

## Computational Physics



# Chrono DEM-Engine: A Discrete Element Method dual-GPU simulator with customizable contact forces and element shape <sup>☆</sup>

Ruochun Zhang <sup>a</sup>, Bonaventura Tagliaferro <sup>b,c,d,\*</sup>, Colin Vanden Heuvel <sup>a</sup>, Shlok Sabarwal <sup>a</sup>,  
Luning Bakke <sup>a</sup>, Yulong Yue <sup>a</sup>, Xin Wei <sup>a</sup>, Radu Serban <sup>a</sup>, Dan Negruț <sup>a</sup>

<sup>a</sup> Department of Mechanical Engineering, University of Wisconsin–Madison, 1513 Engineering Dr, Madison, 53706, WI, USA

<sup>b</sup> Maritime Engineering Laboratory, Universitat Politècnica de Catalunya – BarcelonaTech, C. Jordi/Girona 1-3, Barcelona, 08034, Spain

<sup>c</sup> Department of Civil Engineering, University of Salerno, Via Giovanni Paolo II, Fisciano, 84084, Italy

<sup>d</sup> School of Natural and Built Environment, Queen's University Belfast, Northern Ireland, Belfast, BT9 5AG, UK

## ARTICLE INFO

## Keywords:

Discrete Element Method  
GPU computing  
Physics-based simulation  
Scientific package  
BSD3 open-source

## ABSTRACT

This paper introduces DEM-Engine, a new submodule of Project Chrono, that is designed to carry out Discrete Element Method (DEM) simulations. Based on spherical primitive shapes, DEM-Engine can simulate polydisperse granular materials and handle complex shapes generated as assemblies of primitives, referred to as *clumps*. DEM-Engine has a multi-tier parallelized structure that is optimized to operate simultaneously on two GPUs. The code uses custom-defined data types to reduce memory footprint and increase bandwidth. A novel “delayed contact detection” algorithm allows the decoupling of the contact detection and force computation, thus splitting the workload into two asynchronous GPU streams. DEM-Engine uses just-in-time compilation to support user-defined contact force models. This paper discusses its C++ and Python interfaces and presents a variety of numerical tests, in which impact forces, complex-shaped particle flows, and a custom force model are validated considering well-known benchmark cases. Additionally, the full potential of the simulator is demonstrated for the investigation of extraterrestrial rover mobility on granular terrain. The chosen case study demonstrates that large-scale co-simulations (comprising 11 million elements) spanning 15 seconds, in conjunction with an external multi-body dynamics system, can be efficiently executed within a day. Lastly, a performance test suggests that DEM-Engine displays linear scaling up to 150 million elements on two NVIDIA A100 GPUs.

## Contents

1.	Introduction . . . . .	2
2.	Implementation features . . . . .	3
2.1.	Multi-GPU solution and delayed active-contact set update . . . . .	3
2.1.1.	Comparison against domain decomposition . . . . .	4
2.2.	Just-in-time CUDA kernel compilation . . . . .	5
2.2.1.	Custom force model . . . . .	5
2.2.2.	Family tag . . . . .	5
2.3.	Custom and mixed data type . . . . .	5
2.4.	Geometry hierarchy and tracker . . . . .	6
2.5.	Contact detection algorithm . . . . .	6
2.6.	Python wrapper . . . . .	7
3.	Sample script . . . . .	7

<sup>☆</sup> The review of this paper was arranged by Prof. Andrew Hazel.

\* Corresponding author.

E-mail addresses: [ruochunz@gmail.com](mailto:ruochunz@gmail.com) (R. Zhang), [btagliaferro@unisa.it](mailto:btagliaferro@unisa.it) (B. Tagliaferro), [colin.vandenheuvel@wisc.edu](mailto:colin.vandenheuvel@wisc.edu) (C. Vanden Heuvel), [ssabarwal@wisc.edu](mailto:ssabarwal@wisc.edu) (S. Sabarwal), [lfang9@wisc.edu](mailto:lfang9@wisc.edu) (L. Bakke), [yyue32@wisc.edu](mailto:yyue32@wisc.edu) (Y. Yue), [xwei84@wisc.edu](mailto:xwei84@wisc.edu) (X. Wei), [serban@wisc.edu](mailto:serban@wisc.edu) (R. Serban), [negrut@wisc.edu](mailto:negrut@wisc.edu) (D. Negruț).

<https://doi.org/10.1016/j.cpc.2024.109196>

Received 31 October 2023; Received in revised form 1 March 2024; Accepted 3 April 2024

Available online 8 April 2024

0010-4655/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

3.1.	C++ version . . . . .	7
3.2.	Python version . . . . .	8
4.	DEM model . . . . .	9
4.1.	History-based Hertz–Mindlin model . . . . .	9
4.2.	Providing a custom contact force model . . . . .	10
4.2.1.	Default model implementation explained . . . . .	10
4.3.	Contact model validation . . . . .	12
4.3.1.	Sphere rolling on incline . . . . .	12
4.3.2.	Sphere stacking . . . . .	12
4.3.3.	Contact chain propagation . . . . .	12
5.	Simulator's performance . . . . .	14
6.	Numerical experiments . . . . .	15
6.1.	Ball impact test . . . . .	16
6.2.	Flow sensitivity test . . . . .	16
6.2.1.	Drum tests . . . . .	16
6.2.2.	Hopper tests . . . . .	17
6.3.	Contact modeling for particle breakage . . . . .	18
6.4.	Rover mobility co-simulation . . . . .	20
6.4.1.	Co-simulation . . . . .	21
6.4.2.	Active box scheme . . . . .	21
7.	Conclusions and future directions . . . . .	22
	Code availability . . . . .	23
	CRediT authorship contribution statement . . . . .	23
	Declaration of competing interest . . . . .	23
	Data availability . . . . .	23
	Acknowledgements . . . . .	23
	References . . . . .	23

## 1. Introduction

The Discrete Element Method (DEM) is a numerical technique for predicting the mechanical behavior of granular materials [1]. In DEM, the motion of each individual particle is monitored, and interactions between particles are modeled in a fully detailed manner. Over time, DEM has evolved and is now a popular method for examining the dynamics of extensive granular systems [2], ranging from mixing [3], particulate flows [4], geomechanics events [5–7], to astrophysical scenarios [8]. Applications of DEM include modeling soil dynamics [9], tire–soil interactions [10], and rover movement on extraterrestrial surfaces [11].

Two main challenges make DEM simulations computationally expensive. Firstly, the small and often stiff elements necessitate the time integrator to adopt very small time steps, e.g.,  $1 \times 10^{-6}$ – $1 \times 10^{-5}$  seconds, to ensure numerical stability. Secondly, the collision detection stage of the simulation is computationally demanding. To enhance computational speed, DEM has been accelerated using parallel computing with OpenMP [12] as seen in [13,14]; MPI standard [15] for distributed memory clusters [16]; and combined MPI–OpenMP parallelism [17–21]. The Graphics Processing Unit (GPU) offers another avenue for parallel computations and has been incorporated into DEM, as in [22–26]. Regardless of the computational platform, reported DEM studies typically involve between  $10^3$  and  $10^5$  elements [25,27–43], which is considerably smaller than real-world scenarios. For instance, a cubic meter of sand can contain around two billion particles [44].

LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator) [19] is a widely used open-source software package for molecular dynamics simulations and DEM simulations. LAMMPS is written in C++ and is designed to run efficiently on parallel computing architectures using both MPI and OpenMP, making it suitable for simulating large-scale systems. LAMMPS provides a variety of built-in potentials for modeling interatomic and intermolecular interactions, as well as the ability to define custom potentials. LAMMPS also supports a range of boundary conditions, including periodic, reflecting, and fixed boundaries. The default time stepper in LAMMPS is the Verlet algorithm, which is a symplectic second-order method. LAMMPS supports a range of contact models, including Hertz–Mindlin, linear-spring, cohesive and

inter-particle bond models. LIGGGHTS (LAMMPS Improved for General Granular and Granular Heat Transfer Simulation) is a DEM package that is based on the LAMMPS code. Like LAMMPS, LIGGGHTS is optimized for parallel computing and leverages combined MPI–OpenMP parallelism. While LAMMPS is more versatile, e.g., [45–47], LIGGGHTS focuses specifically on granular material simulations, offering features and capabilities tailored to that end, such as neighbor lists and domain decomposition. These added utilities come into play in granular flows, heat transfer in granular materials, and other DEM-specific concerns.

STAR-CCM+ [20] is a commercial Computational Fluid Dynamics (CFD) software package that includes a DEM solver for simulating the behavior of granular materials. The software also supports a range of contact models. One of the strengths of STAR-CCM+ is its ability to couple DEM simulations with fluid flow simulations, allowing for the simulation of complex multiphase flows. The coupling between the DEM and fluid flow simulations is typically achieved through a two-way coupling algorithm that exchanges information between the two simulations at each time step. The software also includes models for turbulence, heat transfer, and chemically reactive flows, and incorporates design exploration and optimization tools, allowing engineers to not just simulate a given design, but also explore a variety of design possibilities.

A DEM case study anchored by STAR-CCM+ is summarized in [48]. The study aimed to investigate the sand-retention mechanisms that occur at the opening of sand filters under various conditions, such as particle shape, size, and concentration. A coupled CFD–DEM model was used to predict the retention mechanisms under steady flow conditions of the well-bore, where CFD was used to model the fluid flow, and DEM was used to model the particle flow. The coarse grid unresolved and the smoothed unresolved (refined grid unresolved) coupling approaches implemented in STAR-CCM+ were used to transfer data between the fluid and solid phases and calculate the forces. Verification of the CFD–DEM model was then conducted by mesh sensitivity analysis. The growing trend in CFD–DEM coupling research underscores the community's heightened interest in integrating multi-physics into DEM simulations, likely driven by the rapid advancements in computational power.

Compared to the LAMMPS and STAR-CCM+, Chrono::GPU [49], an open-source DEM simulator developed originally as the granular dynamics support for Chrono [50], takes a different path in that it emphasizes efficiency. To maximize performance, Chrono::GPU operates on GPUs and exclusively supports monodisperse spherical DEM elements. Additionally, a custom data type scheme is used to reduce its memory footprint. A recent independent study [51] reveals that Chrono::GPU, while running on an RTX 2060 Mobile NVIDIA GPU card of a laptop, delivers performance that is two orders of magnitude faster than other well-regarded DEM packages operating on clusters with hundreds of CPU cores.

YADE (Yet Another Dynamic Engine) is an open-source DEM simulator for granular materials, powders, and other particulate systems. Written in C++ and Python, it is known for its scripting-imparted extensibility. One notable research study that used YADE is [52], in which the authors were interested in the deformation of the particles under stress. Therein, the particles are modeled as a collection of smaller particles connected by springs. The authors made additional developments to the DEM model, so the volume of the element overlapping area is uniformly redistributed over the particle, the radius of each contact partner is increased, and in the end, the volume and mass are kept constant. Large deformations and complex element geometries are used in this study. Another recent study that used YADE for DEM simulations is reported in [53]. Therein, the authors simulated the process of icing using an Euler–Lagrangian approach. YADE was used to calculate the motion of snow crystals, while the open-source CFD package OpenFOAM was used in conjunction with YADE to simulate flow hydrodynamics.

To circumvent extensive computation times, DEM packages often resort to simplistic element geometries to simplify collision detection. Predominantly, spheres of uniform size are chosen, significantly streamlining collision detection [54]. Yet, certain applications require more intricate geometries, necessitating the usage of nonspherical elements to ensure accurate system dynamics [55–61]. From the aforementioned packages, YADE, along with its extension packages such as SudoDEM [62], enables supports for non-spherical shapes with varying degrees of roundness or sharpness, such as superellipsoid, poly-superellipsoid, and polyhedron. Another approach YADE employs is the use of the “multi-sphere method” [59], meaning grouping simpler particles (like spheres) together to form more complex shapes. Likewise, LAMMPS supports this multi-sphere method, too. LAMMPS also supports ellipsoidal and spherical particles. STAR-CCM+, being an established commercial DEM solution, offers a library of predefined shapes (spheres, cylinders, tetrahedra, etc.), while retaining a general-purpose custom shape support using triangulated surfaces. These custom shapes are treated as rigid bodies within the DEM framework. When these methods to address nonspherical elements are employed, the number of elements in simulations tends to reduce significantly in order to manage the amount of time required to complete a simulation.

Recognizing the characteristics, strengths, and limitations of the existing DEM solvers, the solution presented here, Chrono DEM-Engine [63], aims to strike a balance: (i) it accommodates a large number of discrete elements (into tens of millions); (ii) it employs a composition of multiple spheres to represent nontrivial geometries; (iii) it integrates a rapid collision detection method as per [64] and a novel asynchronous threads management algorithm to ensure a numerical performance ahead of state of the art; (iv) its API design leaves enough room and flexibility for easy integration in co-simulations (explained in Sec. 3 and 6.4), and gives users the freedom to define explicitly the physics they wish to simulate using a custom force model script (explained in Sec. 2.2.1). In this contribution, our emphasis is to thoroughly document the numerical features of this package, and provide guidelines for the user to easily pick up this package and then fully take advantage of its potential.

The structure of this paper is laid out as follows. The literature survey, presented in this section, identifies a prevailing need within the DEM community for an adaptable, efficient solver capable of manag-

ing large-scale simulations. Section 2 explores the distinct numerical capabilities of Chrono DEM-Engine and illustrates how it addresses this identified need. Section 3 offers a breakdown of a sample simulation script, equipping the user with a foundational understanding of the package’s operation. Section 4 unravels the implementation of the default Hertz–Mindlin model and provides guidance on incorporating custom models. Section 5 demonstrates the solver’s efficiency, spotlighting a large-scale simulation involving up to 150 million sphere primitives. Section 6 underscores the validation endeavors, presenting a suite of DEM simulations that emphasize the impact and capabilities of varying element shape representations and force models. Section 7 reiterates the essence of the paper, accentuating the proposed future developments with language models.

## 2. Implementation features

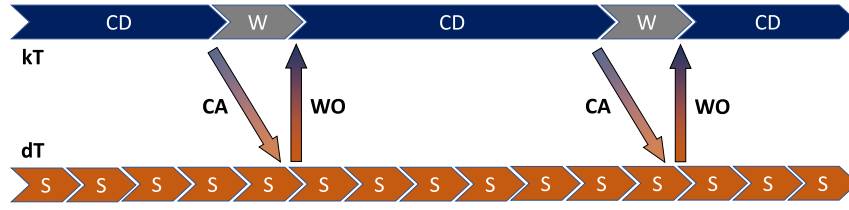
Chrono DEM-Engine is open-source, can run on commodity hardware and it does so fast and at scale. It allows large-scale DEM simulations to be efficiently executed on desktops equipped with one or two graphic cards. Its open-source nature and ability to embed user-defined contact models meet requirements often found in exploratory projects. This section introduces the simulator’s key features.

### 2.1. Multi-GPU solution and delayed active-contact set update

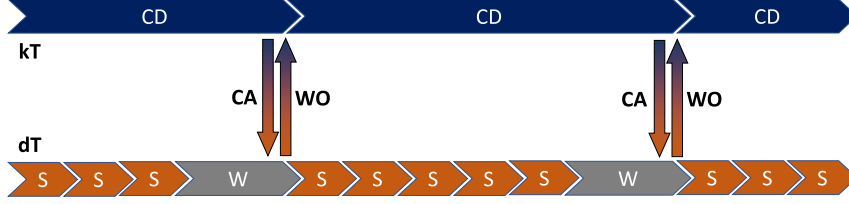
In DEM, the contact detection process is needed to identify the contact pairs in the simulation system before the force calculation step takes place. The contact detection and force calculation are typically done consecutively in each time step. DEM-Engine embraces a different strategy, which uses two distinct and parallel computational threads to update the active contacts set (done by the “kinematics thread”), and the integration of the equations of motion (done by the “dynamics thread”), respectively. The dynamics thread processes each contact in the Active-Contact Set (ACS) *at each time step* to reassess the contact penetration  $\delta_n$  and the ancillary information. The dynamics thread receives an ACS update when the kinematics thread finishes producing it, or if so desired, it can wait for the ACS update when the dynamics thread advances the system state too far ahead of the time stamp of the last ACS update from the kinematics thread. Through this collaboration pattern, the two threads work concurrently and the cost of contact detection is nearly “hidden in the shadow” of computation done by the dynamics thread, which continuously advances the state of the system. To avoid missing mutual contacts that might crop up between the moments the ACS is updated, we artificially enlarge all contact geometries in the DEM system using an approach that is reminiscent of the Verlet table algorithm [65]. This extra margin allows to preemptively detect potential contact pairs that might emerge in the near future. Note that this is done only to include additional potential contacts in the ACS, and does not affect the shape of the elements participating in the simulation.

By adding this artificial margin to all contact geometries, the kinematics thread can report false positives within the provided list of contacts, i.e., a contact between two elements might be in the ACS, yet the two elements are not in contact. This fact will be identified by the dynamics thread when carrying out the force calculation. The thickness of this added margin is determined by the simulation entities’ velocity (which is bounded and known by the solver), the time step size, which is typically small, and  $n_{\max}$ , the maximum number of time steps the dynamics thread is allowed to advance without receiving an ACS update from the kinematics thread. It usually assumes values of the order of tens of microns for millimeter-sized granular material. This is small compared to typical DEM element sizes. Overall, the overhead caused by the false-positive contacts does not offset the benefit of deferring the ACS update.

The synchronization pattern between the kinematics and dynamics threads is illustrated in Fig. 1. There, “S” represents a time step that the dynamics thread executes, where the contact forces are calculated



**Fig. 1.** Ideal collaboration pattern, where the dynamics thread advances the physics continuously while the kinematics thread occasionally waits for updated state information to commence an ACS update.



**Fig. 2.** Non-optimal collaboration pattern, where the dynamics thread waits for the kinematics thread occasionally to generate the ACS. DEM-Engine will automatically avoid this scenario.

(see Sec. 4.1), and the system state is advanced in time. A contact detection step that the kinematics thread executes is marked with “CD”. Periodically, the kinematics thread finishes a contact detection step and sends the signal to the dynamics thread, allowing the dynamics thread to receive the contact array, “CA”, from the kinematics thread. Then the dynamics thread will send a work order “WO” with the current simulation system state, for the kinematics thread to pick up and continue the next contact detection step. Before the next “CA” update is received, the dynamics thread will use this “CA” to execute the time steps.

Because the dynamics thread only receives updates from and sends work orders to the kinematics thread after a time step is finished, the kinematics thread could stay idle between ACS update jobs. This is marked with “W” in Fig. 1. Having the kinematics thread wait occasionally is considered an ideal collaboration pattern since in this case, the dynamics thread runs continuously, therefore the system marches in time uninterruptedly. A less-than-ideal collaboration pattern is illustrated in Fig. 2. There, the dynamics thread occasionally waits for updates from the kinematics thread, reducing the overall efficiency of the solver. This happens when the dynamics thread advances the simulation beyond  $n_{\max}$  time steps without receiving an update from the kinematics thread, and is therefore forced to idle. One could avoid this scenario by increasing  $n_{\max}$ . However, as discussed before this would consequently increase the thickness of the artificial margin added to contact geometries, leading to more undesirable false-positive contacts. Hence,  $n_{\max}$  should be kept at the smallest value that does not cause the dynamics thread to wait. DEM-Engine will automatically use this principle and the execution timing history to adapt  $n_{\max}$  to an appropriate value, and moderate itself so that the collaboration pattern stays as depicted in Fig. 1.

At the implementation level, DEM-Engine is currently optimized for using two GPUs, as each of the two host CPU threads is bound with a CUDA stream, and then each CUDA stream is mapped to a GPU device. If only one GPU device is available, DEM-Engine will still run since both streams are mapped on to that GPU, instead. The kinematics and dynamics thread collaboration pattern is summarized in Fig. 3. After being produced by the kinematics thread, the contact information is transferred to a buffer memory. Then the dynamics thread will be notified and copy the contact information to its working memory. The dynamics thread carries out a similar routine when updating the kinematics thread with new element positions. Neither of them directly modifies the working memory of the other to avoid race conditions. Although logically there are two buffer memory pools and each thread owns one, physically, they are both allocated on the GPU mapped to the

dynamics thread. This allows the dynamics thread to spend minimum time copying from its buffer, speeding up the computation.

#### 2.1.1. Comparison against domain decomposition

This section discusses the advantages and limitations of our proposed multi-GPU solution in comparison to the conventional approaches, such as the domain-decomposition approach. Domain-decomposition methods enhance computational efficiency by partitioning the simulation domain into multiple segments, each processed by a separate GPU. This strategy is evidenced in recent studies [66] and commercial tools like Rocky DEM [67]. Key to this method is its scalable memory usage; as the simulation data is distributed across GPUs, and only the interface regions require inter-GPU communication; each additional GPU effectively enlarges the overall memory capacity. This expansion allows for handling larger-scale problems. Moreover, adding GPUs can potentially reduce computational time by dividing the workload. However, the efficiency gains are often offset by the overheads of array sorting, entity and contact list reorganization, and data transfer at each simulation step [68]. Additionally, the total runtime is heavily dependent on thread workload balance, which requires specific algorithms to correctly predetermine the domain decomposition strategy [68]. The inter-GPU or inter-node data transfer’s latency and bandwidth constraints make time savings through domain decomposition challenging and not always achievable.

In contrast, the asynchronous kinematics–dynamics pattern in Chrono DEM-Engine offers tangible time savings. As discussed, the cost of contact detection is nearly “hidden in the shadow” with this approach. In the proposed implementation, a kinematic step (CD in Fig. 1) is typically 20 times more time-consuming than a dynamic step (S in Fig. 1), translating to over 90% time savings compared to conventional serial contact detection and physics advancement methods. It is important to note that the CD step in DEM-Engine is costly due to its support for unlimited and unbounded element size discrepancy (detailed in Sec. 2.5). Nevertheless, under certain convenient hypotheses, the cost of CD can be reduced. For instance, Chrono::GPU, which exclusively uses monodisperse spherical elements, has a CD step cost comparable to that of a S step [50]. Even with such simplification, a 50% time saving is achievable if the contact detection cost is hidden by the asynchronous pattern. Thus, we anticipate a time saving between 50% and 90% with the kinematics–dynamics asynchronization introduced by DEM-Engine. However, this approach is less memory-efficient due to the buffer arrays, effectively duplicating some simulation data, thereby limiting the scale of the simulation for a given memory capacity.



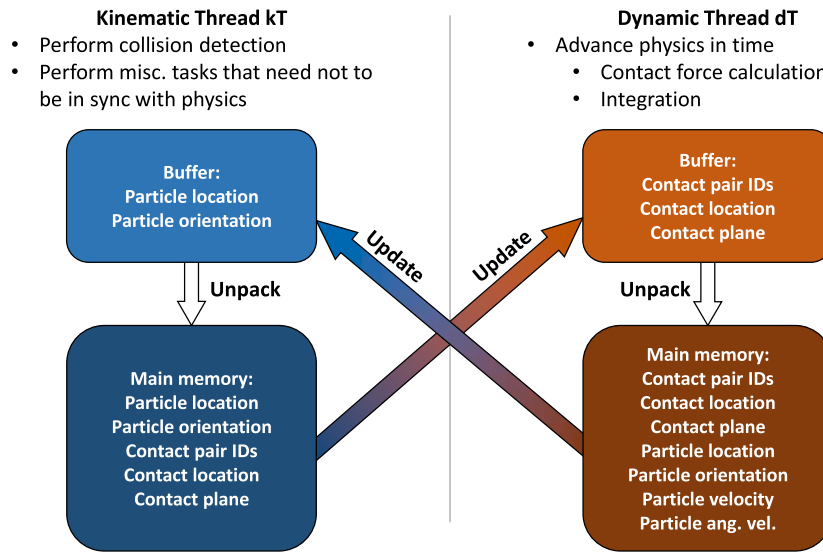


Fig. 3. The collaboration pattern of the kinematic and dynamics thread. They can each run on a dedicated GPU.

In summary, when utilizing two GPU devices, the kinematics–dynamics asynchronous pattern in Chrono DEM-Engine shows a more pronounced time-saving advantage over the domain-decomposition approach. This advantage holds true within the constraints of the available device memory, given the lesser memory efficiency of our approach. Notably, the kinematics–dynamics split in Chrono DEM-Engine does not preclude future integration with domain decomposition, allowing for both **CD** and **S** steps to be executed on multiple GPUs. This potential integration suggests that the kinematics–dynamics split is not an alternative to, but a valuable addition to the DEM algorithm toolkit, complementing the established domain-decomposition approach.

## 2.2. Just-in-time CUDA kernel compilation

The CUDA kernels in DEM-Engine are compiled when the simulation starts being executed, rather than being statically compiled. This is done by leveraging the CUDA runtime compilation tool Jitify [69]. Several benefits come with this software design choice.

With Jitify, the solver can detect the capabilities of the GPU on which it is running and generate code specifically tailored for that device. For instance, if a program is designed to be used across a variety of architectures, just-in-time compilation ensures the utilization of the optimal instruction set for each device, ensuring the generated CUDA code is optimized for an end user’s specific hardware and requirements. At the same time, since the compilation occurs at runtime, the code is not bound to a specific version of the CUDA toolkit. This characteristic can make applications more resilient against changes or updates in the CUDA environment.

It should be mentioned that just-in-time compilation introduces an overhead. The first time a kernel is run, there is a delay due to its compilation. However, assuming the DEM simulations with DEM-Engine are generally large and invoke a time span of typically hours, this cost is negligible.

### 2.2.1. Custom force model

Since Jitify allows for dynamic code generation, we use it for implementing custom force models. The intricate and evolving nature of DEM simulations often requires a higher degree of adaptability to cater to the multifaceted modeling needs of its users, namely the expanding list of approaches in contact and cohesion force modeling [70,71]. Rather than constraining the user to a predefined set of force models, DEM-Engine allows, if so desired, for the force models to be supplied via a user-supplied C++ script, greatly increasing the solver’s applicability. A walk-through of a model implementation can be found in Sec. 4.2.

### 2.2.2. Family tag

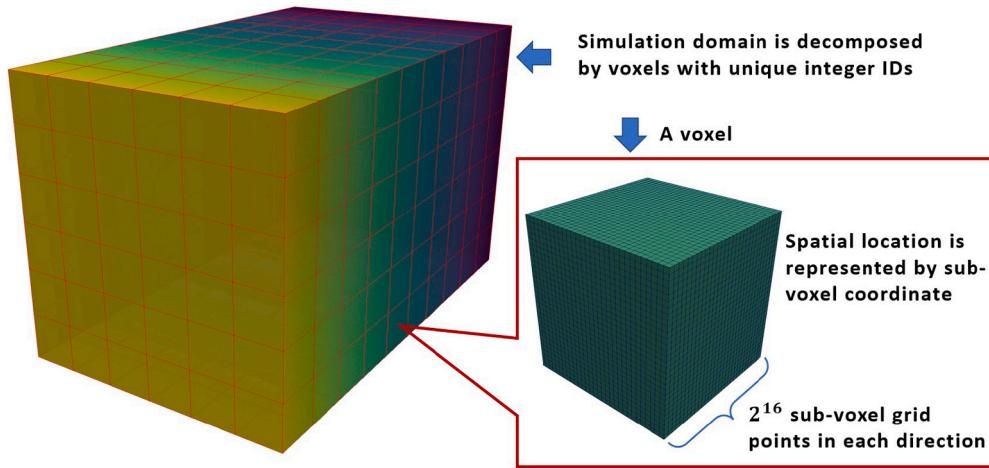
Jitify also allows for a low-cost implementation of prescribed motion. This is done through the family tag utility. Every simulation entity can be assigned an integer family tag between 0 and 255 (this is implemented through a `uint8_t`; though rarely needed, it can be changed to a different data type such as `uint16_t` to expand the range), then the solver can be notified to apply prescribed motions to this family tag. This prescription information is just-in-time compiled as a part of the integration CUDA kernel, thus no branching overhead is introduced. The sample script in Sec. 3 showcases this functionality with the usage of the `SetFamilyPrescribedAngVel` method. On the other hand, if the use case calls for more fine-grain motion control, such as when the velocity of a simulation object is determined by some external process, then the “motion injection” approach detailed in Sec. 2.4 should be used.

As a side note, the family tags can also be used to mask contacts. The user is allowed to specify whether the solver should detect and resolve contacts between simulation entities in certain families. This is a utility used throughout the demos provided along with this package at [63].

## 2.3. Custom and mixed data type

In high-performance computing, memory footprint and bandwidth play a crucial role in determining a code’s performance. As the complexities of the simulations grow, it becomes evident that relying solely on standard data types—such as `double`—might inadvertently lead to sub-optimal memory usage and consequently, potential performance bottlenecks. For instance, the penetration depth of a DEM body in penalty-based DEM is of the order of  $1 \times 10^{-9}$  to  $1 \times 10^{-5}$  m. Why would one use a budget of 64 bits, which is provisioned for the `double` type to capture an extremely broad range of numbers, to represent a very narrow range of the positive real axis that hosts an element’s microscale penetration? This would be a waste of bits, which leads to less accuracy and/or lower bandwidth. Given the hierarchical memory architecture in CUDA, from global to shared memory, the significance of ensuring that the memory bandwidth is utilized effectively and that latency is minimized becomes even more critical.

To this end, DEM-Engine introduces the utilization of custom and mixed data types. Unlike stock data types that come with a predefined bit budget, e.g., 64 bits for `double`, custom data types offer finer control over memory use. For instance, the spatial coordinate in DEM-Engine is represented using integers rather than floating-point numbers. The entire simulation domain is decomposed into cubes with a known edge length, which is adapted based on the domain size. Each of these



**Fig. 4.** The domain decomposition that leads to a compressed spatial coordinate data type. The domain is decomposed into voxels with uint64\_t indices, then each voxel is split into  $2^{16} \times 2^{16} \times 2^{16}$  sub-voxels. The typical precision is estimated to be around  $10^{-11}$  m.

**Table 1**

Various data types in DEM-Engine and their memory location.

Data Type	Variable	Memory Type
uint64_t	Voxel index	Global
uint16_t	Sub-voxel index	Global
int32_t or float	Kinematics quantities, friction history etc.	Global
double	Penetration	Register
float	Contact force calculation	Register
float	Clump types information	Shared Memory

cubes is termed a “voxel” and is assigned an index represented by a uint64\_t data type. Additionally, to specify the location of a body within a voxel, three uint16\_ts are employed, each dividing the voxel uniformly into  $2^{16}$  parts in its respective direction.

For a cubic simulation domain with an edge length of 1 m, the precision (i.e., the smallest discernible length unit within a voxel) is approximately  $10^{-11}$  m. This precision is adequate for capturing micro-deformations. Moreover, this compressed data type requires only 112 bits to represent a spatial location, which is more memory-efficient than using three doubles that would require 192 bits in total. The voxel-based spatial coordinate data type is illustrated in Fig. 4.

The general rule used for the selection of mixed data types is that the data residing in the global memory take a 4-byte or compressed data type. The examples are the state variables such as the quaternions of the elements. The temporary variables used in kernel functions that are essential in governing physics, on the other hand, use 8-byte data types, namely double. An example is the penetration depth between geometries in the Hertzian contact force calculation. The data type usage in DEM-Engine is summarized in Table 1. Since data type conversion is essentially a free operation and the main bottleneck in GPU-based physics simulations is the memory bandwidth limit, the design choice enhances the performance without compromising the physics.

Furthermore, DEM-Engine has a level of encapsulation of the data types in use. Most data types are specified in a file named VariableTypes.h using typedef, including the ones introduced in this section. If the user needs a different selection of data types, such as increasing the size of family tags from 1 byte to 2 bytes to allow for more varieties (see Sec. 2.2.2 for context), or reducing the size of spatial coordinates to allow for faster computation in a low-accuracy setting, they can modify the data types in VariableTypes.h then recompile to conveniently get the updated executable.

#### 2.4. Geometry hierarchy and tracker

DEM-Engine facilitates complex element geometries through a composition of multiple spheres, termed a “clump”. This approach draws inspiration from [72]. A clump denotes a collection of potentially overlapping spheres that together depict a specific element shape. Some examples of these clumps are visually presented in Fig. 14 in Sec. 5. Throughout this paper, the terms “element” and “clump” are used interchangeably to discuss DEM elements with complex shapes. Beyond clumps, DEM-Engine supports integrating triangular meshes and analytical objects (such as rigid objects constructed from analytical planes or cylindrical surfaces) into the simulation framework. However, as a dedicated and performance-centric DEM package, DEM-Engine exclusively handles contacts between clumps and meshes, as well as between clumps and analytical geometries. Should there be a requirement for contacts between meshes or between analytical geometries, users can achieve this through co-simulation, as exemplified in Sec. 6.4.

An important aspect of DEM-Engine’s utilization is understanding its geometry hierarchy, delineating the roles of the “owner” versus the “geometry”. An owner constitutes a simulation entity endowed with mass properties, hence governed by physics. In DEM-Engine’s current implementation, an owner can manifest as a clump, a mesh, or an analytical entity. Conversely, the term geometry is associated with the constituent parts of an owner. A geometric entity can be a sphere (within a clump), a triangular facet (within a mesh), or an analytical component (like a plane in a multi-component analytical object). Each geometric entity carries associated material attributes, granting users flexibility in designing discrete element systems with simulation entities that have spatially varying material properties.

Further, DEM-Engine provides users the control over diverse simulation entities via “tracker” objects. Users can associate trackers with any owner, facilitating real-time status inquiries such as position and velocity or enforcing state modifications, from setting coordinates to applying external loads. Beyond basic operations, trackers offer advanced features: identifying clumps in contact with a tracked owner or, when monitoring a mesh, controlling its deformation. A practical demonstration of tracker usage is encapsulated in Sec. 3.

#### 2.5. Contact detection algorithm

The kinematics thread of the Chrono DEM-Engine (CD in Fig. 1) implements the contact detection algorithm outlined in [64]. This algorithm is optimized for execution on GPU and accommodates simulation entities of vastly differing sizes. We provide a concise overview of its implementation in Chrono DEM-Engine, with more comprehensive details available in [64].

The contact detection process in each step involves a series of tasks, executed sequentially:

1. Initially, all simulation “geometries” (sphere components, triangular facets) are evaluated for potential contact using “bins”. These bins are formed by uniformly segmenting the simulation domain into axis-aligned cubic grids. If a geometry intersects with a bin, this bin–geometry pair is recorded for subsequent processing. It is important to note that due to the allowance for variable sizes of geometries, the maximum number of bins intersecting with a geometry cannot be predetermined. This necessitates two sequential CUDA kernel executions: one to determine the count of intersecting bins per geometry for memory allocation, and another to store the bin–geometry pairs.
2. Next, we sort the bin–geometry pairs based on the bin IDs. This step groups together geometries located within the same bin, effectively clustering entities that are adjacent within the simulation.
3. The final step involves checking geometries within the same bin for potential contacts. This is accomplished through launching CUDA blocks each processing bin to compile the **CA** array (recall Fig. 1), which lists all potential contact pairs. Duplicates are ruled out by discarding all contacts with the exact contact points not inside the current bin. Similar to the previous step, two CUDA kernel calls are required: the first to ascertain the number of potential contacts per bin for array allocation, and the second to populate the **CA** array with these contacts. Given  $n_g$  geometries in a bin,  $(n_g - 1)n_g/2$  checks are necessary to identify all potential contacts. This imposes a limit on  $n_g$ , influencing the bin size. Chrono DEM-Engine dynamically adjusts bin sizes based on execution history for optimal performance.

As described, both the intermediate bin–geometry pair array and the final **CA** array are created using dynamically allocated memory. This approach facilitates handling geometries with unlimited size discrepancies. For an illustrative example showcasing elements of significantly varying sizes, refer to Sec. 6.4.

## 2.6. Python wrapper

DEM-Engine has a Python wrapper, facilitated by the Pybind library. This allows users, irrespective of their CUDA expertise, to tap into DEM-Engine’s features, all within Python’s accessible library ecosystem and widely adopted science tools such as numpy and scikit-learn. The package has been made available on the Python Package Index (PyPI) and can be installed using the familiar pip command. Simply executing pip install DEME ensures that the computational capabilities and functionalities of the package become available within the Python environment, reducing the complexities often associated with software installations in high-performance computing scenarios. An example script is given in Sec. 3.2, where it is compared against its C++ counterpart.

## 3. Sample script

This section discusses a script responsible for the mixer timing analysis discussed in Sec. 5. The focus is placed here on the code implementation. A visual representation of the simulation workflow is provided in Fig. 5. Examples are provided in both C++ and Python. The scripts corresponding to all simulations addressed in this paper can be located in the DEM-Engine’s demo directory [63].

### 3.1. C++ version

The user should first create the DEMSolver object. While the solver comes with default meta-parameters, users have the flexibility to modify them, e.g., verbosity, output detail, and output format.

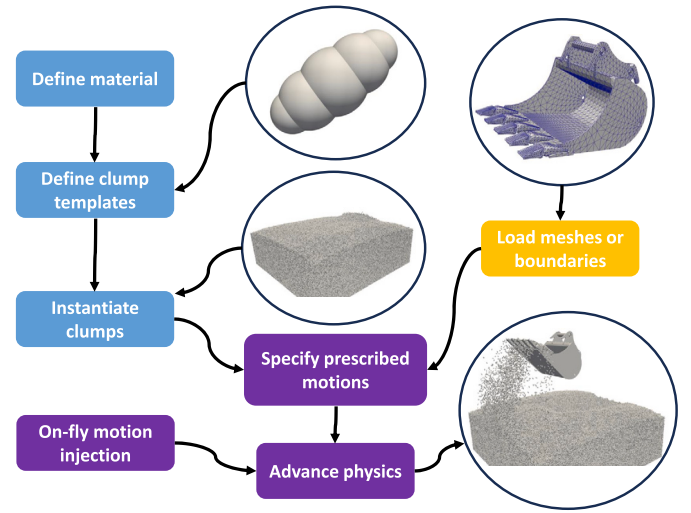


Fig. 5. Typical workflow of running a DEM-Engine simulation.

```

DEMSolver DEMSim;
DEMSim.SetVerbosity("INFO");
DEMSim.SetOutputFormat("CSV");
DEMSim.SetOutputContent("ABSV");
DEMSim.SetMeshOutputFormat("VTK");

```

The following code snippet defines the material types for the mesh geometry and DEM elements. DEM-Engine will return a handle so this material can be used to define clump templates. If a material property, such as the frictional coefficient  $\mu$ , is defined between two materials, the method `SetMaterialPropertyPair` can be used to specify it.

```

auto mat_type_mixer = DEMSim.LoadMaterial({{"E", 1e8}, {"nu", 0.3}, {"CoR", 0.6}, {"mu", 0.5}, {"Crr", 0.0}});
auto mat_type_granular = DEMSim.LoadMaterial({{"E", 1e8}, {"nu", 0.3}, {"CoR", 0.6}, {"mu", 0.2}, {"Crr", 0.0}});
DEMSim.SetMaterialPropertyPair("mu", mat_type_mixer, mat_type_granular, 0.5);

```

The following snippet defines the analytical boundaries of the simulation domain.

```

const double world_size = 1;
DEMSim.InstructBoxDomainDimension(world_size, world_size, world_size);
DEMSim.InstructBoxDomainBoundingBC("all", mat_type_granular);
auto walls = DEMSim.AddExternalObject();
walls->AddCylinder(make_float3(0, make_float3(0, 0, 1), world_size / 2., mat_type_mixer, 0);

```

The following snippet shows the mixer mesh being loaded into the simulation. The stock mixer mesh is then scaled to fit the size of the simulation domain. The mixer is assigned the family code 10, which is subsequently used to prescribe a constant angular velocity  $\pi$  rad/s to the mixer. A “tracker” object is created for the mixer so that we can extract information in real time for this simulation entity, or apply fine-grain motion control, while the simulation is running. In this example, we use it to set the initial location of the mixer to obtain the torque exerted by the DEM elements.

```

const float chamber_height = world_size / 3.;
auto mixer = DEMSim.AddWavefrontMeshObject((GET_DATA_PATH() / "mesh/internal_mixer.obj").string(), mat_type_mixer);
mixer->Scale(make_float3(world_size / 2, world_size / 2, chamber_height));
mixer->SetFamily(10);
DEMSim.SetFamilyPrescribedAngVel(10, "0", "0", "3.14159");
auto mixer_tracker = DEMSim.Track(mixer);

```

The next snippet creates a clump template. It contains mass, shape, and material information. There are stock clump shapes that the user can directly use to reproduce the examples we provide. The user can also easily scale or otherwise modify the template before using it to instantiate more DEM elements.

```
float granular_rad = 0.005;
float mass = 2.6e3 * 5.5886717;
float3 MOI = make_float3(2.928, 2.6029, 3.9908) * 2.6e3;
std::shared_ptr<DEMClumpTemplate> template_granular = DEMSim.
    LoadClumpType(mass, MOI, GetDEMDataFile("clumps/3_clump.csv"),
        mat_type_granular);
template_granular->Scale(granular_rad);
```

When instantiating the DEM elements, the user has the option to leverage the sampler objects that come with the solver, as shown in the following snippet. A sampling region appropriate with respect to the simulation domain is defined, then the hexagonal close-packing sampler is used to create initial elements. These elements are duplicates of the clump template that has just been created.

```
const float fill_height = chamber_height;
const float chamber_bottom = -world_size / 2.;
const float fill_bottom = chamber_bottom + chamber_height;
HCPSampler sampler(3.f * granular_rad);
float3 fill_center = make_float3(0, 0, fill_bottom + fill_height / 2);
const float fill_radius = world_size / 2. - 2. * granular_rad;
auto input_xyz = sampler.SampleCylinderZ(fill_center, fill_radius,
    fill_height / 2);
DEMSim.AddClumps(template_granular, input_xyz);
```

An initialization call is needed to instruct the solver to set up data structures on the GPUs. Before that, several simulation specs should be inputted, e.g., the time step size and metrics that the solver should watch in identifying a diverged simulation, as shown in the following snippet.

```
float step_size = 5e-6;
DEMSim.SetInitTimeStep(step_size);
DEMSim.SetGravitationalAcceleration(make_float3(0, 0, -9.81));
DEMSim.SetErrorOutVelocity(20.);
DEMSim.SetForceCalcThreadsPerBlock(512);
DEMSim.Initialize();
```

Finally, the following code snippet shows the main simulation loop. The output directory is created, the simulation time length is indicated, and the mixer is translated to the correct initial position, before the main loop starts to iteratively make DoDynamics calls, advancing the simulation each time by a frame. The benefit of this design is that the user enjoys free interfacing with the simulation data while it is running. For example, the script writes the simulation status to a file, inspects the torque that the mixer is experiencing, and outputs the execution stats from the kinematics and dynamics threads at the frequency of 20 times per simulation second. Another opportunity this design brings is the ease of co-simulation. A related example is in Sec. 6.4.

```
std::filesystem::path out_dir = current_path();
out_dir += "/DemoOutput_Mixer";
create_directory(out_dir);

float sim_end = 10.0;
unsigned int fps = 20;
float frame_time = 1.0 / fps;
unsigned int curframe = 0;

mixer_tracker->SetPos(make_float3(0, 0, chamber_bottom +
    chamber_height / 2.0));
for (float t = 0; t < sim_end; t += frame_time) {
    std::cout << "Frame:" << curframe << std::endl;
    char filename[200], meshfilename[200];
    sprintf(filename, "%s/DEMDemo_output_%04d.csv", out_dir.c_str(),
        curframe);
    sprintf(meshfilename, "%s/DEMDemo_mesh_%04d.vtk", out_dir.c_str(),
        curframe);
    DEMSim.WriteSphereFile(std::string(filename));
    DEMSim.WriteMeshFile(std::string(meshfilename));

    float3 mixer_moi = mixer_tracker->MOI();
    float3 mixer_acc = mixer_tracker->ContactAngAccLocal();
    float3 mixer_torque = mixer_acc * mixer_moi;
    std::cout << "Contact_torque_on_the_mixer_is" <<
        mixer_torque.x << ", " << mixer_torque.y << ", " <<
        mixer_torque.z << std::endl;

    DEMSim.DoDynamics(frame_time);
    DEMSim.ShowThreadCollaborationStats();
}
```

### 3.2. Python version

A Python version of the same mixer simulation is given in this section. It follows the same workflow as the C++ version, including the material definition, template creation, clump instantiation, mesh loading and motion control, initialization, and a main simulation loop. The names of the methods are not changed in the Python version, and certain data structures are simply converted to their Python counterparts, streamlining the learning experience of the users switching between these programming languages. For example, the C++ version uses an `unordered_map` to define the properties of a material, while the Python version uses a dictionary object; the C++ version takes a `float3` at some places to specify a coordinate, while the Python version uses a list or a numpy array of three floats.

```
import DEME
import numpy as np
import os
import time

if __name__ == "__main__":
    out_dir = "DemoOutput_Mixer/"
    out_dir = os.path.join(os.getcwd(), out_dir)
    os.makedirs(out_dir, exist_ok=True)

    DEMSim = DEME.DEMSolver()
    DEMSim.SetVerbosity("STEP_METRIC")
    DEMSim.SetOutputFormat("CSV")
    DEMSim.SetOutputContent(["ABSV", "XYZ"])
    DEMSim.SetMeshOutputFormat("VTK")

    # E, nu, CoR, mu, Crr... Material properties
    mat_type_mixer = DEMSim.LoadMaterial(
        {"E": 1e8, "nu": 0.3, "CoR": 0.6, "mu": 0.5, "Crr": 0.0})
    mat_type_granular = DEMSim.LoadMaterial(
        {"E": 1e8, "nu": 0.3, "CoR": 0.8, "mu": 0.2, "Crr": 0.0})
    DEMSim.SetMaterialPropertyPair(
        "CoR", mat_type_mixer, mat_type_granular, 0.5)

    # Now define simulation world size and add the analytical boundary
    step_size = 5e-6
    world_size = 1
    chamber_height = world_size / 3.
    fill_height = chamber_height
    chamber_bottom = -world_size / 2.
    fill_bottom = chamber_bottom + chamber_height
    DEMSim.InstructBoxDomainDimension(world_size, world_size,
        world_size)
    DEMSim.InstructBoxDomainBoundingBC("all", mat_type_granular)
    walls = DEMSim.AddExternalObject()
    walls.AddCylinder([0, 0, 0], [0, 0, 1], world_size / 2.,
        mat_type_mixer, 0)

    # Define the meshed mixer and its prescribed motion
    mixer = DEMSim.AddWavefrontMeshObject(
        DEME.GetDEMDataFile("mesh/internal_mixer.obj"),
        mat_type_mixer)
    print(f"Total_num_of_triangles: {mixer.GetNumTriangles()}")
    mixer.Scale([world_size / 2, world_size / 2, chamber_height])
    mixer.SetFamily(10)
    DEMSim.SetFamilyPrescribedAngVel(10, "0", "0", "3.14159")
    # Track the mixer
    mixer_tracker = DEMSim.Track(mixer)

    # Define the clump template used in the simulation
    granular_rad = 0.005
    mass = 2.6e3 * 5.5886717
    MOI = np.array([2.928, 2.6029, 3.9908]) * 2.6e3
    template_granular = DEMSim.LoadClumpType(mass, MOI.tolist(),
        DEME.GetDEMDataFile("clumps/3_clump.csv"),
        mat_type_granular)
    template_granular.Scale(granular_rad)
    # Sampler uses hex close-packing
    sampler = DEME.HCPSampler(3.0 * granular_rad)
    fill_center = [0, 0, fill_bottom + fill_height / 2]
    fill_radius = world_size / 2. - 2. * granular_rad
    input_xyz = sampler.SampleCylinderZ(
        fill_center, fill_radius, fill_height / 2)
    DEMSim.AddClumps(template_granular, input_xyz)
    print(f"Total_num_of_particles: {len(input_xyz)}")
```



```

DEMSim.SetInitTimeStep(step_size)
DEMSim.SetGravitationalAcceleration([0, 0, -9.81])
DEMSim.SetErrorOutVelocity(20.)
DEMSim.SetForceCalcThreadsPerBlock(512)
DEMSim.Initialize()

sim_end = 10.0
fps = 20
frame_time = 1.0 / fps

# Keep a tab of the max velocity in the simulation
max_v_finder = DEMSim.CreateInspector("clump_max_absv")

print(f"Output_at_{fps}_FPS")
currframe = 0

mixer_tracker.SetPos([0, 0, chamber_bottom + chamber_height /
2.0])

t = 0.
start = time.process_time()
while (t < sim_end):
    print(f"Frame:{currframe}", flush=True)
    filename = os.path.join(out_dir, f"DEMDemo_output_{
currframe:04d}.csv")
    meshname = os.path.join(out_dir, f"DEMDemo_mesh_{
currframe:04d}.vtk")
    DEMSim.WriteSphereFile(filename)
    DEMSim.WriteMeshFile(meshname)
    currframe += 1

    max_v = max_v_finder.GetValue()
    print(
        f"Max_velocity_of_any_point_in_simulation_is_{
max_v}", flush=True)

    print(
        f"Solver's_current_update_frequency_{(auto-
adapted):_{DEMSim.GetUpdateFreq()}}", flush=
True)

    print(
        f"Average_contacts_each_sphere_has:{DEMSim.
GetAvgSphContacts()}", flush=True)

    mixer_moi = np.array(mixer_tracker.MOI())
    mixer_acc = np.array(mixer_tracker.ContactAngAccLocal())
    mixer_torque = np.cross(mixer_acc, mixer_moi)
    print(
        f"Contact_torque_on_the_mixer_is_{(mixer_torque
[0]),_{(mixer_torque[1]),_{(mixer_torque[2])}}",
        flush=True)

    DEMSim.DoDynamics(frame_time)
    DEMSim.ShowThreadCollaborationStats()

    t += frame_time

elapsed_time = time.process_time() - start
print(f"{elapsed_time}_seconds_{(wall_time)}_to_finish_this_
simulation")

```

#### 4. DEM model

This section details the default force models in DEM-Engine and the implementation of the geometry representations.

##### 4.1. History-based Hertz–Mindlin model

The default force model is anchored by the Hertzian contact model [73] and integrates the Mindlin friction model [74]. For a comprehensive analysis, readers may refer to [75]. For two bodies, namely  $i$  and  $j$ , when they are in contact, the normal force,  $\mathbf{F}_n$ , operates based on a spring–damper model. The tangential frictional force,  $\mathbf{F}_t$ , is computed considering material attributes and microscopic deformations, ensuring it adheres to the Coulomb limit via the friction coefficient  $\mu$ . The mathematical representation is as follows:

$$\mathbf{F}_n = f(\bar{R}, \delta_n)(k_n \mathbf{u}_n - \gamma_n \bar{m} \mathbf{v}_n), \quad (1a)$$

$$\mathbf{F}_t = f(\bar{R}, \delta_n)(-k_t \mathbf{u}_t - \gamma_t \bar{m} \mathbf{v}_t), \quad \|\mathbf{F}_t\| \leq \mu \|\mathbf{F}_n\|, \quad (1b)$$

$$f(\bar{R}, \delta_n) = \sqrt{\bar{R} \delta_n}, \quad (1c)$$

$$\bar{R} = R_i R_j / (R_i + R_j), \quad (1d)$$

$$\bar{m} = m_i m_j / (m_i + m_j), \quad (1e)$$

where the constants  $k_n$ ,  $k_t$ ,  $\gamma_n$ , and  $\gamma_t$  are inferred from material characteristics, including Young's modulus  $E$ , the Poisson's ratio  $\nu$ , and the restitution coefficient, CoR [76]. The terms  $\bar{m}$  and  $\bar{R}$  depict the effective mass and curvature radius for the specific contact. The foundational premise is that the geometries can undergo small penetration,  $\delta_n$ , at the contact point. The normal penetration vector is  $\mathbf{u}_n = \delta_n \mathbf{n}$ . The relative speed,  $\mathbf{v}_{rel} = \mathbf{v}_n + \mathbf{v}_t$ , at the contact point is defined as:

$$\mathbf{v}_{rel} = \mathbf{v}_j + \boldsymbol{\omega}_j \times \mathbf{r}_j - \mathbf{v}_i - \boldsymbol{\omega}_i \times \mathbf{r}_i, \quad (2a)$$

$$\mathbf{v}_n = (\mathbf{v}_{rel} \cdot \mathbf{n}) \mathbf{n}, \quad (2b)$$

$$\mathbf{v}_t = \mathbf{v}_{rel} - \mathbf{v}_n, \quad (2c)$$

where  $\mathbf{v}_i$ ,  $\boldsymbol{\omega}_i$  and  $\mathbf{v}_j$ ,  $\boldsymbol{\omega}_j$  denote the velocities at the mass centers and angular speeds of entities  $i$  and  $j$ . The position vectors,  $\mathbf{r}_i$  and  $\mathbf{r}_j$ , extend from the mass centers of bodies  $i$  and  $j$  to the shared contact point. The frictional force  $\mathbf{F}_t$  varies based on the historical tangential micro-displacement  $\mathbf{u}_t$ , updated iteratively at each time interval throughout the interaction event based on  $\mathbf{v}_t$ . Let  $\mathbf{u}_t$  be the current tangential micro-displacement and  $\mathbf{u}'_t$  be the updated tangential micro-displacement, then

$$\mathbf{u}'_t = \mathbf{u}_t + h \mathbf{v}_t, \quad (2d)$$

$$\mathbf{u}'_t = \mathbf{u}' - (\mathbf{u}' \cdot \mathbf{n}) \mathbf{n}, \quad (2e)$$

where  $h$  is the time step size. The strategy adopted to update  $\mathbf{u}'_t$  is borrowed from [76]. After the update, we may need to clamp the updated tangential micro-displacement  $\mathbf{u}'_t$  to get the final  $\mathbf{u}_t^{\text{next}}$  for the next time step in order to satisfy the capping condition  $\|\mathbf{F}_t\| \leq \mu \|\mathbf{F}_n\|$ :

$$\mathbf{u}_t^{\text{next}} = \begin{cases} \mathbf{u}'_t & \text{if } \|\mathbf{F}_t\| \leq \mu \|\mathbf{F}_n\|, \\ \frac{\mu \|\mathbf{F}_n\|}{k_t} \frac{\mathbf{u}'_t}{\|\mathbf{u}'_t\|} & \text{otherwise.} \end{cases} \quad (2f)$$

The rolling resistance arises from an asymmetric normal stress profile at the contact patch [77]. In DEM-Engine's default force model, it is implemented as the torque  $\boldsymbol{\tau}_r$ . This torque is induced by a force that has the magnitude of the rolling resistance coefficient  $C_r$  times the normal force. The direction of this force is aligned with the rolling-contributed relative velocity at the contact point. This is summarized in the following equations:

$$\mathbf{F}_r = \frac{\boldsymbol{\omega}_j \times \mathbf{r}_j - \boldsymbol{\omega}_i \times \mathbf{r}_i}{\|\boldsymbol{\omega}_j \times \mathbf{r}_j - \boldsymbol{\omega}_i \times \mathbf{r}_i\|} C_r \mathbf{F}_n, \quad (2g)$$

$$\boldsymbol{\tau}_r = \mathbf{r}_i \times \mathbf{F}_r. \quad (2h)$$

As discussed in Sec. 2.4, a clump has mass properties associated with it, whereas its component spheres have material properties associated with them – in other words, each sphere of the clump that makes up an element can have different material properties. Consequently,  $\mathbf{F}_n$  and  $\mathbf{F}_t$  in Eq. (3a) and (3b) need to be derived from the contacts between component spheres. Then a reduction process is invoked to use these contact forces to update the element  $\mathbf{v}_i$  and  $\boldsymbol{\omega}_i$ , based on each clump's  $m_i$  and  $I_i$ , as well as the location vector for the contact point,  $\mathbf{r}_i$ . This is visualized in Fig. 6, and the equations of motion for entity  $i$  assume the form

$$m_i \frac{d\mathbf{v}_i}{dt} = m_i \mathbf{g} + \sum_{k=1}^{n_c} \mathbf{F}^k, \quad (3a)$$

$$I_i \frac{d\boldsymbol{\omega}_i}{dt} = \sum_{k=1}^{n_c} (\mathbf{r}^k \times \mathbf{F}^k + \boldsymbol{\tau}_r^k), \quad (3b)$$

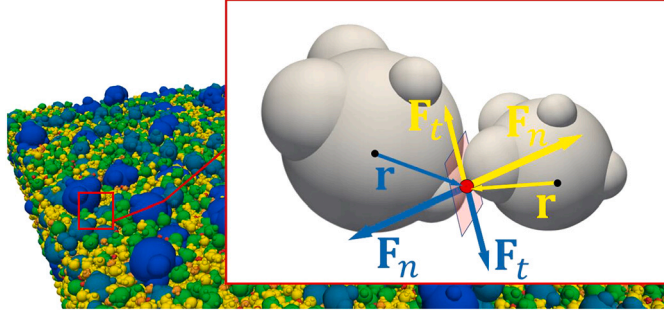


Fig. 6. The normal and tangential contact forces between particles are calculated based on the penetration and displacement history of involved sphere components.

where  $n_c$  is the number of contact spheres that entity  $i$  has, and the superscript  $k$  iterates through each contact. In these equations,  $\mathbf{F}^k = \mathbf{F}_n^k + \mathbf{F}_t^k$  means the total force, containing both the normal and tangential components.

#### 4.2. Providing a custom contact force model

To cater to diverse simulation needs, DEM-Engine supports custom force models through user-provided scripts. This section delves further into this functionality, whose starting point is a custom force model provided as a C++ script. This script undergoes just-in-time compilation at the onset of the simulation (as detailed in Sec. 2.2), replacing the default contact force model. The “ingredients” of a custom force model are called user-referable variables. A comprehensive list of these variables is provided in Table 2. For each contact pair, the solver automatically determines the values for these variables. Users can then harness these referable variables to implement the customized contact force.

Central to scripting the force model is the modification of the user-referable variable force, analogous to  $\mathbf{F}^k$  in Eqs. (3a) and (3b). This variable represents the force that geometry **A** experiences during contact in the global frame. The variable force takes the initial value of (0,0,0). It is worth noting that the solver will auto-apply the corresponding reaction force to geometry **B**. In a similar vein, the user-referable variable torque\_only\_force can be adjusted to store an action–reaction force pair that solely produces torque (without affecting the linear velocity of contact geometries, but only their angular momentum). This is congruent to  $\mathbf{F}_r$  in Eq. (2g). In the default model, the implementation of rolling resistance hinges on this variable. As Eqs. (3a) and (3b) indicate, a subroutine executed by the solver in each iteration, will integrate the motions of simulation entities post the force calculation.

Note that the three “wildcard” type variables in Table 2 are the custom properties that the user is allowed to associate with contacts, owners (clump, mesh, or analytical object), and geometries (sphere, triangle facet, or analytical component), respectively. For the owner wildcards and geometry wildcards, the user can assign their values before or during the simulation, using trackers or family tags. These custom properties can then be used in the custom force model to derive force, or be modified so their values change during simulation according to a user-specified policy.

The contact wildcards, on the other hand, work differently. If the user chooses to associate a wildcard to contacts, then the memory space associated with a contact is allocated when this contact emerges, and deallocated when this contact vanishes. When it is allocated, it always takes the initial value of zero. This is useful for recording some quantities that evolve during the lifespan of a contact. For example, as shown in Sec. 4.2.1, the default force model uses contact wildcards to record the contact history needed for the history-based Hertz–Mindlin model. At the implementation level, the generation of a new ACS by the kinematics thread triggers a CUDA subroutine. This subroutine is responsible for contrasting the new ACS with its predecessor, which is

cached by the kinematics thread, to establish a mapping. Subsequently, this mapping facilitates the transfer of contact wildcard variables associated to the old ACS to the new ACS.

##### 4.2.1. Default model implementation explained

We elaborate on the implementation of the default Hertz–Mindlin model in the remainder of this section, which can be found in the file FullHertzianForceModel.cu from the repository [63]. The code is an appropriate starting point for users to implement their own force model, potentially adding to the existing physics.

The preliminary step, as presented in the ensuing code snippet, involves extracting material properties of the contact geometries. Material property arrays adopt naming conventions consistent with the user-defined property names in the LoadMaterial function call. Consequently, if the default force model is employed, Young’s modulus ( $E$ ), Poisson’s ratio ( $\nu$ ), coefficient of restitution ( $\text{CoR}$ ), friction coefficient ( $\mu$ ), and rolling resistance coefficient ( $\text{Crr}$ ) must be specified in the LoadMaterial invocation. For users implementing a custom force model, the material property names specified during the LoadMaterial function should align with the array names in the force model file. For properties associated singularly with a material type (e.g., Young’s modulus), one should utilize the offset variables bodyAMatType or bodyBMatType to retrieve the property pertinent to the contact material. Conversely, for properties defined between two materials (like the friction coefficient), both offset variables are employed concurrently to obtain the appropriate value for the contact, as illustrated in the subsequent code snippet.

```
// Material properties
float E_cnt, G_cnt, CoR_cnt, mu_cnt, Crr_cnt;
{
    // E and nu are associated with each material, so obtain them
    // this way
    float E_A = E[bodyAMatType];
    float nu_A = nu[bodyAMatType];
    float E_B = E[bodyBMatType];
    float nu_B = nu[bodyBMatType];
    matProxy2ContactParam(E_cnt, G_cnt, E_A, nu_A, E_B, nu_B);
    // CoR, mu and Crr are pair-wise, so obtain them this way
    CoR_cnt = CoR[bodyAMatType][bodyBMatType];
    mu_cnt = mu[bodyAMatType][bodyBMatType];
    Crr_cnt = Crr[bodyAMatType][bodyBMatType];
}
```

In this implementation, because the force is set to be in the global frame, we do the calculation in the global frame. This requires us to compute the global angular velocity of the contact point on both contact geometries (albeit having the same location in space, the contact point on geometry **A** does not have the same velocity as that on geometry **B**, because of the intrinsic velocity that **A** and **B** have), since the user-referable variables ARotVel and BRotVel only give their angular velocity in local frames. This section of the code does this task.

```
float3 rotVelCPA, rotVelCPB;
{
    // This is local rotational velocity (the portion of linear
    // vel contributed by rotation)
    rotVelCPA = cross(ARotVel, locCPA);
    rotVelCPB = cross(BRotVel, locCPB);
    // This is mapping from local rotational velocity to global
    applyOriQToVector3(rotVelCPA.x, rotVelCPA.y, rotVelCPA.z,
        AOriQ.w, AOriQ.x, AOriQ.y, AOriQ.z);
    applyOriQToVector3(rotVelCPB.x, rotVelCPB.y, rotVelCPB.z,
        BOriQ.w, BOriQ.x, BOriQ.y, BOriQ.z);
}
```

Then the model calculates the normal force. Readers are referred to Sec. 4.1 to relate the implementation with the normal contact model. The material properties that are extracted previously, such as  $E_{\text{cnt}}$ , are used here to derive the force. One extra task carried out in this part is the update of the “wildcards”  $\delta_{\text{tan\_x}}$ ,  $\delta_{\text{tan\_y}}$ ,  $\delta_{\text{tan\_z}}$  and  $\delta_{\text{time}}$ , which are used to record the friction history. The contact history is used in the friction and rolling resistance calculation. At the end of this snippet, the variable force is updated to record the normal force.

**Table 2**

The user-referable variables that can be used in composing the custom force model. All data types are the default data type. Some of the data types can be configured in `VariableTypes.h` upon compilation from the source to accommodate the user's specific needs, a concept introduced in Sec. 2.3.

Type	Name	Explanation
double3	contactPnt	Contact point coord in global
float3	B2A	Unit vector pointing from geometry B to geometry A
double	overlapDepth	The length of overlap
float	ts	Time step size
float	time	Current time in simulation
float3	locCPA, locCPB	Positions of the contact point in the contact geometries' frames
double3	AOwnerPos, BOwnerPos	Positions of both owners
double3	bodyAPos, bodyBPos	Positions of both contact geometries
float4	AOriQ, BOriQ	Quaternions of both owners
float	AOwnerMass, BOwnerMass	Masses of both owners
float3	AOwnerMOI, BOwnerMOI	Moment of inertia for both owners
float	ARadius, BRadius	Radius of curvature for both contact geometries at point of contact
uint8_t	bodyAMatType, bodyBMatType	Offset used to query the material properties for both contact geometries
uint8_t	AOwnerFamily, BOwnerFamily	Family number of both owners
float3	ALinVel, BLinVel	Linear velocity of both owners
float3	ARotVel, BRotVel	Angular velocity of both owners, in their local frames
unsigned int	AOwner, BOwner	Offset for both owners in system array
unsigned int	AGeo, BGeo	Offset for both contact geometries in system array
float	User-specified	Contact wildcards: Extra properties associated with contacts
float	User-specified	Owner wildcards: Extra properties associated with owners
float	User-specified	Geometry wildcards: Extra properties associated with geometries
float3	force	Accumulator for contact force (in global)
float3	torque_only_force	Accumulator for contact torque (in global)

```
// A few re-usable variables that might be needed for both the
// tangential and normal force
float mass_eff, sqrt_Rd, beta;
float3 vrel_tan;
float3 delta_tan = make_float3(delta_tan_x, delta_tan_y,
    delta_tan_z);

// Normal force calculation
{
    // The (total) relative linear velocity of A relative to B
    const float3 velB2A = (ALinVel + rotVelCPA) - (BLinVel +
        rotVelCPB);
    const float projection = dot(velB2A, B2A);
    vrel_tan = velB2A - projection * B2A;

    // Update contact history
    {
        delta_tan += ts * vrel_tan;
        const float disp_proj = dot(delta_tan, B2A);
        delta_tan -= disp_proj * B2A;
        delta_time += ts;
    }

    mass_eff = (AOwnerMass * BOwnerMass) / (AOwnerMass +
        BOwnerMass);
    sqrt_Rd = sqrt(overlapDepth * (ARadius * BRadius) / (ARadius
        + BRadius));
    const float Sn = 2. * E_cnt * sqrt_Rd;

    const float loge = (CoR_cnt < 1e-12) ? log(1e-12) : log(
        CoR_cnt);
    beta = loge / sqrt(loge * loge + deme::PI * deme::PI);

    const float k_n = 2. / 3. * Sn;
    const float gamma_n = 2. * sqrt(5. / 6.) * beta * sqrt(Sn *
        mass_eff);

    force += (k_n * overlapDepth + gamma_n * projection) * B2A;
}
```

The snippet below calculates the rolling resistance. At the end of this snippet, the variable `torque_only_force` is updated to record the rolling resistance. Recall that this imaginary “force” contributes only to the contact torque, in agreement with the rolling resistance model in Eq. (2g).

```
if (Crr_cnt > 0.0) {
    bool should_add_rolling_resistance = true;
    {
        float R_eff = sqrtf((ARadius * BRadius) / (ARadius +
            BRadius));
        float kn_simple = 4. / 3. * E_cnt * sqrtf(R_eff);
        float gn_simple = -2.f * sqrtf(5. / 3. * mass_eff *
            E_cnt) * beta * powf(R_eff, 0.25f);

        float d_coeff = gn_simple / (2.f * sqrtf(kn_simple *
            mass_eff));

        if (d_coeff < 1.0) {
            float t_collision = deme::PI * sqrtf(mass_eff /
                (kn_simple * (1.f - d_coeff * d_coeff)));
            if (delta_time <= t_collision) {
                should_add_rolling_resistance = false;
            }
        }
    }
    if (should_add_rolling_resistance) {
        // Tangential velocity (only rolling contribution) of B
        // relative to A, at contact point, in global
        float3 v_rot = rotVelCPB - rotVelCPA;
        // This v_rot is only used for identifying resistance
        // direction
        float v_rot_mag = length(v_rot);
        if (v_rot_mag > 1e-12) {
            torque_only_force = (v_rot / v_rot_mag) * (
                Crr_cnt * length(force));
        }
    }
}
```

**Table 3**

The possible end status of the sphere in the rolling-on-incline test.

Mode	Stationary	Sliding	Rolling	Sliding and rolling
Definition	$\omega = 0, v = 0$	$\omega = 0, v > 0$	$v = \omega r$	$\omega > 0, v > \omega r$

The snippet below implements the friction force. The variable force is updated to record the friction force. Although the contact history variables (delta\_tan\_x, delta\_tan\_y, and delta\_tan\_z) are initially packed into a float3 (delta\_tan) for cleaner code, they are unpacked in the end to allow the solver to detect their modifications and write them back to memory. The contact history variables need modifications due to the potential tangential micro-displacement clamping, as shown in Eq. (2f).

```

if (mu_cnt > 0.0) {
    const float kt = 8. * G_cnt * sqrt_Rd;
    const float gt = -2. * sqrt(5. / 6.) * beta * sqrt(mass_eff *
        kt);
    float3 tangent_force = -kt * delta_tan - gt * vrel_tan;
    const float ft = length(tangent_force);
    if (ft > 1e-12) {
        // Reverse-engineer to get tangential displacement
        const float ft_max = length(force) * mu_cnt;
        if (ft > ft_max) {
            tangent_force = (ft_max / ft) * tangent_force;
            delta_tan = (tangent_force + gt * vrel_tan) / (-
                kt);
        }
    } else {
        tangent_force = make_float3(0, 0, 0);
    }
    force += tangent_force;
}

delta_tan_x = delta_tan.x;
delta_tan_y = delta_tan.y;
delta_tan_z = delta_tan.z;

```

The snippets provided combine to define the complete Hertz-Mindlin contact force model implemented in DEM-Engine. For a practical example of a custom force model in application, see Sec. 6.3 for a material breakage simulation. Users can also refer to the DEMdemo\_Electrostatic.cpp demo within the repository [63]. In that demo, elements are subjected to a contact force and an electrostatic force.

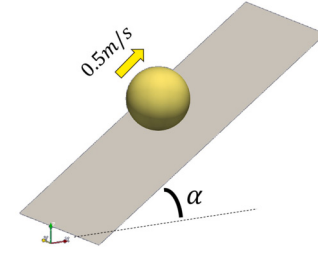
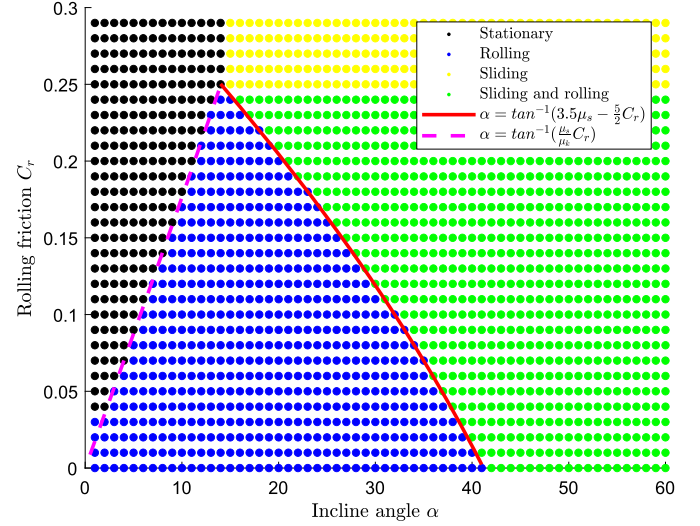
#### 4.3. Contact model validation

In this section, three small-scale tests are introduced to validate the implementation of the default force contact model. For notational brevity, for the rest of the paper, variables have their scopes limited to the respective section.

##### 4.3.1. Sphere rolling on incline

This is a simple but insightful test borrowed from [75], in which a sphere of radius  $r = 0.2$  m and mass 5 kg moves up on an incline with an initial velocity of 0.5 m/s, parallel with the incline and pointing up. In [75], the static friction coefficient  $\mu_s$  and kinetic friction coefficient  $\mu_k$  are allowed to have different values; however, in the default force model that we are validating, they assume the same value, and in this test  $\mu_s = \mu_k = 0.25$ . A test scene is illustrated in Fig. 7. The end status of the sphere can be one of the following modes depending on the incline angle  $\alpha$  and rolling resistance  $C_r$ : stationary; sliding; rolling; sliding and rolling. These modes are defined by the final angular velocity  $\omega$  and linear velocity  $v$  of the sphere, and are summarized in Table 3.

The outcome of this set of simulations is plotted in Fig. 8. It is shown in [75] that for the sphere to be stationary on the incline,  $\alpha \leq \tan^{-1}(\frac{\mu_s}{C_r})$ . For the sphere to roll down the incline without sliding,  $\alpha \leq \tan^{-1}(3.5\mu_s - \frac{5}{2}C_r)$ . These two conditions are plotted in Fig. 8 as the dashed and solid lines respectively, which evidently separate the stationary region, pure rolling region, and sliding-rolling mixed region

**Fig. 7.** A rendering of the sphere moving up an incline.**Fig. 8.** The end status of the sphere's mode, with respect to the incline angle and rolling resistance.

as the theory suggests. The DEM-Engine results confirm the results reported in [75].

##### 4.3.2. Sphere stacking

A set of three-sphere stacking tests was carried out to further validate the friction model implementation. This experiment is borrowed from [49,75]. For each test, two identical spheres of mass  $m_1 = 1$  kg and radius  $R = 0.15$  m with a small gap  $d$  in between were settled on a flat surface. A third sphere of the same radius  $R$  but a different mass,  $m_2$ , was placed between and above the bottom spheres with zero initial velocity, as illustrated in Fig. 9. To minimize the influence of impact, the third sphere was initialized right in contact (zero penetration depth) with the bottom ones. Depending on  $m_1$ , the gap, and rolling resistance coefficient  $C_r$ , two scenarios can occur: the top sphere drops to the ground, or it moves down slightly but the structure eventually stabilizes with the bottom spheres supporting the top sphere. This is a type of physics that also comes into play on a larger scale in angle of repose experiments. For different selections of  $C_r$ , the mass of the top sphere was increased by 0.02 kg to find the critical mass  $m_2$  for the pile to collapse, and the result is demonstrated in Fig. 10. The critical masses found for all initial gap sizes show exact matches with the outcome reported in [75], validating DEM-Engine force model implementation.

##### 4.3.3. Contact chain propagation

A third validation test for the default contact model involves investigating a closed-packed array of spheres arranged on a horizontal surface, initially in equilibrium under an external gravitational acceleration field. A localized force agent is applied to the uppermost symmetric point of this layout, allowing to study the overall sphere-to-sphere contact evolution parametrically by varying force magnitude and internal friction. This proposed setup follows the approach outlined in



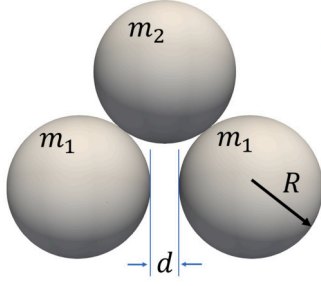


Fig. 9. A rendering of the three-sphere-stacking test.

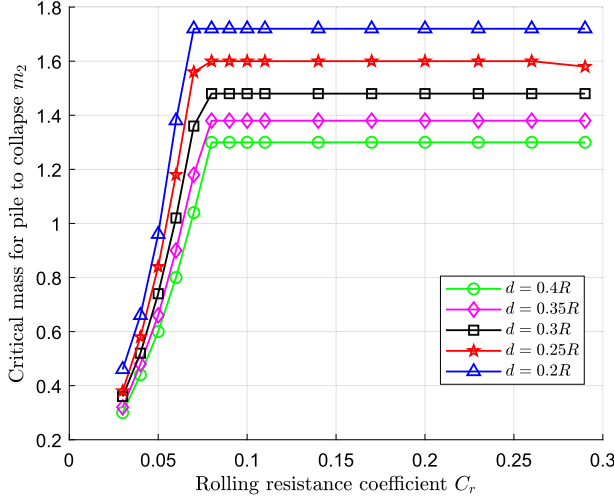


Fig. 10. The critical mass for the three-sphere pile to collapse, given a rolling resistance coefficient.

Table 4

The material and simulation properties used in the contact analysis.

Diameter [m]	Density [kg/m <sup>3</sup> ]	E [Pa]	$\nu$ [-]	CoR [-]
0.020 m	$1.0 \times 10^3$	$1 \times 10^7$	0.33	0.30

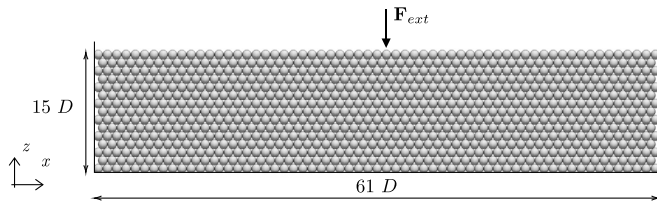


Fig. 11. Initial layout for contact chain benchmark.

[78,79], initially conceived to reveal granular material stress propagation capability; nevertheless, such benchmarks can also be valuable for verifying the sanity of DEM models. Here, we simulate the reference case using monodisperse spheres with a diameter  $D$ , in contrast to cylinders used in [78]. Fig. 11 illustrates the initial geometrical layout of the particle slab, consisting of a  $61 \times 15$  array of spheres layered according to a triangular lattice, thus counting 60 spheres for even layers. Table 4 details the parameters used to initialize the particle geometries and their contact force model. Note that the solid black lines in Fig. 11 indicate the analytical box boundaries, which use the same contact model as the granular material.

Contact chain formation is investigated for three inner friction values,  $\mu_i$ , 0.00, 0.10, and 0.20 respectively, under an external force denoted as  $F_{\text{ext}}$ , as shown in Fig. 11, applied to the CoG of the top row

particle belonging to the symmetry plane in y-z axis. The force magnitude is defined as  $F_{\text{ext}} = kmg$ , where  $k$  is a real number  $\in \mathbb{R}$ ,  $m$  is sphere mass, and  $g$  is the magnitude of the gravitational acceleration vector. For convenience, the weight of the single sphere,  $mg$  is denoted as  $\bar{m}$ . Thus, the DEM-Engine's configuration is completed by ensuring a properly defined time step, chosen in accordance with the critical time step  $\Delta t_{\text{crit}}$  given by the equation [77]

$$\Delta t_{\text{crit}} = \frac{\pi D}{2\beta} \sqrt{\frac{\rho E}{2(1+\nu)}}, \quad (4)$$

where  $\beta$  is approximated as  $0.8766 + 0.163\nu$  [80,81].

Prior to applying  $F_{\text{ext}}$ , to reliably achieve the layout presented in Fig. 11, all the spheres are generated at time 0 following a triangular lattice spaced at  $D$  plus some small clearance ( $1/1000$ th of  $D$ ) to avoid any possible overlap. Soon after, due to the gravitational pull, the particles are allowed to settle for 2 seconds of physical time, during which stationary conditions are reached. To further ensure compliance of the particle positions with the target configuration, the force model is initialized with no rebound factor, that is,  $\text{CoR} = 0.00$ . At this point, phase two begins and the  $F_{\text{ext}}$  is activated. To guarantee a smooth transition, the following function is implemented:

$$F_{\text{ext}}(t) = k\bar{m} \operatorname{erf}\left(\frac{t}{\sqrt{T_{\text{active}}}}\right), \quad (5)$$

where  $\operatorname{erf}$  is the error function,  $t$  is the simulation time, and  $T_{\text{active}} = \sqrt{2Dk\bar{m}}$  is set as the required time to smoothly transmit  $F_{\text{ext}}$  to the rest of particle system [82]. After the force application phase, the simulation is run for 5 more seconds.

Figs. 12a–c display the contact force profiles formed between spheres and the bottom boundary surface for tests with different inner friction coefficients, while varying the external force. These force profiles are constructed by collecting the vertical components of the contact force at the end of the first phase, denoted as  $F_0$ , and then isolating only the force increment, or decrement thereof, induced by  $F_{\text{ext}}$ , yielding  $\Delta F = F(t) - F_0$ , where  $t$  represents time, and here corresponding to the end of the physical time. Such shapes are then normalized with respect to the external force. By contrasting against reference data [78] and another numerical model's results [49], the proposed outcomes exhibit strong agreement, especially in predicting the transition from one-peak shaped curves, where linear elasticity similitude applies, to two-peak curves where contacts get redistributed due to particle displacement. More specifically, for low values of the external force, the one-peak curves totally agree with the results presented in [49,78], showing similar overall shape development and peak values. However, as the external force increases and the particle displacement becomes relevant to achieving new stationary configurations, the peak values begin to diverge from the results provided in [78]. This discrepancy is particularly significant in Panel 12a, where the blue curve corresponding to  $F_{\text{ext}} = 6.0\bar{m}$  exhibits peak values almost twice as high, likely imparted by the differences in the shapes that are involved in the analytical tests. The goodness of the proposed results, however, is further supported by the close agreement shown by the curves in Figs. 12a–c with the results in [49], in which Chrono::GPU was employed.

Fig. 12d offers deeper insights into the system response's dependence on inner friction and applied force. This chart presents the normalized force experienced by the sphere located at  $x/D = 0$ , charted against the magnitude of the external load, here indicated using  $k$ . When compared to the reference paper, this chart provides close matching with the expected behavior for all the studied friction parameters. Minor mismatching, however, affects the profile obtained for frictionless particles (i.e.,  $\mu_i = 0.00$ ). For low values of external force, i.e.,  $F_{\text{ext}} < 1.0\bar{m}$ , our model slightly underestimates the compression load experienced by the spheres located in the middle of the container, whereas for high values, the ratio becomes negative, indicating less pressure around these particles when only gravity is applied. Never-

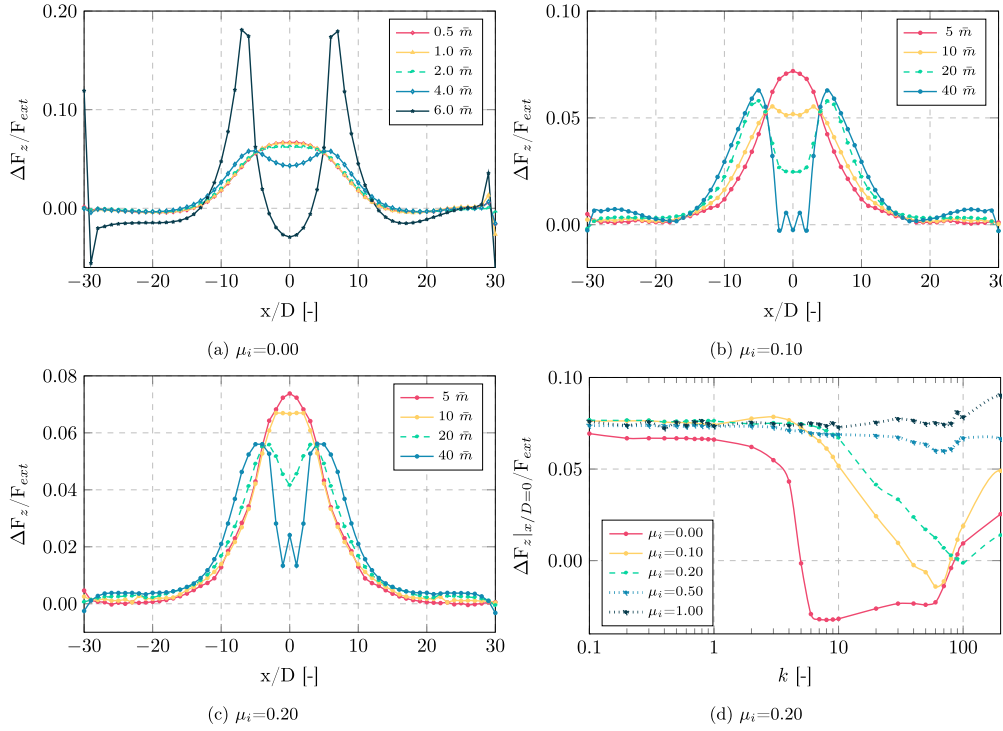


Fig. 12. Contact force distributions at the container bottom surface for different friction conditions and different values for the external force.

theless, the proposed results align very well with the results obtained by Chrono::GPU [49].

To better illustrate the contact distribution process, a series of simulation visuals is proposed in Fig. 13 for  $\mu_i = 0.20$ . In these images, particle contacts are instituted by colored straight lines parallel to the contact direction, with the length of each line proportional to the normalized force magnitude (force contact over  $\bar{m}$ ). Fig. 13a visualizes the contact distribution that forms at the end of the settling phase, showing an initial uniform distribution aligning with the prediction of linear elastic theory. As soon as an external agent is applied, as gradually depicted by Figs. 13b–e, the force diffusion tends to deviate from linearity, as observed in Fig. 12c. In the last case with  $k = 80$ , the sphere displacement becomes more evident, as a slight tilt forms at the horizontal surface of the container.

## 5. Simulator's performance

The scaling analysis in this section seeks to offer insights into the expected simulation performance of DEM-Engine. The chosen test scenario involves a bladed mixer interacting with granular material, where the mixer is modeled using a triangular mesh. Throughout the simulation, the mixer blades maintain a constant angular velocity of  $2\pi$  rad/s. Initially, the elements are positioned within a cylindrical region with a radius of 0.5 m and a height of  $1/3$  m above the mixer, and are subsequently released at the simulation's onset. The test's selection is due to its intensive particle–particle and particle–mesh interactions, demonstrated in Fig. 15. This puts the contact history preservation algorithm to the test, as contacts emerge and vanish in this highly dynamic problem. Material properties and simulation parameters can be found in Table 5.

In this analysis, three clump types are employed: individual spheres, three-sphere clumps, and six-sphere clumps, depicted in Fig. 14. Element sizes are adjusted to regulate the total element count. The mesh representing the mixer blades remains consistent across simulations, comprising 2892 triangular facets. Simulations are run until a pseudo-steady state is achieved at 1 s, after which the wall time required to carry out  $10^6$  time steps is recorded. The time step size is  $5 \times 10^{-7}$  s.

Table 5

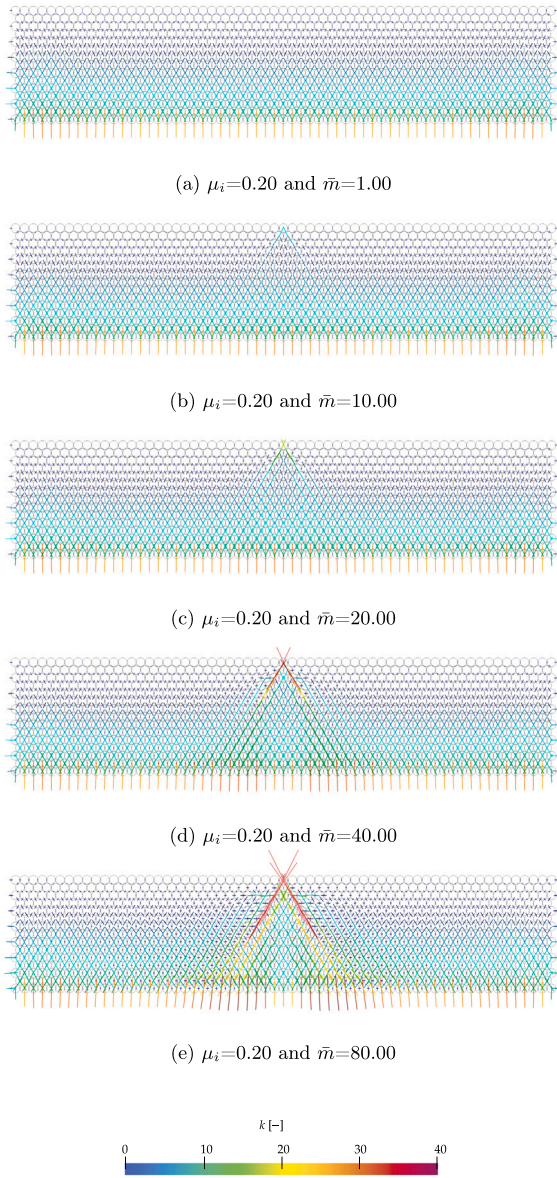
The material and simulation properties used in the mixer scaling analysis.

Density [kg/m <sup>3</sup> ]	$E$ [Pa]	$\nu$ [-]	CoR [-]	Step size [s]
$2.6 \times 10^3$	$1 \times 10^9$	0.3	0.2	$5 \times 10^{-7}$

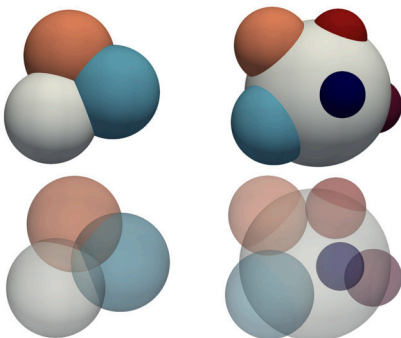
Fig. 16 displays the correlation between wall time and the total number of component spheres (distinct from the number of elements) via blue, green, and black markers. The simulations are performed on two NVIDIA Ampere A100 GPUs. On average, Chrono DEM-Engine takes 0.546, 0.313, and 0.264 hours to complete one million steps for every million component spheres in the simulations for the individual spheres, three-sphere clumps, and six-sphere clumps, respectively. The linear scaling persists to up to 150 million component spheres in the tests.

An identical simulation is also executed with Chrono::GPU (utilizing only one A100 as Chrono::GPU is limited to using a single GPU), and its scaling is represented with red markers. This juxtaposition is pertinent given a recent independent study's findings, which underscored that Chrono::GPU outperforms two other established DEM packages by two orders of magnitude [51]. Therein, for a 420,000-element pebble-packing simulation, Chrono::GPU running on a laptop GPU finished the simulation in an amount of time 261 times shorter than that required by LAMMPS, when the latter ran on 432 CPU cores of a cluster. For a 660,000-element pebble-packing simulation, Chrono::GPU executed 501 times faster than STAR-CCM+, which ran on 160 CPU cores. In both tests, Chrono::GPU ran on the RTX 2060 Mobile NVIDIA GPU card of a laptop. As indicated in Fig. 16, Chrono DEM-Engine demonstrates an additional twofold efficiency boost over Chrono::GPU in the test case of spherical elements. Owing to its ability to handle complex DEM particle shapes, Chrono DEM-Engine expands the modeling capacity of its predecessor without compromising per-GPU efficiency.

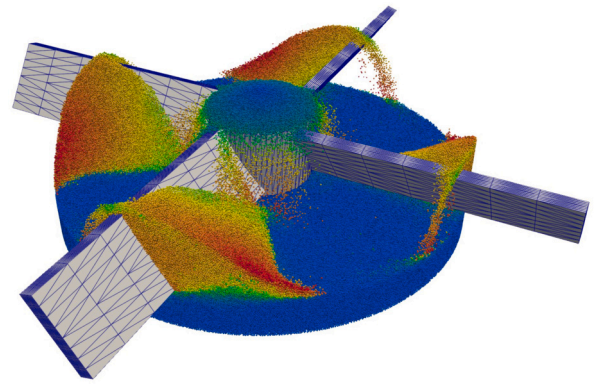
Fig. 17 shows the time spent in the important steps of the kinematics and dynamics threads' work cycles in the largest six-sphere-clump mixer simulation run in the scaling analysis. In that scenario, the amount of mutual contact data produced is relatively large, causing the kinematics



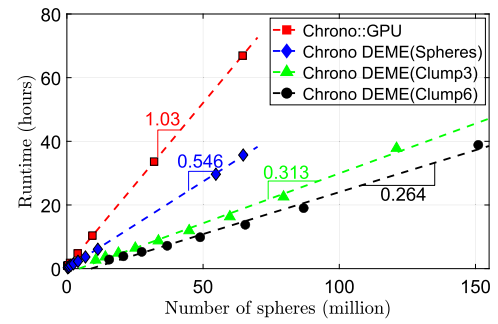
**Fig. 13.** Elaborated visualizations of the contact distribution for the case with  $\mu_i = 0.20$ .



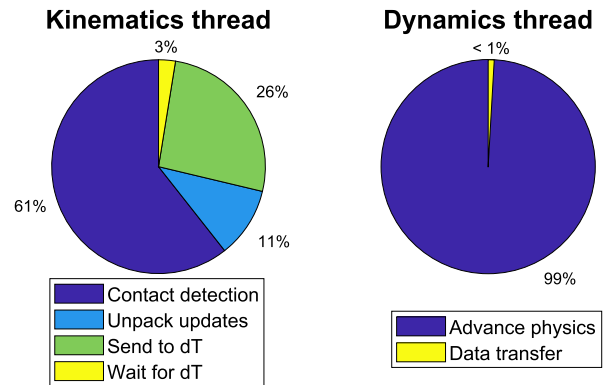
**Fig. 14.** The element shapes for the three-sphere and six-sphere clumps.



**Fig. 15.** A rendering of the mixing process.



**Fig. 16.** The scaling result of the mixer simulation using individual spheres, three-sphere clumps, and six-sphere clumps, on NVIDIA A100s. The wall time to finish simulating  $10^6$  steps is plotted against the number of component spheres in the simulation.



**Fig. 17.** The runtime breakdown for the kinematics and dynamics threads, during the lifespan of the largest six-sphere-clump mixer simulation.

thread to spend a large amount of time transferring it to the dynamics thread, reaching 26% of the former thread's total runtime. The dynamics thread spends minimal time on transferring data. This is done by design to enable the dynamics thread to almost exclusively focus on advancing the state of the system forward in time.

## 6. Numerical experiments

This section introduces a series of numerical tests, from medium-sized hopper flow rate tests to large-scale co-simulation, designed to compare the DEM-Engine simulation results against experimental data.



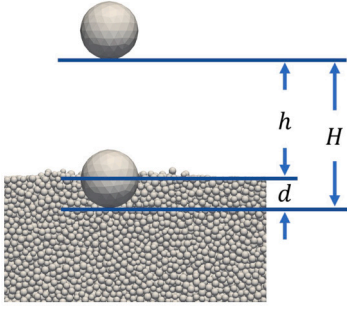


Fig. 18. Diagram of the initial and final projectile positions.

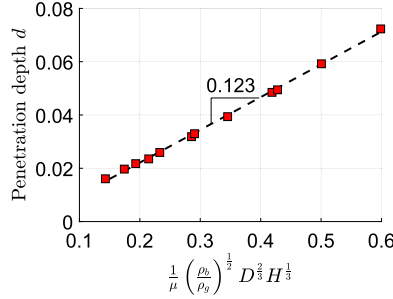


Fig. 19. Penetration depth. Each red square represents a data point in the numerical test.

### 6.1. Ball impact test

This experiment is described in [83]. A spherical projectile characterized by diameter  $D$  and density  $\rho_b$  was released from varying heights,  $h$ , onto a loosely packed pile of granular material, visualized in Fig. 18. The resulting penetration depth  $d$  of this sphere was gauged and set against the empirical model derived from the experimental data in [83], which goes as:

$$d = \frac{C}{\mu} \left( \frac{\rho_b}{\rho_g} \right)^{\frac{1}{2}} D^{\frac{2}{3}} H^{\frac{1}{3}}, \quad (6)$$

where  $\rho_g$  denotes the granular material's bulk density, and  $H = h + d$  is the sum of penetration depth and drop height. In [83], the constant  $C$  is estimated from experiments to be  $C = 0.14$ .

Twelve numerical tests using DEM-Engine were run aiming to reproduce the experiment in [83] as faithfully as possible. These tests incorporate combinations of projectile densities  $\rho_b = 2.2, 3.8, 7.8, 15 \text{ g/cm}^3$ , resembling Teflon, ceramic, steel, and tungsten, respectively. The diameter of the spherical projectile is  $D = 2.54 \text{ cm}$ . The release heights take values  $h = 5, 10, 20 \text{ cm}$ . Each simulation uses eleven types of spherical elements with diameters evenly distributed in the range between 0.25 cm and 0.35 cm (inclusive), and each DEM element has an even chance of spawning as one of them. The grain material in use has density  $\rho_{\text{grain}} = 2.5 \text{ g/cm}^3$ , resembling silica. This is to be differentiated from the bulk density of the granular bed, which is packed at  $\rho_g = 1.46 \text{ g/cm}^3$ , with a sliding friction coefficient of  $\mu = 0.3$ .

The correlation between depth  $d$  and the adjusted total release height  $H$  can be observed in Fig. 19. The line represents a linear regression of the numerical outcomes, showing a slope of 0.123, which confirms the experimentally established empirical model in Eq. (6). Comparable outcomes were also documented in [84] and [49], where both non-smooth and smooth contact dynamics approaches were leveraged for validating the same physics.

### 6.2. Flow sensitivity test

This section investigates the flow behavior exhibited by granular phases characterized by heterogeneous properties, encompassing variations in shape, density, and friction coefficient. Furthermore, relevant details regarding simulation runtimes are provided where applicable. The hardware configuration utilized for these numerical validations features an AMD Ryzen 9 5950X CPU in conjunction with a single NVIDIA A5000 GPU card.

#### 6.2.1. Drum tests

The first test investigates the flowability of particle media comprising four typologies: plastic spheres, plastic cylinders, wooden spheres, and wooden cylinders. The reference data is presented in Cui et al. [85], where experimental and numerical tests were performed on spherical and nonspherical particles. The experimental setup for the estimation of the angle of repose, a schematic of which is proposed in Fig. 20, comprised of a rotating drum made of transparent acrylic with an inner diameter ( $D_d$ ) of 0.19 m and a depth of 0.20 m ( $W_d$ ). For this investigation, the considered physical test outcomes refer to the test performing the drum rotating angular velocity,  $\theta_d$ , of 3.60 revolutions per minute (rpm).

This test is also considered to assess the accuracy of DEM-Engine in simulating complex shapes, which are formed by a compound of spheres, and defined as clumps. In the following, as shown in Fig. 20, the two shapes that characterize the tested particles consist of pure spheres with uniform radii, and five sphere clumps to mimic the geometric outer shape of cylinders.

Fig. 21 illustrates the sensitivity of the angle of repose for the rotating drum experiment. Each plot refers to a different material setup proposed in Table 6, using a test matrix that uses 13 values  $\in [0.00, 0.90]$  for the definition of the inner friction ( $\mu_i$ ) and five values  $\in [0.00, 0.08]$  for the definition of the rolling friction ( $C_r$ ). The material is initialized to fill half of the volume of the drum; then, the drum initiates its rotation at a constant angular velocity of 3.60 rpm, and let run for two seconds, after which it is assumed the system achieves a steady state. For the four different drum configurations, five seconds of simulations took approximately 0.20 h for each case with spheres (i.e., PS and WS), whereas 0.6 h hours for PC and WC. The angle of repose, as reported in the charts, is computed as the mean value of thirty measurements taken at an interval during the three seconds of simulation. For all the cases, very little deviation was observed throughout the post-processing phase.

Fig. 21 illustrates some of the key characteristics exhibited by granular materials when simulated using a DEM-based numerical solver. Firstly, it is evident that as the internal friction assigned to the spheres approaches zero, the system response yields very small angles of repose, ultimately resulting in a near-horizontal surface in the absence of internal friction. Conversely, for cylindrical particles lacking internal friction, the shape itself contributes to the bearing capacity of the system, as expected. Moreover, rolling resistance influences the angle of repose. When  $\mu_i$  is small, the disparity between cylindrical particles with and without  $C_r$  remains consistently lower. Note that, for a given pair of ( $\mu_i$ ,  $C_r$ ), similar particle shapes yield comparable angles of repose, irrespective of particle size or density. These initial observations align with the findings reported in [85], wherein the authors utilized the superquadratic DEM approach implemented in the open-source CFD suite MFIX [86] for simulating these same particles.

By contrasting the numerical solutions against the experimental data presented in [85] and illustrated in Fig. 21 through dashed black lines, one can assess the accuracy of the DEM-Engine in simulating granular materials. First, when considering two simulated spherical particle materials (Fig. 21a and c)), in which the grain shapes align with their physical counterparts, the valid angles of repose exhibit a wide range of values in relation to internal friction (i.e., from  $\mu_i \in 0.25-0.90$ ), while only minimal variability is linked to rolling friction. Secondly,



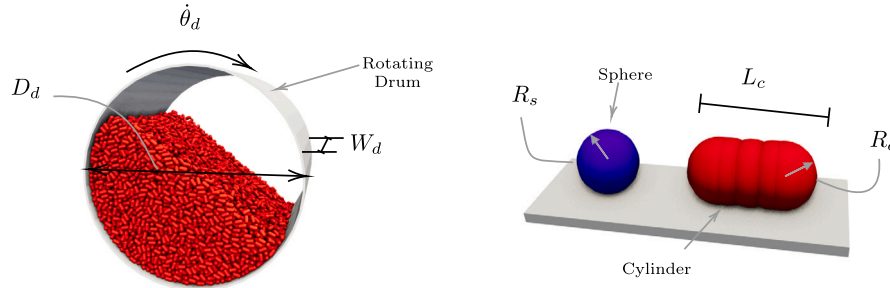
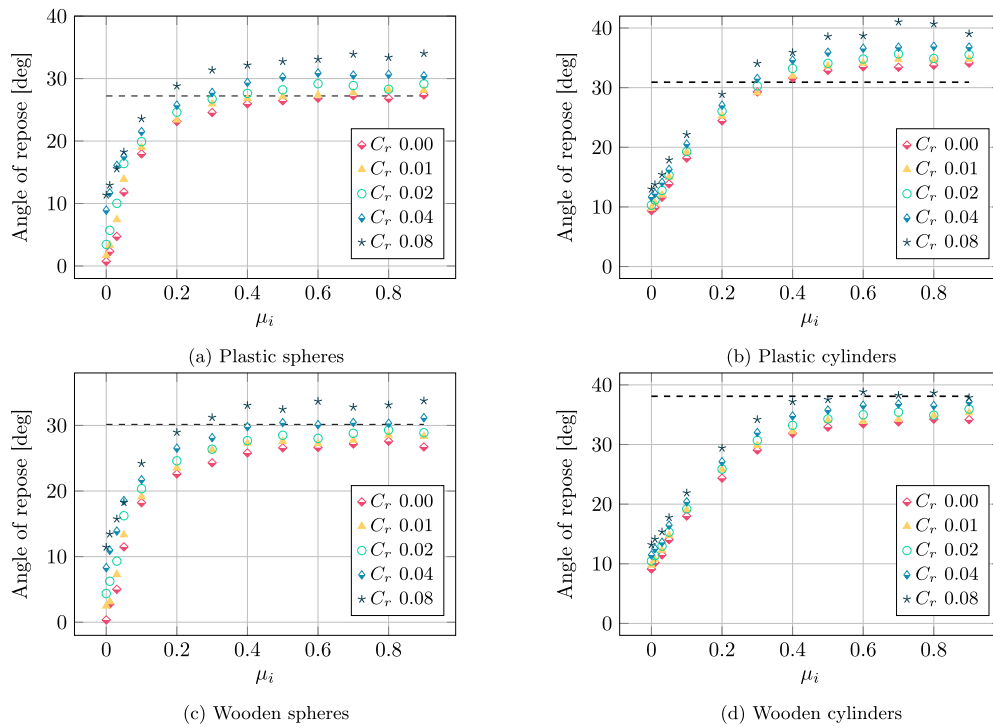


Fig. 20. Schematic visualization of the rotating drum test.

Table 6

Properties of four different particle setups used in this numerical investigation.

ID	Material	Shape	Radius [mm]	Length [mm]	Density [kg/m <sup>3</sup> ]	$E$ [MPa]	$\nu$ [-]	$CoR$ [-]	Clumps [-]	Spheres [-]
PS	Plastic	Sphere	3.0	-	1592	10.0	0.35	0.85	13024	13024
PC	Plastic	Cylinder	2.0	8.0	1128	10.0	0.35	0.85	19036	95180
WS	Wooden	Sphere	2.95	-	674	10.0	0.35	0.55	17112	17112
WC	Wooden	Cylinder	2.0	8.5	476	10.0	0.35	0.55	17016	85080

Fig. 21. Sensitivity of the angle of repose to the inner ( $\mu_i$ ) and rolling friction ( $C_r$ ) for the four granular materials in Table 6. The dashed line in each chart reports the reference value for the corresponding experimental test [85].

employing 5-sphere clumps to emulate plastic and wooden cylinders, as reported in Fig. 21b) and d), offers distinct operational domains for these un-physically consistent cylinders, where both shape and surface properties play pivotal roles. This analysis shows that the combined effects of internal and rolling frictions provide DEM-Engine with greater versatility.

### 6.2.2. Hopper tests

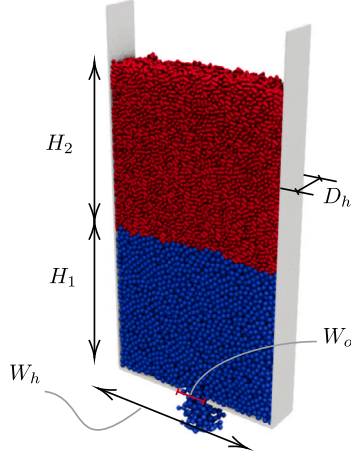
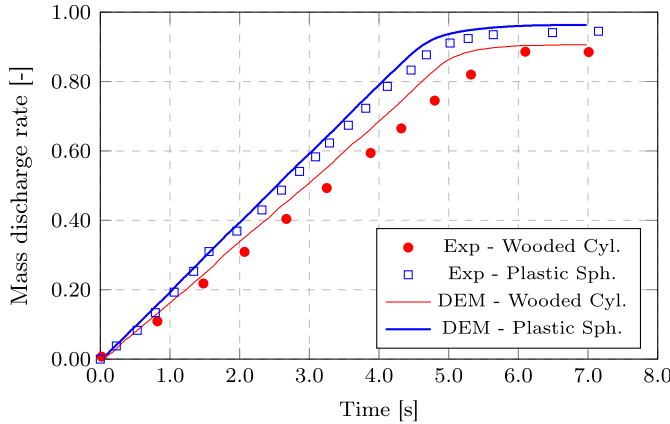
This test assesses the dynamic properties exhibited by a flow of DEM particles when simulated using the DEM-Engine. As reference solutions for this task, data regarding the mass discharge rate for both single and binary component systems are targeted, as made available in [87]. The physical testing was conducted using a flat-bottom hopper, see Fig. 22.

The hopper has a height of 0.40 m, width of 0.20 m, and depth of 0.04 m. An orifice of 0.04 m is symmetrically positioned on the lower surface. For the experimental campaign, various particle configurations were investigated. However, for this numerical validation, only four configurations, which precisely correspond to those outlined in Table 6, are considered. Specifically, Table 7 provides details on the hopper configuration for the tests presented in the subsequent sections. The parameters  $\mu_i$  and  $C_r$  reported in the last two columns are set using the charts in Fig. 21. Note that the first two tests consist of single-component discharge tests, whereas the remaining use binary particle compositions. Each simulation spans a physical time of 7.50 s, with approximate runtimes of: 0.35 h for ID 1; 0.75 for ID 2; and 0.60 for IDs 3 and 4.

**Table 7**

Properties of four different particle combinations used in the hopper numerical investigation.

Test ID	Layer 1	Layer 2	$H_1$ [cm]	$H_2$ [cm]	$\mu_{i,1}$ [-]	$C_{r,1}$ [-]	$\mu_{i,2}$ [-]	$C_{r,2}$ [-]	Clumps [-]	Spheres [-]
1	PS	-	36	-	0.40	0.04	-	-	14058	14058
2	WC	-	36	-	0.70	0.07	-	-	20014	100070
3	PS	PC	18	18	0.40	0.04	0.30	0.03	17545	59565
4	PC	PS	18	18	0.30	0.03	0.40	0.04	17904	61684

**Fig. 22.** Schematic visualization of the flat-bottom hopper.**Fig. 23.** Experimental and numerical comparison of the mass discharge ratio for single component hoppers with plastic spheres (blue) and wooden cylinders (red).

In Fig. 23, the relative mass discharge is presented for Test IDs 1 and 2, which involve plastic sphere and wooden cylinder particles, respectively. The chart depicts a comparison between the experimental and numerical time evolution of the system, showcasing the mass discharge relative to the total mass. For both tests, DEM-Engine demonstrates a fair level of accuracy in predicting the flow evolution. It exhibits an excellent match for purely spherical shapes (PS), while a slight overestimation is shown for the cylinders (PC). This discrepancy, leaning towards a more *fluid* flow, can be attributed to the fact that the clumps of five spheres, used in place of actual cylindrical shapes, do not perfectly replicate the behavior of the physically consistent cylinders.

In Fig. 24, a visual comparison is provided for the binary particle systems: Test IDs 3 and 4 as outlined in Table 6. This comparison contrasts snapshots from both experimental and numerical perspectives, offering lateral views of the hopper at one-second intervals, starting from the initial configuration at Time = 0.00 s. The first and third rows respectively present data from [87], while the second and fourth rows

showcase the results from DEM-Engine's simulation. The two timelines evolve in a remarkably similar fashion, highlighting that the numerical model accurately captures all the pertinent physical phenomena that unfold.

### 6.3. Contact modeling for particle breakage

DEM simulations have often been employed to characterize complex flows, factoring in not only the outer geometries of particles but also specialized features such as flexibility or particle breakage [88]. The DEM-Engine offers an open framework that allows users to implement user-defined constitutive laws. This example details a custom implementation to model the behavior of a cohesive yet highly brittle elastoplastic material. This test involves accounting for the failure of local bonds. To this end, the model outlined in [89] is adopted for defining the constitutive laws and failure modes.

The following implementation leverages the variables presented in Sec. 4.1 for the history-based Hertz–Mindlin model. Pivotal to this implementation is the capability of having stored information regarding the state of the system, as also detailed in Sec. 4.2.1. The material properties that are used to define, in this case, *granite*, are used to define the contact forces. Concerning the parent contact method, two extra “wildcards” are defined: *unbroken* and *initialLength*, which are used to respectively record the contact state (i.e., broken or unbroken) and the initial length of the equivalent spring for the normal force. In the following, the general structure of the contact model is defined. Note that the values of the two wildcard variables are initialized to 1.0 and 0.0, respectively.

```
// DEM force calculation for grain breakage.
// The parameters required for the contact force computation
// are defined.

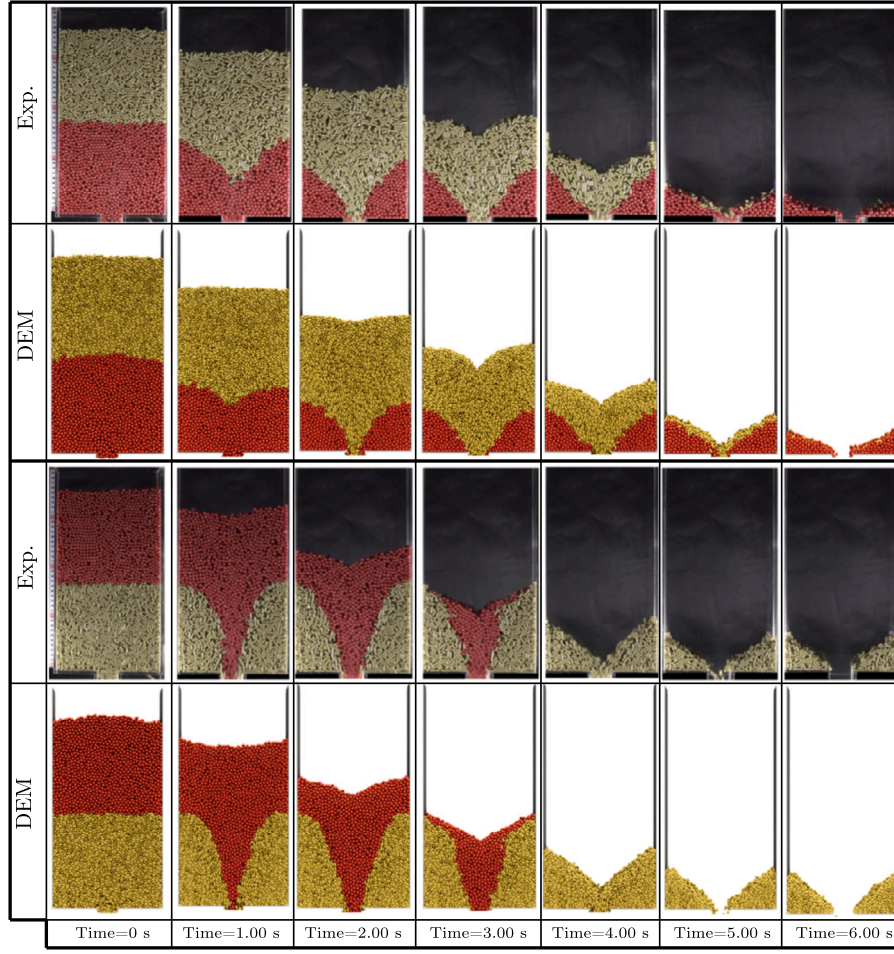
if (unbroken > 1e-12) {
    // Computation of the contact force for the breakage model
    // that accounts for normal and tangential forces, and
    // bending moments.
    // Here goes the implementation
} else {
    if (overlapDepth > 1e-12) {
        // The previously broken contact may still be engaged by
        // compressive force, and this happens especially for
        // compressive tests. The contact is treated with a
        // Hertzian contact law.
        // Here goes the implementation
    }
}
```

The magnitude of the model calculates the normal interactive force  $F_n$  using:

$$\mathbf{F}_n = k_n \mathbf{u}_n - \gamma_n \bar{m} \mathbf{v}_n, \quad (7)$$

where  $\gamma_n = 0.01 \sqrt{k_n / \bar{m}}$ ,  $k_n$  is the normal stiffness and it is defined according to the following cases:

$$k_n = \begin{cases} E_{eq} \bar{R} & \text{if } \text{sign}(\mathbf{u}_n) \|\mathbf{u}_n\| > \delta_y, \\ -E_{eq} \bar{R} & \text{if } \delta_b \leq \text{sign}(\mathbf{u}_n) \|\mathbf{u}_n\| < \delta_y, \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$



**Fig. 24.** Experimental and numerical comparison of the discharging behavior of two different packing patterns of the Plastic Sphere and Cylinders. (reprinted from [87]; copyright (2023), LN 5657690415083, with permission from Elsevier).

where  $E_{eq}$  is the equivalent stiffness of the contact,  $\xi$  is the degrading factor (softening) that accounts for the formation of initial cracks in the material,  $\delta_y$  is the material yielding threshold, and  $\delta_b$  is contact displacement failure, here assumes as three times  $\delta_y$ .

```
float tension = -9.3e6f;

// Normal force calculation
float deltaD = (overlapDepth - initialLength);
float kn = Eeq * (ARadius * BRadius) / ((ARadius + BRadius));

float intialArea = ((ARadius > BRadius) ? ARadius * ARadius :
  BRadius * BRadius) * deme::PI;

float BreakingForce = tension * intialArea;
float deltaY = BreakingForce / kn;
float deltaU = 3.0f * deltaY;

float force_to_A_mag = (deltaD > deltaY) ? kn * deltaD : ((
  deltaU - deltaD) - deltaY) * kn * 0.5f;

float damping = 0.01 * sqrt(mass_eff * kn);

force += B2A * force_to_A_mag - damping * velB2A;
// breaking for excess of tensile force
unbroken = (deltaD < deltaU) ? -1.0 : unbroken;
```

The tangential component of each contact follows from the following relationship:

$$\mathbf{F}_t = -k_t \mathbf{u}_t - \gamma_t \bar{m} \mathbf{v}_t,$$

given:  $\|\mathbf{F}_t\| \leq \begin{cases} \mu \|\mathbf{F}_n\| + c A_{\text{int}} & \text{if } \text{sign}(\mathbf{u}_n) \|\mathbf{u}_n\| > \delta_y, \\ \mu \|\mathbf{F}_n\| & \text{if } \text{sign}(\mathbf{u}_n) \|\mathbf{u}_n\| \leq \delta_y, \end{cases} \quad (9)$

where  $k_t = v_i k_n$ ,  $c$  is the material cohesion, and  $A_{\text{int}}$  is the interacting surface for the contact and defined as  $\pi \cdot \min(R_i, R_j)^2$ . The following snippet of code provides the specific details of the implementation:

```
float cohesion = 200e6;
// Tangential force calculation

float kt = nu_cnt * kn;
float Fsmax = (deltaD > deltaY) ? length(force) * mu_cnt +
  cohesion * intialArea : length(force) * mu_cnt;

const float loge = (CoR_cnt < 1e-12) ? log(1e-12) : log(
  CoR_cnt);
beta = loge / sqrt(loge * loge + deme::PI * deme::PI);
float gt = 2. * sqrt(5. / 6.) * beta * sqrt(mass_eff * kt);

float3 tangent_force = -kt * delta_tan - gt * vrel_tan;
delta_tan = (tangent_force + gt * vrel_tan) / (kt);

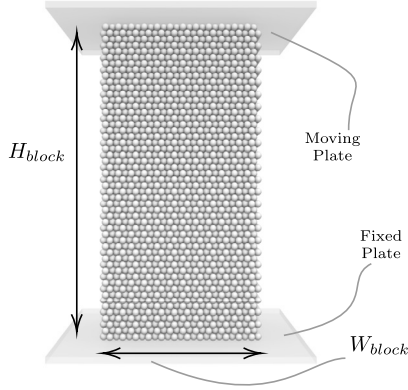
force += tangent_force;
// breaking for excess of tangential stress
unbroken = (length(tangent_force) > Fsmax) ? -1.0 : unbroken;
```

The bending resistance that arises at each contact, being representative of an element of finite size, is computed using Eq. (2g), where the bending stiffness is defined as  $k_r = R_i R_j k_t$  [90,91]. note that the maximum bending moment is capped by  $\min(\eta_i R_i, \eta_j R_j) \|\mathbf{F}_n\|$ , where  $\eta$  is a dimensionless coefficient that controls the rolling behavior of the contact. Lastly, here the code for the implementation of the fictitious bending resistance of the contact is listed. Note that no contact failure is associated with the bending moment value.

**Table 8**

Mechanical properties of the granite block, as defined in [89,92].

Material	Parameter	Value (unit)
Rock	Mass density	2640 kg/m <sup>3</sup>
	Young's modulus	60 × 10 <sup>9</sup> Pa
	Poisson's ratio	0.25
	Internal friction	0.30
	Compressive Strength	200 × 10 <sup>6</sup> Pa
	Tensile strength	9.3 × 10 <sup>6</sup> Pa
Plates	Young's modulus	100 × 10 <sup>9</sup> Pa
	Surface friction	0.50
	Poisson's ratio	0.30

**Fig. 25.** Numerical configuration for the axial compression test of a granite block.

```
// Bending moment induced-force calculation
float kr = ARadius * BRadius * kt;
float eta = 0.1f;

float var_1 = ts * kr / ARadius;
float var_2 = eta * length(force);

float3 torque_force;
if (v_rot_mag > 1e-12) {
    float torque_force_mag = (var_1 < var_2) ? var_1 : var_2;
    torque_force = (v_rot / v_rot_mag) * torque_force_mag;
}
force += torque_force;
```

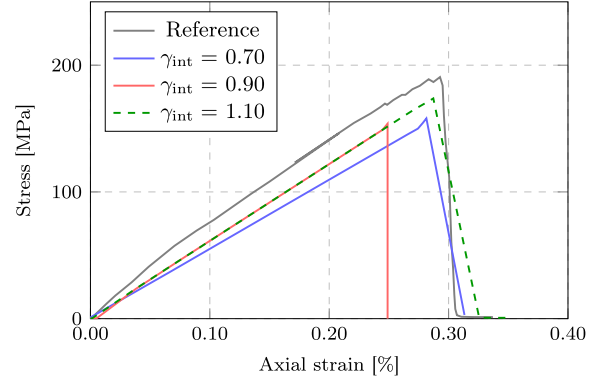
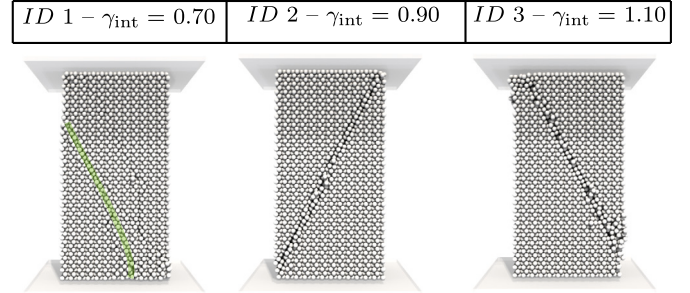
The previous implementation has been validated against experimental data from a uniaxial compression test conducted on a granite block, as defined in [92]. This particular test configuration is commonly utilized in the literature for code validation and calibration. In Fig. 25 and Table 8, we present the numerical test rig along with the mechanical properties of the rock specimen, lower plate, and upper plate. These properties have also been reviewed and interpreted by other studies [89,91,93]. The test rig consists of two rigid plates, with the lower plate fixed to the reference system while the upper plate moves vertically at a constant velocity of 5 mm/s. The tested specimen is constructed as a homogeneous assembly of spheres placed on a regular lattice arranged in a hexagonal close-packed (HCP) configuration, generated using an internal function provided by the DEM-Engine package. The specimen has a base area of  $W_{block} \times W_{block}$  with dimensions of 5.0 cm and a height  $H_{block}$  of 10.0 cm. The chosen sphere radius of 12 mm ensures that there are 20 particles within the width of the specimen.

A crucial parameter that significantly influences the accuracy of the proposed model for contact breaking is the particle interaction range, denoted as  $\gamma_{int} R_i$ , which defines the area of active links around each particle. Essentially, when a particle is initialized as part of the previously defined contact method, it is equipped with contacts that extend to the surrounding particles in accordance with the specified interaction range. In this study, three tests are conducted, considering different val-

**Table 9**

Model parameters description for the simulation of the particle breakage in axial compressive tests.

Test ID	Radius [mm]	$\gamma_{int}$ [-]	Spheres [-]	$\approx N$ [ $\times 10^3$ ]	$N_{i,mod}$ [-]
1	12	0.70	26754	154	6
2	12	0.90	26754	200	8
3	12	1.10	26754	230	9

**Fig. 26.** Strain-stress curves obtained from uniaxial compressive tests performed with three interaction range sizes. The reference solution corresponds to the numerical solution proposed in [89] for  $N_i = 13.8$ .**Fig. 27.** Visualization of the cracked configuration of the three specimens. For test ID 1, the crack has been highlighted using a light green curve.

ues of  $\gamma_{int}$ : [0.70, 0.90, 1.10]. Table 9 provides a summary of the micro properties for these three tests, including the total number of potential contacts, denoted as  $N$ , and the statistical mode of the number of contacts for a single particle, denoted as  $N_{i,mod}$ .

Fig. 26 displays the strain-stress curves for the three tests along with the numerical solution proposed in [89], where the average number of contacts per particle was  $N_i = 13.8$ . The data presented in this chart suggests the excellent agreement achieved by the implemented model compared to the one from the literature. Particularly, an increase in the interaction range leads to a more accurate representation of the specimen's stiffness. Case emphID 3 exhibits the highest level of agreement, with a relative error of less than 8% on the material ultimate resistance and 5% on the elastic modulus. One source of disagreement lies in the relatively small number of links (i.e., 9 compared to 13.8), which is a direct consequence of the uniform pattern used to initialize the particle arrangement and the uniform particle radius. Fig. 27 proposes rendered visualizations for the final instants of the three tests.

#### 6.4. Rover mobility co-simulation

This section discusses a co-simulation between a multi-body system and a DEM system. The rover simulation originally presented in [94] is reproduced herein while adding the usage of the “active box” scheme introduced later in this section. The co-simulation aims to measure the





Fig. 28. MGRU3 climbing a “tilt bed” in NASA’s Glenn Research Center testing facility [95].

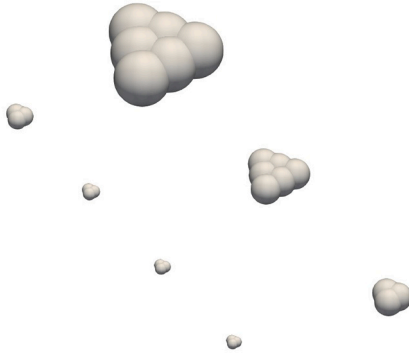


Fig. 29. The seven clump shapes that are used in the rover co-simulation.

Table 10

The weight distribution of the simulant used in the rover test, percent-wise, by clump size. For all element types,  $E = 10^8 \text{ N/m}^2$ ,  $\nu = 0.3$ ,  $\mu_s = 0.4$ , and  $\text{CoR} = 0.5$  in this simulation.

Type	1	2	3	4	5	6	7
Size [mm]	21	11.4	6.6	4.5	3	2.75	2.5
Component radius [mm]	3.6	1.95	1.81	1.24	0.82	0.75	0.7
%, by weight	17	21	14	19	16	5	8

slip ratios of a rover when operating on a “tilt bed” under Earth’s gravitational pull. The experimental data used for comparison are obtained using NASA’s Moon Gravitation Representative Unit 3 (MGRU3), see Fig. 28 (obtained from a publicly available video of the test [95]) for a photo of the test scene. However, in the co-simulation presented herein, since the MGRU3 CAD model is inaccessible, a similar VIPER rover model publicly available in the latest Chrono distribution [50] is used. The rover moves around by prescribing all its four wheels a  $0.8 \text{ rad/s}$  angular velocity on inclines of  $0^\circ$ ,  $5^\circ$ ,  $10^\circ$ ,  $15^\circ$ ,  $20^\circ$ , and  $25^\circ$ , where the inclines are modeled in simulation by adjusting the direction of the gravitational pull.

The experiment shown in Fig. 28 was done at Glenn Research Center, where the terrain simulant used is called GRC-1 [96]. In the co-simulation presented herein, the numerical representation of the terrain is inherited from [94], where seven different DEM element types are used (rendered in Fig. 29), each with a specific size and percentage of the total weight, see Table 10. The size distribution is plotted in Fig. 30, showing the DEM representation is uniformly increased by a factor of 20 the actual particle sizes encountered in GRC-1. For more details and the validation of this terrain representation, see [94]. A rendering of the co-simulation is shown in Fig. 32.

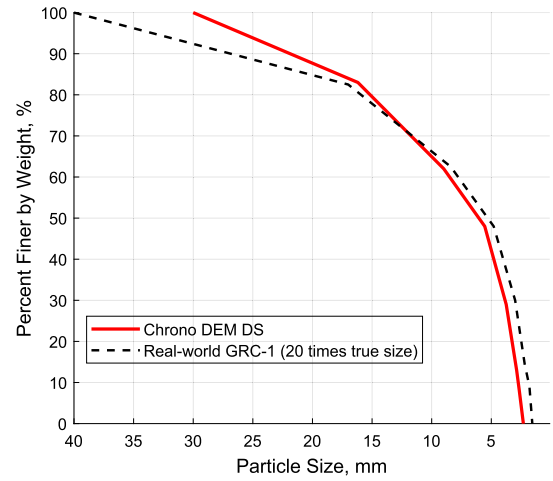


Fig. 30. The size distribution of the DEM elements used in the rover co-simulation, plotted against a scaled real-world GRC-1 simulant size distribution.

#### 6.4.1. Co-simulation

The co-simulation setup is depicted in Fig. 31. DEM-Engine handles the evolution of the granular terrain, while Chrono manages the rover dynamics. The two simulators are connected through the meshes representing the wheels. DEM-Engine calculates the force exerted by the terrain on the wheel mesh. This force information is employed when the Chrono numerical integrator propels the evolution of the meshes forward in time. Subsequently, the updated position of the wheels will serve as new boundary conditions for the granular material. The rover’s mobility is also influenced by forces that originate in the chassis and suspension, independent of the motion of the granular terrain. In this co-simulation, the rover system progresses with a time step size of  $2 \times 10^{-5} \text{ s}$ , whereas the DEM system uses a smaller time step of  $2 \times 10^{-6} \text{ s}$ . This means for every ten DEM time steps, the multi-body system in Chrono advances by just one step.

#### 6.4.2. Active box scheme

Using DEM-Engine’s API, the user can implement a partially active simulation domain to reduce computational cost. The user can assign different family tags (introduced in Sec. 2.2.2) to the elements inside and outside certain regions in the simulation domain to distinguish them. In this use case, no assigned motions are prescribed to the elements inside the  $1 \text{ m} \times 0.5 \text{ m}$  boxes centered around each wheel, as shown in Fig. 32 – their motion is to be determined by the simulator. These boxes are called active boxes. The DEM elements outside the active boxes are fixed in position and do not participate in the contact detection, i.e., remain dormant and contribute no computational cost. Note that the locations of the active boxes are updated (based on the locations of the wheel) 10 times per simulation second in this test.

The full-simulation data shown in Fig. 33 displays no notable difference compared to the active box-based counterpart. In [94], it is reported that the 15-second simulation requires approximately 109 hours of run time on two NVIDIA A100 GPUs. At the same time, the active box-based simulation presented herein, which involves 11,336,638 DEM elements (34,691,952 component spheres) takes around 30 hours. This suggests that the more expeditious active box-based tests are likely sufficient to gain insights into the rover’s mobility attributes, while costing a fraction of the computational cost of a full simulation. The numerical simulations also show good agreement with the experimental data (black line) from NASA’s Glenn Research Center MGRU3 experiments with the real-world GRC-1 simulant. The slip ratio increases relatively slowly with the slope angle in the interval between  $0^\circ$  and  $10^\circ$ . Past  $10^\circ$ , this rate of increase escalates, and the rover almost fails to climb on a  $25^\circ$  incline.

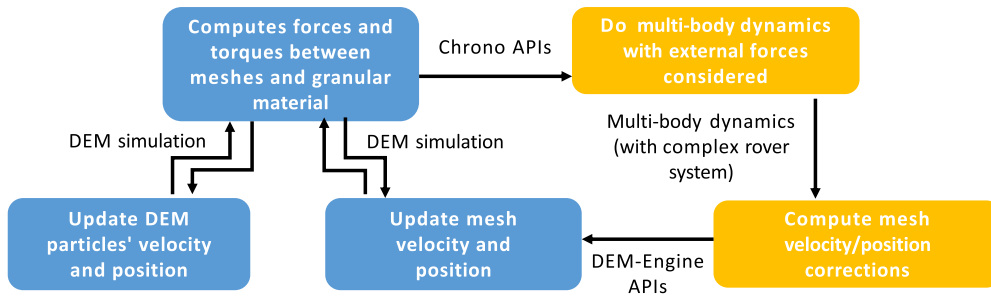


Fig. 31. The co-simulation workflow between the multi-body system simulated by Chrono and DEM-Engine.

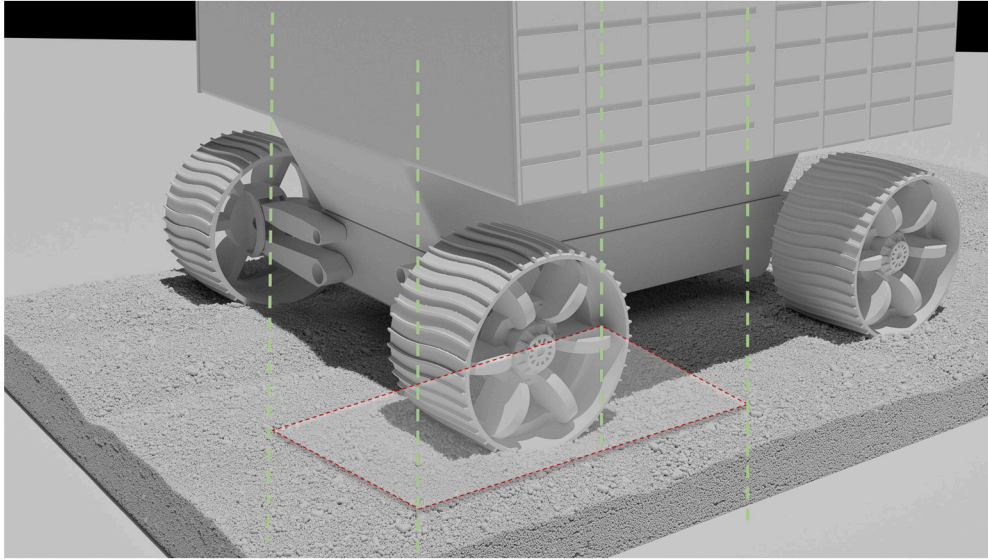


Fig. 32. A rendering of the VIPER rover operating on a 20° incline. The active box is marked and only the elements in that region are subject to the simulation physics; the rest are fixed.

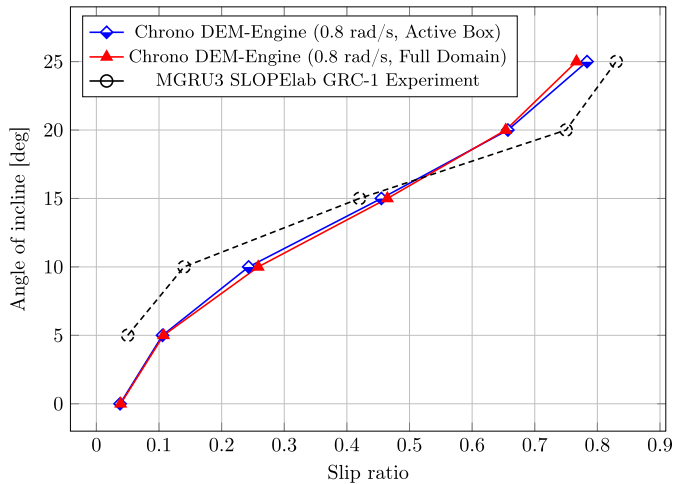


Fig. 33. The comparison between the full-domain and active box-based VIPER rover slip test results. The experimental data used for comparison (black line) are from Glenn Research Center's MGRU3 experiments with the GRC-1 simulant.

## 7. Conclusions and future directions

This paper has introduced Chrono DEM-Engine, an open-source, physics-based, dual-GPU DEM package that supports complex element shapes, positioning it as an enhancement to the existing Chrono::GPU simulator. The most distinctive implementation feature of Chrono DEM-

Engine is its partitioning of kinematic processes, such as contact detection, and dynamic computations, e.g., computation of the contact forces and carrying out numerical integrations. The resulting two computational threads operate asynchronously and share data only when necessary. Chrono DEM-Engine supports custom force models through just-in-time CUDA kernel compilation.

This manuscript first presents the C++ and Python code implementations. They are detailed to highlight the primary code features and specialized software components. From the default force model, Hertz–Mindlin, which possesses the capability to trace the history of contact interactions, the paper focuses on the code structure. Emphasis is placed on the data handling, accompanied by an overview of the procedure for customizing the force model. Following a rigorous contact validation against analytical solutions, the computational performance of DEM-Engine's core implementation is evaluated. This new DEM simulator can process tens of millions of elements on two A100 GPUs, achieving a throughput of one million time steps for one million DEM elements within an hour. In contrast to its predecessor, Chrono::GPU, which demonstrated in third-party studies to be two orders of magnitude faster than established DEM packages, the scaling analysis in this paper reveals that the new solver further increases this performance by a factor of 2×. Furthermore, the new simulator demonstrates linear scalability for up to 150 million component spheres using two GPUs.

The paper validates the solver's implementation through a comprehensive set of tests, including fine-grain force model evaluations and macro-scale experiments, such as ball drop, hopper flow rate, and rover climbing. The software is designed to handle complex particle geometries using clump models. This feature is validated through comparisons

with physical data for the flow discharge of spheres, cylinders, and combinations thereof from a rectangular hopper. Moreover, the software integrates with the multi-physics simulation engine Chrono, facilitating co-simulations with mechanical and multi-body systems, as evidenced by the proposed test case of simulating the rover operation.

Chrono DEM-Engine is an open-source, BSD3-distributed research code. As such, there is an inherent learning curve associated with its use. Users are required to sift through numerous APIs. Identifying and addressing the tool's limitations can also be daunting and may require time-consuming customization. This challenge becomes even more pronounced in modern cross-disciplinary research, where researchers are simultaneously handling a range of tools. However, the emergence of Large Language Models (LLMs) [97] offers a potential solution. As a future development thrust, it remains to investigate the use of LLMs to design assistant AIs that can translate users' natural language directives into executable DEM-Engine scripts. If this research trajectory proves successful, the resulting tool will be made available as open-source.

### Code availability

Chrono DEM-Engine is accessible as part of Project Chrono at <https://github.com/projectchrono/DEM-Engine>. All numerical examples discussed in this paper are provided as demo simulations.

### CRedit authorship contribution statement

**Ruochun Zhang:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Bonaventura Tagliaferro:** Investigation, Validation, Visualization, Writing – original draft, Writing – review & editing, Methodology, Software. **Colin Vanden Heuvel:** Writing – original draft, Methodology, Software. **Shlok Sabarwal:** Writing – original draft, Methodology, Software, Validation. **Luning Bakke:** Formal analysis, Investigation, Software, Writing – original draft, Writing – review & editing. **Yulong Yue:** Software, Methodology, Writing – original draft. **Xin Wei:** Software, Methodology, Writing – original draft. **Radu Serban:** Supervision, Methodology, Writing – original draft, Writing – review & editing, Conceptualization, Funding acquisition, Software. **Dan Negrut:** Funding acquisition, Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Data availability

The code and data used for the research is openly available on GitHub, as indicated in the Code Availability section.

### Acknowledgements

Open Access Funding provided by Università degli Studi di Salerno within the CRUI-CARE Read&Publish agreement. B. Tagliaferro gratefully acknowledges financial support for this publication by the Fulbright Schuman Program, which is administered by the Fulbright Commission in Brussels and jointly financed by the U.S. Department of State, and the Directorate-General for Education, Youth, Sport and Culture (DG.EAC) of the European Commission. The content of this manuscript does not represent the official views of the Fulbright Program, the Government of the United States, or the Fulbright Commission in Brussels. This work has been partially supported by NSF project OAC2209791 and the US Army Research Office project W911NF1910431.

### References

- [1] P.A. Cundall, O.D. Strack, A discrete numerical model for granular assemblies, *Geotechnique* 29 (1) (1979) 47–65.
- [2] T. Pöschel, T. Schwager, *Computational Granular Dynamics: Models and Algorithms*, Springer, Berlin, Heidelberg, 2005.
- [3] M. Lemieux, G. Léonard, J. Doucet, L.-A. Leclaire, F. Viens, J. Chaouki, F. Bertrand, Large-scale numerical investigation of solids mixing in a v-blender using the discrete element method, *Powder Technol.* 181 (2) (2008) 205–216.
- [4] K. Apostolou, A. Hrymak, Discrete element simulation of liquid-particle flows, *Comput. Chem. Eng.* 32 (4–5) (2008) 841–856.
- [5] C.-L. Tang, J.-C. Hu, M.-L. Lin, J. Angelier, C.-Y. Lu, Y.-C. Chan, H.-T. Chu, The Tsaoiling landslide triggered by the Chi-Chi earthquake, Taiwan: insights from a discrete element simulation, *Eng. Geol.* 106 (1–2) (2009) 1–19.
- [6] D. Salciarini, C. Tamagnini, P. Conversini, Discrete element modeling of debris-avalanche impact on earthfill barriers, *Phys. Chem. Earth Parts A/B/C* 35 (3–5) (2010) 172–181.
- [7] C. O'Sullivan, Particle-based discrete element modeling: geomechanics perspective, *Int. J. Geomech.* 11 (6) (2011) 449–464.
- [8] P. Sánchez, D.J. Scheeres, Simulating asteroid rubble piles with a self-gravitating soft-sphere distinct element method model, *Astrophys. J.* 727 (2) (2011) 120.
- [9] F.F. Foldager, L.J. Munkholm, O. Balling, R. Serban, D. Negrut, R.J. Heck, O. Green, Modeling soil aggregate fracture using the discrete element method, *Soil Tillage Res.* 218 (2022) 105295.
- [10] A.M. Recuero, R. Serban, B. Peterson, H. Sugiyama, P. Jayakumar, D. Negrut, A high-fidelity approach for vehicle mobility simulation: nonlinear finite element tires operating on granular material, *J. Terramech.* 72 (2017) 39–54, <https://doi.org/10.1016/j.jterra.2017.04.002>.
- [11] J.B. Johnson, A.V. Kulchitsky, P. Duvo, K. Iagnemma, C. Senatore, R.E. Arvidson, J. Moore, Discrete element method simulations of Mars exploration rover wheel performance, *J. Terramech.* 62 (2015) 31–40.
- [12] OpenMP: Specification Standard 5.2. Available online at, <http://openmp.org/>, 2021.
- [13] A. Amritkar, S. Deb, D. Tafti, Efficient parallel CFD-DEM simulations using OpenMP, *J. Comput. Phys.* 256 (2014) 501–519.
- [14] M.A. Knuth, J. Johnson, M. Hopkins, R. Sullivan, J. Moore, Discrete element modeling of a Mars exploration rover wheel in granular material, *J. Terramech.* 49 (1) (2012) 27–36.
- [15] Message Passing Interface Forum: MPI: A Message-Passing Interface Standard Version 3.0. Chapter author for Collective Communication, Process Topologies, and One Sided Communications (2012).
- [16] B. Yan, R.A. Regueiro, A comprehensive study of MPI parallelism in three-dimensional discrete element method (dem) simulation of complex-shaped granular particles, *Comput. Part. Mech.* 5 (4) (2018) 553–577.
- [17] A.W.M. Checkaraou, A. Rousset, X. Besseron, S. Varrette, B. Peters, Hybrid MPI+openMP implementation of extended discrete element method, in: 2018 30th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD), IEEE, 2018, pp. 450–457.
- [18] LIGGGHTS, Open source discrete element method particle simulation code, <http://cfdem.dcs.computing.com/?q=OpenSourceDEM>, 2013.
- [19] LAMMPS, A molecular dynamics simulator, <http://lammps.sandia.gov/>, 2013.
- [20] Simcenter STAR-CCM+ software website, <https://plm.sw.siemens.com/en-US/simcenter/fluids-thermal-simulation/star-ccm/>, 2023. (Accessed 25 September 2023).
- [21] R. Serban, N. Olsen, D. Negrut, High performance computing framework for co-simulation of vehicle-terrain interaction, in: NDIA Ground Vehicle Systems Engineering and Technology Symposium, 2017.
- [22] J. Xu, H. Qi, X. Fang, L. Lu, W. Ge, X. Wang, M. Xu, F. Chen, X. He, J. Li, Quasi-real-time simulation of rotating drum using discrete element method with parallel GPU computing, *Particuology* 9 (4) (2011) 446–450.
- [23] N. Govender, D. Wilke, S. Kok, Blaze-DEMGPU: modular high performance DEM framework for the GPU architecture, *SoftwareX* 5 (2016) 62–66.
- [24] J. Gan, Z. Zhou, A. Yu, A GPU-based DEM approach for modeling of particulate systems, *Powder Technol.* 301 (2016) 1172–1182.
- [25] Y. He, T. Evans, A. Yu, R. Yang, A GPU-based DEM for modeling large scale powder compaction with wide size distributions, *Powder Technol.* 333 (2018) 219–228.
- [26] C. Kelly, N. Olsen, C. Vanden Heuvel, R. Serban, D. Negrut, Towards the democratization of many-body dynamics: billion degree of freedom simulation of granular material on commodity hardware, in: Proceeding of the ECCOMAS Multibody Dynamics Conference, Duisburg, Germany, 2019.
- [27] K. Iwashita, M. Oda, Rolling resistance at contacts in simulation of shear band development by DEM, *J. Eng. Mech.* 124 (3) (1998) 285–292.
- [28] A.D. Renzo, F.P.D. Maio, Comparison of contact-force models for the simulation of collisions in DEM-based granular flow codes, *Chem. Eng. Sci.* 59 (3) (2004) 525–541.
- [29] F. Cruz, S. Emam, M. Prochnow, J.N. Roux, F. Chevoir, Rheophysics of dense granular materials: discrete simulation of plane shear flows, *Phys. Rev. E* 72 (2005) 021309, <https://doi.org/10.1103/PhysRevE.72.021309>.
- [30] C.H. Rycroft, G.S. Grest, J.W. Landry, M.Z. Bazant, Analysis of granular flow in a pebble-bed nuclear reactor, *Phys. Rev. E* 74 (2006) 021306.



- [31] H. Kruggel-Emden, M. Sturm, S. Wirtz, V. Scherer, Selection of an appropriate time integration scheme for the discrete element method (DEM), *Comput. Chem. Eng.* 32 (10) (2008) 2263–2279.
- [32] T.M. Wasfy, H.M. Wasfy, J.M. Peters, Coupled multibody dynamics and discrete element modeling of vehicle mobility on cohesive granular terrains, in: ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, 2014, V006T10A050 <http://proceedings.asmedigitalcollection.asme.org/proceeding.aspx?articleid=2091049>.
- [33] S. Lommen, DEM speedup: stiffness effects on behavior of bulk material, *Particuology* (2014) 107–112.
- [34] S. Utili, T. Zhao, G.T. Houlsby, 3D DEM investigation of granular column collapse: evaluation of debris motion and its destructive power, *Eng. Geol.* 186 (2015) 3–16.
- [35] M. Potticary, A. Zervos, J. Harkness, An investigation into the effect of particle playness on the strength of granular material using the discrete element method, in: IV International Conference on Particle-based Methods - Fundamentals and Applications, 2015, <https://eprints.soton.ac.uk/394117/1/particles2015.pdf>.
- [36] M. Michael, F. Vogel, B. Peters, DEM-FEM coupling simulations of the interactions between a tire tread and granular terrain, *Comput. Methods Appl. Mech. Eng.* (2015).
- [37] M. Ciantia, M. Arroyo, J. Butlanska, A. Gens, DEM modelling of cone penetration tests in a double-porosity crushable granular material, *Comput. Geotech.* 73 (2016) 109–127.
- [38] Z. Zheng, M. Zang, Numerical simulations of the interactions between a pneumatic tire and granular sand by 3D DEM-FEM, in: 7th International Conference on Discrete Element Methods, 2017, pp. 289–300, [https://link.springer.com/chapter/10.1007/978-981-10-1926-5\\_32](https://link.springer.com/chapter/10.1007/978-981-10-1926-5_32).
- [39] E. Parteli, T. Poschel, Particle-based simulation of powder application in additive manufacturing, *Powder Technol.* (2016) 96–102.
- [40] R. Kivugo, Tire-Soil Interaction for Off-Road Vehicle Applications, Phd, Politecnico di Milano, 2017, <https://www.politesi.polimi.it/handle/10589/136229>.
- [41] F. Calvetti, C. Prisco, E. Vairaktaris, DEM assessment of impact forces of dry granular masses on rigid barriers, *Acta Geotech.* (2016).
- [42] T. Wen, X. Zeng, Accelerating Polyhedral Discrete Element Method with CUDA, *Astron. J.* 166 (2023), <https://doi.org/10.3847/1538-3881/acfc46>.
- [43] G. Liu, W. Xu, A GPU-based DEM framework for simulation of polyhedral particulate system, *Granul. Matter* 25 (2023), <https://doi.org/10.1007/s10035-023-01321-2>.
- [44] M. Furuichi, D. Nishiura, O. Kuwano, A. Bauville, T. Hori, H. Sakaguchi, Arcuate stress state in accretionary prisms from real-scale numerical sandbox experiments, *Nat. Sci. Rep.* 8 (2018), [www.nature.com/scientificreports/](http://www.nature.com/scientificreports/).
- [45] O. Henrich, Y.A. Gutierrez Fosado, T. Curk, T. Ouldrige, Coarse-grained simulation of DNA using LAMMPS, 2018.
- [46] C.S. Dias, Molecular dynamics simulations of active matter using LAMMPS, *arXiv:2102.10399 [cond-mat.soft]*, 2021.
- [47] R. Li, Z. Liu, Z. Feng, J. Liang, L.-G. Zhang, High-fidelity MC-DEM modeling and uncertainty analysis of HTR-PM first criticality, *Front. Energy Res.* 9 (2022), <https://doi.org/10.3389/fenrg.2021.822780>.
- [48] F. Razavi, A. Komrakova, C.F. Lange, CFD—DEM simulation of sand-retention mechanisms in slurry flow, *Energies* 14 (13) (2021), <https://doi.org/10.3390/en14133797>.
- [49] L. Fang, R. Zhang, C. Vanden Heuvel, R. Serban, D. Negrut, Chrono::GPU: an open-source simulation package for granular dynamics using the discrete element method, *Processes* 9 (10) (2021), <https://doi.org/10.3390/pr9101813>.
- [50] A. Tasora, R. Serban, H. Mazhar, A. Pazouki, D. Melanz, J. Fleischmann, M. Taylor, H. Sugiyama, D. Negrut, Chrono: an open source multi-physics dynamics engine, in: T. Kozubek (Ed.), *High Performance Computing in Science and Engineering*, in: *Lecture Notes in Computer Science*, Springer, 2016, pp. 19–49.
- [51] D. Reger, E. Merzari, P. Balestra, R. Stewart, G. Strydom, Discrete element simulation of pebble bed reactors on graphics processing units, *Ann. Nucl. Energy* 190 (2023) 109896, <https://doi.org/10.1016/j.anucene.2023.109896>.
- [52] M. Hauastein, A. Gladky, R. Schwarze, Discrete element modeling of deformable particles in YADE, *SoftwareX* 6 (2017) 118–123, <https://doi.org/10.1016/j.softx.2017.05.001>.
- [53] D. Romanova, S. Strijhak, M. Kraposhin, Development of snowYadeFoam solver for snow particles simulation, in: 2020 Ivannikov Ispras Open Conference (ISPRAS), 2020, pp. 166–169.
- [54] C. Ericson, Real Time Collision Detection, Morgan Kaufmann, San Francisco, CA, 2005.
- [55] J. Favier, M. Abbaspour-Fard, M. Kremmer, A. Raji, Shape representation of axisymmetrical, non-spherical particles in discrete element simulation using multi-element model particles, *Eng. Comput.* (1999).
- [56] J. Hilton, P. Cleary, The influence of particle shape on flow modes in pneumatic conveying, *Chem. Eng. Sci.* 66 (3) (2011) 231–240.
- [57] K. Kiangi, A. Potapov, M. Moys, DEM validation of media shape effects on the load behaviour and power in a dry pilot mill, *Miner. Eng.* 46 (2013) 52–59.
- [58] B. Ren, W. Zhong, B. Jin, Y. Shao, Z. Yuan, Numerical simulation on the mixing behavior of corn-shaped particles in a spouted bed, *Powder Technol.* 234 (2013) 58–66.
- [59] W. Zhong, A. Yu, X. Liu, Z. Tong, H. Zhang, DEM/CFD-DEM modelling of non-spherical particulate systems: theoretical developments and applications, *Powder Technol.* 302 (2016) 108–152.
- [60] R. Kawamoto, E. Andò, G. Viggiani, J.E. Andrade, All you need is shape: predicting shear banding in sand with LS-DEM, *J. Mech. Phys. Solids* 111 (2018) 375–392.
- [61] E. Marteau, J.E. Andrade, An experimental study of the effect of particle shape on force transmission and mobilized strength of granular materials, *J. Appl. Mech.* 88 (11) (2021).
- [62] S. Zhao, J. Zhao, SudoDEM: unleashing the predictive power of the discrete element method on simulation for non-spherical granular particles, *Comput. Phys. Commun.* 259 (2023), <https://doi.org/10.1016/j.cpc.2020.107670>.
- [63] R. Zhang, C. Vanden Heuvel, D. Negrut, DEM-Engine, a multi-GPU DEM solver with complex geometry support, <https://github.com/projectchrono/DEM-Engine>, 2022, Simulation-Based Engineering Laboratory, University of Wisconsin-Madison.
- [64] H. Mazhar, T. Heyn, D. Negrut, A scalable parallel method for large collision detection problems, *Multibody Syst. Dyn.* 26 (37–55) (2011), <https://doi.org/10.1007/s11044-011-9246-y>.
- [65] L. Verlet, Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules, *Phys. Rev.* 159 (98–103) (1967), <https://doi.org/10.1103/PhysRev.159.98>.
- [66] Q. Zhou, W. Xu, Y. Chen, Multi-GPUs DEM algorithm and its application in the simulation of granular materials, *Powder Technol.* 430 (118969) (2023), <https://doi.org/10.1016/j.powtec.2023.118969>.
- [67] Rocky DEM, <https://rocky.esss.co/>. (Accessed 29 December 2021).
- [68] Q. Liu, W. Wang, H. Ma, Parallelized combined finite-discrete element (FDEM) procedure using multi-GPU with CUDA, *Int. J. Numer. Anal. Methods Geomech.* 44 (2) (2023) 208–238, <https://doi.org/10.1002/nag.3011>.
- [69] B. Barsdell, K. Clark, A single-header C++ library for simplifying the use of CUDA runtime compilation, <https://github.com/NVIDIA/jitify>. (Accessed 24 August 2023).
- [70] N. Berry, Y. Zhang, S. Haeri, Contact models for the multi-sphere discrete element method, *Powder Technol.* 416 (2023) 118209, <https://doi.org/10.1016/j.powtec.2022.118209>.
- [71] C.J. Coetzee, O.C. Scheffler, Review: the calibration of dem parameters for the bulk modelling of cohesive materials, *Processes* 11 (1) (2023), <https://doi.org/10.3390/pr11010005>.
- [72] M. Price, V. Murariu, G. Morrison, Sphere clump generation and trajectory comparison for real particles, in: *Proceedings of Discrete Element Modelling*, vol. 2007, 2007.
- [73] H. Hertz, Ueber die Verdunstung der Flüssigkeiten, insbesondere des Quecksilbers, im luftleeren Raume, *Ann. Phys.* 253 (10) (1882) 177–193, <https://doi.org/10.1002/andp.18822531002>, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.18822531002>.
- [74] R. Mindlin, H. Deresiewicz, Elastic spheres in contact under varying oblique forces, *J. Appl. Mech.* 20 (1953) 327–344.
- [75] L. Fang, D. Negrut, Producing 3D friction loads by tracking the motion of the contact point on bodies in mutual contact, *Comput. Part. Mech.* 8 (2021) 905–929, <https://doi.org/10.1007/s40571-020-00376-9>.
- [76] J. Fleischmann, R. Serban, D. Negrut, P. Jayakumar, On the importance of displacement history in soft-body contact models, *J. Comput. Nonlinear Dyn.* 11 (4) (2016) 044502.
- [77] K.L. Johnson, *Contact Mechanics*, Cambridge University Press, 1987.
- [78] C. Goldenberg, I. Goldhirsch, Friction enhances elasticity in granular solids, *Nature* 435 (7039) (2005) 188–191.
- [79] C. Goldenberg, I. Goldhirsch, Force chains, microelasticity, and macroelasticity, *Phys. Rev. Lett.* 89 (8) (2002) 084302.
- [80] P. Thornton, C.W. Randall, *Applications of Theoretical Contact Mechanics to Solid Particle System Simulation*, Elsevier, Cambridge, MA, USA, 1988.
- [81] Y. Li, Y. Xu, C. Thornton, A comparison of discrete element simulations and experiments for ‘sandpiles’ composed of spherical particles, *Powder Technol.* 160 (219–228) (2005), <https://doi.org/10.1016/j.powtec.2005.09.002>.
- [82] T. Qu, Y.T. Feng, T.Z. Min Wang, Calibration of linear contact stiffnesses in discrete element models using a hybrid analytical-computational framework, *Powder Technol.* 356 (795–807) (2019), <https://doi.org/10.1016/j.powtec.2019.09.016>.
- [83] M.A. Ambroso, C.R. Santore, A.R. Abate, D.J. Durian, Penetration depth for shallow impact cratering, *Phys. Rev. E* 71 (2005) 051305, <https://doi.org/10.1103/PhysRevE.71.051305>.
- [84] T. Heyn, On the modeling, simulation, and visualization of many-body dynamics problems with friction and contact, PhD thesis, Department of Mechanical Engineering, University of Wisconsin–Madison, 2013, [http://sbel.wisc.edu/documents/TobyHeynThesis\\_PhDfinal.pdf](http://sbel.wisc.edu/documents/TobyHeynThesis_PhDfinal.pdf).
- [85] X. Cui, J. Dai, H. Xu, X. Gao, Superdem simulation and experiment validation of non-spherical particles flows in a rotating drum, *Ind. Eng. Chem. Res.* 62 (16) (2023) 6525–6535, <https://doi.org/10.1021/acs.iecr.3c00919>.
- [86] X. Gao, J. Yu, R.J.F. Portal, J.-F. Dietiker, M. Shahnam, W.A. Rogers, Development and validation of superdem for non-spherical particulate systems using a superquadric particle method, *Particuology* 61 (2022) 74–90, <https://doi.org/10.1016/j.partic.2020.11.007>.
- [87] B. Jian, X. Gao, Investigation of spherical and non-spherical binary particles flow characteristics in a discharge hopper, *Adv. Powder Technol.* 34 (5) (2023) 104011, <https://doi.org/10.1016/j.apt.2023.104011>.
- [88] Y. Guo, J.S. Curtis, Discrete element method simulations for complex granular flows, *Annu. Rev. Fluid Mech.* 47 (1) (2015) 21–46, <https://doi.org/10.1146/annurev-fluid-010814-014644>.



- [89] L. Scholtès, F.-V. Donzé, A dem model for soft and hard rocks: role of grain interlocking on strength, *J. Mech. Phys. Solids* 61 (2) (2013) 352–369, <https://doi.org/10.1016/j.jmps.2012.10.005>.
- [90] N. Belheine, J.-P. Plassiard, F.-V. Donzé, F. Darve, A. Seridi, Numerical simulation of drained triaxial test using 3d discrete element modeling, *Comput. Geotech.* 36 (1) (2009) 320–331, <https://doi.org/10.1016/j.compgeo.2008.02.003>.
- [91] G.-Y. Liu, W.-J. Xu, Q.-C. Sun, N. Govender, Study on the particle breakage of ballast based on a GPU accelerated discrete element method, *Geosci. Front.* 11 (2) (2020) 461–471, <https://doi.org/10.1016/j.gsf.2019.06.006>.
- [92] D.O. Potyondy, P.A. Cundall, A bonded-particle model for rock, *Int. J. Rock Mech. Min. Sci.* 41 (8 Spec.Iss.) (2004) 1329–1364, <https://doi.org/10.1016/j.ijrmmms.2004.09.011>.
- [93] Y. Wang, F. Tonon, Modeling Lac du Bonnet granite using a discrete element model, *Int. J. Rock Mech. Min. Sci.* 46 (7) (2009) 1124–1135, <https://doi.org/10.1016/j.ijrmmms.2009.05.008>.
- [94] R. Zhang, C.V. Heuvel, A. Schepelmann, A. Rogg, D. Apostolopoulos, S. Chandler, R. Serban, D. Negrut, A GPU-accelerated Simulator for the DEM Analysis of Granular Systems Composed of Clump-shaped Elements.
- [95] Simulated Lunar Operations Laboratory: NASA's VIPER Prototype Motors Through Moon-like Obstacle Course, <https://www.nasa.gov/feature/ames/nasas-viper-prototype-motors-through-moon-like-obstacle-course>. (Accessed 2 April 2023).
- [96] H.A. Oravec, X. Zeng, V.M. Asnani, Design and characterization of GRC-1: a soil for lunar terramechanics testing in Earth-ambient conditions, *J. Terramech.* 47 (6) (2010) 361–377, <https://doi.org/10.1016/j.jterra.2010.04.006>.
- [97] OpenAI (2023), ChatGPT (Sep 25 version), <https://chat.openai.com>.