



Title	Development of a 3D Hybrid Finite-Discrete Element Simulator Based on GPGPU-Parallelized Computation for Modelling Rock Fracturing Under Quasi-Static and Dynamic Loading Conditions
Author(s)	Fukuda, Daisuke; Mohammadnejad, Mojtaba; Liu, Hongyuan; Zhang, Qianbing; Zhao, Jian; Dehhoda, Sevda; Chan, Andrew; Kodama, Jun-ichi; Fujii, Yoshiaki
Citation	Rock Mechanics and Rock Engineering, 53(3), 1079-1112 https://doi.org/10.1007/s00603-019-01960-z
Issue Date	2020-03
Doc URL	http://hdl.handle.net/2115/80506
Rights	This is a post-peer-review, pre-copyedit version of an article published in Rock Mechanics and Rock Engineering. The final authenticated version is available online at: http://dx.doi.org/10.1007/s00603-019-01960-z
Type	article (author version)
File Information	RMRE_Text - 2nd Revision - Final.pdf



[Instructions for use](#)

1 **Development of a 3D Hybrid Finite-Discrete Element Simulator based on GPGPU-**
2 **Parallelized Computation for Modelling Rock Fracturing under Quasi-Static and**
3 **Dynamic Loading Conditions**

4
5 D. Fukuda ^{a, b}, M. Mohammadnejad ^{b, c}, H.Y. Liu ^{b, *}, Q.B. Zhang ^d, J. Zhao ^d, S. Dehkhoda ^{b, c},
6 A. Chan ^b, J. Kodama ^a and Y. Fujii ^a

7 ^a Faculty of Engineering, Hokkaido University, Hokkaido 060-8628, Japan

8 ^b College of Sciences and Engineering, University of Tasmania, TAS 7001, Australia

9 ^c CSIRO Minerals Resources Business Unit, Queensland Centre for Advanced Technologies,
10 Brisbane, QLD 4069, Australia

11 ^d Department of Civil Engineering, Monash University, VIC3800, Australia

12

13

14 **Abstract**

15 As a state-of-the-art computational method for simulating rock fracturing and fragmentation,
16 the combined finite-discrete element method (FDEM) has become widely accepted since
17 Munzija (2004) published his comprehensive book of FDEM. This study developed a general-
18 purpose graphic-processing-unit (GPGPU)-parallelized FDEM using the compute unified
19 device architecture (CUDA) C/C++ based on the authors' former sequential two-dimensional
20 (2D) and three-dimensional (3D) Y-HFDEM IDE (integrated development environment) code.
21 The theory and algorithm of the GPGPU-parallelized 3D Y-HFDEM IDE code are first
22 introduced by focusing on the implementation of the contact detection algorithm, which is
23 different from that in the sequential code, contact damping and contact friction. 3D modelling
24 of the failure process of limestone under quasi-static loading conditions in uniaxial
25 compressive strength (UCS) tests and Brazilian tensile strength (BTS) tests are then conducted
26 using the GPGPU-parallelized 3D Y-HFDEM IDE code. The 3D FDEM modelling results
27 show that mixed-mode I-II failures are the dominant failure mechanisms along the shear and
28 splitting failure planes in the UCS and BTS models, respectively, with unstructured meshes.
29 Pure mode I splitting failure planes and pure mode II shear failure planes are only possible in
30 the UCS and BTS models, respectively, with structured meshes. Subsequently, 3D modelling
31 of the dynamic fracturing of marble in dynamic Brazilian tests with a split Hopkinson pressure

* Corresponding author: Dr H.Y. Liu (Email: Hong.Liu@utas.edu.au and Tel: 0061 6226 2113): If you are interested in using our GPGPU-parallelized hybrid finite-discrete element code "3D Y-HFDEM IDE" for your research or verification, please send an email to the corresponding author. However, we do not allow the use of the code for military and commercial purposes.

32 bar (SHPB) apparatus is conducted using the GPGPU-parallelized 3D HFDEM IDE code
33 considering the entire SHPB testing system. The modelled failure process, final fracture pattern
34 and time histories of the dynamic compressive wave, reflective tensile wave and transmitted
35 compressive wave are compared quantitatively and qualitatively with those from experiments,
36 and good agreements are achieved between them. The computing performance analysis shows
37 the GPGPU-parallelized 3D HFDEM IDE code is 284 times faster than its sequential version
38 and can achieve the computational complexity of $O(N)$. The results demonstrate that the
39 GPGPU-parallelized 3D Y-HFDEM IDE code is a valuable and powerful numerical tool for
40 investigating rock fracturing under quasi-static and dynamic loading conditions in rock
41 engineering applications although very fine elements with maximum element size no bigger
42 than the length of the fracture process zone must be used in the area where fracturing process
43 is modelled.

44

45 **Keywords:** Rocks, 3D fracture process analysis, FDEM, Quasi-static loading, Dynamic
46 loading, Parallel computation, GPGPU, CUDA C/C++

47

48 **List of acronyms used in the paper**

49 BEM: boundary element method

50 BTS: Brazilian tensile strength

51 CE6: 6-node initially zero-thickness cohesive element

52 CPU: central-processing-unit

53 CUDA: compute unified device architecture

54 CZM: cohesive zone model

55 DEM: distinct element method

56 DDA: discontinuous deformation analysis

57 DFPA: dynamic fracture process analysis

58 ECDA: efficient contact detection activation

59 ECZM: extrinsic cohesive zone model

60 FDEM: combined finite-discrete element method

61 FDM: finite difference method

62 FEM: finite element method

63 FPZ: fracture process zone

64 GPGPU: general-purpose graphic-processing-unit

65 HPC: high-performance-computing

66 ICZM: intrinsic cohesive zone model
67 IB: incident bar
68 ISRM: International Society for Rock Mechanics
69 MPI: message-passing interface
70 1D: one-dimensional
71 OpenCL: open computing language
72 OpenMP: open multi-processing
73 PC: personal computer
74 SBFEM: scaled boundary finite element method
75 SHPB: split Hopkinson pressure bar
76 SPH: smoothed particle hydrodynamics
77 TET4: 4–node tetrahedral finite element
78 TB: transmission bar
79 3D: three-dimensional
80 TtoP: TET4 to point - contact interaction between the contactor 4–node tetrahedral finite
81 element and the target point
82 TtoT: TET4 to TET4 - contact interaction between the contactor 4–node tetrahedral finite
83 element and the target 4–node tetrahedral finite element
84 2D: two-dimensional
85 UCS: uniaxial compressive strength
86

87 **1. Introduction**

88 Understanding the mechanism of the fracturing process in rocks is important in civil and
89 mining engineering and several other fields, such as geothermal, hydraulic, and oil and gas
90 engineering, in which rock fractures play an important role. Numerical methods have been
91 increasingly applied to analyze the fracturing process of rocks (e.g., Mohammadnejad et al.
92 2018). Due to the limitations of computing power/environments and the difficulty in extending
93 some numerical techniques to three-dimension, numerous previous studies of rock fracture
94 modelling using various numerical methods have been limited to two-dimensional (2D)
95 analyses. However, successful simulation of the three-dimensional (3D) fracturing process is
96 essential for better understanding and solving many practical rock engineering problems
97 because rock fracturing essentially involves complex 3D processes. Moreover, meaningful
98 developments and applications of 3D numerical methods for 3D rock fracture modelling have

99 been limited in the past, except those that had full access to the most advanced high-
100 performance-computing (HPC) environments, such as supercomputers. This situation has been
101 dramatically improved due to the recent advancements in computer technology, such as
102 general-purpose graphic-processing-unit (GPGPU) accelerators and the many integrated cores
103 architecture, which can be installed even in a personal computer (PC) or a workstation and can
104 achieve HPC environments in ordinary computer environments with relatively low costs. These
105 improvements have allowed scientists and engineers to apply and develop 3D rock fracture
106 computational methods that were not tractable previously.

107 Recent advances in computational mechanics have resulted in modelling complex rock
108 fracturing processes using various numerical approaches. Generally, these approaches can be
109 classified mainly into continuous and discontinuous methods. In the framework of rock fracture
110 process analysis, the continuum-based methods include the finite element method (FEM), finite
111 difference method (FDM), boundary element method (BEM), scaled boundary finite element
112 method (SBFEM), extended finite element method (XFEM), several mesh-less/mesh-free
113 methods, such as smoothed particle hydrodynamics (SPH), and those based on peridynamics
114 and phase-field method. The discontinuum-based methods include the distinct element method
115 (DEM), discontinuous deformation analysis (DDA), lattice model and molecular dynamics.
116 Comprehensive reviews of recent advances in computational fracture mechanics of rocks can
117 be found in recent review articles (Lisjak and Grasselli 2014; Mohammadnejad et al. 2018).
118 For a realistic simulation of the rock fracturing process, numerical techniques must be capable
119 of capturing the transition of rock from a continuum to a discontinuum through crack initiation,
120 growth and coalescence. Therefore, increasing attention has been paid in recent years to these
121 techniques, which can combine/couple the advantages of the aforementioned continuum-based
122 and discontinuum-based methods while overcoming the disadvantages of each method (Lisjak
123 and Grasselli 2014; Mohammadnejad et al. 2018).

124 The combined finite-discrete element method (FDEM) proposed by Munjiza (1995) has
125 recently attracted the attention of many engineers and researchers in the field of rock
126 engineering which deals with rock fracturing and fragmentation. The method incorporates the
127 advantages of both continuous and discontinuous methods and is able to simulate the transition
128 from a continuum to a discontinuum caused by rock fracturing. The two main implementations
129 of the FDEM include the open-source research code, Y code (Munjiza 2004), and the
130 commercial code, ELFEN (Rockfield 2005). Due to the open-source nature of the Y code,
131 several attempts have been made to actively extend it, such as Y-Geo (Mahabadi et al. 2014),
132 Y-Flow (Yan and Jiao 2018; Yan and Zheng 2017), Irazu (Lisjak et al. 2018; Mahabadi et al.

133 2016), Solidity (Guo 2014; Solidity 2017) and HOSS with MUNROU (Rougier et al. 2014;
134 Rougier et al. 2011). Moreover, the authors have developed Y-HFDEM IDE (An et al. 2017;
135 Liu et al. 2016; Liu et al. 2015; Mohammadnejad et al. 2017). In addition, using user-defined
136 subroutines in the explicit module of the commercial software ABAQUS, Ma et al. (2018)
137 recently implemented the FDEM to investigate the effects of different fracture mechanisms on
138 impact fragmentation of brittle rock-like materials. The principles of all of the FDEM codes
139 are based on continuum mechanics, the cohesive zone model (CZM) and contact mechanics,
140 which make the FDEM extremely computationally expensive. Thus, few practical rock
141 engineering problems can be solved using the 3D FDEM based on sequential central-
142 processing-unit (CPU)-based implementations. Therefore, it is imperative to develop robust
143 parallel computation schemes to handle large-scale 3D FDEM simulations with massive
144 numbers of nodes, elements and contact interactions.

145 To date, several successful parallel implementations of FDEM codes have been reported
146 using the message-passing interface (MPI) (Elmo and Stead 2010; Hamdi et al. 2014; Lei et al.
147 2014; Lukas et al. 2014; Rockfield 2005; Rogers et al. 2015; Rougier et al. 2014) and shared
148 memory programming, such as OpenMP (Xiang et al. 2016). Among these, Lukas et al. (2014)
149 proposed a novel approach for the parallelization of 2D FDEM using MPI and dynamic domain
150 decomposition-based parallelization solvers, and they successfully applied the parallelized Y
151 code to a large-scale 2D problem on a PC cluster, which should be able to be applied to practical
152 3D problems, although future developments are needed. In addition, Lei et al. (2014)
153 successfully developed the concept of the virtual parallel machine for the FDEM using MPI,
154 which can be adapted to different computer architectures ranging from several to thousands of
155 CPU cores. Rougier et al. (2014) introduced the HOSS with MUNROU code, in which they
156 used 208 processors for parallel computation controlled by MPI and developed novel contact
157 detection and contact force calculation algorithms (Munjiza et al. 2011). The MUNROU code
158 was then successfully applied to perform 3D simulations of a dynamic Brazilian tensile test of
159 rock with a SHPB apparatus. ELFEN (Elmo and Stead 2010; Hamdi et al. 2014; Rockfield
160 2005; Rogers et al. 2015) uses MPI in its parallelization scheme and has been employed
161 successfully in 2D and 3D simulations of rock fracturing process. For example, analyses of the
162 3D fracturing process in conventional laboratory tests using up to 3 million elements have been
163 reported (Hamdi et al. 2014). Xiang et al. (2016) optimized the contact detection algorithm in
164 their Solidity code and parallelized the code using OpenMP. Although they modelled a packing
165 system with 288 rock-like boulders and achieved a speedup of 9 times on 12 CPU threads, the
166 details of the applied algorithm and its implementation were not given. In general, MPI requires

167 large and expensive CPU clusters to achieve the best performance. In addition, the application
168 of shared memory programming such as OpenMP is limited by the total number of
169 multiprocessors that can reside in a single computer; thus, MPI is still required for large-scale
170 problems, in which each computer uses both OpenMP and MPI to transfer the data between
171 multiple computers. This means that the hybrid MPI/OpenMP is necessary. In all of these
172 approaches, more than 100 CPUs are necessary to achieve a speed-up of more than 100 times
173 compared with sequential CPU-based FDEM simulations, which results in the need for a larger
174 space or expensive HPC environments.

175 In addition to the CPU-based parallelization schemes, a GPGPU accelerator controlled by
176 either the Open Computing Language (OpenCL) (Munshi et al. 2011) or the Compute Unified
177 Device Architecture (CUDA) (NVIDIA 2018) can be considered as another promising method
178 for the parallelization of FDEM codes. Hundreds and thousands of GPU-core processors can
179 reside and concurrently work in a small GPGPU accelerator within an ordinary laptop/desktop
180 PC or a workstation, which also has lower energy consumption than CPU-based clusters. Zhang
181 et al. (2013) developed a CUDA-based GPGPU parallel version of the Y code (2D) without
182 considering the fracturing process and contact friction. Batinić et al. (2018) implemented a
183 GPGPU-based parallel FEM/DEM that is based on the Y code to analyze cable structures using
184 CUDA. However, none of these implementations have been employed to simulate rock
185 fracturing. In this regard, a GPGPU-based FDEM commercial code, namely Irazu (Lisjak et al.
186 2017; Lisjak et al. 2018), has just been developed with OpenCL and used successfully in rock
187 fracture simulations. Irazu is the only available commercial GPGPU-based FDEM code with
188 OpenCL that is currently capable of modelling the rock fracturing process (Lisjak et al. 2017;
189 Lisjak et al. 2018). In addition, the authors have developed a free FDEM research code, Y-
190 HFDEM IDE (An et al. 2017; Liu et al. 2016; Liu et al. 2015; Mohammadnejad et al. 2017),
191 and parallelized its 2D implementation using GPGPU with CUDA C/C++ (Fukuda et al.,
192 2019). This paper focuses on parallelizing the 3D implementation of the Y-HFDEM IDE code
193 using GPGPU with CUDA C/C++, which is completely different from Irazu's GPGPU
194 parallelization using OpenCL. Additional studies are required to verify and validate the
195 GPGPU-based 3D FDEM code. Furthermore, for any newly implemented GPGPU-based
196 codes, it is desirable to describe their complete details because the implementation of any
197 GPGPU-based code is generally different from that of CPU-based sequential codes. Most
198 importantly, there are no freely available GPGPU-based FDEM codes, whereas the GPGPU-
199 based Y-HFDEM IDE is free to use, and the freely available GPGPU-parallelized 2D/3D Y-

200 HFDEM IDE software may significantly contribute to researches in the field of rock
201 engineering.

202 To validate and calibrate newly developed codes in the field of rock mechanics, two
203 standard rock mechanics laboratory tests, the uniaxial compressive strength (UCS) tests and
204 Brazilian tensile strength (BTS) tests, have often been modelled to simulate the fracturing
205 process and associated failure mechanisms of rock materials under quasi-static loading
206 conditions. Although the UCS and BTS tests have been actively modelled using 2D FDEM,
207 their modelling in the framework of 3D FDEM has been very limited and less well explained.
208 For example, UCS and BTS tests were three-dimensionally simulated using Y-Geo (Mahabadi
209 et al. 2014) with a relatively large element size (2 mm for both tests) and a very high loading
210 velocity (1 m/s) resulting in the appearance of dynamic effects in the model, such as multi-
211 fracture propagation around the center of the BTS models, which is also shown in the dynamic
212 BTS simulation discussed in Section 3.3. Moreover, although Mahabadi et al. (2014) observed
213 splitting fractures in their BTS modelling, the post-peak behavior of the stress-strain curve was
214 not well demonstrated. Later, Lisjak et al. (2018) introduced IRAZU based on GPGPU-
215 parallelization and three-dimensionally simulated UCS and BTS tests using a finer element
216 size (1.5 mm for both tests) and slower loading rate (0.1 m/s). In comparison with the
217 aforementioned simulations by Mahabadi et al. (2014), many reasonable results were achieved
218 by Lisjak et al. (2018). Mahabadi et al. (2014) may have chosen the relatively high loading rate
219 because 3D FDEM modelling using Y-Geo is computationally demanding, which may have
220 required them to find ways to reduce the running time and accordingly the time required for
221 calibration, although any remedies should not significantly affect the obtained results.
222 Moreover, Lisjak et al. (2018) claimed that the effect of their high loading rates was
223 compensated for by the critical damping scheme. However, it should be noted that the critical
224 damping scheme implemented in almost all FDEM simulations is intended to model the quasi-
225 static loading conditions using the dynamic relaxation method, which cannot reduce the effect
226 of the loading rate, unlike what was claimed by Lisjak et al. (2018). Guo (2014) investigated
227 3D FDEM modelling of BTS tests at different loading rates using the critical damping scheme
228 and showed that both the fracture pattern and obtained peak load were affected by the loading
229 rate when the velocity of the loading platens is higher than 0.01 m/s (i.e. loading rate = 0.02
230 m/s). Correspondingly, the findings of Guo (2014) support the statement that “the critical
231 damping scheme used in almost all FDEM simulations cannot reduce the effect of the loading
232 rate”. Therefore, for an accurate simulation with quasi-static loading conditions in the
233 framework of the FDEM, the loading rate should be selected correctly to avoid the dynamic

234 effects of the loading rate before selecting any input parameters, which is another strong
235 motivation for this study to conduct 3D simulations of UCS and BTS tests using the GPGPU-
236 parallelized 3D Y-HFDEM IDE.

237 Furthermore, 3D FDEM simulations of dynamic fracturing processes of rock materials under
238 dynamic loads, such as SHPB tests (e.g. Zhang and Zhao 2014), have been much more limited;
239 in particular, 3D FDEM simulations of the full system of the SHPB test are rare. However, for
240 any qualitative and quantitative discussion, accurate modelling and reasonable calibration of
241 the SHPB test are paramount for any meaningful numerical simulations of the dynamic
242 fracturing process in rock such as blasting. In fact, to date, only four peer-reviewed
243 international journal papers have been reported for modelling SHPB tests in the framework of
244 the 2D/3D FDEM. Using the 2D FDEM, Mahabadi et al. (2010) modelled the dynamic
245 fracturing process of Barre granite in a dynamic BTS test with an SHPB apparatus and found
246 good agreement between the numerical simulations and experiments. In their modelling, each
247 of the SHPB bars was modelled as a large single triangular element, and not only was the
248 element assigned mechanical properties, but velocities were prescribed to all nodes. This may
249 cause the given mechanical properties except for the contact penalty and contact friction to
250 have no meaning and the element used to model the SHPB bar to become rigid. Moreover, the
251 mode I and mode II fracture energies cannot be distinguished, which has generally been
252 recognized as important for reasonable simulations of rock fracturing by the current FDEM
253 community. Using the 3D FDEM (HOSS), Rougier et al. (2014) modelled the dynamic BTS
254 tests of weathered granite documented in Broome et al. (2012) by explicitly considering
255 elastically deformable SHPB bars. The results showed remarkably good agreement with those
256 from the experiment. Subsequently, although 2D HOSS was applied, Osthus et al. (2018)
257 proposed a novel and detailed calibration procedure based on a general and probabilistic
258 approach for numerical simulations of dynamic BTS tests of the weathered granite modelled
259 by Rougier et al. (2014) with the SHPB apparatus. Furthermore, by targeting the same SHPB-
260 based dynamic BTS tests of the same weathered granite, Godinez et al. (2018) conducted
261 several sensitivity analyses with different combinations of input parameters using 2D HOSS.
262 They showed that the simulation results are most sensitive to the parameters related to the
263 tensile and shear strengths and the fracture energies, which are valuable information. In this
264 way, additional knowledge has gradually been accumulated for the realistic modelling of
265 dynamic BTS tests with the SHPB apparatus in the FDEM community. However, although the
266 target rock in previous studies (Mahabadi et al. 2010; Osthus et al. 2018; Rougier et al. 2014)
267 was granite, in which heterogeneity and anisotropy generally play important roles, none of the

268 studies considered or discussed these important characteristics. In light of these results, further
269 investigation of FDEM modelling is needed, especially for 3D modelling of dynamic fracturing
270 of rocks of various types, because only one case of the 3D dynamic fracturing of granite in a
271 dynamic BTS test has been modelled using the FDEM to date.

272 Based on this background, this paper aims to first explain the theory and algorithm of the
273 recently developed GPGPU-parallelized FDEM implemented in the 3D Y-HFDEM IDE code.
274 The capability of the 3D Y-HFDEM IDE code in rock engineering applications is then
275 demonstrated by modelling the fracturing process of rocks and the movements of the resultant
276 rock fragments under a series of quasi-static and dynamic loading conditions. Thus, this paper
277 may provide a basis for further improvement and development of the FDEM codes based on
278 GPGPU parallelization for modelling rock fracturing, especially the dynamic fracturing of
279 rock. Many previous publications have focused on the advantages of the FDEM. However,
280 although the FDEM is very useful, there is currently no universal/perfect method to simulate
281 rock fracturing/fragmentation processes, so the disadvantages of this method other than the
282 high computational burden must also be carefully addressed. In fact, the calibration of the
283 FDEM, which is the most fundamental procedure for any meaningful numerical simulation,
284 tends to be more complex than in non-combined methods. Although the calibration method for
285 the 2D FDEM has been reported in the literature (e.g., Tatone and Grasselli 2015), it was found
286 during the development of the 3D FEM/DEM code that the calibration of 3D simulations
287 requires more careful treatment and is more sensitive to the input parameters than the 2D
288 counterparts. Especially in 3D simulations of the fracturing process of rocks due to dynamic
289 loading involving significant fragmentation of hard rock, special care must be paid to avoid
290 spurious fracture modes, which have sometimes been misunderstood or omitted/hidden in the
291 literature.

292 This paper is organized as follows. The theory used in the 3D FDEM, i.e. Y-HFDEM IDE
293 code, is first introduced, and its implementation in the framework of the GPGPU parallel
294 computation is then explained in detail. Subsequently, the accuracy and capability of the
295 developed code are investigated by modelling several common examples in rock mechanics,
296 including modelling the 3D fracturing process of rocks in BTS and UCS tests, which have been
297 used to benchmark new computational methods for rock fracturing. The entire SHPB testing
298 system of 3D dynamic BTS tests is then simulated using the newly-developed GPGPU-
299 parallelized hybrid FDEM, and the numerical results are compared with those from SHPB
300 experiments by focusing on the dynamic fracturing process of rock in the 3D dynamic Brazilian
301 tests. Finally, conclusions are drawn from this study.

302 **2. GPGPU-parallelized 3D FDEM**

303 The FDEM code “Y-HFDEM 2D/3D IDE” was originally developed using object-oriented
304 programming with visual C++ (Liu et al. 2015) based on the CPU-based sequential open-source
305 Y 2D/3D libraries (Munjiza 2004; Munjiza et al. 2010) and OpenGL. The Y-HFDEM 2D/3D
306 IDE code can significantly simplify the process of building and manipulating the input models
307 and greatly reduce the possibility of erroneous model setup, and it can also display the
308 calculated results graphically in real time with OpenGL. The code has been successfully
309 employed in simulations of rock fracturing in various geotechnical engineering problems (An
310 et al. 2017; Liu et al. 2016; Liu et al. 2015; Mohammadnejad et al. 2017). Because of the nature
311 of sequential programming, it has mainly been applied to small-scale 2D problems using
312 relatively rough meshes. To overcome this limitation, the parallel programming scheme using
313 the GPGPU controlled by CUDA C/C++ was implemented in the code in a recent study by the
314 authors (Fukuda et al. 2019) for 2D modelling and in this study for 3D modelling. Because the
315 various FDEM-based codes reviewed in Section 1 were independently developed by each
316 research institute/organization and have different features, the fundamental features of the 3D
317 Y-HFDEM IDE code and its GPGPU-based parallelization scheme are explained in detail in
318 the following subsections.

319 **2.1. Fundamental theory of 3D Y-HFDEM IDE**

320 The principles of the FDEM are based on continuum mechanics, nonlinear fracture
321 mechanics based on the CZM and contact mechanics, all of which are formulated in the
322 framework of explicit FEM (Munjiza 2004). Therefore, this section focuses on introducing the
323 features of 3D Y-HFDEM IDE unavailable in other FDEMs such as the hyperelastic model,
324 the irreversible damage during unloading, and the extrinsic cohesive zone model. Of course,
325 these features need to be introduced in context and some of the fundamental FDEM theory is
326 reviewed here to provide the context, which is also motivated by the fact that there are some
327 poor, unclear and even incorrect descriptions in some literatures of the FDEM community.

328 The continuum behavior of materials, including rocks, is modelled in 3D by an assembly of
329 continuum 4–node tetrahedral finite elements (TET4s) (Fig. 1(a)). Two types of isotropic
330 elastic constitutive models have been implemented. In the first type, which was implemented
331 in the original Y-code and has been widely used, the isotropic elastic solid obeys Eq. (1) of the
332 neo-Hookean elastic model:

333
$$\sigma_{ij} = \frac{\lambda}{2} \left(J - \frac{1}{J} \right) \delta_{ij} + \frac{\mu}{J} (B_{ij} - \delta_{ij}) + \eta D_{ij} \quad (i, j = 1, 2, 3) \quad (1)$$

334 where σ_{ij} denotes the Cauchy stress tensor, B_{ij} is the left Cauchy-Green strain, λ and μ are the
 335 Lamé constants, J is the determinant of the deformation gradient, η is the viscous damping
 336 coefficient, δ_{ij} is the Kronecker delta, and D_{ij} is the rate of deformation tensor. However, Eq.
 337 (1) cannot model anisotropic elasticity, which is important in the field of rock engineering.
 338 Thus, in the second type, a hyperelastic solid obeying Eqs. (2) and (3) is also implemented:

339
$$S_{KL} = C_{KLMN} E_{MN} \quad (K, L, M, N = 1, 2, 3) \quad (2)$$

340
$$\sigma_{ij} = \frac{1}{J} F_{iK} S_{KL} F_{jL} + \eta D_{ij} \quad (i, j, K, L = 1, 2, 3) \quad (3)$$

341 where S_{KL} denotes the 2nd Piola-Kirchhoff stress tensor, C_{KLMN} is the effective elastic stiffness
 342 tensor, E_{MN} is the Green-Lagrange strain tensor, and F_{iK} is the deformation gradient. The
 343 Einstein's summation convention applies in Eqs. (2) and (3). By properly setting C_{KLMN} in Eq.
 344 (2), both isotropic and anisotropic elastic behaviors can be simulated although only isotropic
 345 behavior is considered in this study since the rocks used in this study are better modelled as
 346 isotropic materials. The small strain tensor is not used in Eqs. (1) and (2); therefore, large
 347 displacements and large rotations can be simulated. Eq. (1) is used for the 3D UCS and BTS
 348 modellings presented in Section 3.2 while both Eq. (1) and Eqs. (2-3) are used for the 3D
 349 dynamic BTS modelling presented in Section 3.3 and no noticeable differences between Eq. 1
 350 and Eqs. (2-3) are observed since rock deformation is small. To simulate the deformation
 351 process of materials under quasi-static loading, $\eta = \eta_{crit} = 2h\sqrt{\rho E}$ is used to achieve critical
 352 damping (Munjiza 2004), where h , ρ and E are the element length, density and Young's
 353 modulus, respectively, of the target material. The value of σ_{ij} within each TET4 is converted to
 354 the equivalent nodal force f_{int} (e.g., Munjiza et al. 2015).

355 Fracturing of rock under mode I and mode II loading conditions (i.e., opening and sliding
 356 cracks, respectively) is modelled using the CZM with the concept of a smeared crack (Munjiza
 357 et al. 1999). To model the behavior of the fracture process zone (FPZ) in front of the crack tips,
 358 tensile and shear softening is applied using an assembly of 6-node initially zero-thickness
 359 cohesive elements (CE6s) (Fig. 1(a)) as a function of the crack opening and sliding
 360 displacements, (o, s) , respectively (Fig. 1(b)). Two methods can be used for the insertion of the
 361 CE6s. One is to insert the CE6s into all of the boundaries of the TET4s at the beginning of the
 362 analysis, which is known as the intrinsic cohesive zone model (ICZM), and the second is to
 363 adaptively insert the CE6s into particular boundaries of the TET4s with the help of adaptive
 364 remeshing techniques where a given failure criterion is met, which is referred to as the extrinsic
 365 cohesive zone model (ECZM) (Zhang et al. 2007; Fukuda et al. 2017). Many existing FDEM

366 codes, such as the family of Y codes that includes 3D Y-HFDEM IDE, have employed the
 367 ICZM, whereas some codes, such as ELFEN, use the ECZM. One of the advantages of the
 368 ICZM is that the implementation and application of parallel computing algorithms is
 369 straightforward, but an “artificial” intact and elastic behavior of CE6s before the onset of
 370 fracturing must be specified, which requires the introduction and correct estimation of penalty
 371 terms for CE6s and the careful selection of the time step increment, Δt , to avoid numerical
 372 instability. In the GPGPU-based 3D Y-HFDEM IDE, the normal and shear cohesive tractions,
 373 (σ^{coh} and τ^{coh} , respectively), acting on each face of the CE6s are computed using Eqs. (4) and
 374 (5) assuming tensile and shear softening behaviors, respectively:

$$375 \quad \sigma^{\text{coh}} = \begin{cases} \frac{2o}{o_{\text{overlap}}} T_s & \text{if } o < 0 \\ \left[\frac{2o}{o_p} - \left(\frac{o}{o_p} \right)^2 \right] f(D) T_s & \text{if } 0 \leq o \leq o_p \\ f(D) T_s & \text{if } o_p < o \end{cases} \quad (4)$$

$$376 \quad \tau^{\text{coh}} = \begin{cases} \left[\frac{2|s|}{s_p} - \left(\frac{|s|}{s_p} \right)^2 \right] (-\sigma^{\text{coh}} \tan(\phi) + f(D)c) & \text{if } 0 \leq |s| \leq s_p \\ -\sigma^{\text{coh}} \tan(\phi) + f(D)c & \text{if } s_p < |s| \end{cases} \quad (5)$$

377 where o_p and s_p are the “artificial” elastic limits of o and s , respectively, o_{overlap} is the
 378 representative overlap when o is negative, T_s is the tensile strength of a CE6, c is the cohesion
 379 of a CE6, and ϕ is the internal friction angle of a CE6. Positive o and σ^{coh} values indicate crack
 380 opening and a tensile cohesive traction, respectively. Eq. (5) corresponds to the Mohr-Coulomb
 381 shear strength model with a tension cut-off. The cohesive tractions σ^{coh} and τ^{coh} are applied to
 382 the opposite directions of the relative opening and sliding in a CE6, respectively. The artificial
 383 elastic behavior of each CE6 characterized by o_p and s_p along with o_{overlap} is necessary when
 384 the ICZM is used to connect the TET4s to express the intact deformation process, which is
 385 given as follows and has been used in most FDEM codes (Munjiza et al. 1999):

$$386 \quad o_p = 2hT_s / P_{\text{open}} \quad (6)$$

$$387 \quad s_p = 2hc / P_{\text{tan}} \quad (7)$$

$$388 \quad o_{\text{overlap}} = 2hT_s / P_{\text{overlap}} \quad (8)$$

389 where P_{open} , P_{tan} , and P_{overlap} are the artificial penalty terms of the CE6 for opening in the normal
 390 direction, sliding in the tangential direction and overlapping in the normal direction,
 391 respectively, and h is the element length. In this paper, the terminologies “fracture penalties”
 392 for P_{open} , P_{tan} and P_{overlap} used in previous publications are intentionally avoided because they

393 should not be considered as penalties for the fracturing behavior but rather as those for
394 controlling the artificial elastic (intact) regime of CE6. The values of P_{open} , P_{tan} and P_{overlap} can
395 be considered as the artificial stiffnesses of the CE6 for opening, sliding and overlapping,
396 respectively. Ideally, their values should be infinity to satisfy the elastic (intact) behavior of
397 rocks according to Eqs. (1) or (2), resulting in the requirement of the infinitesimal Δt .
398 Therefore, reasonably large values of the artificial penalty terms of CE6s compared to the
399 Young's modulus or Lamé constants are required because it is impossible to use infinity in
400 actual numerical simulations. Otherwise, the intact behavior of the bulk rock shows
401 significantly different (i.e., softer) behavior from that specified by Eqs. (1) or (2), and the elastic
402 constants used in these elastic constitutive equations completely lose their meanings. In
403 addition, during the development of the 3D Y-HFDEM IDE code, the authors found that the
404 sufficiently large values for the penalty terms (10 times the Young's modulus of rock, E_{rock} , in
405 most cases) recommended in many previous studies that applied the FDEM are not sufficient
406 to satisfy the continuum elastic behavior of Eqs. (1) or (2); this topic is investigated and
407 discussed in Section 3 for both quasi-static and dynamic loading problems. Some studies argue
408 that the artificial penalty terms are mechanical properties. However, in that case, the CE6s must
409 be considered as joint elements, and the artificial penalty terms of the CE6s should be called
410 joint stiffnesses to describe discontinuous media, such as preexisting joints. The function, $f(D)$,
411 in Eqs. (4) and (5) is the characteristic function for the tensile and shear softening curves (Fig.
412 1(b)) and depends on a damage value D of the CE6. When $0 < D < 1$ or $D = 1$ for a CE6, the CE6
413 can be considered to be a microscopic or macroscopic crack, respectively. The following
414 definitions of D and $f(D)$ are used to consider not only the mode I and II fracturing modes but
415 also a mixed mode I-II fracturing mode (Mahabadi et al. 2012; Munjiza et al. 1999):

$$416 \quad D = \text{Minimum} \left(1, \sqrt{\left(\frac{o - o_p}{o_t} \right)^2 + \left(\frac{|s| - s_p}{s_t} \right)^2} \right) \text{ if } o \geq o_p \text{ or } |s| > s_p, \text{ otherwise } 0 \quad (9)$$

$$417 \quad f(D) = \left[1 - \frac{A+B-1}{A+B} \exp \left(D \frac{A+CB}{(A+B)(1-A-B)} \right) \right] [A(1-D) + B(1-D)^C] \quad (0 \leq D \leq 1) \quad (10)$$

418 where A , B and C are intrinsic rock properties that determine the shapes of the softening curves,
419 and o_t and s_t are the critical values of o and s , respectively, at which a CE6 breaks and becomes
420 a macroscopic fracture. To avoid unrealistic damage recovery (i.e., an increase of f), the
421 following treatment has been implemented in the code. If the trial f computed from Eq. (10) at
422 the current time step becomes larger than that at the previous time step, f_{pre} , a condition of $f =$

423 f_{pre} is assigned to avoid unrealistic damage recovery. The o_t and s_t in Eq. (9) satisfy the Mode I
 424 and II fracture energies G_{II} and G_{III} (Fig. 1(b)) specified in Eqs. (11) and (12), respectively:

$$425 \quad G_{II} = \int_{o_p}^{o_t} \sigma^{coh}(o) do \quad (11)$$

$$426 \quad G_{III} + W_{res} = \int_{s_p}^{s_t} \{ \tau^{coh}(|s|) \} d|s| \quad (12)$$

427 where W_{res} is the amount of work per area of a CE6 done by the residual stress term in the
 428 Mohr-Coulomb shear strength model. Note that in the current formulation, mode II and III
 429 fracturing modes are not distinguished, and it is assumed that in-plane (mode II) and out-plane
 430 (mode III) responses of the micro cracks (i.e. CE6s) are simply described by the parameter G_{III}
 431 because the clear definition of crack tips and conducting reproducible/reliable mode II and
 432 Mode III fracture toughness tests are challenging. This paper uses the same $f(D)$ with A , B and
 433 C equal to 0.63, 1.8 and 6.0, respectively (Munjiza et al. 1999), for both mode I and II fracture
 434 processes because of the lack of experimental data. However, it is worth mentioning that the
 435 recent studies by Osthus et al. (2018) and Godinez et al. (2018) using the FDEM code “2D
 436 HOSS” showed that the shape of the softening curve had minor influences on the obtained
 437 results and that the tensile and shear strengths and fracture energies are the main affecting
 438 factors. Unloading (i.e., a decrease of o or $|s|$) can also occur during the softening regime (i.e.,
 439 $o > o_p$ or $|s| > s_p$) (see Fig. 1(b)), which is modelled based on Eqs. (13) and (14) (Camacho and
 440 Ortiz 1996):

$$441 \quad \sigma^{coh} = f(D_{max}) T_s \frac{o}{o_{max}} \quad \text{if } 0 < o < o_{max} \quad \text{and } o_{max} > o_p \quad (13)$$

$$442 \quad \tau^{coh} = \left\{ -\sigma^{coh} \tan(\phi) + f(D_{max}) c \right\} \frac{|s|}{s_{max}} \quad \text{if } |s| < s_{max} \quad \text{and } s_{max} > s_p \quad (14)$$

443 In each CE6, the computed σ^{coh} and τ^{coh} are converted to the equivalent nodal force \mathbf{f}_{coh} using a
 444 3-point or 7-point Gaussian integration scheme depending on the required precision of the
 445 simulation. When either o_t or s_t is achieved in a CE6, the CE6 is deactivated, and its surfaces
 446 are considered as new macroscopic fracture surfaces that are subjected to contact processes.

447 The contact processes between the material surfaces, including the new macroscopic
 448 fractures created by the separation of each CE6, are modelled by the penalty method (Munjiza
 449 2004); a complete and excellent explanation of the method is given in the literature (Munjiza
 450 2004). As a brief explanation, when any two TET4s subjected to contact detection (see
 451 Subsection 2.2 for the implementation of contact detection in the framework of GPGPU) are

452 found to overlap each other, the contact potential due to the overlapping of the two TET4s (i.e.,
 453 the contacting couple) is exactly computed. The normal contact force, f_{con_n} , is then computed
 454 for each contacting couple, which acts normally to the contact surface and is proportional to
 455 the contact potential. The proportional factor is called the normal “contact penalty”, P_{n_con} .
 456 After the normal contact force, f_{con_n} , and its acting point are obtained, the nominal normal
 457 overlap, o_n , and relative displacement vector, $\Delta \mathbf{u}_{\text{slide}}$, at the acting point of f_{con_n} are readily
 458 computed. The contact damping model proposed by An and Tannant (2007) (Fig. 2) can also
 459 be applied if the role of contact damping is very important. When this scheme is applied, the
 460 normal contact force, f_{con_n} , described above is regarded as a trial contact force, $(f_{\text{con}_n})^{\text{try}}$, and a
 461 trial contact stress $(\sigma_{\text{con}_n})^{\text{try}}$ is then computed by dividing $(f_{\text{con}_n})^{\text{try}}$ by the contact area, A_{con} . Eq.
 462 (15) is then used to determine the contact stress σ_{con_n} :

$$463 \quad \sigma_{\text{con}_n} = \begin{cases} \text{Minimum} \left((\sigma_{\text{con}_n})^{\text{try}}, T \right) & \text{during the increase of } o_n \text{ (Loading)} \\ T \left(o_n / o_{n_max} \right)^b & \text{during the decrease of } o_n \text{ (Unloading)} \end{cases} \quad (15)$$

464 where T is the transition force, b is the exponent, and o_{n_max} is the maximum value of o_n
 465 experienced during the loading process at the contact. T limits σ_{con_n} and defines the transition
 466 between a linearly elastic stress–displacement relationship and a ‘recoverable’ displacement at
 467 a constant contact stress. The values of T may be related to the physical properties of rocks,
 468 such as the uniaxial compressive strength. The exponent b adjusts the power of the damping
 469 function that is applied to the rebound or extension phase of the contact. The value of the
 470 exponent has an effect on the energy loss during an impact event. A similar contact damping
 471 model is implemented in the 2D Y-Geo code in the framework of the FDEM, in which only b
 472 is considered (Mahabadi et al. 2012). After σ_{con_n} is computed using Eq. (15), it is converted to
 473 f_{con_n} ($= A_{\text{con}} \times \sigma_{\text{con}_n}$). The verification of the implemented contact damping is discussed in
 474 Section 3.1. After f_{con_n} is determined, the magnitude of the tangential contact force vector,
 475 $\|\mathbf{f}_{\text{con_tan}}\|$, is computed according to the classical Coulomb friction law. The $\|\mathbf{f}_{\text{con_tan}}\|$ is
 476 computed based on Eq. (16):

$$477 \quad \|\mathbf{f}_{\text{con_tan}}\| = \mu_{\text{fric}} f_{\text{con}_n} \quad (16)$$

478 where μ_{fric} is the friction coefficient between the contact surfaces. The tangential contact force,
 479 $\mathbf{f}_{\text{con_tan}}$, is applied parallel to the contact surface in the opposite direction to $\Delta \mathbf{u}_{\text{slide}}$. The
 480 verification of the implementation of the contact friction, which is important in any simulation
 481 of fracturing due to quasi-static loading, is discussed in Section 3.1. In each contacting couple,
 482 the contact force is converted to the equivalent nodal force \mathbf{f}_{con} (Munjiza 2004).

483 By computing the nodal forces described above, the following equation of motion, Eq. (17),
484 is obtained and solved in the framework of the explicit FEM (Munjiza 2004):

$$485 \quad \mathbf{M} \frac{\partial^2 \mathbf{u}}{\partial t^2} = \mathbf{f}_{\text{ext}} + \mathbf{f}_{\text{int}} + \mathbf{f}_{\text{coh}} + \mathbf{f}_{\text{con}} \quad (17)$$

486 where \mathbf{M} is a lumped nodal mass computed from the initial TET4 volume and element mass
487 density ρ , \mathbf{u} is the nodal displacement, and \mathbf{f}_{ext} is the nodal force corresponding to the external
488 load. The central difference scheme is employed for the explicit time integration to solve Eq.
489 (17). A careful selection of the time step, Δt , is necessary to avoid numerical instability and
490 spurious fracture modes. An excellent explanation of the reasonable selection of Δt in the
491 ordinary FDEM can be found in Guo (2014).

492 **2.2. GPGPU-based Parallelization of 3D Y-HFDEM IDE by CUDA C/C++**

493 To speed up the simulation process of the 3D Y-HFDEM IDE code, a parallel computation
494 scheme based on the NVIDIA® GPGPU accelerator is incorporated. In our case, the
495 computation on the GPGPU device is controlled through NVIDIA’s CUDA C/C++ (NVIDIA
496 2018), which is essentially an ordinary C/C++ programming language with several extensions
497 that make it possible to leverage the power of the GPGPU in the computations. The CUDA
498 programming model uses abstractions of “*threads*”, “*blocks*” and “*grids*” (Fig. 3). A greater
499 degree of parallelism occurs within the GPGPU device itself. Functions, also known as
500 “*kernels*”, are launched on the GPGPU device and are executed by many “*threads*” in parallel.
501 A “*thread*” is just an execution of a “*kernel*” with a given “*thread index*” within a particular
502 “*block*”. As shown in Fig. 3, a “*block*” is a group of threads, and a unique “*block index*” is
503 given to each “*block*”. The “*block index*” and “*thread index*” enable each thread to use its
504 unique “*index*” to globally access elements in the GPGPU data array such that the collection
505 of all threads processes the entire data set in massively parallel manner. The “*grid*” is just a
506 group of “*blocks*”. Only a single “*grid*” system is used in this study. The concept of a GPGPU
507 cluster with a massive number of GPGPU accelerators is also possible, although this is beyond
508 the scope of this paper. The “*blocks*” can execute concurrently or serially depending on the
509 number of streaming processors available in a GPGPU accelerator. Synchronization between
510 “*threads*” within the same “*block*” is possible, but no synchronization is possible between the
511 “*blocks*”. In each “*thread*” level, the corresponding code that the “*threads*” execute is very
512 similar to the CPU-based sequential code (see Fukuda et al. 2019), which is one of the
513 advantages of the application of CUDA C/C++. For example, the Quadro GP100 accelerator
514 (in Pascal generation) used in this paper contains 56 and 3584 streaming processors and CUDA

515 cores (NVIDIA 2018), respectively. Higher computational performance of the GPGPU-
516 parallelized code running on the same GPU accelerator can be achieved than that of ordinary
517 CPU-based sequential codes. The number of “blocks” per “grid” (N_{BpG}) and the number of
518 “threads” per “block” (N_{TpB}) can be changed to speed up the GPGPU (Fig. 3). The current
519 version of 3D Y-HFDEM IDE normally sets N_{TpB} to either 256 or 512, and N_{BpG} is
520 automatically computed by dividing the total number of threads (N_{thread}) in each “kernel” by
521 N_{TpB} , in which an additional block is needed if $N_{\text{thread}}/N_{TpB}$ is not a multiple of N_{TpB} . The value
522 of N_{thread} is set to be equal to the total number of TET4s, CE6s, contact couples or nodes
523 depending on the purpose of each “kernel”.

524 In the GPGPU implementation of 3D Y-HFDEM IDE, the computations for each TET4 (\mathbf{f}_{int}
525 and \mathbf{M}), CE6 (\mathbf{f}_{coh}), contact couple (\mathbf{f}_{con}) or nodal equation of motion (Eq. (17)) are assigned to
526 each GPGPU “kernel” as shown in Fig. 4 and processed in a massively parallel manner. The
527 CUDA code used in each “kernel” is similar to the functions/subroutines in CPU-based
528 sequential codes, which also holds true for the computations shown in Fig. 4. Thus, most parts
529 of the original sequential CPU-based code can be used with minimum modifications. For the
530 computation of the contact force, \mathbf{f}_{con} , “TET4 to TET4 (TtoT)” contact interaction kinematics
531 are used in the earliest versions of the Y3D code (Munjiza 2004). This TtoT approach exactly
532 considers the geometries of both the contactor and target TET4s, and the integration of the
533 contact force distributed along the surfaces of the TET4s is performed analytically. Because
534 this approach integrates the contact forces exactly, it is precise although quite time-consuming.
535 As pointed out in the literature (Lei et al. 2014), the contact interaction in 3D can be further
536 simplified by “TET4 to point (TtoP)” contact interaction kinematics, which make the
537 implementation simpler and more time efficient. However, the precision of the computed
538 contact force using the TtoP approach is less accurate unless a sufficient number of target points
539 per TET4 is used. Thus, the TtoT approach is intentionally applied for all of the numerical
540 simulations in this paper instead of the TtoP approach to ensure the precision of the computed
541 contact force.

542 A flowchart of the GPGPU-based 3D Y-HFDEM IDE is shown in Fig. 5. One of the
543 challenging tasks in Fig. 5 is the implementation of contact detection to identify each contacting
544 couple only through the GPGPU without a sequential computational procedure. For sequential
545 CPU implementation, powerful and efficient contact detection algorithms, such as No Binary
546 Search and Munjiza-Rougier contact detection algorithms, have been proposed (Munjiza 2004;
547 Munjiza et al. 2011), and these can achieve the fastest (i.e., linear) neighbor searches with the
548 computational complexity of $O(N)$, in which N is the number of elements and the required

549 computation for the contact detection is proportional to the number of TET4 candidates
550 subjected to contact detection. However, these contact detection algorithms are not
551 straightforward to be implemented in the GPGPU-based code. In the GPGPU-based 3D Y-
552 HFDEM IDE code, because the FDEM modelling requires a fine mesh that often consists of
553 TET4s with similar sizes, the following contact detection algorithm is implemented. In this
554 algorithm, the analysis domain comprising a massive number of TET4s is subdivided into
555 multiple equal-sized (n_x, n_y, n_z) cubic sub-cells (Fig. 6) along each direction, so the largest
556 TET4 in the analysis domain is completely included in a single sub-cell. In this way, the center
557 point of every TET4 can always belong to a unique sub-cell. Using integer coordinates $(ix, iy,$
558 $iz)$ ($ix = 0, \dots, n_x-1, iy = 0, \dots, n_y-1, iz = 0, \dots, n_z-1$) for the location of each sub-cell (Fig. 6),
559 unique hash values, $h (= iz \times n_x \times n_y + iy \times n_x + ix)$, are assigned to each sub-cell. The subsequent
560 contact detection procedure is explained using a simplified example, as shown in Fig. 7, where
561 ten TET4 candidates with similar sizes are subjected to contact detection. First, all of the TET4s
562 are mapped into integer coordinates ($ix = 0, 1$ and $2, iy = 0, 1$ and $2,$ and $iz = 0, 1$ and 2) with
563 $n_x = n_y = n_z = 3$ along with the computation of the hash values in each sub-cell. In this way,
564 the list L-1 is readily constructed using the massively parallel computation based on the similar
565 concept shown in Fig. 4 by assigning the computation of each TET4 to each CUDA “*thread*”.
566 The IDs of the TET4s in the list L-1 in Fig. 7 are then sorted from the smallest to largest
567 according to the hash values h as keys, which generates the list L-2 in Fig. 7. The radix sorting
568 algorithm optimized for CUDA (Satish et al. 2009) and implemented in the open-source
569 “*thrust*” library is used for the key-sorting by the hash values; therefore, this procedure can
570 also be processed in a massively parallel manner. Utilizing the list L-2 and the GPGPU device’s
571 shared memory (NVIDIA 2018), the list L-3 in Fig. 7 is further constructed in a GPGPU
572 “*kernel*”, which makes it possible to identify the first and last indices for the particular hash
573 value in the list L-2. Therefore, with the lists L-2 and L-3, the IDs of all the TET4s included in
574 a particular sub-cell with its unique hash value are readily available. Finally, for a particular
575 TET4 with its sub-cell position, it is sufficient to only search its adjacent 27 sub-cells (or 14
576 sub-cells using the concept of a contact mask (Munjiza 2004)) for contact detection, which
577 makes it possible to achieve efficient contact detection using only the GPGPU device and
578 without a sequential CPU procedure. When the sizes of the TET4s are completely different,
579 the size of the cubic sub-cell is very large, resulting in very inefficient contact detection; thus,
580 other parallel neighbor search schemes, such as the Barnes-Hut tree algorithm (Burtscher and
581 Pingali 2011), should be used for efficient contact detection. The contact detection algorithm
582 used in the GPGPU-based 3D Y-HFDEM IDE is similar to that applied by Lisjak et al. (2018)

583 in terms of the applied hash-based contact detection algorithm. However, Lisjak et al. (2018)
584 further incorporated the hyperplane separation theorem as the 2nd step of contact detection.
585 Unfortunately, Lisjak et al. (2018) did not discuss the algorithm used to generate the hash-table
586 in the framework of OpenCL, which is also one of very important aspects of the performance
587 gain, and the effect of introducing the 2nd step. Furthermore, no discussion was given for the
588 speed up performance of Irazu for the 3D simulation. Instead, the computing performance of
589 our GPGPU-parallelized 3D Y-HFDEM IDE is discussed in Section 3.4 of this study proving
590 that the GPGPU-parallelized 3D Y-HFDEM IDE is very efficient in terms of contact detection
591 and has the computational complexity of $O(N)$.

592 Therefore, the GPGPU-parallelized 3D Y-HFDEM IDE code can run in a completely parallel
593 manner on the GPGPU device, and no sequential processing is necessary, except for the input
594 and output procedures. The data transfer from the GPGPU device to the host computer is
595 always necessary to output the analysis results, the time of which is often negligible compared
596 to the entire simulation time for most Y-HFDEM IDE simulations. The results can be
597 visualized in either OpenGL implemented in the 3D Y-HFDEM IDE code (Liu et al. 2015) or
598 in the open source visualization software Paraview (Ayachit 2015).

599 Finally, it should be noted that an efficient contact calculation activation approach has been
600 applied in some publications about the FDEM in the framework of ICZM (e.g., Section 2.3.3.2
601 along with Fig. 2.14 in Guo (2014)). In this approach, only the TET4s in the vicinity of newly
602 broken/failed CE6s become contact candidates and are added to the contact detection list. One
603 advantage of this approach is that the contact detection and contact force calculations are
604 necessary only for the initial material surfaces by the time when the broken/failed CE6s are
605 generated; thus, dramatic savings in the computational time for the contact detection are
606 possible. This approach is called an efficient contact detection activation (ECDA) approach
607 hereafter. However, in the ICZM-based FDEM simulation of hard rocks under compressive
608 loading conditions, most of the TET4s can overlap during the progress of compression even
609 before the generation of broken/failed CE6s. In this case, if the amount of overlap is not
610 negligible when broken CE6s are generated, this ECDA approach generally results in the
611 sudden application of the contact force like a step-function, which can easily cause numerical
612 instability and result in unrealistic/spurious fragmentation. To avoid the numerical instability
613 and spurious fracturing modes, an infinitesimally small Δt must be used, which makes the
614 simulation intractable. One way to avoid this instability is to monitor the overlapping of the
615 CE6s. If a significant overlap is detected in a CE6, the TET4s in the vicinity of the CE6 are
616 immediately considered as new candidates for the contact detection although the threshold of

617 this “significant overlap” is problem-dependent. Another simple remedy is to add all of the
618 TET4s as contact candidates. This approach is called the brute-force contact detection
619 activation approach. Although it is too time consuming, the brute-force contact detection
620 activation approach can work as a remedy for a wide range of rock conditions, including very
621 hard rock, which is used for the 3D dynamic BTS modelling of marble with the split Hopkinson
622 pressure bar (SHPB) testing system. It should be noted that all FDEM simulations must deal
623 with this problem carefully to avoid inaccurate simulation results, although it has not been
624 reported in the literature. Otherwise, the obtained fracture patterns may be spurious.

625 **3. Numerical tests and code validation**

626 This section aims to verify and validate the GPGPU-parallelized 3D Y-HFDEM IDE code by
627 conducting several numerical simulations. All of the numerical simulations in this section are
628 conducted using the GPGPU-based code.

629 **3.1. Verifications of contact damping and contact friction**

630 To assess the accuracy of the contact damping model implemented in Section 2.1, a simple
631 impact test is modelled (Fig. 8(a)) using the GPGPU-parallelized 3D Y-HFDEM IDE. The
632 model is the extension of the 2D model reported by Mahabadi et al. (2012) to 3D, and the
633 obtained results are discussed. The model consists of a spherical elastic body with a radius of
634 0.1 m impacting a fixed rigid surface vertically. The elastic body is not allowed to fracture in
635 this model. Following the study (Mahabadi et al. 2012), gravitational acceleration is neglected,
636 the density of the elastic body is $2,700 \text{ kg/m}^3$, and the initial total kinetic energy of the elastic
637 body, E^0_{kin} , before the impact event is 565.5 J. Because the Lamé constants λ and μ for the
638 elastic body are not available in that study (Mahabadi et al. 2012), it is simply assumed that λ
639 $= \mu = 5.0 \text{ GPa}$ and that the viscous damping coefficient $\eta = 0$ for internal viscous damping.
640 Contact friction is also neglected. Thus, energy dissipation is only due to the contact damping
641 to make it simple to discuss the effect of the contact damping. Parametric analyses are
642 conducted by changing the exponent b , the transition force T in Eq. (15) and the normal contact
643 penalty P_{con_n} between the elastic body and the rigid surface. The normalized total kinetic
644 energy of the elastic body by E^0_{kin} as a function of time is monitored during the parametric
645 analyses.

646 Fig. 8(b) compares five cases with b values equal to 1, 2, 5, 20 and 30 when $T = \infty$ (very
647 large value, i.e. $1.0\text{e}+30 \text{ Pa}$) and $P_{\text{con}_n} = 0.1 \text{ GPa}$. The case with $b = 1$ corresponds to an elastic
648 contact; thus, no energy dissipation occurs due to the contact, although a very small decrease

649 in the kinetic energy actually occurs after the impact because a small amount of the kinetic
650 energy is converted to the strain energy of the elastic body. As the value of b increases, the
651 amount of kinetic energy dissipates from the system increases. This behavior is similar to that
652 reported in the literature (Mahabadi et al. 2012) using a sequential 2D FDEM. The cases with
653 different values of P_{con_n} ($= 0.1$ GPa and 10 GPa) but constant $b = 2$ and $T = \infty$ show that the
654 same b does not result in the same energy dissipation when P_{con_n} is different. This is a
655 reasonable outcome because the maximum value of the nominal normal overlap o_{n_max} in Eq.
656 (15) during the impact event changes for different values of P_{con_n} (Munjiza 2004). However,
657 this important fact has not been reported in the literature (Mahabadi et al. 2012). Likewise, the
658 two cases with different values of T ($= \infty$ and 1 MPa) but constant $b = 2$ and $P_{con_n} = 0.1$ GPa
659 show different amounts of energy dissipation, which can be also explained by the change in
660 o_{n_max} . These expected results verify that the contact detection and computation of \mathbf{f}_{con} are
661 properly processed in the GPGPU-based code, although this paper does not consider contact
662 damping in the following numerical simulations because the calibration of these parameters
663 against rock fall experiments is beyond the scope of this paper.

664 To assess the accuracy of the contact friction model implemented in Section 2.1, a simple
665 sliding test, which was originally suggested by Xiang et al. (2009) as a 2D problem, is modelled
666 as a 3D problem, and the obtained results are compared with those from theoretical analyses.
667 The model consists of a simple cube sliding along a fixed plane with a friction coefficient of
668 $\mu_{fric} = 0.5$. The cube is assigned an initial velocity, which varies from 1 m/s to 6 m/s. With each
669 initial velocity, the cube slows and stops due to the friction between the sliding cube and the
670 rigid base. Theoretically, the sliding distance can be defined as a function of the initial velocity
671 (v_i), gravitational acceleration (g) and the friction coefficient (μ_{fric}) through Eq. (18).

$$672 \quad L = v_i^2 / (2\mu_{fric}g) \quad (18)$$

673 Figure 9 shows an excellent agreement between the numerical simulation and the theoretical
674 solution from Eq. (18), which validates the accuracy of the implemented contact friction model.

675 **3.2 3D FDEM modelling of the failure process of rock under quasi-static loading** 676 **conditions**

677 In this subsection, two standard rock mechanics laboratory tests, the UCS test and BTS test,
678 of a relatively homogeneous limestone are modelled to investigate the capabilities of the
679 GPGPU-parallelized 3D Y-HFDEM IDE for simulating the fracturing process and associated
680 failure mechanism of the rock under quasi-static loading conditions.

681 The numerical models for the 3D FDEM simulations of the UCS and BTS tests are shown in
682 Fig. 10, in which the diameter of both specimens is 51.7 mm, and the height and thickness of
683 the specimens are 129.5 mm and 25.95 mm, respectively. The rock specimens are placed
684 between two moving rigid loading platens. Flat rigid loading platens are used in the UCS
685 model, whereas curved rigid loading platens are used in the BTS model, whose curvature is 1.5
686 times the diameter of the BTS disk, as suggested by the International Society for Rock
687 Mechanics (ISRM). It is important to note that in the FDEM simulation with the CZM, it is
688 essential to use an unstructured mesh to obtain reasonable rock fracture patterns because the
689 FDEM only allows the fractures to initiate and propagate along the boundaries of the solid
690 elements. Moreover, the use of a very fine mesh is the key to reducing the mesh dependency
691 of the crack propagation paths, which is why most 2D FDEM simulations in previous studies
692 used very fine meshes. However, if very fine meshes are used in 3D models, the number of
693 TET4s and CE6s can easily exceed several million, which makes the 3D modelling become
694 intractable in terms of memory limitations and the simulation time required for the calibration
695 process using FDEM codes parallelized by a single GPGPU accelerator. Thus, fine meshes
696 comparable to those investigated by Lisjak et al. (2018) are used in the 3D FDEM models.
697 Accordingly, the average edge length of the TET4s in both models is set to 1.5 mm, and the
698 UCS and BTS models contain 695,428 TET4s and 1,298,343 CE6s and 187,852 TET4s and
699 348,152 CE6s, respectively. According to a recent study on 2D FDEM modelling conducted
700 by Liu and Deng (2019), the effect of the element size can be negligible if there are not less
701 than 27-28 meshes in the length corresponding to the diameter of the specimen and the
702 maximum element size should not be longer than the length of fracture process zone of the
703 CZM. The UCS and BTS models used in this study satisfy these requirements. Based on the
704 review of the effects of the loading rate presented above, a constant velocity of 0.01 m/s is
705 applied on the loading platens to satisfy the quasi-static loading conditions, as suggested by
706 Guo (2014). Moreover, our preliminary sensitivity study of 3D UCS and BTS modellings under
707 various loading rates confirms the quasi-static loading conditions can be achieved with the
708 platen velocity of 0.01 m/s, which further shows little difference can be noticed from 3D UCS
709 modelling even if the platen velocity is increased to 0.05 m/s although slightly higher peak
710 loads and spurious fragmentations around the loading areas are observed from 3D BTS
711 modelling with the platen velocity increasing.

712 The physical/mechanical properties of the limestone are obtained from laboratory
713 measurements, which are used to determine the input parameters for the numerical modelling
714 (Table 1). As introduced in Section 2, the intact behavior of the numerical model follows Eq.

715 (1), and the elastic parameters are determined in such a way that the elastic region of the stress-
 716 strain curve from the 3D FDEM simulation of the UCS test agrees well with that of the
 717 laboratory experiment. Our preliminary investigation showed that the experimentally obtained
 718 elastic parameters can be directly used as the input parameters if sufficiently large artificial
 719 penalty terms of the CE6s are used. The penalty terms for the contacts and CE6s are determined
 720 as multiples of the elastic modulus of the rock (i.e., E_{rock}). The value of the contact penalty
 721 (P_{n_con}) is set as $10E_{\text{rock}}$. However, higher values are required for the artificial penalty terms of
 722 the CE6s (i.e., $100E_{\text{rock}}$ for P_{open} and P_{tan} and $1000E_{\text{rock}}$ for P_{overlap}) to allow the artificial
 723 increase in the compliance of the bulk rock to become negligible. Because the artificial elastic
 724 regime of the CE6 in Eqs. (6)-(8) can also be influenced by the strength parameters, an artificial
 725 increase in the bulk compliance of the rock becomes non-negligible if the artificial penalty
 726 terms for the CE6s are not set to be sufficiently large. In other words, smaller artificial penalty
 727 terms for the CE6s can result in non-negligible changes in the artificial elastic response of the
 728 CE6s during the calibration process, in which the strength parameters are also varied. This fact
 729 has not been pointed out by any studies of the development and/or application of the FDEM.
 730 Then, after setting the penalty terms of the CE6s, the strength parameters (i.e., tensile strength
 731 T_{s_rock} , cohesion c_{rock} and internal friction angle φ_{rock}) and the fracture energies, $G_{\text{fl_rock}}$ and
 732 $G_{\text{nl_rock}}$, of the numerical model are calibrated by trial and error. To do this, a series of FDEM
 733 simulations was conducted to achieve a reasonable match between the numerical and
 734 experimental results so the peak load and fracture patterns from the numerical simulations
 735 agree well with those from the experiments. The same input parameters shown in Table 1 are
 736 also used for the numerical modelling of the BTS test. The friction coefficients, μ_{fric} , of the
 737 contact between the platens and the rock, and, that between rock surfaces generated by broken
 738 CE6s are assumed to be 0.1 and 0.5, respectively, as in Mahabadi (2012). The ECDA approach
 739 introduced in Section 2.2 is used in these simulations because the rock can be considered to be
 740 relatively soft, and no spurious modes are observed in this case. The concept of mass scaling
 741 (Heinze et al. 2016) with the mass-scaling factor = 5 is applied to increase Δt in such a way
 742 that the quasi-static loading condition is still satisfied. Thus, $\Delta t = 4.5$ and 9 nanoseconds are
 743 used for the simulations of the BTS and UCS tests, respectively. Further details about the mass
 744 scaling concept can be found in the literature (Heinze et al. 2016). The critical damping scheme
 745 ($\eta = \eta_{\text{crit}} = 2h\sqrt{\rho E}$ in Eq. (1)) is also used in all of the FDEM simulations in this section.
 746 Hereafter, compressive stresses are considered negative (cold colours), whereas tensile stresses
 747 are regarded as positive (warm colours).

748 Figure 11(a)-(c) shows the modelled 3D progressive rock failure process in terms of the
749 distributions of the minor principal (mostly compressive) stress (upper row) and the damage
750 variable (i.e., D in Eq. (9); lower row) at different loading stages (points A, B and C in Fig.
751 11(d)) in the FDEM simulation of the UCS test. Fig. 11(d) shows the obtained axial stress
752 versus axial strain curve. Fig. 11(a) shows the stress and damage distribution in the sample at
753 the stage before the onset of nonlinearity in the axial stress versus axial strain curve (point A
754 in Fig. 11(d)). As the loading displacement continues, the growth of unstable microscopic
755 cracks commences and continues until the peak stress of the stress-strain curve is reached (point
756 B in Fig. 11(d)). Subsequently, the microscopic cracks coalesce to form macroscopic cracks,
757 which results in the loss of bearing capacity of the bulk rock, and the axial stress begins to
758 decrease with increasing strain. Finally, the formed macroscopic cracks propagate further,
759 resulting in the complete loss of the bearing capacity of the rock (point C in Fig. 11(d)). Fig.
760 12(a) compares the final fracture patterns obtained from the FDEM simulation and the
761 laboratory experiment. The resulting fracture patterns (Fig. 12(a)) and the peak loads (Fig.
762 11(d)) from the numerical simulation and laboratory experiment are in good agreement.
763 Although the formation of the shearing planes is evident in Fig. 12(b), it must be noted that
764 mixed mode I-II fractures are the dominant mechanism of rock fracturing, as shown in Fig.
765 12(c), which will be explained in detail later. Therefore, the obtained results demonstrate that
766 the developed GPGPU-parallelized 3D Y-HFDEM IDE is able to reasonably model the
767 fracturing process of rock in UCS tests.

768 Figure 13(a)-(c) illustrates the modelled 3D progressive rock failure process in terms of the
769 distributions of the horizontal stress, σ_{xx} , (upper row) and the damage variable D in Eq. (9)
770 (lower row) at different stages (points A, B and C in Fig. 13(d)) in the FDEM simulation of the
771 BTS test. Fig. 13(d) shows the obtained indirect tensile stress versus axial strain curve. As the
772 loading displacement gradually increases, a uniform horizontal (tensile) stress (σ_{xx}) field
773 gradually builds up around the central line of the rock disk. Fig. 13(a) shows that although
774 some microscopic damage, $D \ll 1$, appears in the rock disk near the loading platens due to stress
775 concentrations, there is no macroscopic crack (i.e., CE6s with $D = 1$). Once the peak indirect
776 tensile strength of the rock (point B in Fig. 13(d)) is reached, macroscopic cracks form around
777 the central diametrical line of the rock disk due to the coalescence and propagation of
778 microscopic cracks. As shown in Fig. 13(b), the macroscopic crack that causes the splitting
779 failure of the rock disk nucleates slightly away from the exact center of the disk. The reason
780 for the nucleation of this vertically off-center macroscopic crack is that the curved loading
781 platens provide a relatively narrow contact strip; accordingly, an off-center horizontal stress

782 concentration first develops within the rock disk, which is consistent with the location of the
783 macroscopic crack nucleation in Fig. 13(b). This phenomenon was also addressed by Fairhurst
784 (1964) and Erarslan et al. (2012), who further pointed out that a wider contact strip was required
785 to ensure near-center crack initiation in a BTS test with curved loading platens. Moreover, Li
786 and Wong (2013) conducted strain-stress analyses of a 50-mm-diameter rock disk in BTS tests
787 using FLAC3D and pointed out that the maximum indirect tensile stress and strain were located
788 approximately 5 mm away from the two loading points along the central loading diametrical
789 line of the rock disk. Nearly the same results are observed in Fig. 13(b); most importantly, the
790 vertically off-center macroscopic cracks are captured explicitly. Moreover, following Lisjak et
791 al. (2018), Fig. 13(e) compares the simulated stress distributions along the diameter of the rock
792 disk in the middle and on the surface (i.e., lines AB and CD, respectively) with the analytical
793 solution of Hondros (1959), where y and r represent the vertical distance from the center and
794 the radius of the rock disk, respectively. It can be seen from Fig. 13(e) that the stress distribution
795 on the surface of the rock disk differs from that in the middle plane, which is consistent with
796 Hondros' solution based on the plane strain assumption. In other words, Hondros' solution is
797 invalid for the stress distributions on the surface of the rock disk, especially for the tensile stress
798 concentrations in the regions near the loading platens, which are clearly depicted in Fig. 13(e).
799 The local tensile stress concentrations in the regions near the loading platens on the surface of
800 the rock disk explain the nucleation of the off-center macroscopic cracks modelled in Fig 13(b),
801 which will not be possible if 2D plane-strain modelling is conducted. As the loading platens
802 continue to move toward each other, the resultant macroscopic cracks propagate and coalesce
803 to split the rock disk into two halves (Fig. 13(c)), and the stress-strain curve decreases toward
804 zero during the post-peak stage (i.e., line BC in Fig. 13(d)). These results show that mixed
805 mode I-II failure is the dominant mechanism during the nucleation, propagation and
806 coalescence of the splitting macroscopic cracks (Fig. 14(a)), which is due to the unstructured
807 mesh used in the FDEM modelling. To clarify this point, Fig. 14(b) shows the modelled failure
808 pattern of the rock disk using the FDEM simulation with a structured mesh, in which the
809 loading diametrical line aligns with the boundaries of the TET4 elemental mesh. The splitting
810 fracture forms exactly along the loading diametrical line, and the mode I failure is the only
811 failure mechanism. Fig. 14(c) and (d) show the topological relationships between the horizontal
812 indirect tensile stress (blue arrows in the x direction) and the TET4s in the cases of the
813 unstructured and structured meshes used in Fig. 14(a) and (b), respectively. In Fig. 14(d), the
814 normal directions of the planes A and A' of the CE6 located between the two TET4s exactly
815 aligns with the direction of the indirect tensile stress (i.e., the x direction). Therefore, pure mode

816 I cracks preferably develop along the CE6s on the loading diametrical line in terms of the most
817 efficient energy release due to fracturing. On the other hand, when the unstructured mesh in
818 Fig. 14(a) is used, few CE6s have planes exactly on the loading diametrical line, as illustrated
819 in Fig. 14(c), in which two CE6s (i.e., planes A-A' and B-B') contributing to the fracturing
820 process are depicted, and none of their normal directions are aligned with that of the indirect
821 tensile stress (blue arrows in the x direction). In this case, it is obvious that a pure mode I crack
822 cannot form due to the topological restriction, and a combination of mode I (opening) and mode
823 II (sliding) cracks can always form resulting in a macroscopic fracture. This is why mixed
824 mode I-II fracturing is the main failure mechanism in the FDEM simulations of the BTS tests
825 with unstructured meshes. Tijssens et al. (2000) conducted a comprehensive mesh sensitivity
826 analysis for the CZM and concluded that the fractures tended to propagate along dominant
827 directions of local mesh alignment. Guo (2014) further commented that unstructured meshes
828 should be used in the numerical simulation using the CZM to reduce mesh dependency but
829 fracture paths were still dependent on local mesh orientation in the unstructured meshes.
830 Accordingly, when the CZM is applied to model material failure, it is *unrealistic* to pursue a
831 pure mode I splitting fracture in the simulation of the BTS test with an unstructured mesh
832 regardless of the numerical approach and 2D/3D modelling. In this sense, any intentional
833 reduction of the mode-I fracture energy G_{fl_rock} and tensile strength T_{s_rock} to capture the
834 unreasonable pure mode I fracture pattern prevalent in studies of the FDEM should be
835 considered a manipulation of the input parameters. The same explanation is valid for the
836 dominant mixed-mode I-II failures along the macroscopic shear fracture plane modelled in the
837 UCS test. The boundaries of the TET4s will not exactly align with the macroscopic shear stress
838 direction at each location; thus, mixed mode I-II failure along the macroscopic shear fracture
839 plane, rather than pure mode II failure, is the natural consequence.

840

841 **3.3. Full 3D modelling of the fracturing process of rock under dynamic loads in SHPB** 842 **tests**

843 In this subsection, the GPGPU-parallelized Y-HFDEM IDE is applied to model the dynamic
844 fracturing of Fangshan marble, which is much more isotropic and homogeneous than granite,
845 in dynamic BTS tests while considering the entire SHPB testing system.

846 The 4th and 5th authors of this paper conducted dynamic BTS tests of Fangshan marble with a
847 SHPB apparatus (Zhang and Zhao 2013). The marble consists of dolomite (98%) and quartz
848 (2%), and the size of the minerals ranges from 10 μm to 200 μm with an average dolomite size

849 of 100 μm and an average quartz size of 200 μm (Zhang and Zhao 2013). The marble can be
850 considered as a homogeneous and isotropic rock, which is ideal to avoid the complexity
851 intrinsic to highly anisotropic rocks such as granite. The detailed procedure of the dynamic
852 BTS test can be found in the literature (Zhang and Zhao 2013), and the test is briefly
853 summarized here. As illustrated in Fig. 15(a), a metal projectile called a striker is first
854 accelerated by a gas gun, and the striker impacts one end of a long cylindrical metal bar called
855 an incident bar (IB). Upon the impact of the striker on the IB, a dynamic compressive strain
856 wave (ϵ_{inci}) is induced in the IB. The ϵ_{inci} propagates toward the other end of the IB, on which
857 the target marble disk is placed. When the ϵ_{inci} arrives at the interface between the IB and the
858 marble disk, some portion is reflected as a tensile strain wave (ϵ_{refl}), and remaining portion is
859 transmitted into the marble disk as a compressive strain wave ($\epsilon_{\text{tans_rock}}$). The $\epsilon_{\text{tans_rock}}$ then
860 propagates toward the interface between the marble disk and one end of another long
861 cylindrical metal bar called a transmission bar (TB). When the $\epsilon_{\text{tans_rock}}$ arrives at the interface,
862 the marble disk is subjected to dynamic loading (i.e., compressed by the IB and the TB). In
863 addition, a compressive strain wave (ϵ_{tans}) generated in the TB propagates toward the other end
864 of the TB. The diameter and thickness of the marble disk used in the experiment were 50 mm
865 and 20 mm, respectively. The lengths of the IB and the TB were 2 m and 1.5 m, respectively,
866 and the diameter of both the IB and the TB was 50 mm. Strain gauges are attached on the
867 surfaces of the IB and TB at 1 m from the interfaces between the marble disk and each bar to
868 measure the time history of the axial strain in the IB (to measure ϵ_{inci} and ϵ_{refl}) and the TB (to
869 measure ϵ_{tans}). Assuming one-dimensional (1D) stress wave propagation in each bar without
870 wave attenuation, the axial stresses on the metal bars (i.e., σ_{inci} , σ_{refl} and σ_{tans}) are calculated by
871 multiplication of the measured axial strains (ϵ_{inci} , ϵ_{refl} and ϵ_{tans}) by the Young's modulus of
872 each bar. In practice, the axial compressive force f_{IB} in the IB is calculated from the
873 superposition of the wave shapes corresponding to σ_{inci} and σ_{refl} , whereas the axial compressive
874 force f_{TB} in the TB is directly calculated from σ_{tans} (Fig. 15(b)). Thus, the axial compressive
875 forces f_{IB} and f_{TB} can be obtained by multiplication of $(\sigma_{\text{inci}} - \sigma_{\text{refl}})$ and σ_{tans} by the cross-sectional
876 areas of each bar. By ensuring that the time histories of the axial compressive forces f_{IB} and f_{TB}
877 are nearly equal up to the peak, the dynamic indirect tensile stress can be defined at the center
878 of the marble disk using the theory applied to the BTS test due to quasi-static loading
879 conditions. Satisfying these conditions is equivalent to achieving dynamic stress equilibrium
880 in the marble disk. The experimental results in Fig. 15(b) satisfy the dynamic stress equilibrium
881 state, and the corresponding loading rate is approximately 830 GPa/s (Zhang and Zhao 2013).
882 The peak value of the dynamic indirect tensile stress is called the dynamic indirect tensile

883 strength. Due to this dynamic indirect tensile stress, the marble disk is dynamically split into
884 two halves due to the formation of blocky fragments near the diametrical center line and
885 numerous shear fractures near the impact region by the IB and the TB. The dynamic indirect
886 tensile strength calculated under these conditions was 32 MPa, which was significantly higher
887 than the quasi-static indirect tensile strength of 9.5 MPa (Zhang and Zhao 2013). The fracture
888 pattern obtained after the test is shown in Fig. 15(c). This paper attempts to model the test
889 condition.

890 Fig. 16 shows the 3D FDEM numerical model for the dynamic BTS test of the marble with
891 the entire SHPB testing apparatus modelled explicitly, which consists of 647,456 nodes,
892 179,022 TET4s and 314,689 CE6s. The average edge length of the mesh for the rock disk is
893 1.3 mm. The experimental evaluation by Brooks et al. (2012) indicates that the sizes of the
894 FPZs for Carrara marble (typical grain size 300 μm) and Danby marble (typical grain size 520
895 μm) are approximately 2~3 mm and 6 mm, respectively, based on nanomechanical and
896 environmental scanning electron microscopic observations. In other words, the size of the FPZ
897 is correlated well with the typical grain size. Thus, it can be estimated that the FPZ size of
898 Fangshan marble is on the order of 1 mm because the grain size of the dominant mineral (i.e.,
899 dolomite) is approximately 100 μm , as described above. Thus, the FDEM mesh with an average
900 edge length of 1.3 mm used to simulate the Fangshan marble should be fine. Moreover, it
901 should be noted that this mesh is finer than the 90,000 TET4s used by Rougier et al. (2014) in
902 a 3D numerical simulation of SHPB-based dynamic BTS tests, and the size of the rock disk is
903 similar. In the model, the striker is not explicitly modelled, and the prescribed velocity
904 corresponding to the impact of the striker is instead applied to the nodes on the left end of the
905 IB (Fig. 16). Once the prescribed velocity reaches zero after the peak, the corresponding nodes
906 are considered to be free nodes; otherwise, an unrealistic stress wave is generated in the IB.
907 The shape of the prescribed velocity profile is determined by the time history of the measured
908 axial strain profile, $\varepsilon_{\text{inci}}$, of the strain gauge on the IB. Based on the measured density (7,697
909 kg/m^3) and dilatational wave speed (5,600 m/s) of the SHPB bars, the dynamic Young's
910 modulus used for the simulation is determined assuming that 1D stress wave theory (i.e., zero
911 Poisson's ratio) is approximately applicable for stress propagation in the IB and the TB. This
912 assumption should be acceptable to the first approximation; otherwise, all of the previous
913 publications on dynamic rock experiments using the SHPB lose their meanings. For the elastic
914 parameters of the marble, homogeneous and isotropic elasticity is assumed in Eq. (2); therefore,
915 the effective elastic stiffness C_{ijkl} in Eq. (2) is merely given by the Lamé constants λ and μ .
916 The measured wave speeds of the dilatational wave (6,000 m/s) and shear wave (2,800 m/s) of

917 the marble along with its measured mass density ($2,800 \text{ kg/m}^3$) are used to determine λ and μ .
918 Thus, the approach for the determination of the elastodynamic parameters is clearly different
919 from that used by Mahabadi et al. (2010) and Rougier et al. (2014), who used the Young's
920 modulus and Poisson's ratio of rock obtained from the quasi-static tests reported by Iqbal and
921 Mohanty (2006) and Broome and Pfeifle (2012), respectively, as the input parameters for their
922 FDEM simulations. In the case of dynamic simulations, the intact stress wave and its wave
923 speed are the most important, and it must be emphasized that the dynamic fracturing is just the
924 outcome of the intact stress wave propagation. Thus, the most reasonable approach to
925 determine the input elastic parameters should be based on the measured wave speeds of the
926 target rock instead of quasi-statically obtained elastic parameters, which often result in
927 incorrect wave speeds. In addition, the values of the artificial penalty terms used for the CE6s
928 are not explicitly given in the publications described above. As mentioned in Sections 2 and
929 3.2, the artificial penalty terms P_f , P_{\tan} , and P_{overlap} of the CE6s must be set reasonably high;
930 otherwise, the input elastic properties lose their meaning. Based on a recent study conducted
931 by the authors (Fukuda et al., 2019), the conditions of $P_{\text{overlap}} = 100E_{\text{dyn}}$ and $P_f = P_{\tan} = 50E_{\text{dyn}}$
932 (E_{dyn} is the dynamic Young's modulus of the marble calculated from the aforementioned
933 measured wave speeds of the marble) are used. The reason for using the smaller penalty terms
934 of the CE6s than those used for the case of quasi-static loading discussed in Section 3.2 is that
935 using the high values can cause spurious modes in dynamic fracturing simulations, whereas the
936 adopted penalty values of the CE6s here ensure that the aforementioned measured wave speeds
937 of the marble result in a negligible numerical error. Following Mahabadi (2012), the contact
938 penalty $P_{\text{con}_n} = 10E_{\text{dyn}}$ is used for the rock surfaces generated by broken CE6s and the surfaces
939 between rocks and loading platens. The damping factor η in Eq. (3) is assumed to be $10^4 \text{ Pa}\cdot\text{s}$
940 to filter out waves with excessively high frequencies that result in spurious modes. It should be
941 noted that even $\eta = \eta_{\text{crit}}$ does not result in significant damping in the time range of interest in
942 the current dynamic simulation, and the effect of η can be considered as a filtering function of
943 high-frequency noise that results in spurious modes. Similar to Rougier et al. (2014), a dynamic
944 Coulomb type friction law is used for the contact friction. The dynamic friction coefficients
945 $\mu_{\text{fric_dyn}}$ on the rock surfaces generated by broken CE6s and the surfaces between rocks and
946 loading platens are assumed to be 0.6 and 0.1, respectively, according to Mahabadi et al. (2010)
947 and Rougier et al. (2014).

948 In the FDEM simulation of dynamic fracturing, the most challenging task is to correctly set
949 the parameters governing the rock fracturing process, and the lack of experimental evidence is
950 still significant, especially for the Mode II parameters. The approach used in some previous

951 studies (Mahabadi et al. 2010; Rougier et al. 2014) for modelling the dynamic fracturing of
952 rock using the FDEM, in which the input parameters such as strengths were obtained from
953 quasi-static tests, should be reviewed carefully. In fact, a significant amount of experimental
954 evidence (e.g., Zhang, 2016) clearly shows that with increasing loading rate, both inter-grain
955 fractures (weaker portions of the rock) and intra-grain fractures (stronger portions of the rock)
956 occur, which indicates that the nominal input strengths for the FDEM simulation should be
957 increased. On the other hand, increasing the values of the critical opening and sliding
958 displacements σ_t and s_t in Eqs. (11) and (12) too much can result in physically unrealistic
959 situations in which cohesive tractions still act on the macroscopic (i.e., almost visible) fracture
960 surfaces, whereas the CZM is mainly for the modelling of the micro-cracking in the FPZ of
961 rocks. In fact, the experimental evidence reported by Sato and Hashida (2006) showed that σ_t
962 is roughly 100 μm in the case of Idate granite. For marble, the linear elastic fracture mechanics
963 relation $G_{\text{fl}} = (K_{\text{IC}})^2/E_{\text{static}}$ can be used to estimate the value of G_{fl} to be 26.5 J/m^2 considering
964 its quasi-static Young's modulus (85 MPa) and quasi-static mode I fracture toughness K_{IC} (1.5
965 $\text{MPa}\cdot\text{m}^{1/2}$). Hence, with the aforementioned quasi-static tensile strength (9.5 MPa), a rough
966 estimate of σ_t in Eq. (11) is 10 μm in the quasi-static loading case. Then, it is simply assumed
967 that the range of σ_t is expected to be on the order of 10 $\mu\text{m} \sim 100 \mu\text{m}$ in the calibration. Because
968 no experimental evidence is available for the value of s_t in Eq. (12) to the best of the authors'
969 knowledge, the relation $\sigma_t = s_t$ is assumed following Rougier et al. (2014). Based on the loading
970 rate independence of φ (e.g., Yao et al. (2017)), the internal friction angle φ is set to 59° from
971 the quasi-static test. Finally, we varied the values of the tensile strength T_s and cohesion c under
972 the conditions of various $\sigma_t (= s_t = 10 \mu\text{m} \sim 100 \mu\text{m})$ so the values of $\varepsilon_{\text{inci}}$, $\varepsilon_{\text{refl}}$ and $\varepsilon_{\text{trans}}$ in the
973 SHPB bars obtained from the FDEM simulation reasonably match those in the experiment
974 (Zhang and Zhao 2013). For the initial guesses of T_s and c , we started from $T_s = 32 \text{ MPa}$ from
975 the dynamic BTS (Zhang and Zhao 2013) and $c = 42.5 \text{ MPa}$ from the dynamic shear strength
976 (Yao et al. 2017), both of which correspond to a loading rate of 830 GPa/s. Here, the
977 experimental evidence is used, in which the dynamic shear strength shows a similar loading
978 rate dependency to the dynamic BTS (i.e., dynamic shear strength $\approx 0.025 \times \text{loading rate} \times$
979 static shear strength (= 21.7 MPa)). By trial and error, we found that the calibrated conditions
980 of $T_s = 30 \text{ MPa}$ and $c = 65 \text{ MPa}$ with $\sigma_t = s_t = 10 \mu\text{m}$ results in a good match, including the
981 obtained fracture pattern between the FDEM simulation and the experiment; therefore, only
982 the simulation results for these conditions are presented. The ECDA approach described in
983 Section 2.2 is not used because it can easily result in spurious modes, as shown in the following
984 paragraphs. Thus, the brute-force contact detection activation approach is used instead; i.e., all

985 of the TET4s corresponding to the marble disk are added to the contact list from the beginning
986 of the simulation. In this case, approximately 8,242,000 contact couples (i.e., a tremendous
987 number of contact force calculations) must be processed in each time step due to the GPGPU
988 parallelization.

989 Figure 17 shows the modelled dynamic fracturing process of Fangshan marble in the SHPB-
990 based dynamic BTS test using the GPGPU-based 3D Y-HFDEM IDE with the calibrated input
991 parameters. The stress wave propagation in the bars is not shown here to save space. In Fig.
992 17, the spatial distribution of σ_{zz} (left column), the macroscopic fracture pattern (damage $D =$
993 1; middle column) and the spatial distribution of D (right column) in the rock disk are shown,
994 in which $t = 0$ is set when the non-negligible strain/stress waves in the IB arrive at the interface
995 between the IB and the rock disk. The warmer and colder colors of σ_{zz} correspond to tensile
996 and compressive stresses, respectively. The impact between the IB and the rock disk results in
997 stress propagation from the IB side of the rock disk toward the TB side ($t = 35 \mu\text{s}$). After the
998 arrival of the stress wave at the interface between the rock disk and the TB, the dynamic indirect
999 tensile stress field begins to develop due to the dynamically increasing loads from both the IB
1000 and TB sides, and the stress field shows approximate symmetries with respect to the y and z
1001 directions across the center of the disk ($t = 70 \mu\text{s}$). It is notable that the micro-cracking is
1002 initiated at this stage in the vicinity of the IB and TB. Then, due to the dynamically induced
1003 indirect tensile stress field, macroscopic fractures (tension-dominant mode I-II fractures) begin
1004 to nucleate away from the center of the disk (i.e., the IB side) and propagate approximately in
1005 the y -direction ($t = 83 \mu\text{s}$). In fact, using the digital image correlation technique, Zhang and
1006 Zhao (2013) evaluated the dynamic strain field that developed on one surface of a marble disk,
1007 which showed that the large strains (i.e., a macroscopic crack) began to develop not from the
1008 exact center of the disk but rather slightly away from the center toward the IB side (see Fig.
1009 11(f) for the result of the digital image correlation data at $48 \mu\text{s}$ in Zhang and Zhao (2013)).
1010 The FDEM modelling shows a similar trend to the experimental observations. Local crack
1011 branching also occurred during microscopic crack propagation from the IB and TB sides ($t =$
1012 $83 \mu\text{s}$) along with the commencement of macroscopic crack propagation from the TB side. It
1013 can be seen that the front of the propagating cracks from the IB and TB are far from flat which
1014 shows another importance of the 3D modelling. In the 2D dynamic simulation conducted by
1015 Mahabadi et al. (2010), Osthus et al. (2018) and Godinez et al. (2018), all the surface waves
1016 on the positive and negative x -planes in Fig. 17 cannot be considered, which makes the
1017 interpretation of the 2D simulations very difficult. Finally, these microscopic and macroscopic
1018 cracks coalesced and formed the resultant splitting macroscopic fracture plane ($t = 98 \mu\text{s}$).

1019 Furthermore, in contrast to the BTS test under quasi-static loading, in which stress can be
1020 released through the opening of splitting fractures, the induced stress in the dynamic BTS test
1021 cannot be released completely in such a short time due to the high-speed loading by the IB and
1022 TB. In addition, the shear-dominant mixed mode I-II fractures result in the formation of
1023 crushed zones near the IB and TB, which contribute to the stress release process. The reason
1024 for the formations of shear-dominant mixed mode I-II fractures in the unstructured mesh was
1025 discussed in Section 3.2. With the splitting fractures interacting with the crushed zones, a
1026 blocky rock fragment is also generated ($t = 98 \mu\text{s}$), which was not found in the BTS test under
1027 quasi-static loading conditions. The comparison of the resultant fracture pattern at $t = 98 \mu\text{s}$
1028 with the experimental pattern in Fig. 15(c) indicates that the FDEM modelling shows a good
1029 correspondence with the experiment in terms of the formation of the splitting central cracks,
1030 the crushed zones near the IB and TB and the blocky rock fragments along the centerline.

1031 Figs. 18(a) and (b) compare the time histories of $\varepsilon_{\text{inci}}$ and $\varepsilon_{\text{refl}}$ in the IB and $\varepsilon_{\text{trans}}$ in the TB
1032 obtained from the experiment and the calibrated FDEM simulation. Because the momentum
1033 bar is not modelled in the FDEM simulation, the profile of $\varepsilon_{\text{trans}}$ in the TB is only shown at the
1034 time when $\varepsilon_{\text{trans}}$ drops to zero after the 1st peak; a comparison between the FDEM simulation
1035 and the experiment after this time has no meaning. The calibration and comparison are
1036 performed using the axial strains instead of the axial stress/force because the authors have
1037 difficulty in interpreting the Young's modulus used for the conversion from the strain to stress
1038 in the IB and TB. In the FDEM, the elastodynamic parameters (i.e., the Lamé constants) that
1039 satisfy the wave speed of the SHPB bars are used; thus, the stress should be interpreted as the
1040 dynamic stress. In contrast, the quasi-statically obtained Young's moduli of the IB and the TB
1041 have conventionally been used for the most SHPB experiments to convert the measured axial
1042 strain to the axial stress. However, if we use the quasi-statically obtained Young's moduli of
1043 the IB and the TB for the FDEM simulation, the wave speed in the FDEM simulation becomes
1044 incorrect. Therefore, the axial strains in the IB and the TB are used for the calibration. The
1045 results show that the FDEM simulation is well calibrated against the measured strain profile,
1046 which demonstrates that the modelled results of the SHPB-based dynamic BTS test from the
1047 GPGPU-parallelized Y-HFDEM agree well with those from the experiment.

1048 Finally, Fig. 19 shows the results of applying the ECDA approach (Guo 2014) described in
1049 Section 2.2 to model the SHPB-based dynamic BTS test. In this model, only the TET4s
1050 corresponding to the surfaces of the IB and the TB as well as the rock disk surface in the vicinity
1051 of the IB and TB are registered as contact candidates at the onset of the FDEM simulation (red
1052 regions in Fig. 19(a)). Then, with the generation of macroscopic cracks, the TET4s in the

1053 vicinity of the macroscopic cracks are adaptively registered as contact candidates. This
1054 approach is much faster in terms of the total run time compared with the case using the brute-
1055 force contact detection activation approach. However, when the same mesh with the same input
1056 parameters except for the ECDA approach is applied, the simulation can easily result in
1057 spurious modes, as shown in Fig. 19(b), in which too many unrealistic fragmentations are
1058 generated. Using a smaller time step Δt may solve this problem. However, the authors could
1059 not find a time step Δt in which the total run time of the ECDA approach is less than that for
1060 the brute-force contact detection activation approach. This spurious mode (i.e., numerical
1061 instability) could be due to the use of a very high value of the internal friction angle as well as
1062 the target loading rate being very high. In fact, our quasi-static simulation presented in Section
1063 3.2 used the ECDA approach, but no unstable results were obtained. Interestingly, this finding
1064 has not been pointed out in any previous studies that used the 3D FDEM and can be regarded
1065 as the most valuable information. For example, in the case of the FDEM simulation of the
1066 “penetration problem”, which is important in impact engineering to understand the dynamic
1067 fracturing of rock, the aforementioned spurious mode can easily occur if the ECDA approach
1068 is used. One of the most serious problems is that the fragmentation due to penetration always
1069 occurs in this kind of simulation; therefore, judicious judgement must be used in evaluating
1070 whether the obtained fragmentation is an artificial fracture due to the spurious mode or not. In
1071 this sense, the brute-force contact detection activation approach can solve or alleviate the
1072 spurious mode.

1073

1074 **3.4 Computing performance of the GPGPU-parallelized 3D FDEM**

1075 This section discusses the computing performance of the GPGPU-parallelized 3D Y-
1076 HFDEM IDE, mainly in terms of its improvement compared with the sequential
1077 implementation of the 3D Y-HFDEM IDE and its performance on several GPGPU accelerators.
1078 To accomplish this goal, the modeling of the rock failure process in the 3D UCS test, as
1079 discussed in subsection 3.2, is selected as a benchmark because it is a computationally
1080 demanding simulation. To make the computation become further intensive, the brute-force
1081 contact detection activation approach introduced in Section 2.2 is used in the 3D modelling of
1082 the UCS test to evaluate the computing performance of the GPGPU-parallelized 3D Y-HFDEM
1083 IDE. Moreover, since the performance of GPGPU-parallelized code is significantly dependent
1084 on the applied GPGPU accelerators, the GPGPU-parallelized 3D Y-HFDEM IDE is run using
1085 several NVIDIA® GPGPU accelerators, i.e., Quadro GV100, Titan V and Quadro GP100 to
1086 investigate its performance. The developed GPGPU-parallelized 3D Y-HFDEM IDE can be

1087 run in all these GPGPU accelerators without any modifications. At the same time, an Intel®
1088 Xeon® Silver 4112 processor (2.60 GHz and 32.0 RAM) is used to run our sequential CPU-
1089 based 3D Y-HFDEM IDE. Table 2 shows the number of TET4s, CE4s and nodes, and the
1090 initial number of contact couples in each model. It is evident that the mesh discretization with
1091 the average element size $h_{ave} = 1.3$ mm results in tremendously massive computation. The list
1092 of actual runtimes required for 500 calculation time steps is shown at the bottom of Table 2 for
1093 several values of h_{ave} , in the cases of the GPGPU-based code using the Quadro GV100
1094 accelerator and the sequential CPU-based code (i.e. the original Y3D code). The results show
1095 that 134,943 s (37.5 hours) are required to solve the 500 time steps in the sequential CPU-based
1096 code for $h_{ave} = 1.3$ mm, which means that solving the problem with this level of fine
1097 discretization is too computationally expensive using the sequential code. Fig. 20 shows the
1098 speed-up times of the GPGPU-parallelized 3D Y-HFDEM IDE relative to the CPU sequential
1099 code running on a single thread for the 3D modelling of the UCS test with 695,428 TET4s and
1100 187,852 CE6s. In Fig. 20, the vertical axis shows the quotients of the total run time using each
1101 GPGPU accelerator divided by that using the CPU sequential code, which, thus, correspond to
1102 the speed-up times of the GPGPU-parallelized 3D Y-HFDEM IDE relative to the CPU
1103 sequential code. Clearly, the GPGPU-parallelized 3D Y-HFDEM IDE run on all GPU
1104 accelerators has achieved significant speed-up times compared with the CPU sequential code,
1105 and the Quadro GV100 accelerator shows the maximum speed-up times of 284, which is even
1106 better than the maximum speed-up times of 128.6 achieved by the authors' GPGPU-
1107 parallelized 2D Y-HFDEM IDE (Fukuda et al., 2019) and much higher than the maximum
1108 speed-up times of 100 achieved by Irazu's parallelization using OpenCL (Lisjak et al., 2018).
1109 Moreover, the relative speed-up times between the GPGPU-parallelized code and the CPU
1110 sequential code depend on the number of TET4s used in the numerical model, which are
1111 illustrated in Fig. 21(a). As can be seen from Fig. 21 (a), the relative speed-up times initially
1112 increase with the number of TET4s increasing, which reveals that, in this case, keeping all the
1113 GPGPU cores busy is the most important factor in achieving the best performance of the
1114 GPGPU-parallelized code. However, after the number of TET4s increases to a certain value,
1115 which is about 430,000 TET4s in this study, the relative speed-up times slightly decrease with
1116 the number of TET4s increasing. In other words, there is a model limit of the maximum speed-
1117 up times for a single GPU accelerator to achieve the best performance and multiple GPU
1118 accelerators may be needed to lift the limit. Compared with the limit of the model size of
1119 294,840 TET4s in Irazu's parallelization using OpenCL (Lisjak et al., 2018), the limit of the
1120 model size for the maximum speed-up times of the GPGPU-parallelized 3D FDEM increases

1121 significantly. Therefore, as long as currently available publications are concerned, our 3D
1122 GPGPU-parallelized FDEM using CUDA has much better performance than Irazu's
1123 parallelization using OpenCL in terms of both the maximum speed-up times and the limits of
1124 the model sizes. Furthermore, no performance report of CUDA version of Irazu-software is
1125 unavailable in any publications. In addition, as illustrated in Fig. 21 (b), the computation time
1126 of the GPGPU-parallelized 3D Y-HFDEM IDE linearly increases with respect to the number
1127 of TET4s. Therefore, the computational complexity of the GPGPU-parallelized 3D Y-HFDEM
1128 IDE is $O(N)$, i.e. the amount of computation is proportional to the number of elements, which
1129 proves the high computational efficiency of the GPGPU-parallelized 3D Y-HFDEM IDE. The
1130 implementation of the hyper separating theorem into the contact detection algorithm (e.g.
1131 Lisjak et al., 2018) may be able to further enhance the computational performance of the
1132 GPGPU-parallelized 3D Y-HFDEM IDE.

1133 **4. Conclusion and future work**

1134 This paper developed a general-purpose graphics-processing-unit (GPGPU) - parallelized
1135 combined finite-discrete element method (FDEM) based on the authors' former sequential two-
1136 dimensional (2D) and three-dimensional (3D) Y-HFDEM IDE codes using compute unified
1137 device architecture (CUDA) C/C++. The algorithm of the developed 3D GPGPU-parallelized
1138 code was first presented in detail, which can provide a basis for further improvement and
1139 progress of the FDEM codes reviewed in the introduction section based on GPGPU
1140 parallelization. It should be noted that a different contact detection algorithm from that used in
1141 the sequential code was implemented in the 3D GPGPU-parallelized code because the
1142 algorithm in the sequential code is not suitable for GPGPU parallelization. Contact damping
1143 and contact friction were then implemented and verified although further verification may be
1144 needed. After that, the GPGPU-parallelized 3D Y-HFDEM IDE code was applied to 3D
1145 modelling of the failure process of limestone in a uniaxial compression strength (UCS) test and
1146 a Brazilian tensile strength (BTS) test to demonstrate its capability in modelling rock
1147 engineering applications under quasi-static loading conditions. The 3D modelling results
1148 demonstrate that for the FDEM simulations with the cohesive zone model (CZM) using
1149 unstructured meshes, mixed-mode I-II failures are the dominant failure mechanisms along the
1150 shear and splitting failure planes in the UCS and BTS tests, respectively, whereas pure mode I
1151 failure along the splitting failure plane in the BTS test and pure mode II failure along the shear
1152 failure in the UCS test are only possible in models with structured meshes. Moreover,
1153 compared with 2D models of the BTS test, new insights about the nucleation locations of

1154 macroscopic tensile splitting cracks were gained from the 3D model of the BTS test. Moreover,
1155 the GPGPU-parallelized 3D Y-HFDEM IDE code was implemented to model dynamic
1156 fracturing of a relatively isotropic and homogeneous marble in a split Hopkinson pressure bar
1157 (SHPB) - based dynamic Brazilian test to investigate its applicability in modelling rock
1158 engineering problems under dynamic loading conditions. Thanks to the GPGPU
1159 parallelization, the entire SHPB system was modelled using the 3D Y-HFDEM IDE code. The
1160 physical-mechanical parameters and computing parameters, including the penalty terms and
1161 dynamic strengths, were carefully selected for the 3D FDEM simulation, and the limitations of
1162 previous studies in the parameter selection were discussed. The modelled failure process, final
1163 fracture pattern and time histories of the dynamic compressive strain wave, reflective tensile
1164 strain wave and transmitted compressive strain wave were compared with those from
1165 experiments, and good agreements were achieved between them. In addition, the spurious
1166 fracturing mode in the form of unrealistic fragmentation, which can occur when the efficient
1167 contact detection activation approach is used, is highlighted because it has not been pointed out
1168 in previous FDEM studies. Finally, the computing performance of the GPGPU-parallelized 3D
1169 Y-HFDEM IDE run on various GPGPU accelerators was discussed against that of the
1170 sequential CPU-based 3D Y-HFDEM and other GPGPU parallelization of 2D FDEM using
1171 OpenCL since the computing performance of the GPGPU parallelization of 3D FDEM using
1172 OpenCL is not available in any publications.

1173 The following conclusions can be drawn from this study:

1174 – A GPGPU-parallelized 3D Y-HFDEM IDE code was developed to model the fracturing
1175 process of rock under quasi-static and dynamic loading conditions. In addition to GPGPU
1176 parallelization, robust contact detection, contact damping and contact friction were
1177 implemented in the 3D Y-HFDEM IDE code and were verified and validated through a series
1178 of numerical simulations under quasi-static and dynamic loading conditions. The computing
1179 performance analysis shows the GPGPU-parallelized 3D HFDEM IDE code is 284 times faster
1180 than its sequential version and can achieve the computational complexity of $O(N)$.

1181 – The 3D models of the failure processes of limestone in UCS and BTS tests demonstrated
1182 the capability of the GPGPU-parallelized 3D HFDEM IDE code in simulating rock engineering
1183 applications under quasi-static loading conditions. Moreover, important findings and new
1184 insights were obtained from the 3D modelling: 1) the selection of penalty terms for cohesive
1185 elements in any FDEM simulation with the intrinsic cohesive zone model (ICZM) is crucial to
1186 reasonably model rock continuous behaviour before fracturing, which has been overlooked in
1187 the literature. 2) For all FDEM simulations with the CZM using unstructured meshes, mixed-

1188 mode I-II failures are the dominant failure mechanisms along the shear and splitting failure
1189 planes in the UCS and BTS tests, respectively, whereas pure mode I failure along the splitting
1190 failure plane in the BTS test and pure mode II failure with shear failure in the UCS test are only
1191 possible with structured meshes. Because rock is a collection of mineral grains, whose structure
1192 generally corresponds to that of an unstructured mesh in FDEM simulations, the results
1193 indicated that the mixed-mode I-II failures may be the possible failure mechanisms in
1194 experiments of the UCS and BTS tests. 3) Compared with 2D models of the BTS test, the 3D
1195 model of the BTS test results in different stress distributions along the loading diametrical lines
1196 in the middle plane and on the surface of the Brazilian disk as well as different nucleation
1197 locations of macroscopic tensile splitting cracks.

1198 – The GPGPU-parallelized 3D Y-HFDEM IDE code can consider the entire SHPB testing
1199 system and successfully model the dynamic fracturing of marble in the SHPB-based dynamic
1200 BTS test. The modelled failure process, final fracture pattern and time histories of the dynamic
1201 compressive strain wave, reflective tensile strain wave and transmitted compressive strain
1202 wave are compared with those from experiments, and good agreements are achieved between
1203 them. Therefore, the 3D modelling of dynamic fracturing of marble in the SHPB-based
1204 dynamic BTS test demonstrates the capability of the GPGPU-parallelized 3D Y-HFDEM IDE
1205 code in simulating rock engineering applications under dynamic loading conditions.

1206 Therefore, with careful calibration and insights, the FDEM, including the newly developed
1207 GPGPU-parallelized 3D Y-HFDEM IDE code in this study, is a valuable and powerful
1208 numerical tool for investigating the failure process of rock under quasi-static and dynamic
1209 loading conditions in rock engineering applications although very fine elements with maximum
1210 element size no bigger than the length of the fracture process zone must be used in the area
1211 where fracturing process is modelled.

1212

1213 **Acknowledgements**

1214 This work was supported by JSPS KAKENHI for Grant-in-Aid for Young Scientists (Grant
1215 Numbers 18K14165) for the first author. The corresponding author would like to acknowledge
1216 the support of Australia-Japan Foundation (AJF Grant No. 17/20470), Australia Academy of
1217 Science (AAS Grant No. RI8) and Institution Research Grant Scheme (IRGS Grant No.
1218 L0018929) and Natural Science Foundation of China (NSFC Grant Nos. 51574060 and No.
1219 51079017), in which the corresponding author is the international collaborator. Moreover, all
1220 authors would like to thank the editor-in-chief, i.e. Prof G. Barla, and the three anonymous

1221 reviewers for their helpful and constructive comments that significantly contributed to
1222 improving the final version of the paper.

1223

1224 **Conflicts of Interest**

1225 The first author and the corresponding author declare that the received funds do not lead to any
1226 conflicts of interest regarding the publication of this manuscript.

1227

1228 **References**

1229 An B, Tannant DD (2007) Discrete element method contact model for dynamic simulation of
1230 inelastic rock impact *Computers & Geosciences* 33:513-521

1231 doi:10.1016/j.cageo.2006.07.006

1232 An HM, Liu HY, Han H, Zheng X, Wang XG (2017) Hybrid finite-discrete element
1233 modelling of dynamic fracture and resultant fragment casting and muck-piling by

1234 rock blast *Computers and Geotechnics* 81:322-345 doi:

1235 10.1016/j.compgeo.2016.09.007

1236 Ayachit U (2015) *The ParaView Guide: A Parallel Visualization Application*. Kitware, Inc.,

1237 Batinić M, Smoljanović H, Munjiza A, Mihanović A (2018) GPU based parallel FDEM for
1238 analysis of cable structures *Građevinar* 69:1085-1092

1239 Brooks Z, Ulm FJ, Einstein HH (2012) Role of Microstructure Size In Fracture Process Zone
1240 Development of Marble. Paper presented at the 46th U.S. Rock

1241 Mechanics/Geomechanics Symposium, Chicago, Illinois, 2012/1/1/

1242 Broome ST, Pfeifle TW, Sussman AJ. (2012) Phase1 and 2 mechanical property test results
1243 for borehole U-15n in support of NCNS source physics experiment. National Centre

1244 for Nuclear Security Briefing.

1245 Burtscher M, Pingali K (2011) An Efficient CUDA Implementation of the Tree-Based
1246 Barnes Hut n-Body Algorithm. In: *GPU Computing Gems Emerald Edition*.

1247 doi:10.1016/B978-0-12-384988-5.00006-1.

1248 Camacho GT, Ortiz M (1996) Computational modelling of impact damage in brittle materials
1249 *International Journal of Solids and Structures* 33:2899-2938 doi:10.1016/0020-

1250 7683(95)00255-3

1251 Elmo D, Stead D (2010) An Integrated Numerical Modelling–Discrete Fracture Network
1252 Approach Applied to the Characterisation of Rock Mass Strength of Naturally

1253 Fractured Pillars *Rock Mechanics and Rock Engineering* 43:3-19

1254 doi:10.1007/s00603-009-0027-3

1255 Erarslan N, Liang ZZ, Williams DJ (2012) Experimental and Numerical Studies on
1256 Determination of Indirect Tensile Strength of Rocks *Rock Mechanics and Rock*

1257 *Engineering* 45:739-751 doi:10.1007/s00603-011-0205-y

1258 Fairhurst C (1964) On the validity of the ‘Brazilian’ test for brittle materials *International*
1259 *Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts* 1:535-

1260 546 doi: 10.1016/0148-9062(64)90060-9

1261 Fukuda D, Mohammadnejad M, Liu HY, Dehkhoda S, Chan A, Cho SH, Min GJ, Han H,
1262 Kodama J, Fujii Y. (2019) Development of a GPGPU-parallelized hybrid finite-

1263 discrete element method for modelling rock fracture. *International Journal for*

1264 *Numerical and Analytical Methods in Geomechanics*. DOI: 10.1002/nag.2934

1265 Fukuda D, Nihei E, Cho SH, Kodama J, Fujii Y Comparison of FEM-based 3-D dynamic
1266 fracturing simulations using intrinsic and extrinsic cohesive zone models. In:

1267 Proceedings of 2017 ISRM Young Scholars' Symposium on Rock Mechanics (YSRM
1268 2017) & NDRMGE 2017, 2017. International Society for Rock Mechanics, pp 100-
1269 104

1270 Godinez HC, Rougier E, Osthus D, Lei Z, Knight E, Srinivasan G (2018) Fourier amplitude
1271 sensitivity test applied to dynamic combined finite-discrete element methods-based
1272 simulations International Journal for Numerical and Analytical Methods in
1273 Geomechanics doi:10.1002/nag.2852

1274 Guo L (2014) Development of a three-dimensional fracture model for the combined finite-
1275 discrete element method. Imperial College London

1276 Hamdi P, Stead D, Elmo D (2014) Damage characterization during laboratory strength
1277 testing: A 3D-finite-discrete element approach Computers and Geotechnics 60:33-46
1278 doi:10.1016/j.compgeo.2014.03.011

1279 Heinze T, Jansen G, Galvan B, Miller SA (2016) Systematic study of the effects of mass and
1280 time scaling techniques applied in numerical rock mechanics simulations
1281 Tectonophysics 684:4-11 doi:10.1016/j.tecto.2015.10.013

1282 Hondros G (1959) The evaluation of Poisson's ratio and the modulus of materials of low
1283 tensile resistance by the Brazilian (Indirect tensile) test with particular reference to
1284 concrete Australian J Appl Sci 10:243-268

1285 Iqbal MJ, Mohanty B (2006) Experimental Calibration of ISRM Suggested Fracture
1286 Toughness Measurement Techniques in Selected Brittle Rocks Rock Mechanics and
1287 Rock Engineering 40:453-475 doi:10.1007/s00603-006-0107-6

1288 Lei Q, Latham J-P, Tsang C-F (2017) The use of discrete fracture networks for modelling
1289 coupled geomechanical and hydrological behaviour of fractured rocks Computers and
1290 Geotechnics 85:151-176 doi:10.1016/j.compgeo.2016.12.024

1291 Lei Z, Rougier E, Knight EE, Munjiza A (2014) A framework for grand scale parallelization
1292 of the combined finite discrete element method in 2d Computational Particle
1293 Mechanics 1:307-319 doi:10.1007/s40571-014-0026-3

1294 Li D, Wong LNY (2013) The Brazilian Disc Test for Rock Mechanics Applications: Review
1295 and New Insights Rock Mechanics and Rock Engineering 46:269-287
1296 doi:10.1007/s00603-012-0257-7

1297 Lisjak A, Grasselli G (2014) A review of discrete modeling techniques for fracturing
1298 processes in discontinuous rock masses Journal of Rock Mechanics and Geotechnical
1299 Engineering 6:301-314 doi: 10.1016/j.jrmge.2013.12.007

1300 Lisjak A, Kaifosh P, He L, Tatone BSA, Mahabadi OK, Grasselli G (2017) A 2D, fully-
1301 coupled, hydro-mechanical, FDEM formulation for modelling fracturing processes in
1302 discontinuous, porous rock masses Computers and Geotechnics 81:1-18 doi:
1303 10.1016/j.compgeo.2016.07.009

1304 Lisjak A, Mahabadi OK, He L, Tatone BSA, Kaifosh P, Haque SA, Grasselli G (2018)
1305 Acceleration of a 2D/3D finite-discrete element code for geomechanical simulations
1306 using General Purpose GPU computing Computers and Geotechnics 100:84-96 doi:
1307 10.1016/j.compgeo.2018.04.011

1308 Liu HY, Han H, An HM, Shi JJ (2016) Hybrid finite-discrete element modelling of asperity
1309 degradation and gouge grinding during direct shearing of rough rock joints
1310 International Journal of Coal Science & Technology 3:295-310 doi:10.1007/s40789-
1311 016-0142-1

1312 Liu HY, Kang YM, Lin P (2015) Hybrid finite-discrete element modeling of geomaterials
1313 fracture and fragment muck-piling International Journal of Geotechnical Engineering
1314 9:115-131 doi:10.1179/1939787913Y.0000000035

1315 Liu Q, Deng P (2019) A numerical investigation of element size and loading/unloading rate
1316 for intact rock in laboratory-scale and field-scale based on the combined finite-
1317 discrete element method. *Engineering Fracture Mechanics* 211: 442-462

1318 Lukas T, Schiava D'Albano GG, Munjiza A (2014) Space decomposition based
1319 parallelization solutions for the combined finite–discrete element method in 2D
1320 *Journal of Rock Mechanics and Geotechnical Engineering* 6:607-615 doi:
1321 10.1016/j.jrmge.2014.10.001

1322 Ma G, Zhang Y, Zhou W, Ng T-T, Wang Q, Chen X (2018) The effect of different fracture
1323 mechanisms on impact fragmentation of brittle heterogeneous solid *International*
1324 *Journal of Impact Engineering* 113:132-143 doi:10.1016/j.ijimpeng.2017.11.016

1325 Mahabadi O, Kaifosh P, Marschall P, Vietor T (2014) Three-dimensional FDEM numerical
1326 simulation of failure processes observed in Opalinus Clay laboratory samples *Journal*
1327 *of Rock Mechanics and Geotechnical Engineering* 6:591-606 doi:
1328 10.1016/j.jrmge.2014.10.005

1329 Mahabadi O, Lisjak A, He L, Tatone B, Kaifosh P, Grasselli G Development of a new fully-
1330 parallel finite-discrete element code: Irazu. In: 50th US Rock
1331 Mechanics/Geomechanics Symposium, 2016. American Rock Mechanics
1332 Association,

1333 Mahabadi OK (2012) Investigating the Influence of Micro-scale Heterogeneity and
1334 Microstructure on the Failure and Mechanical Behaviour of Geomaterials. University
1335 of Toronto

1336 Mahabadi OK, Cottrell BE, Grasselli G (2010) An Example of Realistic Modelling of Rock
1337 Dynamics Problems: FEM/DEM Simulation of Dynamic Brazilian Test on Barre
1338 Granite *Rock Mechanics and Rock Engineering* 43:707-716 doi:10.1007/s00603-010-
1339 0092-7

1340 Mahabadi OK, Lisjak A, Munjiza A, Grasselli G (2012) Y-Geo: New Combined Finite-
1341 Discrete Element Numerical Code for Geomechanical Applications *International*
1342 *Journal of Geomechanics* 12:676-688 doi:10.1061/(asce)gm.1943-5622.0000216

1343 Mohammadnejad M, Liu H, Chan A, Dehkhoda S, Fukuda D (2018) An overview on
1344 advances in computational fracture mechanics of rock *Geosystem Engineering*:1-24
1345 doi:10.1080/12269328.2018.1448006

1346 Mohammadnejad M, Liu H, Dehkhoda S, Chan A (2017) Numerical Investigation of
1347 Dynamic Rock Fragmentation in Mechanical Cutting Using Combined FEM/DEM.
1348 Paper presented at the 3rd Nordic Rock Mechanics Symposium - NRMS 2017,
1349 Helsinki, Finland, 2017/11/3/

1350 Munjiza A (2004) *The Combined Finite-Discrete Element Method*. Wiley,

1351 Munjiza A, Andrews KRF, White JK (1999) Combined single and smeared crack model in
1352 combined finite-discrete element analysis *International Journal for Numerical*
1353 *Methods in Engineering* 44:41-57 doi:10.1002/(SICI)1097-
1354 0207(19990110)44:1<41::AID-NME487>3.0.CO;2-A

1355 Munjiza AA, Knight EE, Rougier E (2011) *Computational mechanics of discontinua*. John
1356 Wiley & Sons

1357 Munjiza A, Knight E, Rougier E (2015) *Large Strain Finite Element Method: A Practical*
1358 *Course*. Wiley,

1359 Munjiza A, Owen DRJ, Bicanic N (1995) A combined finite-discrete element method in
1360 transient dynamics of fracturing solids. *Engineering Computations* 12(2): 145-174.
1361 <https://doi.org/10.1108/02644409510799532>

1362 Munjiza A, Xiang J, Garcia X, Latham JP, D'Albano GGS, John NWM (2010) The Virtual
1363 Geoscience Workbench, VGW: Open Source tools for discontinuous systems
1364 *Particuology* 8:100-105 doi:10.1016/j.partic.2009.04.008

1365
1366 Munshi A, Gaster B, Mattson TG, Fung J, Ginsburg D (2011) OpenCL Programming Guide.
1367 Addison-Wesley Professional,
1368 NVIDIA (2018) Cuda C Programming Guide. <http://docs.nvidia.com/cuda/>.
1369 Osthus D, Godinez HC, Rougier E, Srinivasan G (2018) Calibrating the stress-time curve of a
1370 combined finite-discrete element method to a Split Hopkinson Pressure Bar
1371 experiment International Journal of Rock Mechanics and Mining Sciences 106:278-
1372 288 doi:10.1016/j.ijrmms.2018.03.016
1373 Rockfield (2005) Rockfield Software Ltd. <http://www.rockfield.co.uk/elfen.htm>.
1374 Rogers S, Elmo D, Webb G, Catalan A (2015) Volumetric Fracture Intensity Measurement
1375 for Improved Rock Mass Characterisation and Fragmentation Assessment in Block
1376 Caving Operations Rock Mechanics and Rock Engineering 48:633-649
1377 doi:10.1007/s00603-014-0592-y
1378 Rougier E, Knight EE, Broome ST, Sussman AJ, Munjiza A (2014) Validation of a three-
1379 dimensional Finite-Discrete Element Method using experimental results of the Split
1380 Hopkinson Pressure Bar test International Journal of Rock Mechanics and Mining
1381 Sciences 70:101-108 doi:10.1016/j.ijrmms.2014.03.011
1382 Rougier E, Knight EE, Sussman AJ, Swift RP, Bradley CR, Munjiza A, Broome ST (2011)
1383 The Combined Finite-Discrete Element Method Applied to the Study of Rock
1384 Fracturing Behavior In 3D. 2011/1/1/
1385 Satish N, Harris M, Garland M Designing efficient sorting algorithms for manycore GPUs.
1386 In: Parallel & Distributed Processing, 2009. IPDPS 2009. IEEE International
1387 Symposium on, 2009. IEEE, pp 1-10
1388 Sato K, Hashida T (2006) Fracture Toughness Evaluation Based on Tension-softening Model
1389 and its Application to Hydraulic Fracturing Pure and Applied Geophysics 163:1073-
1390 1089 doi:10.1007/s00024-006-0066-6
1391 Solidity (2017) Solidity Project. <http://solidityproject.com/>.
1392 Tatone BSA, Grasselli G (2015) A calibration procedure for two-dimensional laboratory-
1393 scale hybrid finite–discrete element simulations International Journal of Rock
1394 Mechanics and Mining Sciences 75:56-72 doi: 10.1016/j.ijrmms.2015.01.011
1395 Tijssens MGA, Sluys BLJ, van der Giessen E (2000) Numerical simulation of quasi-brittle
1396 fracture using damaging cohesive surfaces. European Journal of Mechanics –
1397 A/Solids 19(5): 761-779. [https://doi.org/10.1016/S0997-7538\(00\)00190-X](https://doi.org/10.1016/S0997-7538(00)00190-X)
1398 Xiang J, Latham J-P, Farsi A Algorithms and Capabilities of Solidity to Simulate Interactions
1399 and Packing of Complex Shapes. In, Singapore, 2016. Proceedings of the 7th
1400 International Conference on Discrete Element Methods. Springer Singapore, pp 139-
1401 149
1402 Xiang J, Munjiza A, Latham J-P, Guises R (2009) On the validation of DEM and FEM/DEM
1403 models in 2D and 3D Engineering Computations 26:673-687
1404 Yan C, Jiao Y-Y (2018) A 2D fully coupled hydro-mechanical finite-discrete element model
1405 with real pore seepage for simulating the deformation and fracture of porous medium
1406 driven by fluid Computers & Structures 196:311-326 doi:
1407 10.1016/j.compstruc.2017.10.005
1408 Yan C, Zheng H (2017) FDEM-flow3D: A 3D hydro-mechanical coupled model considering
1409 the pore seepage of rock matrix for simulating three-dimensional hydraulic fracturing
1410 Computers and Geotechnics 81:212-228 doi: 10.1016/j.compgeo.2016.08.014
1411 Yao W, He T, Xia K (2017) Dynamic mechanical behaviors of Fangshan marble Journal of
1412 Rock Mechanics and Geotechnical Engineering 9:807-817
1413 doi:10.1016/j.jrmge.2017.03.019

1414 Zhang L, Quigley SF, Chan AHC (2013) A fast scalable implementation of the two-
1415 dimensional triangular Discrete Element Method on a GPU platform *Advances in*
1416 *Engineering Software* 60-61:70-80 doi: 10.1016/j.advensoft.2012.10.006
1417 Zhang QB, Zhao J (2013) Determination of mechanical properties and full-field strain
1418 measurements of rock material under dynamic loads *International Journal of Rock*
1419 *Mechanics and Mining Sciences* 60:423-439 doi:10.1016/j.ijrmms.2013.01.005
1420 Zhang QB, Zhao J (2014) A review of dynamic experimental techniques and mechanical
1421 behaviour of rock materials *Rock Mechanics and Rock Engineering* 47(4):1411-1478
1422 doi:10.1007/s00603-013-0463-y
1423 Zhang Z, Paulino G, Celes W (2007) Extrinsic cohesive modelling of dynamic fracture and
1424 microbranching instability in brittle materials *International Journal for Numerical*
1425 *Methods in Engineering* 72:893-923 doi:10.1002/nme.2030
1426

1427 **Table caption**

1428 **Table 1.** Physical-mechanical properties of rock and computational parameters used in the
1429 FDEM simulations of the UCS and BTS tests (Subsection 3.2).

1430 **Table 2.** Model details for several h_{ave} values.

1431

1432 **Figure captions**

1433 **Fig. 1** Two types of finite elements (TET4 and CE6) and the mechanical behaviors of CE6
1434 during the failure process: (a) two TET4s surrounding a CE6, (b) tensile/shear softening curves
1435 in the ICZM.

1436 **Fig. 2.** Elastic–inelastic power function model for contact damping (modified after An and
1437 Tannant (2007))

1438 **Fig. 3.** Concept of the CUDA programming model using the abstractions of “*threads*”,
1439 “*blocks*” and “*grid*”.

1440 **Fig. 4.** Concept of massively parallel computation for each CUDA “*kernel*” for particular
1441 computational purposes.

1442 **Fig. 5.** Flowchart of the GPGPU-parallelized 3D Y-HFDEM IDE code.

1443 **Fig. 6.** Concept of hash values assigned to each cubic sub-cell (left) and a TET4 included in a
1444 particular sub-cell (right).

1445 **Fig. 7.** GPGPU-based contact detection algorithm implemented in the GPGPU-parallelized
1446 3D Y-HFDEM IDE code.

1447 **Fig. 8.** Contact damping verification: (a) 3D model configuration (modified after Mahabadi et
1448 al. (2012)), (b) comparison between numerical and theoretical results.

1449 **Fig. 9.** Contact friction verification: (a) 3D model configuration (modified after Xiang et al.
1450 (2009)), (b) comparison between numerical and theoretical results.

1451 **Fig. 10.** 3D numerical models for analysis of the fracturing process of rock under quasi-static
1452 loading. (a) UCS test, (b) BTS test.

1453 **Fig. 11.** 3D modelling of the UCS test under quasi-static loading. (a) Initiation and
1454 propagation of microscopic cracks before the peak stress, (b) unstable crack propagation at
1455 the peak stress, (c) post-failure fracture pattern, and (d) axial stress versus axial strain curve.

1456 **Fig. 12.** Comparison of the final rock fracture patterns in the UCS test from the numerical
1457 simulation and experiment: (a) comparison of the final fracture patterns, (b) final fracture
1458 pattern with pure mode II damages highlighted, (c) final fracture pattern with mixed mode I-
1459 II damages highlighted.

1460 **Fig. 13.** 3D simulation of the fracturing process of rock in the BTS test under quasi-static
1461 loading: (a) distributions of the horizontal stress and microscopic damage before the peak
1462 stress, (b) distributions of the horizontal stress, microscopic damage and macroscopic cracks
1463 at the peak stress, (c) distributions of the horizontal stress and macroscopic fracture pattern in
1464 the post-failure stage, (d) Brazilian indirect tensile stress versus axial strain curve and (e)
1465 comparison between the simulated stress distributions and those from Hondros' solution
1466 along the loading diametrical line in the middle plane and on the surface of the disk.

1467 **Fig. 14.** Modelled fracture patterns of rock in the BTS tests with unstructured and structured
1468 meshes. (a) Dominant mixed mode I-II fractures in the BTS test with an unstructured mesh,
1469 (a) dominant pure mode I fracture in the BTS test with a structured mesh, (c) schematic
1470 sketch of the tensile fracturing mechanism in a structured mesh, (d) schematic sketch of the
1471 tensile fracturing mechanism in an unstructured mesh.

1472 **Fig. 15.** Overview of the SHPB-based dynamic BTS test simulated by the GPGPU-
1473 parallelized 3D Y-HFDEM IDE code. (a) Configuration of the SHPB system (after Fig. 3 in
1474 Zhang and Zhao (2013) with minor additions), (b) achievement of dynamic stress equilibrium
1475 between the axial forces in the IB and the TB (after Fig. 11(d) in Zhang and Zhao (2013) with
1476 minor additions) and (c) failure patterns of the marble specimen at a dynamic loading rate of
1477 830 GPa/s (after Fig. 12(b) in Zhang and Zhao (2013)).

1478 **Fig. 16.** 3D FDEM model for the SHPB-based dynamic BTS test with velocity boundary
1479 conditions applied on the one end of the IB.

1480 **Fig. 17.** 3D modelling of the dynamic fracturing process of marble in the SHPB-based
1481 dynamic BTS test: distribution of σ_{zz} (left column), distribution of macroscopic cracks
1482 (damage $D = 1$) (middle column) and distribution of damage D (right column). Note that $t = 0$
1483 is set when the stress wave in the IB arrives at the interface between the IB and the rock disk.

1484 **Fig. 18.** Comparison between numerically simulated and experimentally measured axial

1485 strains (ϵ_{inci} , ϵ_{refl} and ϵ_{trans}) in the IB and the TB.

1486 **Fig. 19.** Spurious fracture caused when the ECDA approach is used instead of the brute-force
 1487 contact detection activation approach. (a) Highlighted contact candidates from the onset of
 1488 the simulation for the ECDA, (b) unrealistic fragmentation due to the ECDA approach.

1489 **Fig. 20.** Relative speed-up of the GPGPU-parallelized 3D Y-HFDEM IDE compared with the
 1490 sequential CPU-based code.

1491 **Fig. 21** Computing performance of the GPGPU-parallelized 3D Y-HFDEM IDE: (a) Variation
 1492 of relative speed-up with the increase of the number of TET4s and (b) Variation of
 1493 computation time with the increase of the number of TET4s.

1494
 1495
 1496
 1497
 1498

Table 1

<i>Parameter</i>	<i>Unit</i>	<i>Value</i>
Density (ρ_{rock})	kg/m ³	1,800
Young's modulus (E_{rock})	GPa	12.2
Poisson's ratio (ν_{rock})	-	0.25
Tensile strength (T_{s_rock})	MPa	1.2
Cohesion (C_{rock})	MPa	4.2
Internal friction angle (ϕ_{rock})	°	25
Mode I fracture energy (G_{fl_rock})	J/m ²	2
Mode II fracture energy (G_{fll_rock})	J/m ²	26.5
Normal contact penalty (P_{n_con})	GPa	122
Artificial stiffness penalty (P_{open} , P_{tan})	GPa	1,220
Artificial stiffness penalty ($P_{overlap}$)	GPa	12,200

1499

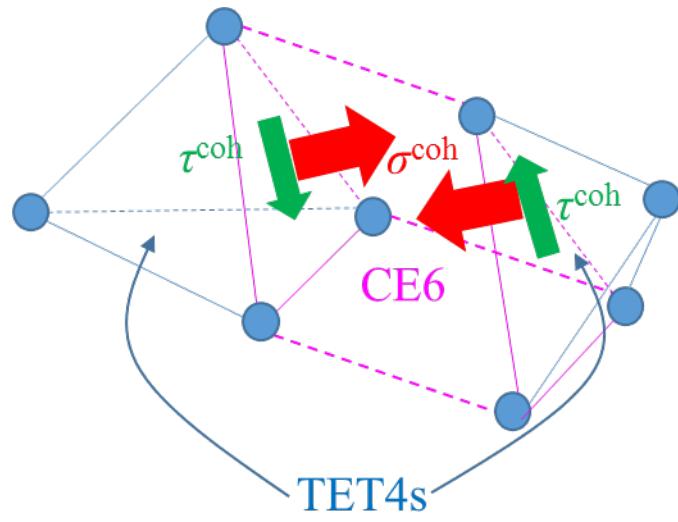
1500

1501 Table 2

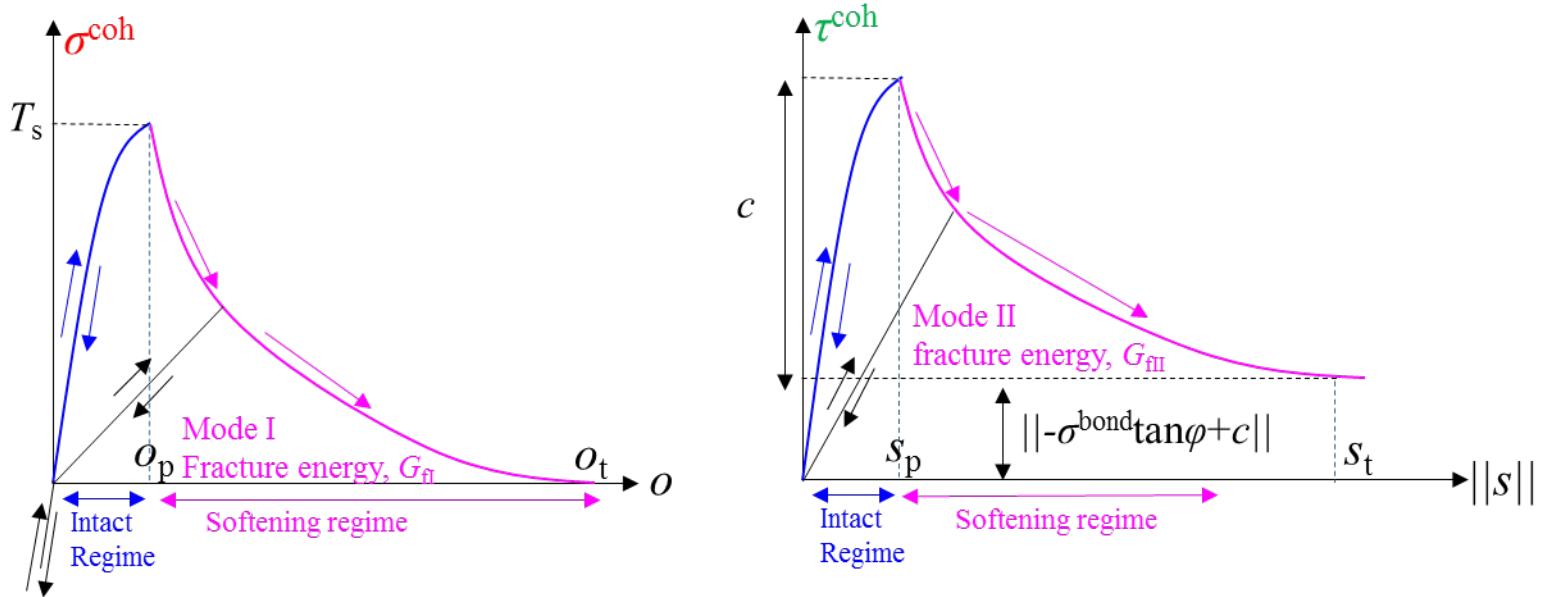
$h_{ave}(mm)$	1.3	1.5	1.9	2.9	3.9	8
The number of nodes	200,525	129,244	82,382	31,327	19,471	1,741
The number of TET4s	1,106,054	695,428	430,290	146,554	81,932	7,383
The number of CE6s	2,113,582	1,298,343	773,545	213,792	86,740	12,266
The initial number of contact couples.	65,447,009	40,561,982	24,753,231	7,881,957	4,093,472	384,582

Simulation time (sec) for sequential CPU- based code /500 steps	134,943	82,755	51,598	16,563	8,318	8,318
Simulation time (sec) for GPUGU-based code (Quadro GV100) /500 steps	491	291	178	59	34	8

1502



(a)



(b)

Fig. 1.

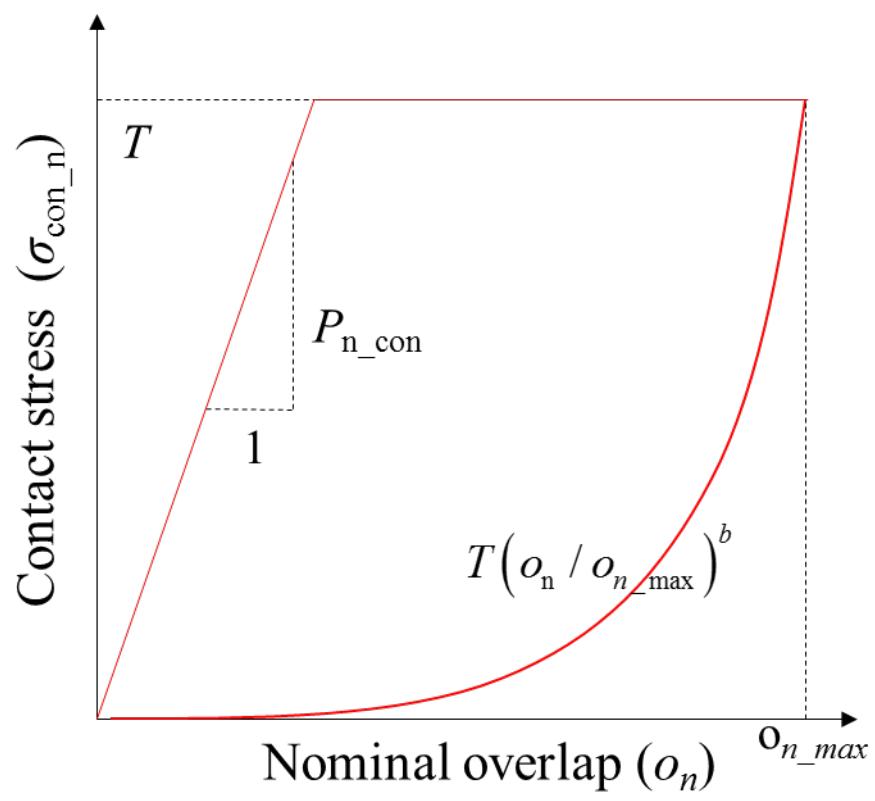


Fig. 2.

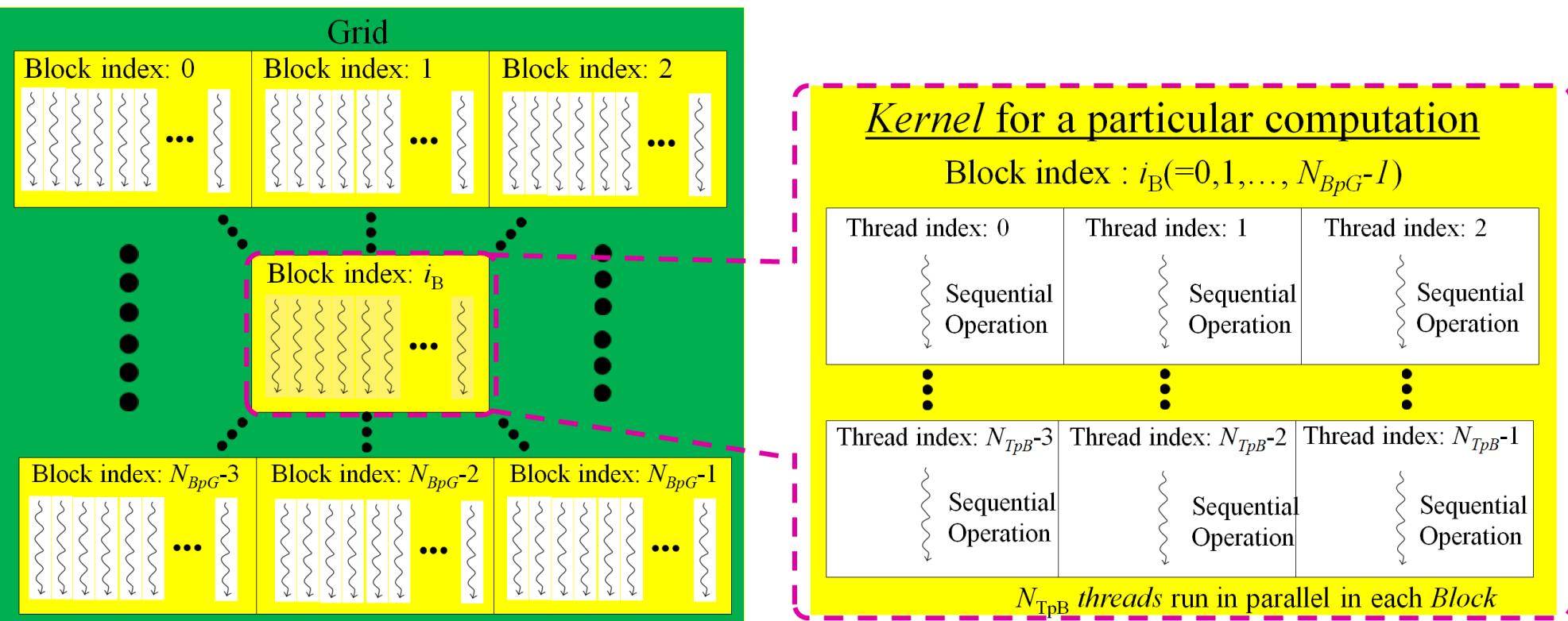


Fig. 3.

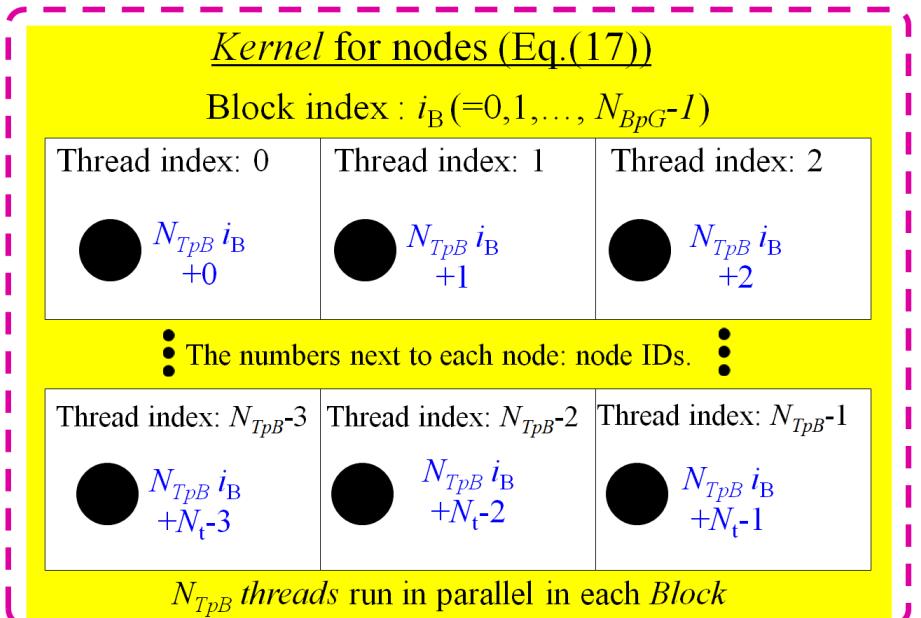
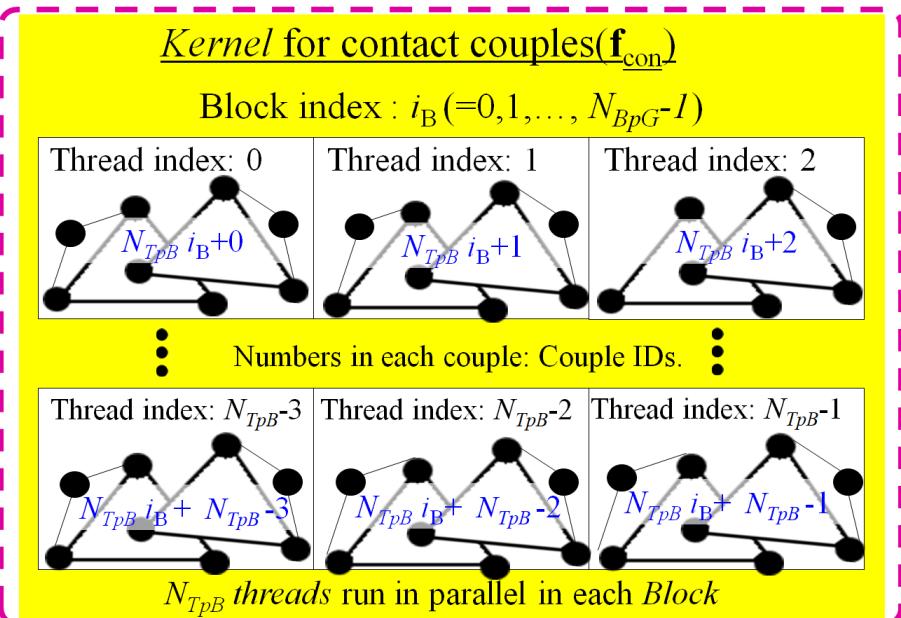
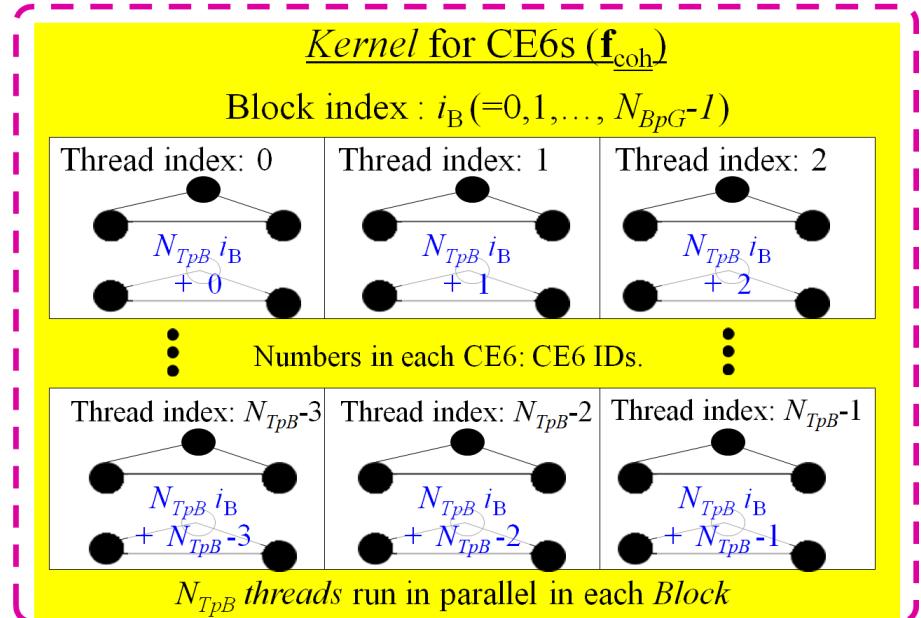
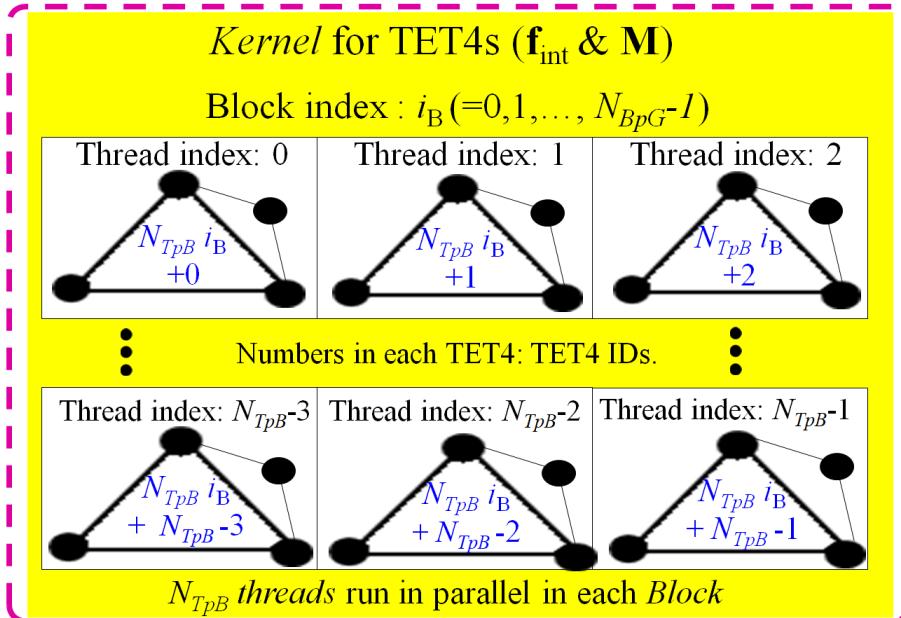


Fig. 4.

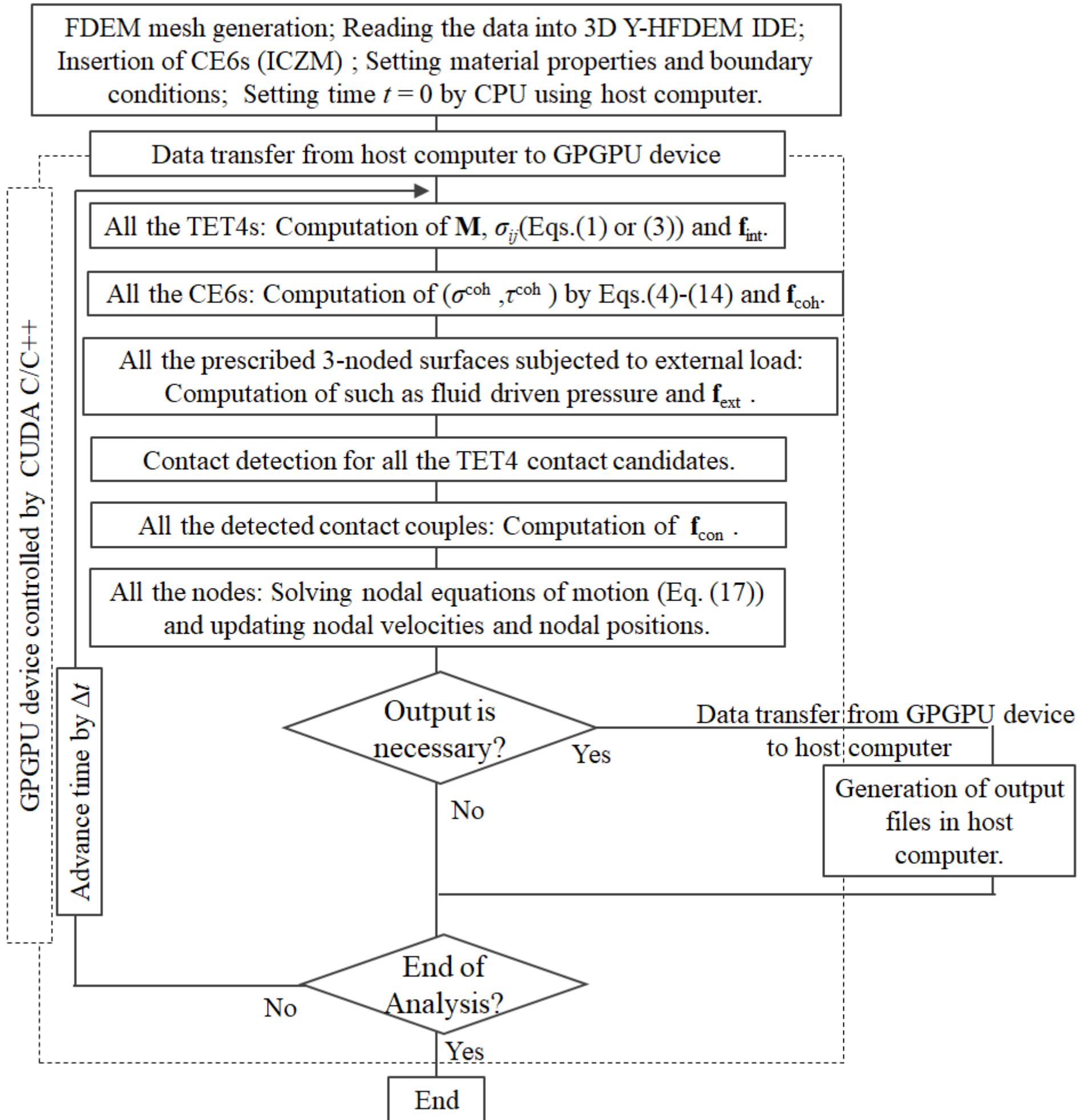


Fig. 5.

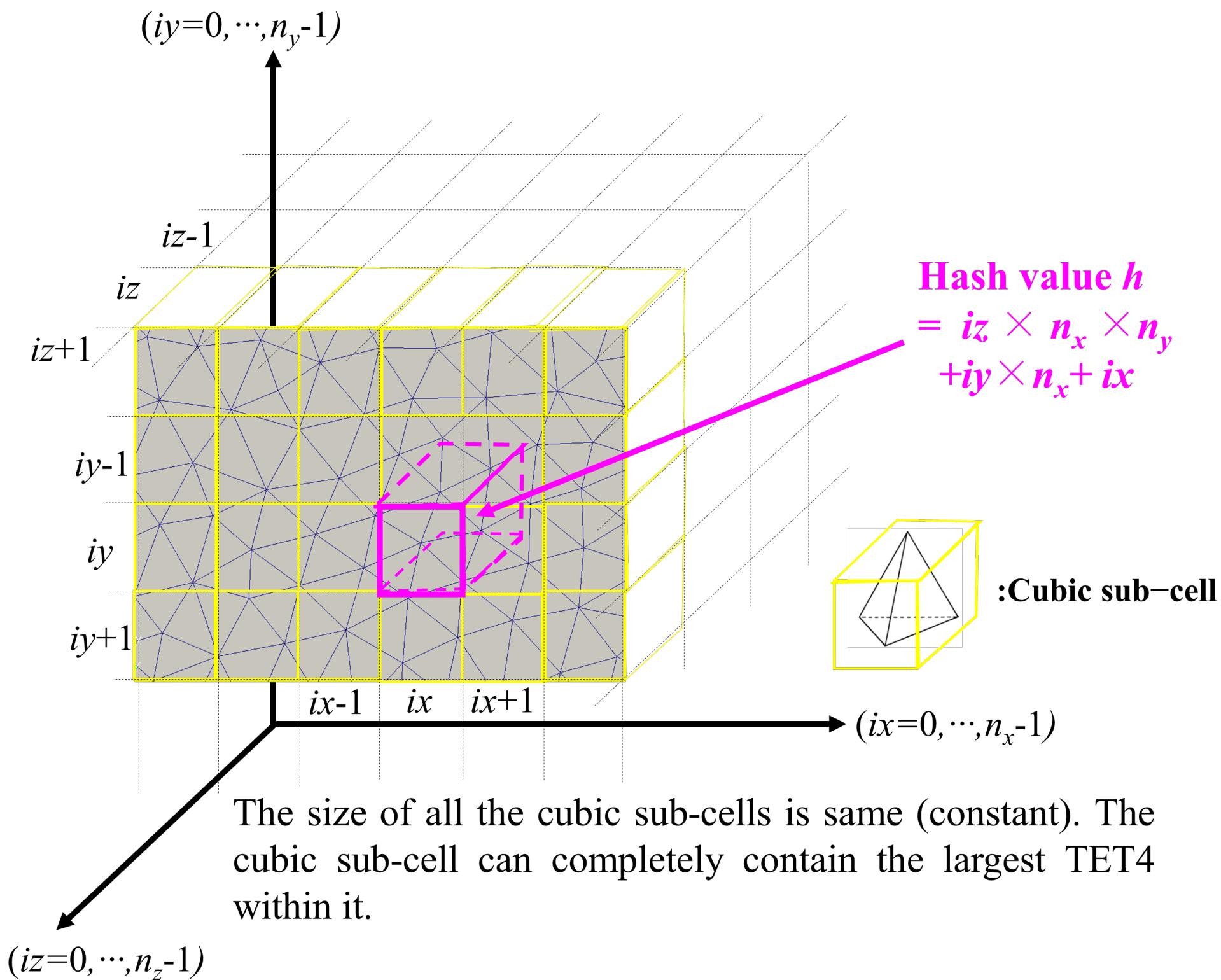
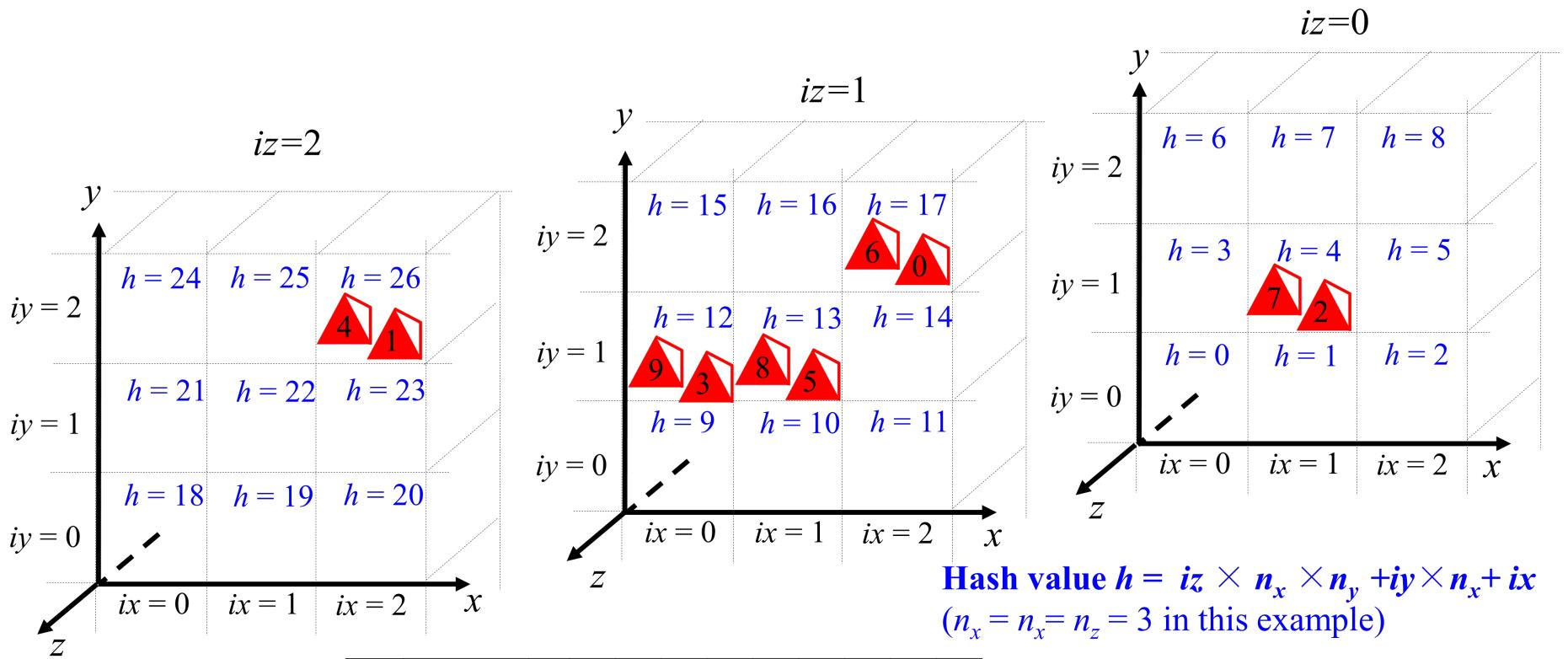


Fig. 6.



List L-1	TET4 ID List	0	1	2	3	4	5	6	7	8	9
	<i>h</i>	17	26	4	12	26	13	17	4	13	12



Radix sorting of TET4 IDs using h as key.

List L-2	Cell index	0	1	2	3	4	5	6	7	8	9
	Sorted TET4 ID List	2	7	3	9	5	8	0	6	1	4
	Sorted <i>h</i>	4	4	12	12	13	13	17	17	26	26

List L-3	<i>h</i>	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
	First cell Index	E	E	E	E	0	E	E	E	E	E	E	E	2	4	E	E	E	6	E	8							
	Last cell Index	E	E	E	E	1	E	E	E	E	E	E	E	3	5	E	E	E	7	E	9							

E: Empty(There are no TET4s for the value of h)

Fig. 7.

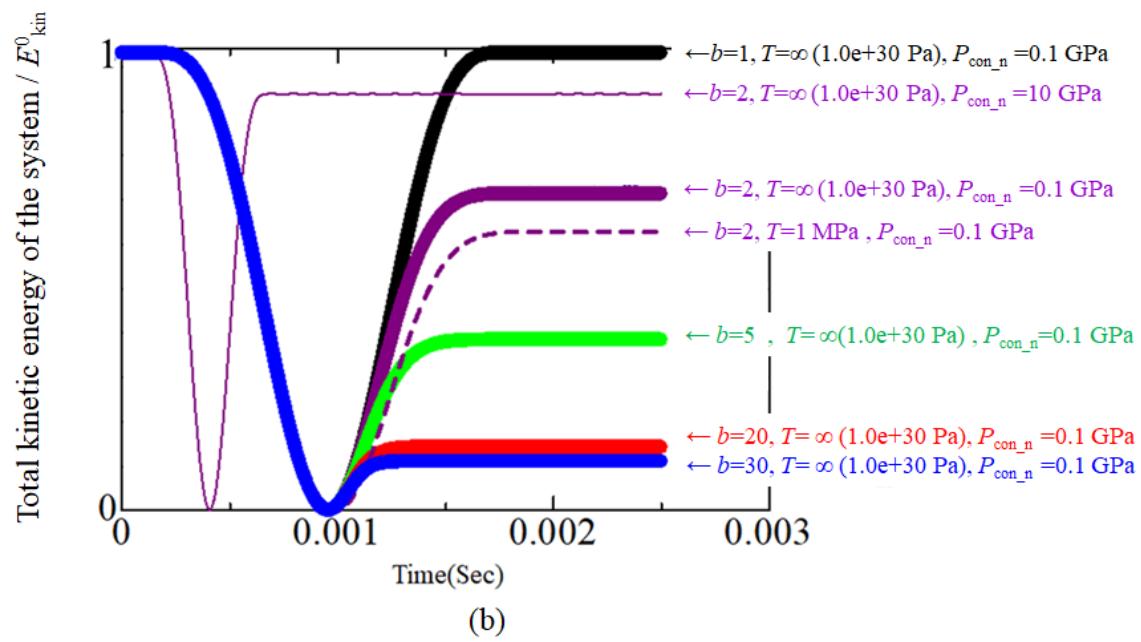
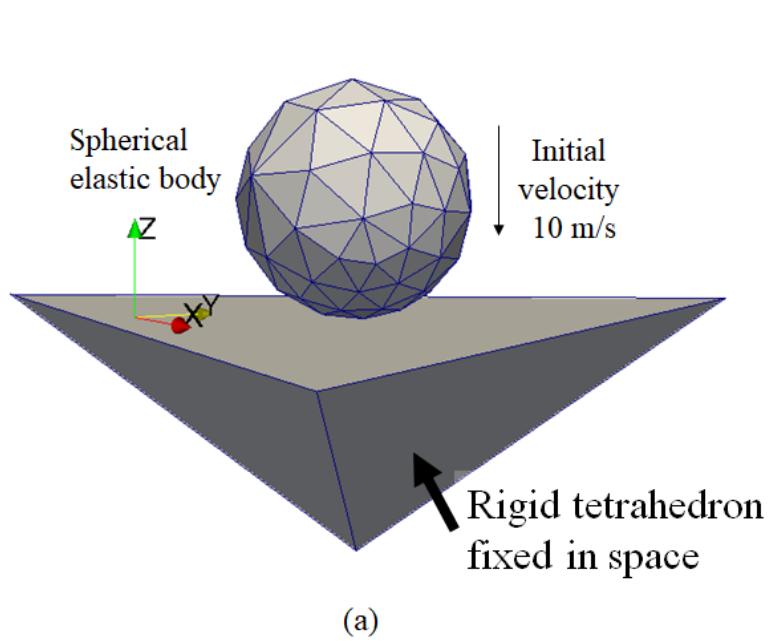
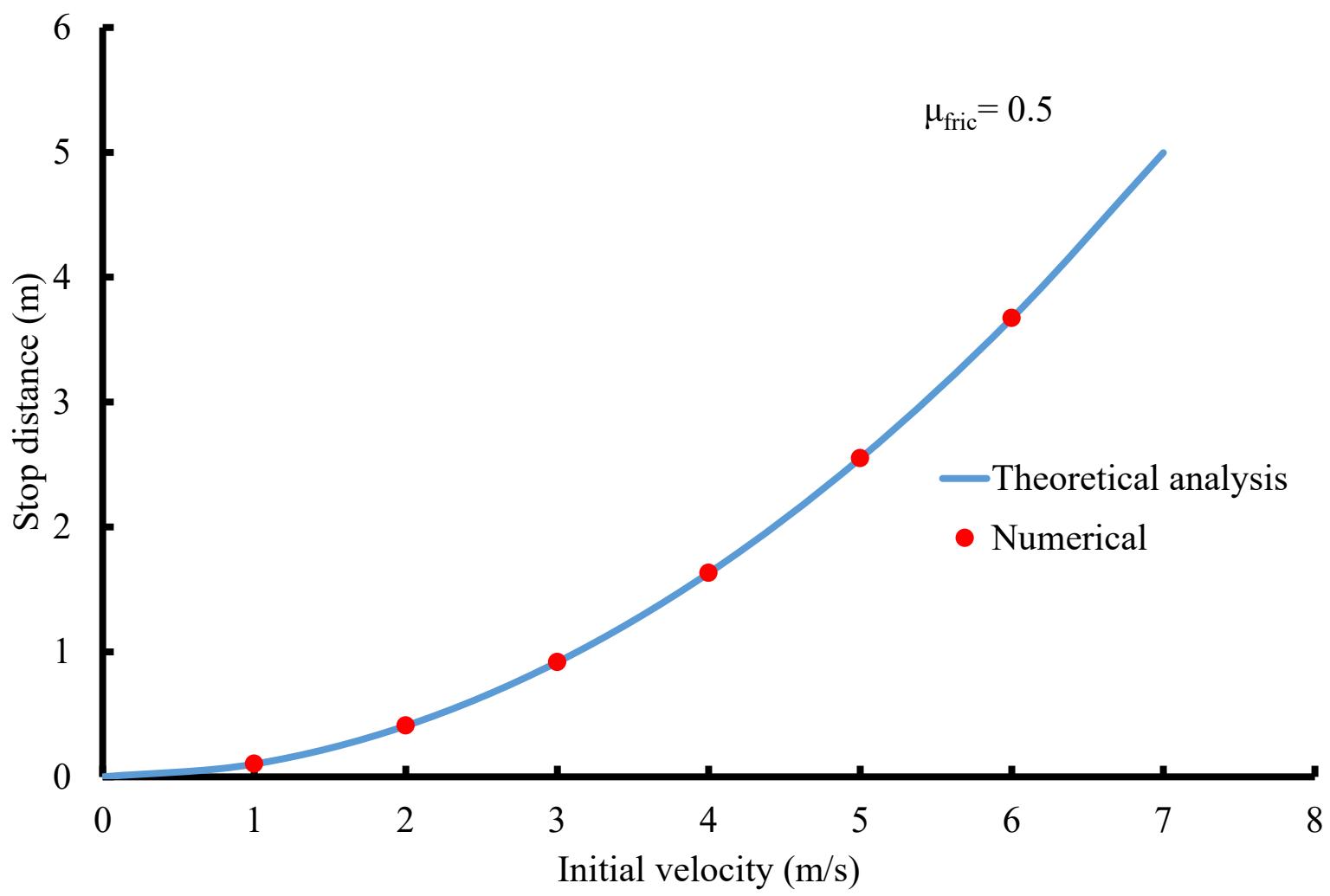
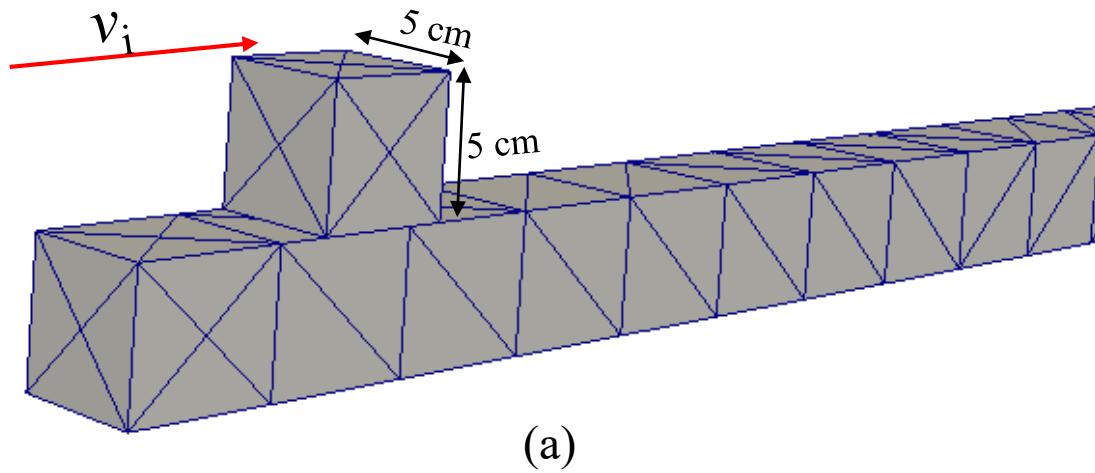
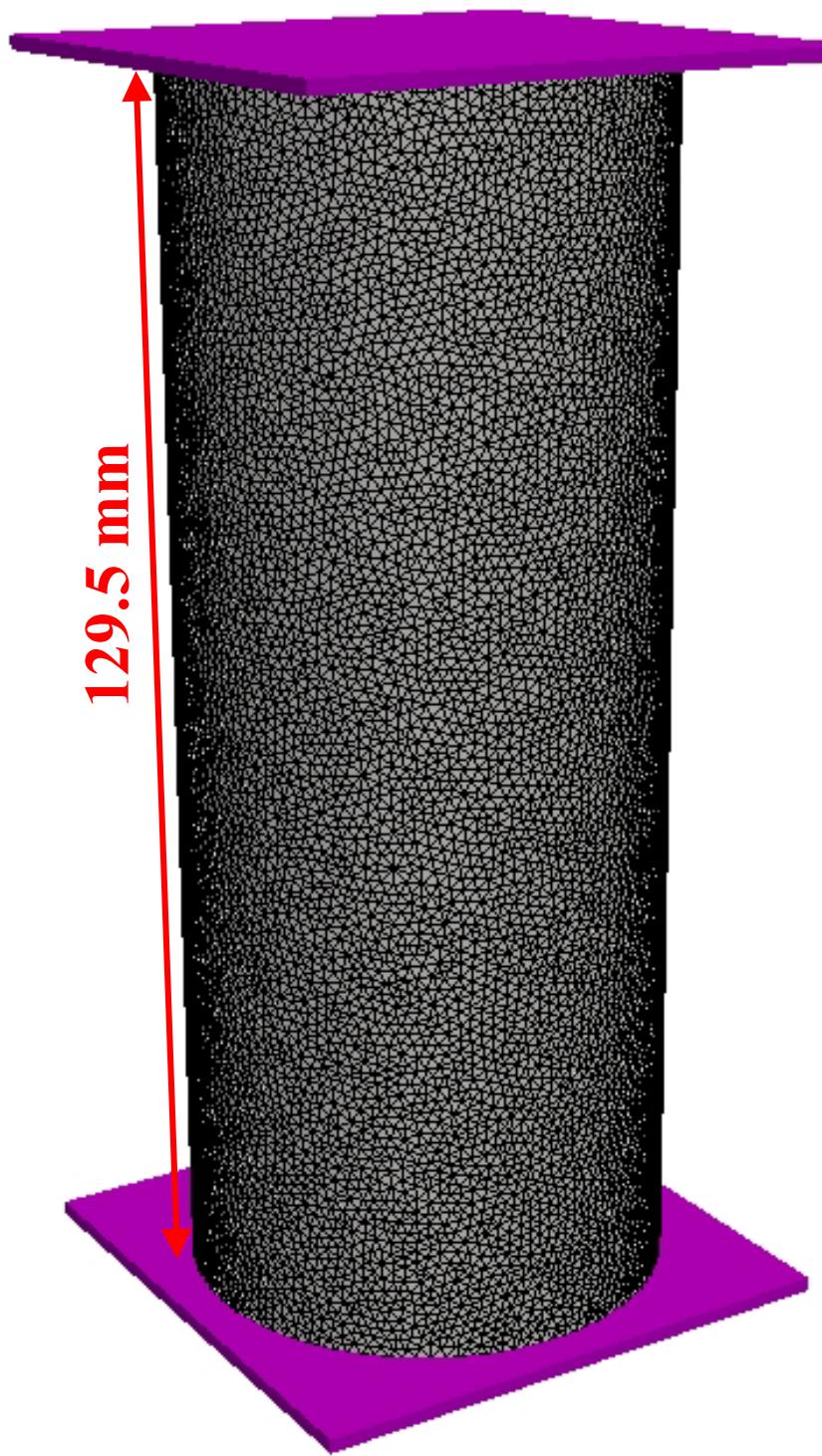


Fig. 8.

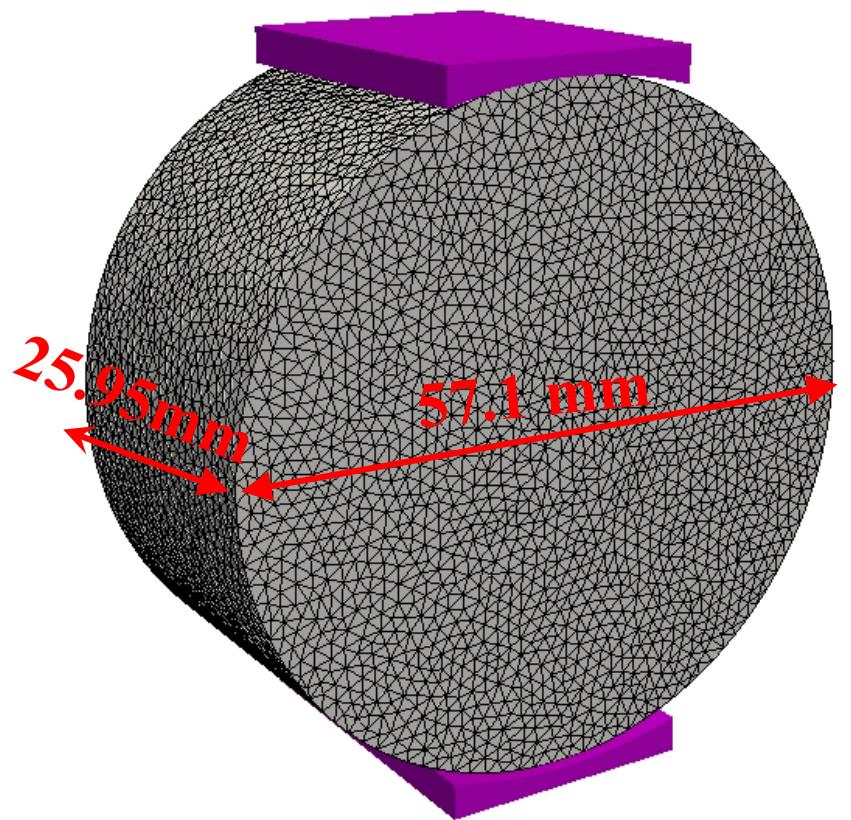


(b)

Fig. 9.

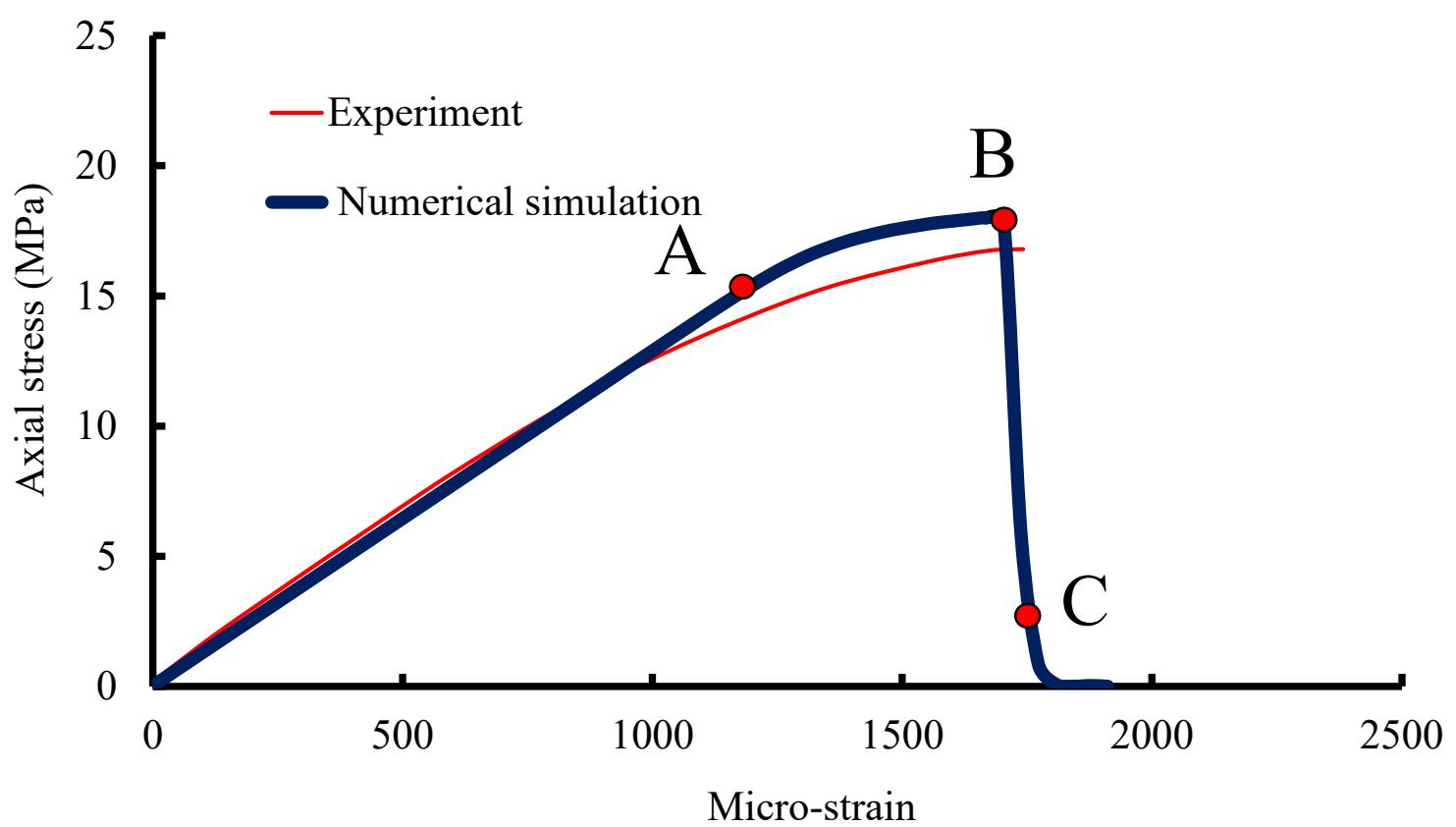
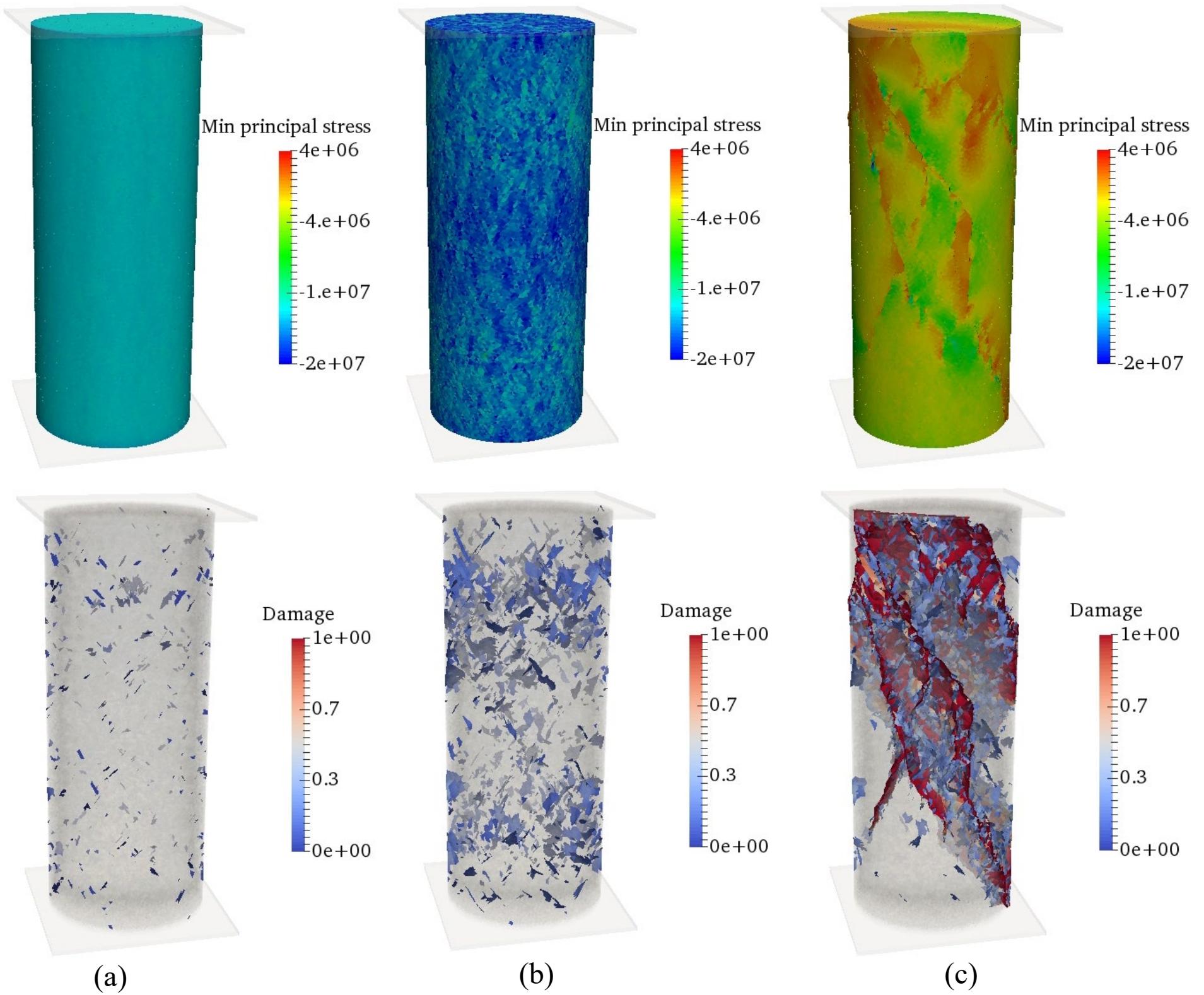


(a)



(b)

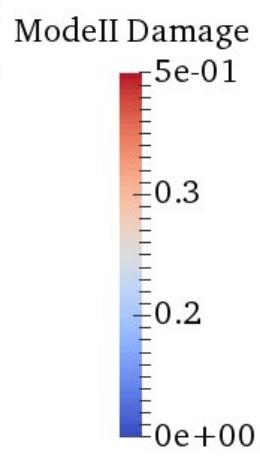
Fig. 10.



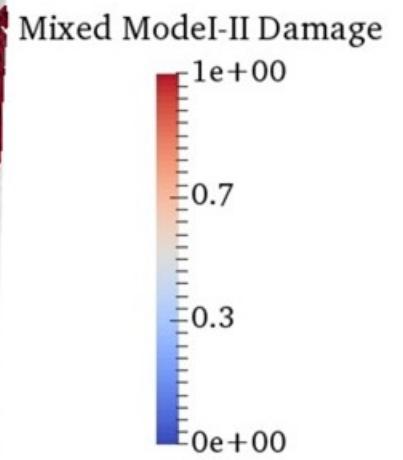
(d)
Fig. 11.



(a)

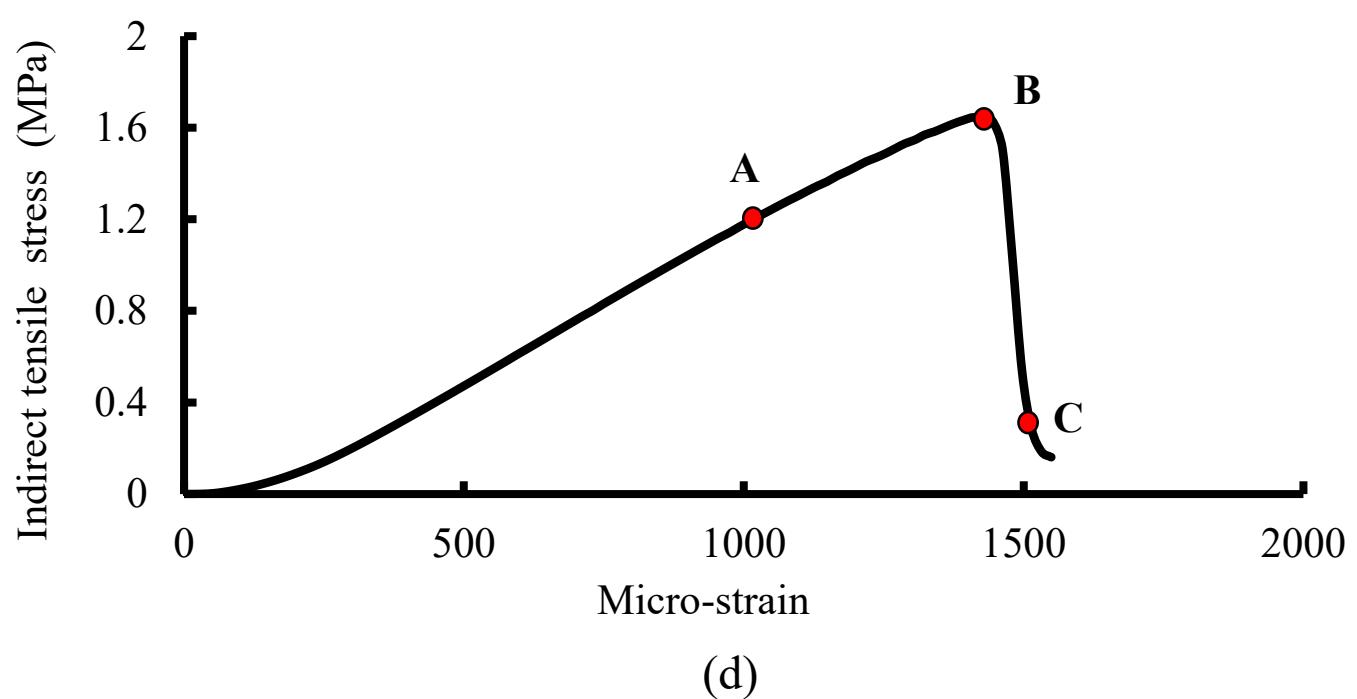
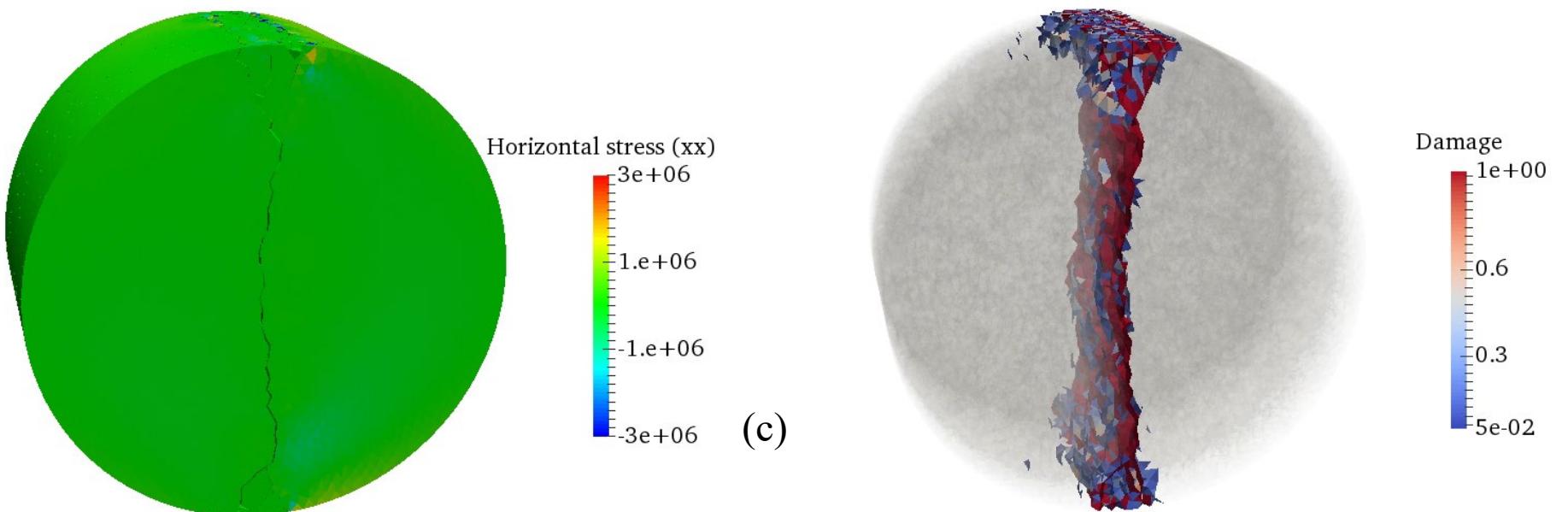
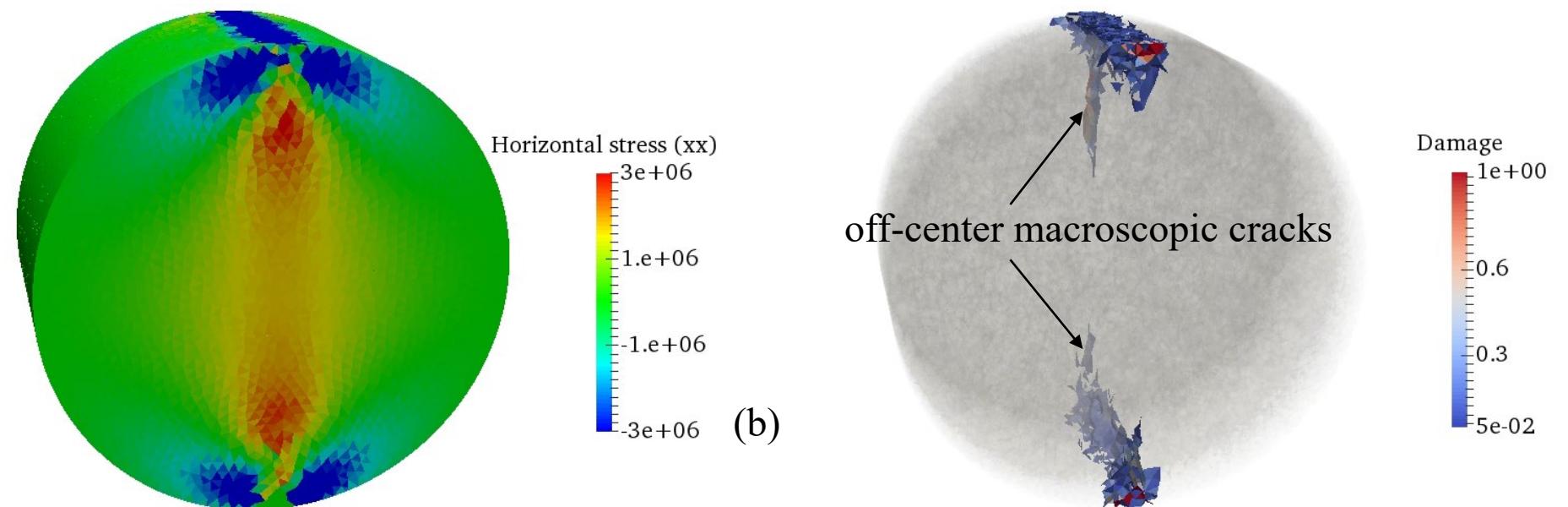
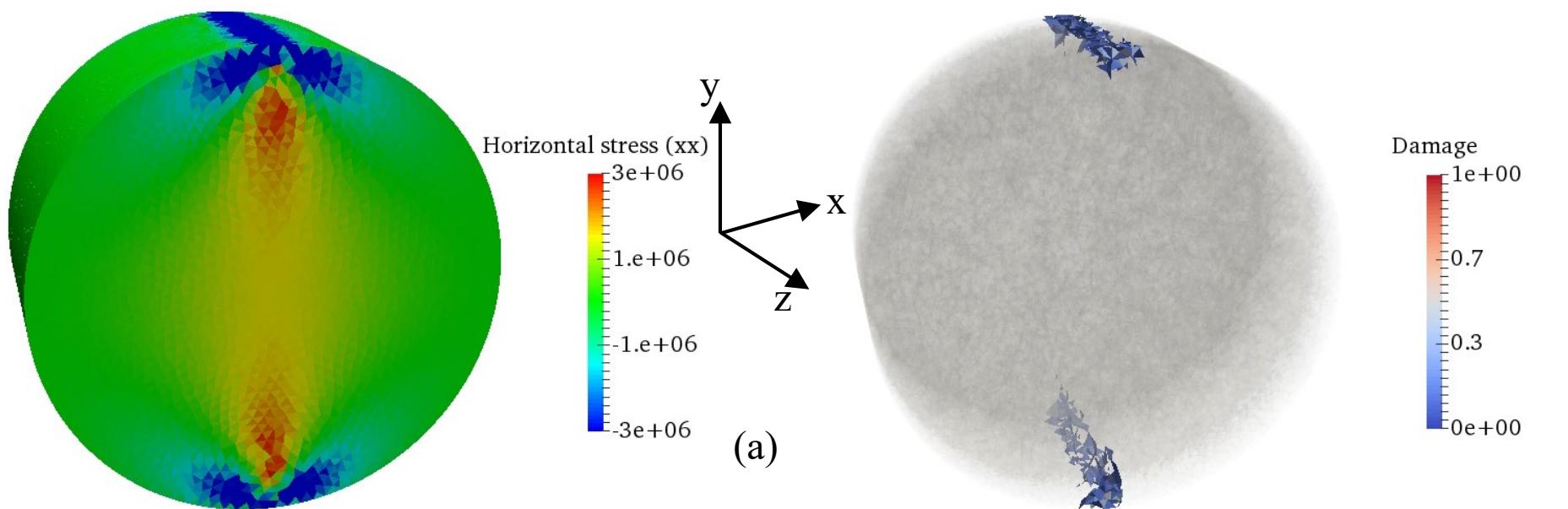


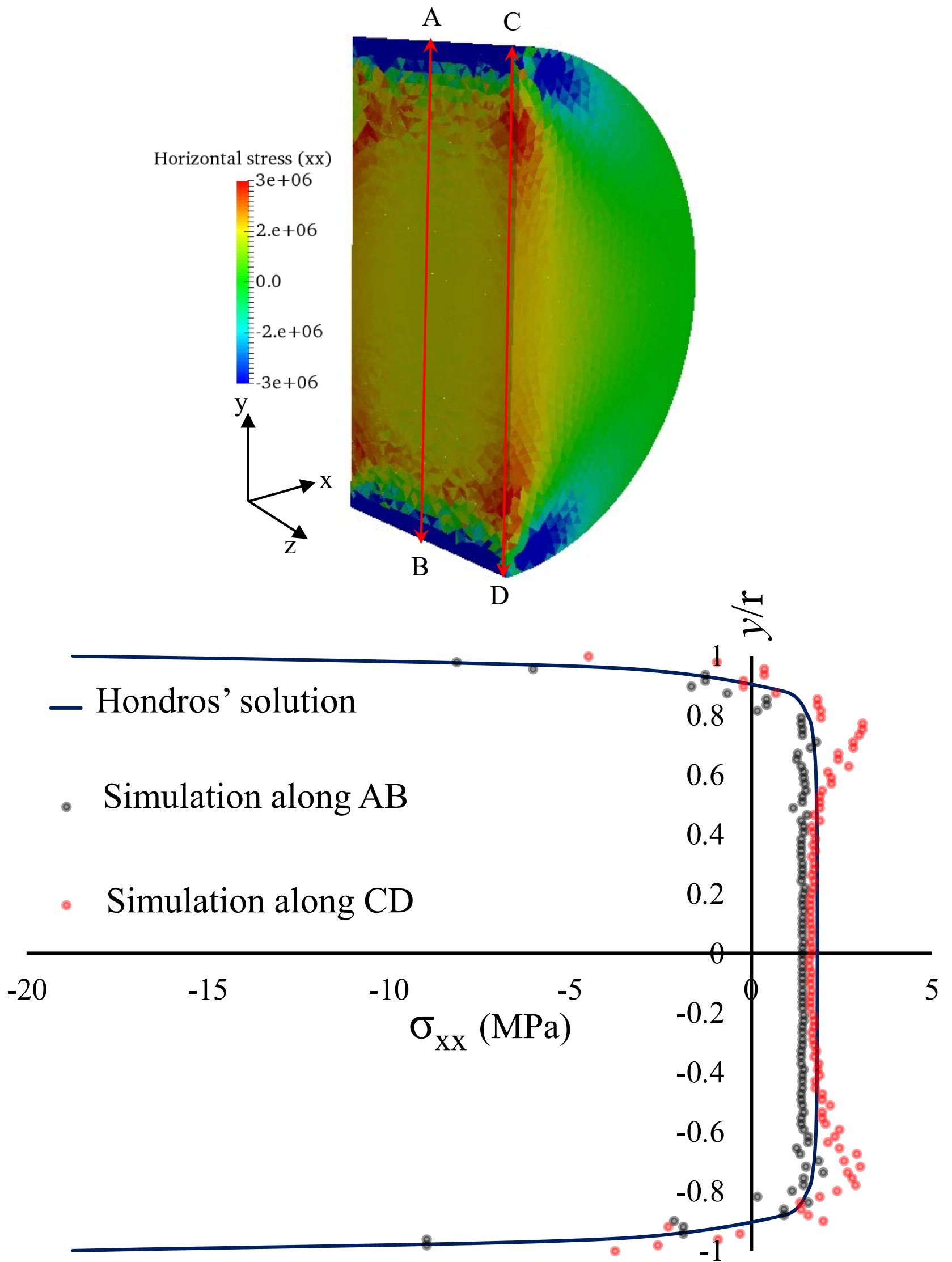
(b)



(c)

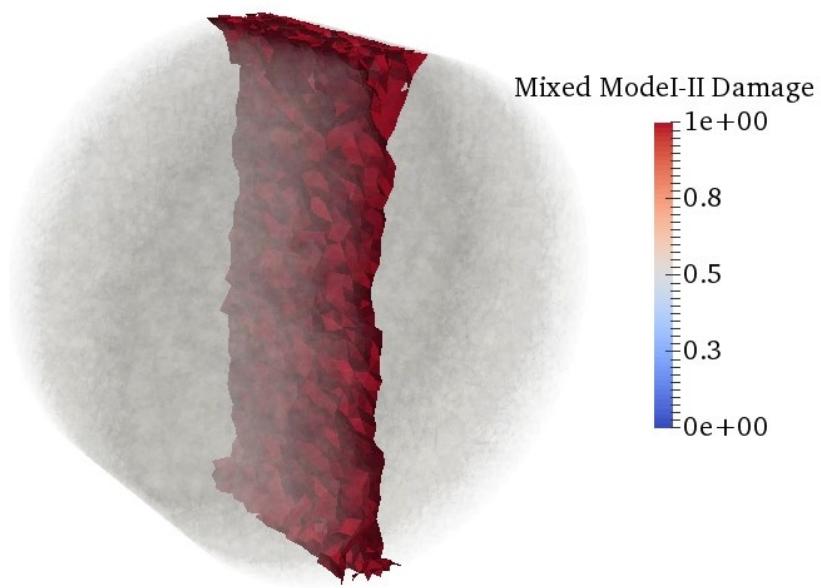
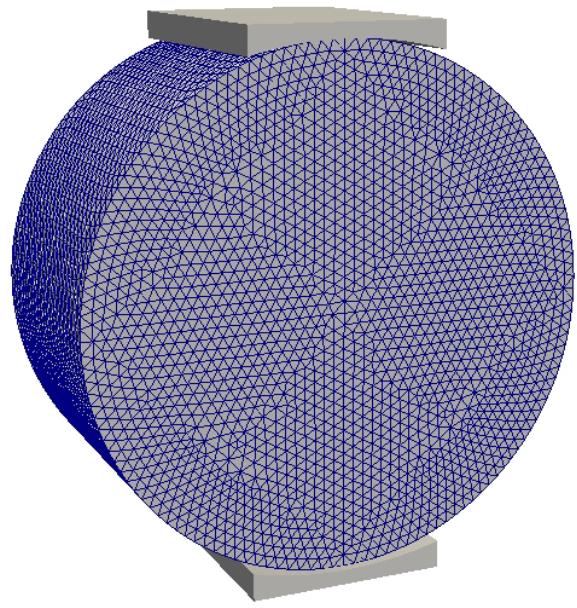
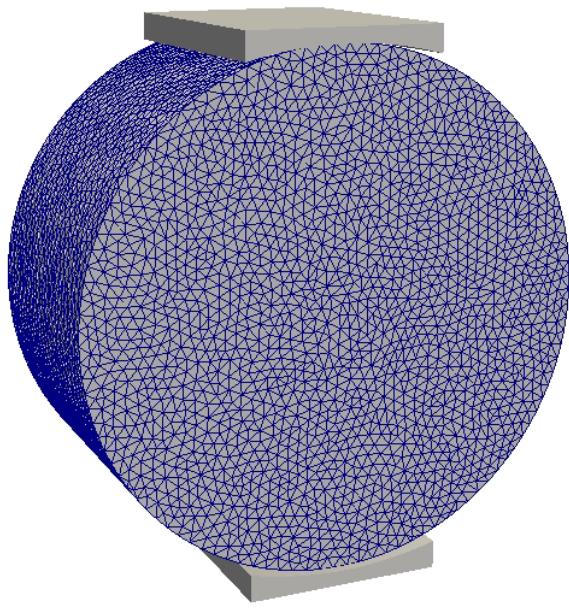
Fig. 12.



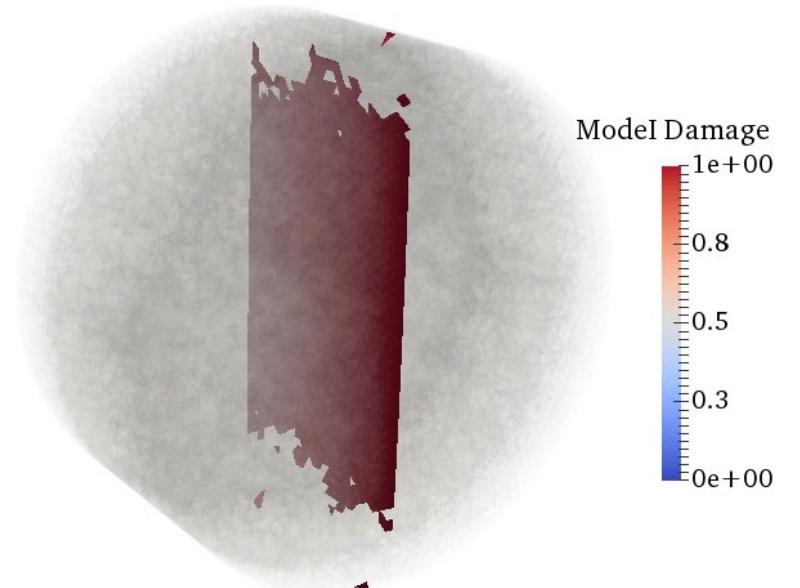


(e)

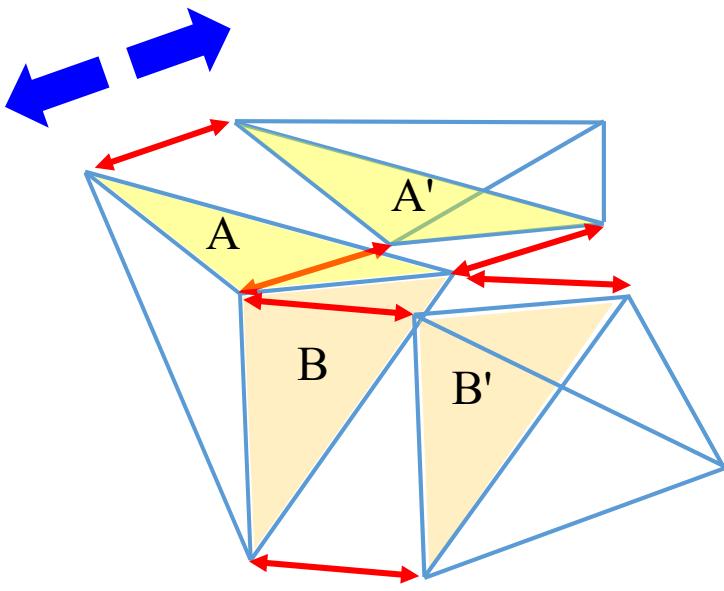
Fig. 13.



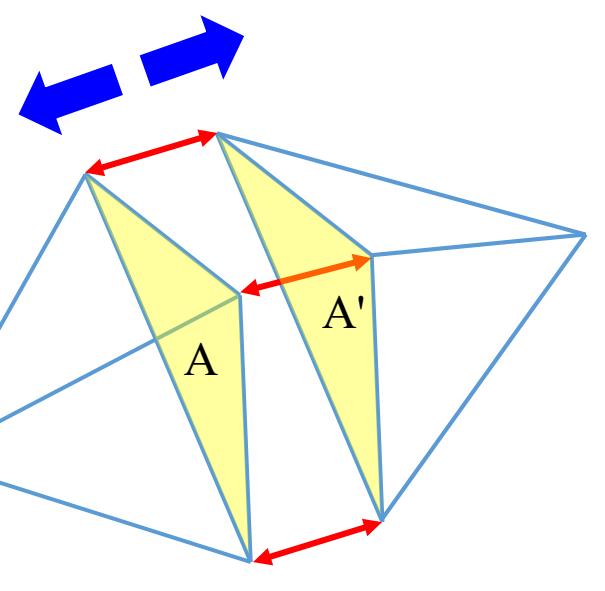
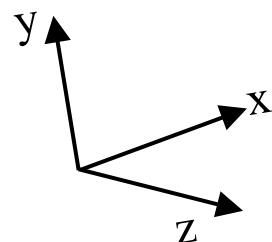
(a)



(b)



(c)



(d)

Fig. 14.

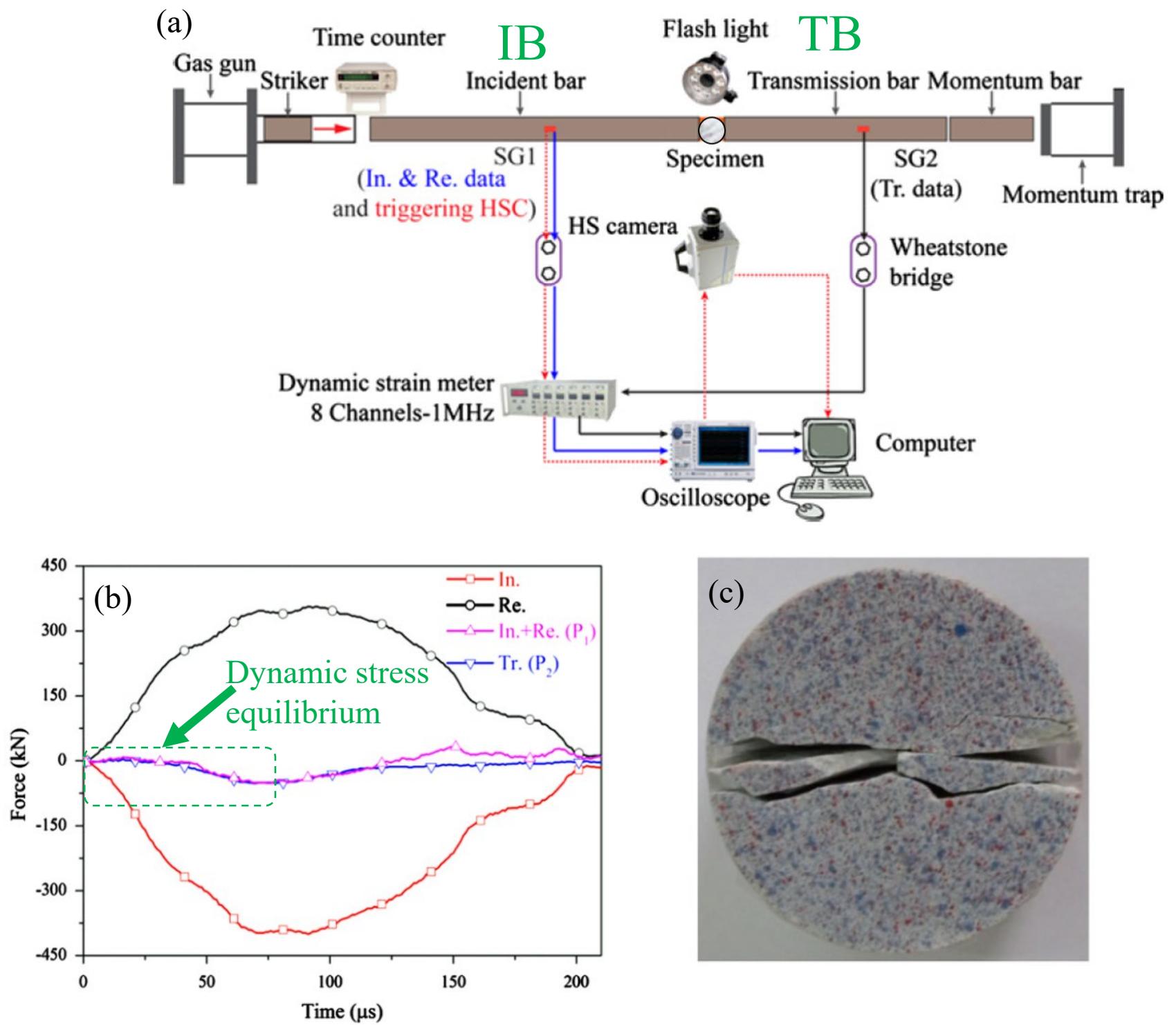


Fig. 15.

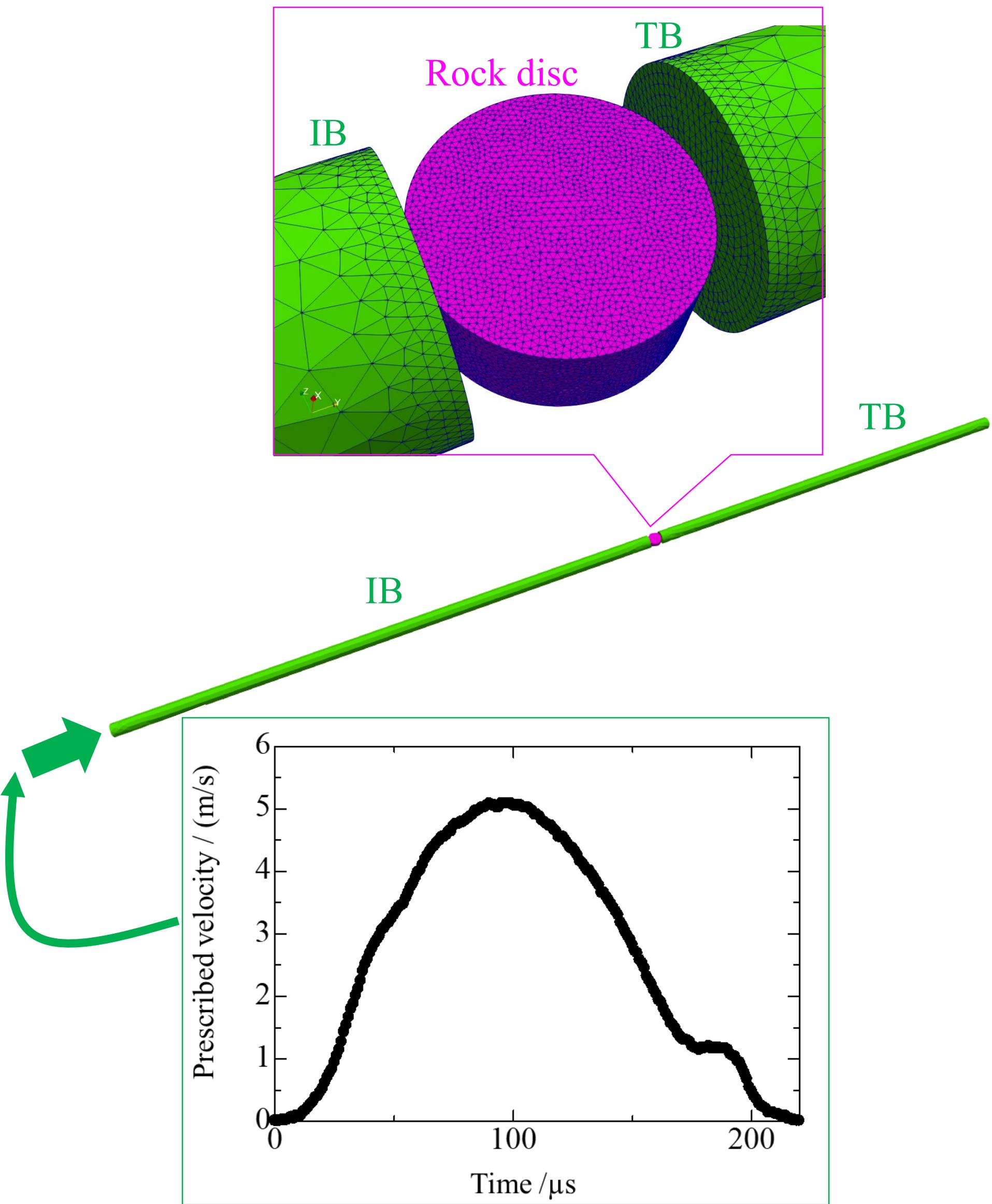


Fig. 16.

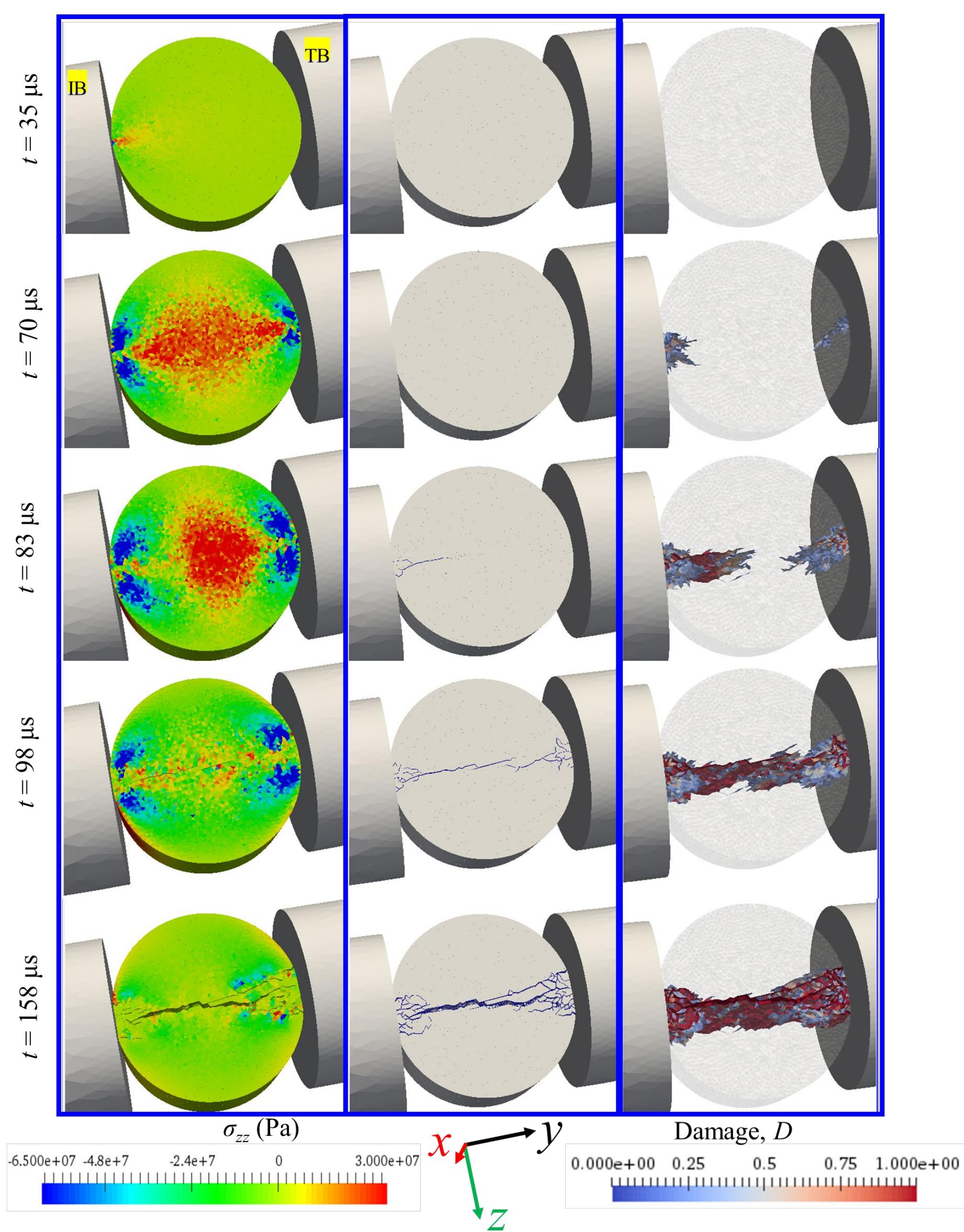


Fig. 17.

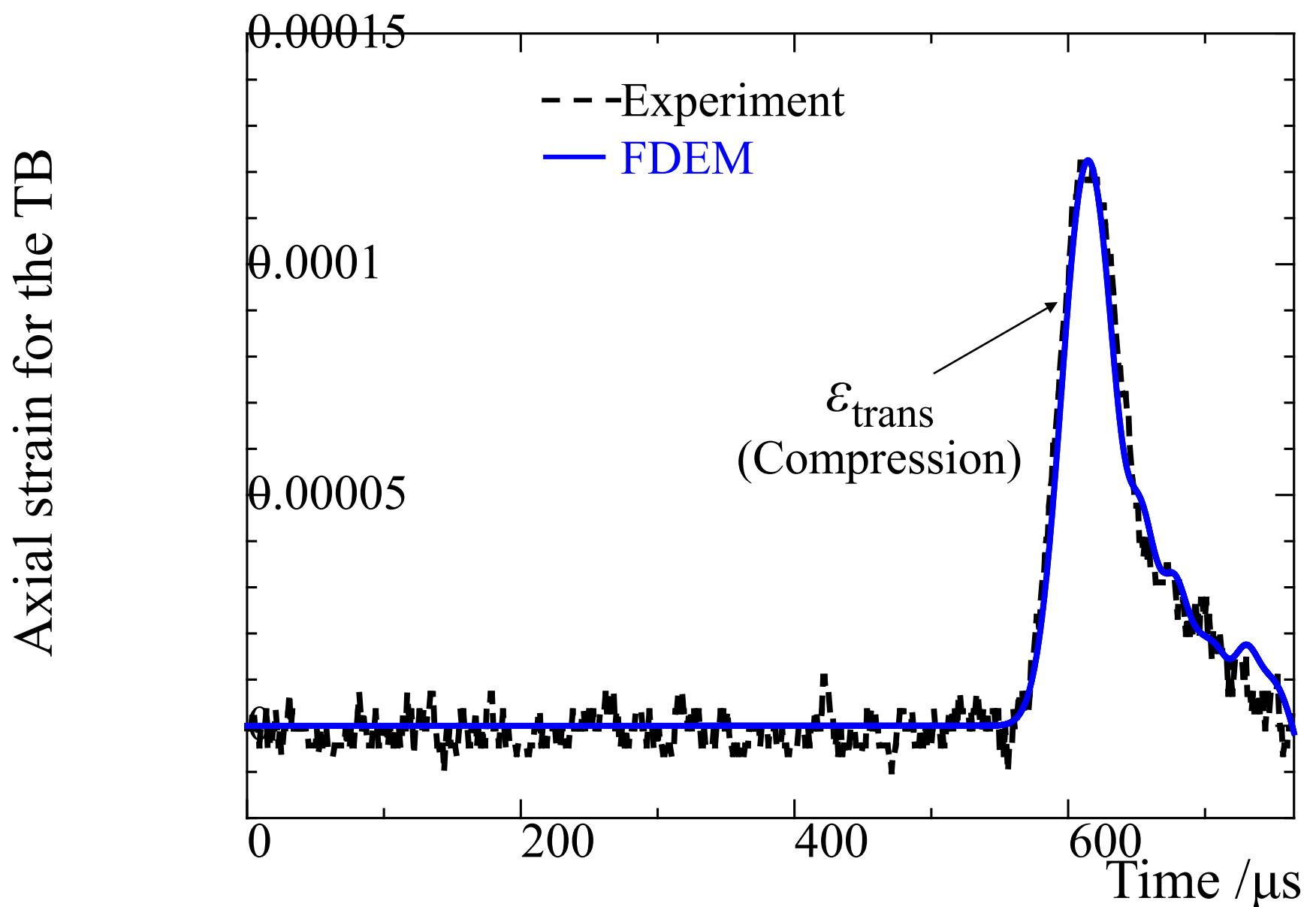
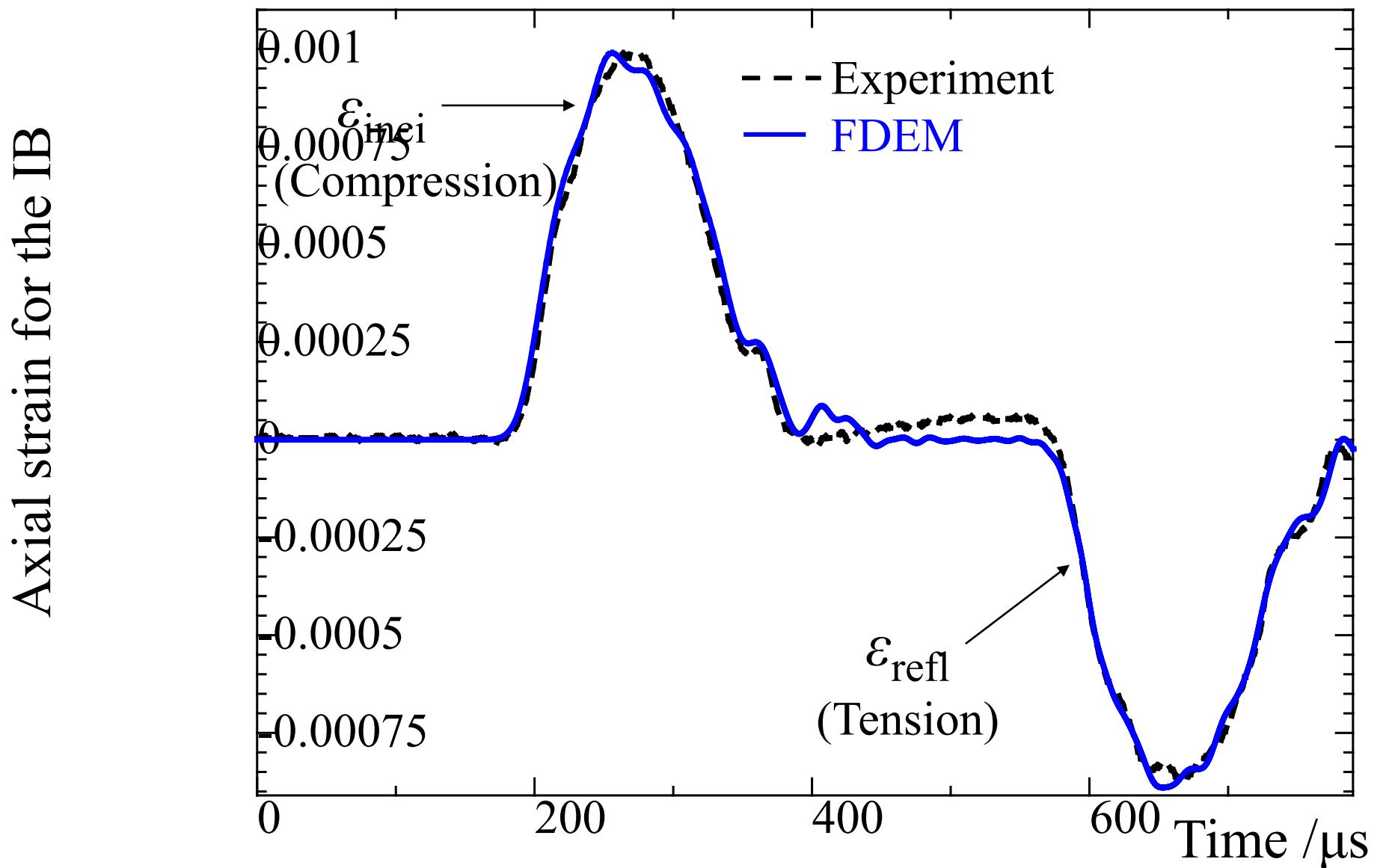


Fig. 18.

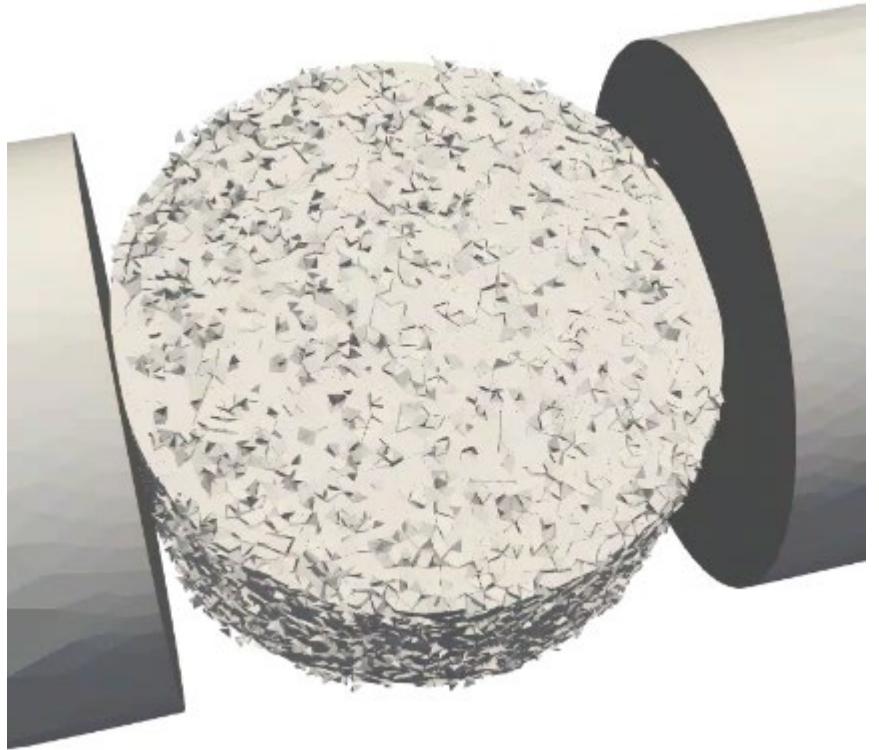
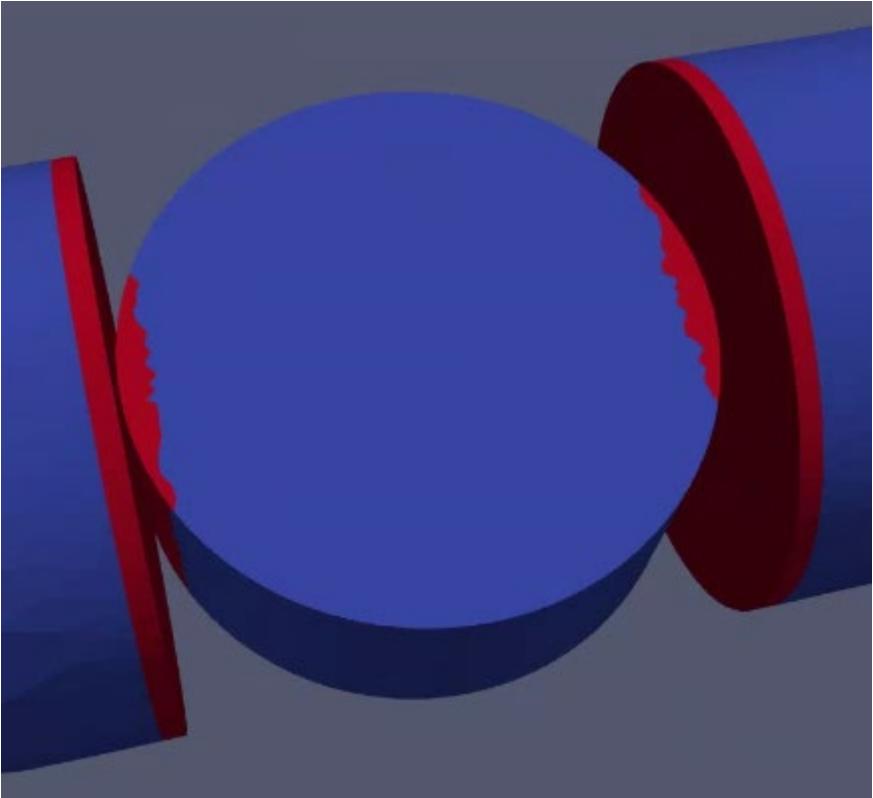


Fig. 19.

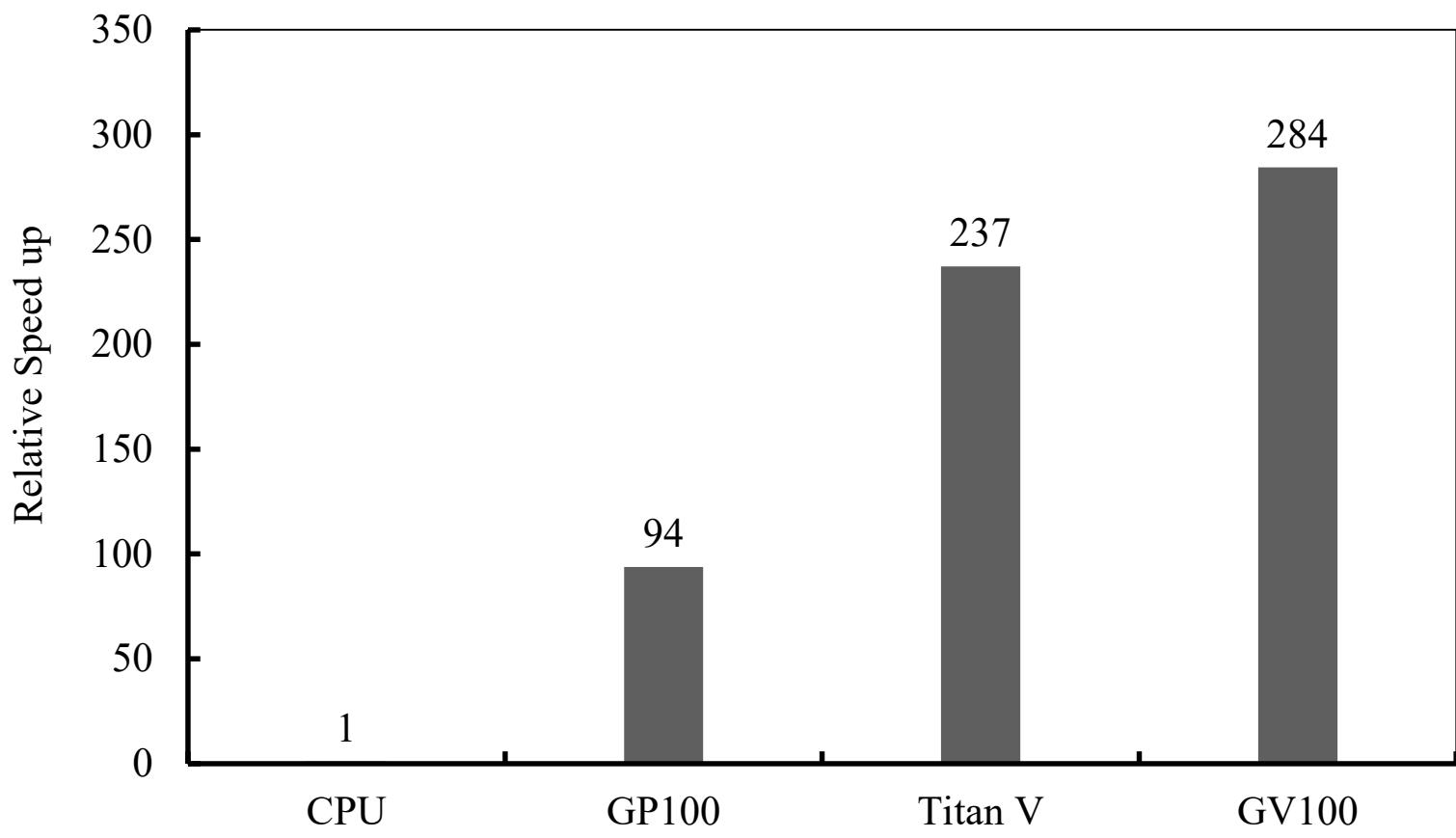
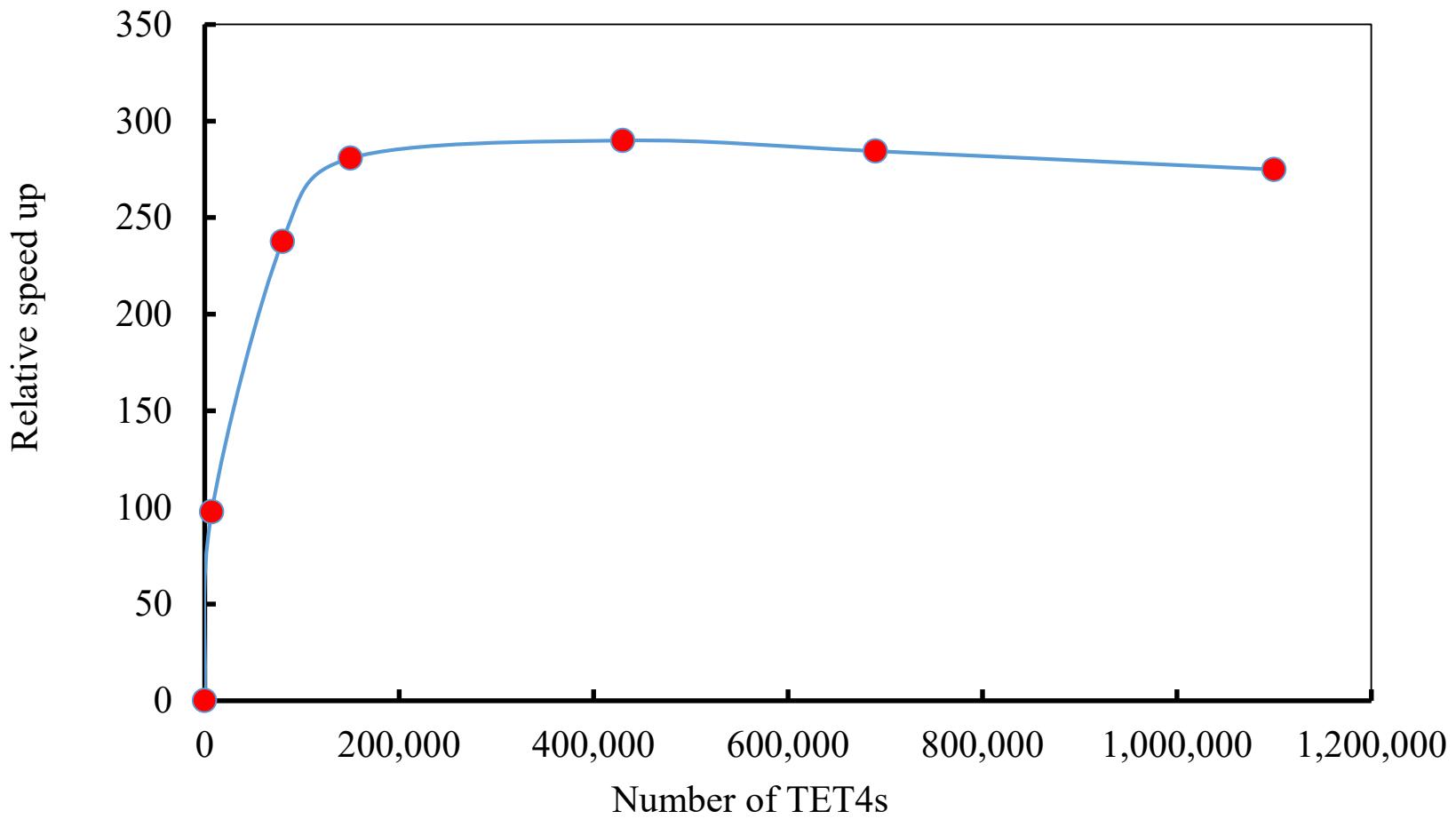
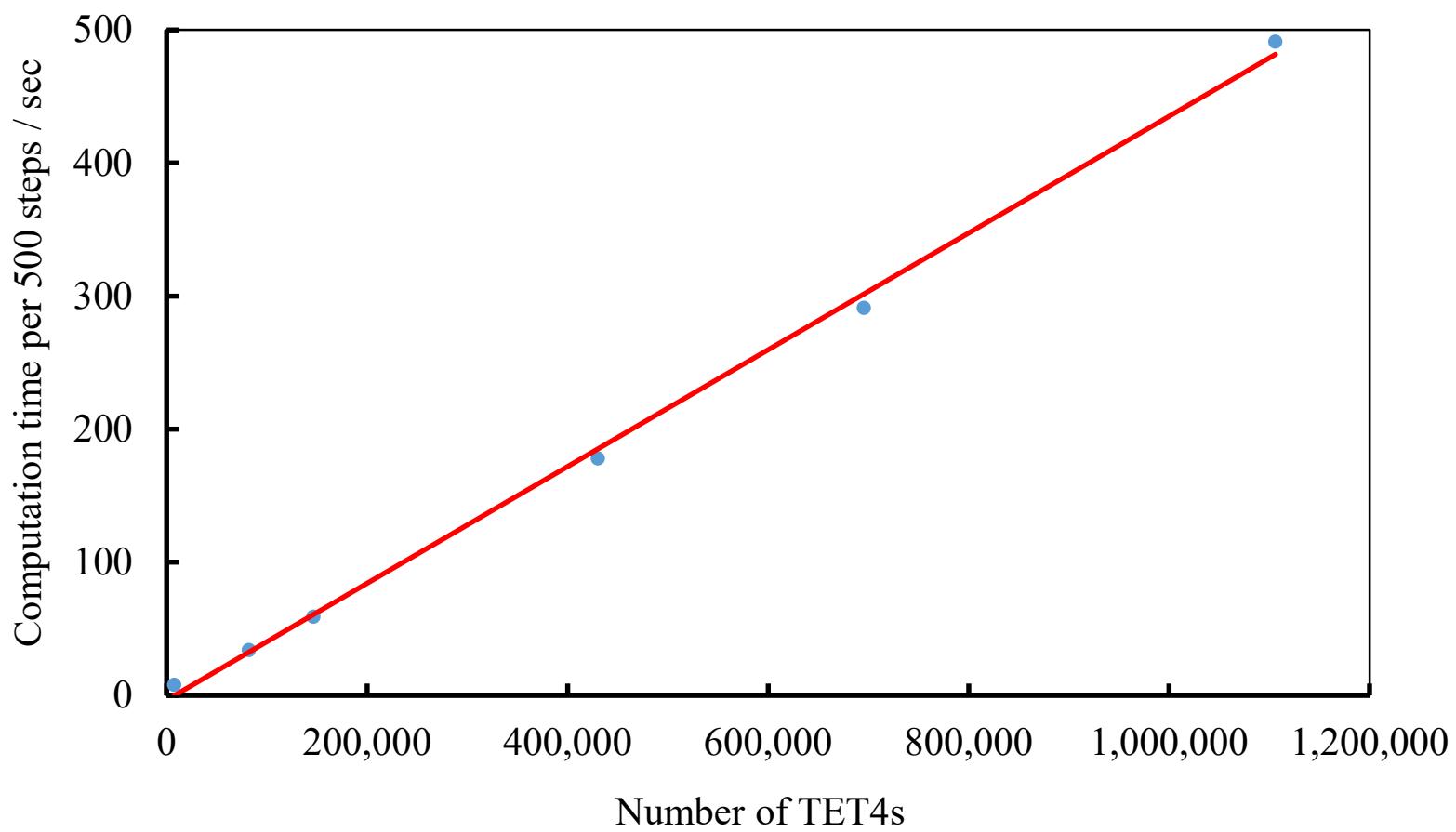


Fig. 20.



(a)



(b)

Fig. 21.