linearly decreases with increasing number of nodes for MFiX runs. Including the PIC calculation, the maximum speed gain for constant parcels and variable parcels is approximately seven for both implicit and explicit coupling. Speed gain steeply decreases until ten nodes and further decreases linearly as the number of nodes increases. It is noted that the PIC calculation of MFiX-AI with constant parcels carried out on a single GPU is even faster than the corresponding calculation in MFiX on 25 nodes. MFiX-AI runs with variable parcels are still beneficial as compared to the corresponding cases in MFiX on less than 15 nodes and almost equivalent to those on 22 and 25 nodes. In most cases, MFiX-AI with just four GPUs scales almost as well as weak scaling on CPU nodes with MFiX. The performance in MFiX-AI on four GPUs is similar to MFiX runs carried out on many hundreds to thousands of CPU cores.

5 Energy consumption efficiency

The dramatic reduction in computation time leads to significant energy savings. Multi-CPU MFiX computations were conducted on the JOULE supercomputer equipped with Intel Omni-Path interconnects. With 64 Omni-Path Edge and two Omni-Path Director switches, each consuming 187 W and 229 W of electricity, respectively, the power consumed by Omni-path interconnects corresponding to each of the 1856 nodes on JOULE is estimated as 6.7 W. Thus, the total power per each node, defined by the sum of the CPU node (800 W) power requirement and the interconnect per node (6.7 W) power requirement, is 806.7 W. Energy consumption of MFiX-AI is estimated from the maximum power requirement of the DGX A100 system, which is 6.5 kW [33]. By using four A100 GPUs (each of which consumes 400 W [34]) out of eight GPUs, the estimated total power for running MFiX-AI is 4.9 kW. Figure 11a, b compare the energy consumption of both codes for single SIMPLE iteration in single-phase and MP-PIC calculations. In strong scaling analysis, the energy consumption of MFiX-AI with four GPUs is the lowest for the single phase, while slightly increasing from one to four GPUs for MP-PIC. This evidence consistently shows that the use of four GPUs is optimal. The energy saving was estimated by comparing the energy consumption of four GPUs against that of 25 CPU nodes. Despite the higher power consumption rate of DGX A100 compared to a single CPU node, the MFiX-AI highly benefits from fast computation time. The energy savings of four GPUs against 25 nodes is 82% for the single phase and 67% for the MP-PIC calculations. In weak scaling, the energy consumption of both codes increases monotonically with an increasing cell count. The

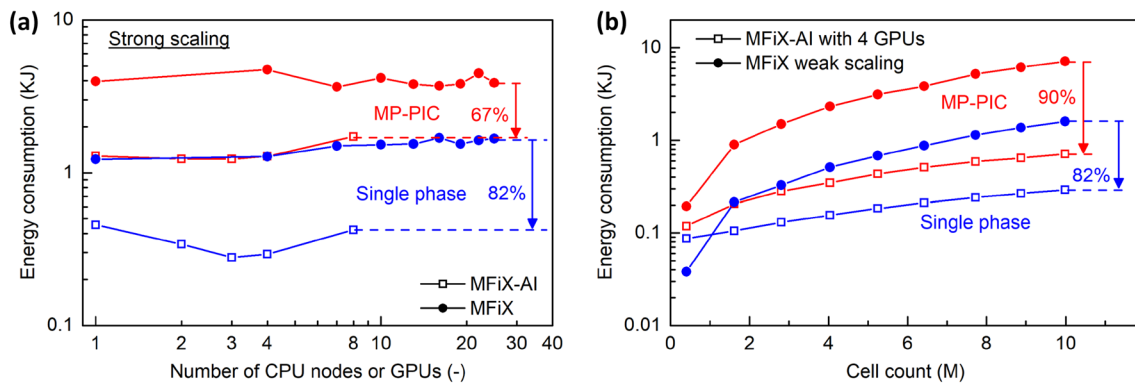


Fig. 11 **a** Estimated energy consumption for a single SIMPLE iteration in the strong scaling analysis for single phase with 10 M cells and MP-PIC with implicit coupling. **b** The energy consumption of weak scaling analysis for MFiX and MFiX-AI with four GPUs

energy savings for weak scaling analysis is shown to be up to 82% for single phase and 90% for MP-PIC calculations. This study indicates that there are distinct benefits of using multi-GPU computations on a single node relative to as many as 25 CPU nodes. Although it is not straightforward to convert a CPU-based CFD code to a new platform that enables effective GPU acceleration, the benefits in speed gain and energy consumption will be a strong motivating factor to adopt the new paradigm.

6 Conclusion and outlook

This study is aimed at developing an equation-based GPU parallelization strategy that enables acceleration of CFD simulations with a multiphase PIC model. For this purpose, MFiX, a well-established and thoroughly validated computer code for multiphase flow, and TF, a powerful open-source library for optimized matrix calculations on CPUs and GPUs, were coupled by replacing the core of a CFD solution procedure in MFiX by a TF-based computer code, MFiX-AI. MFiX-AI distributes the variables into each assigned device, which allows straightforward independent calculations for each transport equation, simultaneously. Thus, an intuitive equation-based parallelization was accomplished with no domain decomposition. The developed MFiX-AI code produces results that are not bit perfect, but match within many significant digits.

The equation-based parallelization strategy within MFiX-AI was found to work exceptionally well. Time parities for MFiX-AI were found to be above 15 CPU nodes in all cases and were most typically well above 25 nodes. For single-phase cases, MFiX-AI with just four GPUs with variable work per device was found to scale almost as well as MFiX in weak scaling at near optimal workload per core. MFiX-AI is able to scale out to many tens of millions of cells in single phase. At all tested conditions, MFiX-AI was faster than

MFiX by a significant margin. Extending the scaling curves shows that MFiX-AI is preferable to MFiX at scales all the way up to approximately 15 million cells. As hardware and code improvements are made, this balance will shift out significantly.

For most multiphase simulations, particle/parcel load imbalances greatly affect the time to solution for domain-decomposed Eulerian Lagrangian software architectures. When particles stack up or are naturally imbalanced, as with a fluid bed simulation, fixed domain decomposition methods struggle to scale well as some computing units get overloaded with more particle calculations than others. Load imbalance is even more severe in complex systems such as recirculating fluidized beds. MFiX-AI does not have this issue and shows promise to be an effective method for multiphase calculations in a very similar manner to MFiX. MFiX-AI is already a competitive software platform on GPUs for explicit coupling with solids.

Acknowledgements This work was supported by Science-based Artificial Intelligence (AI)/Machine Learning (ML) Institute—SAMI, and the project, CFD for Advanced Reactor Design—CARD. We appreciate NVIDIA providing access to NVIDIA Solutions Lab (NSL-B) for computations on DGX A100s. We also thank Dr. Madhava Syamlal for his inspiration and useful comments. Dr. Syamlal provided significant inspiration, motivation, and support for this work. We congratulate Dr. Syamlal on his retirement and wish him the very best on this new chapter of his life. We will never forget all that he has done for us in your long and prosperous career.

Funding This project was funded by the Department of Energy, National Energy Technology Laboratory, an agency of the United States Government, through an appointment administered by the Oak Ridge Institute for Science and Education. Neither the United States Government nor any agency thereof, nor any of its employees, nor the support contractor, nor any of their employees, makes any warranty, expressor implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark,

manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Code availability The MFIX-AI code is available as part of the MFIX suite of codes that are provided through NETL's Multiphase Flow Science (MFS) web portal (<https://mfix.netl.doe.gov>).

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

- Howard M, Fisher T, Hoemmen M, Dinzl D, Overfelt J, Bradley A, et al (2018) Employing multiple levels of parallelism for CFD at large scales on next generation high-performance computing platforms. In: Tenth International Conference on Computational Fluid Dynamics (ICCFD10); ICCFD10–079
- Appa J, Turner M, Ashton N (2021) Performance of CPU and GPU HPC Architectures for off-design aircraft simulations. AIAA Scitech 2021 Forum
- Castagna J, Guo X, Seaton M, O'Cais A (2020) Towards extreme scale dissipative particle dynamics simulations using multiple GPGPUs. *Comput Phys Commun* 251:107159
- Friedrichs MS, Eastman P, Vaidyanathan V, Houston M, Legrand S, Beberg AL et al (2009) Accelerating molecular dynamic simulation on graphics processing units. *J Comput Chem* 30(6):864–872
- Chen T, Ning Y, Amritkar A, Qin G (2018) Multi-GPU solution to the lattice Boltzmann method: an application in multiscale digital rock simulation for shale formation. *Concurr Comput Pract Experience* 30(19):e4530
- Nguyen G, Dlugolinsky S, Bobák M, Tran V, López García Á, Heredia I et al (2019) Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. *Artif Intell Rev* 52(1):77–124
- Awan AA, Manian KV, Chu C-H, Subramoni H, Panda DK (2019) Optimized large-message broadcast for deep learning workloads: MPI, MPI+NCCL, or NCCL2? *Parallel Comput* 85:141–152
- Mudigere D, Sridharan S, Deshpande A, Park J, Heinecke A, Smelyanskiy M et al (2015) Exploring shared-memory optimizations for an unstructured mesh CFD application on modern parallel systems. *IEEE International Parallel and Distributed Processing Symposium* 2015:723–732
- Selvam M, Hoffmann K (2015) MPI/Open-MP Hybridization of Higher Order WENO Scheme for the Incompressible Navier-Stokes Equations
- Crespo AC, Dominguez JM, Barreiro A, Gómez-Gesteira M, Rogers BD (2011) GPUs, a new tool of acceleration in CFD: efficiency and reliability on smoothed particle hydrodynamics methods. *PLoS ONE* 6(6):e20685
- Weller HG, Tabor G, Jasak H, Fureby C (1998) A tensorial approach to computational continuum mechanics using object-oriented techniques. *Comput Phys* 12(6):620–631
- Fernandes C, Faroughi SA, Ribeiro R, Isabel A, McKinley GH (2022) Finite volume simulations of particle-laden viscoelastic fluid flows: application to hydraulic fracture processes. *Eng Comput*. <https://doi.org/10.1007/s00366-022-01626-5>
- Afzal A, Ansari Z, Faizabadi AR, Ramis MK (2017) Parallelization strategies for computational fluid dynamics software: state of the art review. *Arch Comput Methods Eng* 24(2):337–363
- Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al (2015) TensorFlow: Large-scale machine learning on heterogeneous systems
- Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al (2016) TensorFlow: a system for large-scale machine learning. In: Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation. Savannah, GA, USA: USENIX Association, pp 265–283
- Zhao X-z, Xu T-y, Ye Z-t, Liu W-j (2020) A TensorFlow-based new high-performance computational framework for CFD. *J Hydrodyn* 32(4):735–746
- Clarke M, Musser J (2020) The MFIX Particle-in-Cell Method (MFIX-PIC) Theory Guide. NETL Technical Report Series, DOE/NETL-2020/2115. U.S. Department of Energy, National Energy Technology Laboratory: Morgantown, WV
- Li T, Dietiker J-F, Shahnam M (2012) MFIX simulation of NETL/PSRI challenge problem of circulating fluidized bed. *Chem Eng Sci* 84:746–760
- Herzog N, Schreiber M, Egbers C, Krautz HJ (2012) A comparative study of different CFD-codes for numerical simulation of gas–solid fluidized bed hydrodynamics. *Comput Chem Eng* 39:41–46
- Garg R, Galvin J, Li T, Pannala S (2010) Documentation of open-source MFIX–DEM software for gas-solids flows
- Li T, Garg R, Galvin J, Pannala S (2012) Open-source MFIX-DEM software for gas-solids flows: Part II — validation studies. *Powder Technol* 220:138–150
- Garg R, Galvin J, Li T, Pannala S (2012) Open-source MFIX-DEM software for gas–solids flows: Part I—verification studies. *Powder Technol* 220:122–137
- Li F, Song F, Benyahia S, Wang W, Li J (2012) MP-PIC simulation of CFB riser with EMMS-based drag model. *Chem Eng Sci* 82:104–113
- Kadyrov T, Li F, Wang W (2019) Impacts of solid stress model on MP-PIC simulation of a CFB riser with EMMS drag. *Powder Technol* 354:517–528
- Vaidheeswaran A, Musser J, Clarke MA (2020) Verification and Validation of MFIX-PIC. NETL Technical Report Series, NETL-TRS-2–2020. U.S. Department of Energy, National Energy Technology Laboratory: Morgantown, WV
- Xiong Q, Li B, Chen F, Ma J, Ge W, Li J (2010) Direct numerical simulation of sub-grid structures in gas–solid flow—GPU implementation of macro-scale pseudo-particle modeling. *Chem Eng Sci* 65(19):5356–5365
- Snider DM (2001) An incompressible three-dimensional multiphase particle-in-cell model for dense particle flows. *J Comput Phys* 170(2):523–549
- Syamlal M, O'Brien TJ (1989) Computer simulation of bubbles in a fluidized bed. *AIChE Symp Ser* 85:22–31
- Isakov M, Rosario Ed, Madireddy S, Balaprakash P, Carns P, Ross RB, et al (2020) HPC I/O throughput bottleneck analysis with explainable local models. SC20. In: International Conference for High Performance Computing, Networking, Storage and Analysis, pp 1–13
- Harrington P (2017) Diagnosing Parallel I/O Bottlenecks in HPC Applications. SC17, The International Conference for High Performance Computing, Networking, Storage and Analysis
- Shantharam M, Tatineni M, Choi D, Majumdar A (2018) Understanding I/O bottlenecks and tuning for high performance I/O on large HPC Systems: a case study. In: Proceedings of the Practice and Experience on Advanced Research Computing. Pittsburgh, PA, USA: Association for Computing Machinery; Article 54

32. Magni A, Dubach C, O'Boyle M (2014) Exploiting GPU hardware saturation for fast compiler optimization. In: Proceedings of Workshop on General Purpose Processing Using GPUs
33. NVIDIA DGX A100 THE UNIVERSAL SYSTEM FOR AI INFRASTRUCTURE; 2020. Available from: <https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/nvidia-dgx-a100-datasheet.pdf>
34. NVIDIA. Whitepaper: NVIDIA A100 Tensor Core GPU Architecture. 2020

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.