Development of a more accurate Geant4 quantum molecular dynamics model for hadron therapy

Yoshi-hide Sato[1], Dousatsu Sakata[2,3,*], David Bolst[4], Edward C. Simpson[5], Susanna Guatelli[4], and Akihiro Haga[1,*]

1 Department of Biological Sciences, Tokushima University, Tokushima 770-8503, Japan

2 Department of Accelerator and Medical Physics, Institute for Quantum Medical Science, QST, Chiba 263-8555, Japan

3 Division of Health Sciences, Graduate School of Medicine, Osaka University, Osaka 565-0871, Japan

4 Centre For Medical and Radiation Physics, University of Wollongong, Wollongong NSW 2522, Australia

5 Department of Nuclear Physics and Accelerator Applications, Research School of Physics, The Australian National University, Canberra ACT 2601, Australia

E-mail: `haga@tokushima-u.ac.jp, dosatsu.sakata@cern.ch`

Abstract.

Objective: Although in heavy-ion therapy, the quantum molecular dynamics (QMD) model is one of the most fundamental physics models providing an accurate daughter-ion production yield in the final state, there are still non-negligible differences with the experimental results. The aim of this study is to improve fragment production in water phantoms by developing a more accurate QMD model in Geant4.

Approach: A QMD model was developed by implementing modern Skyrme interaction parameter sets, as well as by incorporating with an *ad hoc* α-cluster model in the initial nuclear state. Two adjusting parameters were selected that can significantly affect the fragment productions in the QMD model: the radius to discriminate a cluster to which nucleons belong after the nucleus-nucleus reaction, denoted by R, and the squared standard deviation of the Gaussian packet, denoted by L. Squared Mahalanobis's distance of fragment yields and angular distributions with 1, 2, and the higher atomic

number for the produced fragments were employed as objective functions, and multi-objective optimization (MOO), which make it possible to compare quantitatively the simulated production yields with the reference experimental data, was performed.

Main results: The MOO analysis showed that the QMD model with modern Skyrme parameters coupled with the proposed α-cluster model, denoted as SkM$^*\alpha$, can drastically improve light fragments yields in water. In addition, the proposed model reproduced the kinetic energy distribution of the fragments accurately. The optimized L in SkM$^*\alpha$ was confirmed to be realistic by the charge radii analysis in the ground state formation.

Significance: The proposed framework using MOO was demonstrated to be very useful in judging the superiority of the proposed nuclear model. The optimized QMD model is expected to improve the accuracy of heavy-ion therapy dosimetry.

1. Introduction

Heavy-ion therapy (such as carbon ion therapy) is currently regarded as one of the most advanced clinical methods for cancer treatment because of its successful clinical results (Schulz-Ertner et al. 2004, Komatsu et al. 2011, Habermehl et al. 2013, Kawashiro et al. 2018). The efficacy of the treatment results from a high relative biological effectiveness (RBE) close to the Bragg peak, followed by a low-dose profile in the distal region beyond the Bragg peak, offering a more localized dose distribution to the tumor than that in the photon therapy (Schardt et al. 2010, Torikoshi et al. 2007, Ando & Kase 2009, Elsässer et al. 2010, Karger & Peschke 2017). Calculating the dose accurately for the tumor target and organs at risk is crucial to improve the delivery of carbon ion therapy further. Ideally, the dose calculations should have an accuracy of $\pm$ 3%, with a submillimeter spatial resolution, especially close to the distal edge, where organs at risk may be set (Dudouet et al. 2013, Tran et al. 2018). In carbon ion therapy, where the kinetic energy of carbon ions is between 50 and 430 MeV/u, approximately 15% of the physical dose at the Bragg peak originates from fragments (Schardt et al. 1996, Matsufuji et al. 2003, Haettner et al. 2006, Bolst et al. 2017).

Nuclear fragmentation of both incident carbon ions and target nuclei in the human body leads to the production of secondary fragments of lower Z, characterised by larger ranges and wider angular distributions compared with carbon ions, affecting the energy deposition laterally to the beam and after the Bragg peak, ultimately affecting the radiobiological effectiveness. Therefore, the accurate simulation of nuclear fragmentation reactions is important for the evaluation of the biological dose (Krämer & Durante 2010, Schwaab et al. 2011, Inaniwa et al. 2014, 2020). Moreover, an accurate modelling of nuclear fragmentation is essential to develop novel imaging applications used for hadron therapy, such as positron emission tomography (Rahmim et al. 2013, Bertolli et al. 2016, Chacon et al. 2019).

In carbon ion therapy, Monte Carlo (MC) simulations are extensively used to perform dosimetric calculations and to develop associated quality assurance technology. Among other general-purpose MC codes, Geant4 (Geant4 Collaboration 2020, Agostinelli et al. 2003, Allison et al. 2006, 2016) has been extensively used in carbon ion therapy studies. Previous studies have shown that there is space for improvement of the Geant4 quantum molecular dynamics (QMD) model (Arce et al. 2021, Chen et al. 2018, Bolst et al. 2017), which has led to developments aimed at improving the inelastic nuclear interaction cross section (Sakata et al. 2020).

In the present study, an attempt was made to improve QMD nuclear fragment reaction model implemented in Geant4 in response to the need to improve the accuracy of this MC code for dosimetric and RBE calculations of interest for carbon ion therapy. The QMD model provides the most fundamental description of nuclear fragmentation available in Geant4 since release 9.1. It is a microscopic nuclear reaction model that makes it possible to simulate the time evolution of each nucleon participating as a Gaussian wave packet. In contrast to the other inelastic reaction model provided in Geant4 of interest for carbon ion therapy (e.g. the Binary Ion Cascade, BIC (Folger et al. 2004)), the propagation of each participating nucleon is simulated according to the equation of motion (EOM) in a many-body system with a self-consistent mean

field potential. This approach makes it possible to evaluate fragment production with a more fundamental approach, because the QMD can calculate the formation of the secondary nuclei by classifying the nucleon cluster as fragments at the end of the time evolution. In past studies, many variations of the QMD model have been proposed. They have been widely and successfully applied in the study of heavy-ion collisions in the range from low to relativistic energies to understand nuclear phenomena, such as the multi-fragmentation mechanism as well as collective flows, particle emissions and fusion-fission phenomena (Zhang et al. (2020) and references therein). Thus, the QMD model is considered a sophisticated approach for evaluating of fragments yields in hadron therapy.

Despite successful applications of the QMD to explain nuclear phenomena thanks to its fundamental reaction mechanism approach, the QMD model implemented in Geant4 (hereinafter, "G4QMD") can be further improved to reproduce fragment yields with a better accuracy. For example, the angular distributions of light fragments in thin targets for 62- or 95-MeV/u and in liquid water for 400-MeV/u carbon ions obtained with the G4QMD could not reproduce adequately the measured distributions, especially in the forward direction (Napoli et al. 2012, Dudouet et al. 2014, Bolst et al. 2017). In this study, three methods to improve G4QMD were considered: (i) updating Skyrme-type nuclear interaction, (ii) forming a realistic initial state of the nuclei involved in the collision, and (iii) finding the best model parameter set for the QMD model.

(i) Skyrme interaction: The first aim of this study is to assess the G4QMD with modern Skyrme models for carbon ions propagating in a water target, using a therapeutic beam energy. The G4QMD employs the relativistic JAERI QMD model (JQMD), which is the Lorentz covariant version of JQMD employing the simplest Skyrme-type interaction (Niita et al. 1995). Since the release of G4QMD, and even before that, many versions of Skyrme interaction models have been proposed for heavy-ion collisions using the QMD model following the available knowledge on nuclear

states (Zhang et al. 2020). Therefore, it is an essential task for hadron therapy applications to update the G4QMD with these improved Skyrme interaction models.

(ii) Initial state of nuclei: Establishing appropriate initial conditions for the ions involved in the collision is of vital importance for correctly describing the ion collisions (Zhang et al. 2020). In the G4QMD model, the nucleon positions and momenta in the reaction system are firstly sampled according to the density distribution and the Fermi momentum distribution of the projectile and target nuclei, respectively, then, the nucleon positions and their momenta in the initial ground states are finely tuned by solving the damped EOM in the QMD model as the calculated binding energy becomes realistic (Niita et al. 1995, 1999). One shortcoming of this model is that it can provide an unrealistic nuclear ground state owing to the random sampling in the phase space, which enables identical particles to occupy the same phase space simultaneously, and as a consequence, the nuclear ground state partially violates the Fermionic feature. To compensate for this, the two-body Pauli potential has been introduced by several authors to mimic the Pauli principle (Ohnishi et al. 1992, Huang & Ma 2021). Although the Pauli potential successfully provides the α-cluster structure in the ground state, it requires additional computing time to calculate the ground state. In this study, for the first time, a ground state nuclear model was introduced to create an α-cluster structure without additional computing by setting the regular tetrahedron formation as the initial positions of the nucleons.

(iii) The best model parameter set: In this work, an attempt was made to improve G4QMD with the modern Skyrme interactions by using the framework of the improved QMD approach (ImQMD) (Wang et al. 2002, 2004). The main improvement in ImQMD with respect to the current G4QMD is that the surface symmetry energy terms, which are expected to be more important in such light nuclei as ^{12}C, are introduced in the potential energy part based on the Skyrme interaction. Furthermore, an *ad hoc* α-cluster structure model was implemented for the initial positions of the nucleons that computes

a realistic nuclei structure, as described above. The developed models also depend on such model parameters as the width of the nucleon wave packet. However, the model parameters in G4QMD were adjusted with the original QMD model (Niita et al. 1995), which is not guaranteed to be the best parameter set for the developed QMD model. Thus, in the present work, certain model parameters of G4QMD were optimized to be consistent with the measured fragments yields and their angular distribution. For the optimization, the multi-objective optimization (MOO) approach was applied, in which two optimization parameters, the radius to judge a cluster formation after the nucleus-nucleus reaction (hereinafter denoted by R) and squared standard deviation of the Gaussian packet commonly used in the QMD model (hereinafter denoted by L) were selected that can sensitively affect the fragment productions.

2. Materials and Methods

In this section, the QMD model developed in this study is described first. Subsequently, the parameter optimization method is described. The Geant4 physics list used for the simulation of hadronic interactions in this study is included in Appendix A, and the developed code is found in Github,
https://github.com/hagaakihiro/G4QMDprojectV1/blob/main/README.md
for use. Geant4 10.7.p01 was used, and the new QMD model was then activated in the Geant4 physics list.

2.1. QMD model

2.1.1. Skyrme nucleon-nucleon interaction: The basic assumption of the QMD is that each nucleon state $\varphi_i(\boldsymbol{r})$ is represented by a Gaussian wave function with squared standard deviation L:

$$\varphi_i(\boldsymbol{r}) \equiv \frac{1}{(2\pi L)^{3/4}} \exp\left(-\frac{(\boldsymbol{r}-\boldsymbol{r}_i)^2}{4L} + \frac{i}{\hbar}\boldsymbol{r}\cdot\boldsymbol{p}_i\right) \tag{1}$$

where $\boldsymbol{r}_i$ and $\boldsymbol{p}_i$ represent the centroid of position and momentum of the ith particle, respectively. The equations of the motion of the particle derived from the time dependent variation principle are,

$$\dot{\boldsymbol{r}}_i = \frac{\partial H}{\partial \boldsymbol{p}_i}, \quad \dot{\boldsymbol{p}}_i = -\frac{\partial H}{\partial \boldsymbol{r}_i}, \tag{2}$$

where H is the Hamiltonian with the Lorentz covariant form (Geant4 Collaboration 2020, Mancusi et al. 2009),

$$H = \sum_i \sqrt{\boldsymbol{p}_i^2 + m_i^2 + 2m_i V_i} \tag{3}$$

where V_i is the effective potential felt by the ith particle, which essentially controls the time evolution of the system. In the current work, the Skyrme model is used; adapting Eq. (1) as a Hartree-Fock basis in the Skyrme interaction gives the following expression for V_i:

$$\begin{aligned} V_i = & \frac{A}{2\rho_0} \sum_{j \neq i} \rho_{ij} + \frac{B}{\gamma+1} \frac{1}{\rho_0^\gamma} \left(\sum_{j \neq i} \rho_{ij} \right)^\gamma \\ & + \frac{g_0}{2\rho_0} \sum_{j \neq i} f_{sij} \rho_{ij} + \frac{g_\tau}{\rho_0^\eta} \left(\sum_{j \neq i} \rho_{ij} \right)^\eta \\ & + \frac{C_s}{2\rho_0} \sum_{j \neq i} \tau_{3,i} \tau_{3,j} \rho_{ij} (1 - \kappa_s f_{sij}) \\ & + \frac{\alpha \hbar c}{2} \sum_{j \neq i} \frac{c_i c_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \mathrm{erf} \left(\frac{|\boldsymbol{r}_i - \boldsymbol{r}_j|}{\sqrt{4L}} \right), \end{aligned} \tag{4}$$

where index j runs on all participating particles (nucleons) except for the ith particle, and $f_{sij} = \frac{3}{2L} - \left(\frac{\boldsymbol{r}_i - \boldsymbol{r}_j}{2L}\right)^2$. The quantities $A, B, g_0, g_\tau, C_s, \kappa_s, \gamma$, and η are the parameters of the Skyrme-QMD interaction, and the last term indicates the Coulomb interaction (Niita et al. 1995, 1999) with $c_i = \frac{1+\tau_{3,i}}{2}$ ($c_i = 1$ for proton and $c_i = 0$ for neutron), where the fine-structure constant $\alpha = 1/137$. Here, ρ_{ij} is defined as,

$$\rho_{ij} = \int d\boldsymbol{r} \rho_i(\boldsymbol{r}) \rho_j(\boldsymbol{r}) = \frac{1}{(4\pi L)^{3/2}} e^{-\frac{1}{4L}(\boldsymbol{r}_i - \boldsymbol{r}_j)^2}, \tag{5}$$

for the overlap between the ith and the jth nucleons. The third term on the right-hand side of Eq. (4) is the symmetry potential, where $\tau_{3,i} = 1$ when the ith particle is a proton and $\tau_{3,i} = -1$ when the ith particle is a neutron. The parameter sets considered in this

study are shown in Table 1, which shows that the original G4QMD (Niita et al. 1995) is regarded as one of the Skyrme-type interactions neglecting the terms with parameters g_0, g_τ, and κ_s, in which g_0 and κ_s are related to the surface energy, and g_τ is the $\rho\tau$ term obtained from the momentum-dependent term of the Skyrme potential by applying the Thomas-Fermi approximation, which gives $\eta = 5/3$ (Wang et al. 2002). The parameters shown were validated with the saturation density ρ_0 [fm^{-3}], the binding energy BE [MeV], and incompressibility K_0 [MeV] based on the nuclear matter properties (Dutra et al. 2012). In particular, the parameter sets SLy4 and SkM* have been demonstrated to describe the heavy ion collision relatively well (Zhang & Li 2006, Zhang et al. 2014) as well as the nuclear structure (Dobaczewski et al. 2003, Bayram et al. 2014, Kean et al. 2020). In this study, the Skyrme parameter sets were applied for carbon ion beam simulation using Geant4.

2.1.2. α-cluster model: Another development is the involvement of the α-cluster structure in the initialization of the nuclei. The cluster structure may be important for the production of α particles during/after collisions for a carbon interacting with an oxygen nucleus of the target in water, of which the ground states are known to partly form a 3-α cluster and 4-α cluster, respectively. However, the normal QMD is not anti-symmetrized and therefore, identical particles can exist in the phase space where Pauli's exclusion principle is broken. This makes it lose the cluster structure in the ground state nuclei, which has been successfully reproduced in the anti-symmetrized molecular dynamics (AMD) (Ono 2019, Kanada-En'yo et al. 2012), although AMD requires the additional computational effort owing to the inclusion of the exchange process among identical particles. Approximating the Pauli blocking phenomenon by introducing an effective potential has also been attempted (Ohnishi et al. 1992, Huang & Ma 2021). The inclusion of the effective potential, however, still increases the simulation time significantly to obtain the appropriate initial nuclear state, because the initial state created is frequently wasted because of the violation of the initial condition. Therefore,

Table 1. Parameters in the Skyrme-QMD model.

	JQMD[1] (G4QMD)	SLy4[2]	SkM*[2]	SIII[2]
A[MeV]	-219.4	-297.82	-318	-122.921
B[MeV]	165.3	219.21	249.5	55.343
g_0[MeVfm^2]	–	24.569	21.86	18.286
g_τ[MeV]	–	9.70	5.9357	6.439
C_s[MeV]	25	32	32	32
κ_s[fm^2]	–	0.08	0.08	0.08
γ	4/3	7/6	7/6	2
$\eta^\dagger$	–	5/3	5/3	5/3
ρ_0[fm^{-3}]†	0.168	0.160	0.165	0.1452
BE[MeV]†	-16.00	-15.97	-15.77	-15.83
K_0[MeV]†	237.8	230.2	216.8	355.9

† Resulting from nuclear matter simulation: $\eta = 5/3$, which is given by the Thomas-Fermi approximation in the momentum-dependent term of the Skyrme potential, the saturation density ρ_0 [fm^{-3}], the binding energy BE [MeV], and incompressibility K_0 [MeV].

1. Niita et al. (1995)
2. Kean et al. (2020)

a simpler approach was introduced to form the α cluster within the QMD model by sampling with the α structure in ground state nuclei. In this method, the nucleons in an alpha particle are first located at:

$$\begin{aligned}
\boldsymbol{R}_1^\alpha &= \frac{a}{\sqrt{3}}(1,1,1),\\
\boldsymbol{R}_2^\alpha &= \frac{a}{\sqrt{3}}(-1,-1,1),\\
\boldsymbol{R}_3^\alpha &= \frac{a}{\sqrt{3}}(1,-1,-1),\\
\boldsymbol{R}_4^\alpha &= \frac{a}{\sqrt{3}}(-1,1,-1),
\end{aligned} \tag{6}$$

where the center-of-mass (COM) is the origin. The distance between the particles can be controlled by a ($a = 0.5$ fm was used in this study). Then, the Euler rotation matrix,

$$R(\theta_\alpha\theta_\beta\theta_\gamma) = \begin{pmatrix} \cos\theta_\alpha\cos\theta_\beta\cos\theta_\gamma - \sin\theta_\alpha\sin\theta_\gamma & -\cos\theta_\alpha\cos\theta_\beta\sin\theta_\gamma - \sin\theta_\alpha\cos\theta_\gamma & \cos\theta_\alpha\sin\theta_\beta \\ \sin\theta_\alpha\cos\theta_\beta\cos\theta_\gamma + \cos\theta_\alpha\sin\theta_\gamma & -\sin\theta_\alpha\cos\theta_\beta\sin\theta_\gamma + \cos\theta_\alpha\cos\theta_\gamma & \sin\theta_\alpha\sin\theta_\beta \\ -\sin\theta_\beta\cos\theta_\gamma & \sin\theta_\beta\sin\theta_\gamma & \cos\theta_\beta \end{pmatrix}, \tag{7}$$

with the Euler angles $(\theta_\alpha, \theta_\beta, \theta_\gamma)$ randomly sampled from the uniform distribution is operated to $\boldsymbol{R}_i^\alpha$.

$$\boldsymbol{r}_1^{C12} = (\frac{b}{\sqrt{3}}, 0, 0),\ \boldsymbol{r}_2^{C12} = (-\frac{b}{2\sqrt{3}}, \frac{b}{2}, 0),\ \boldsymbol{r}_3^{C12} = (-\frac{b}{2\sqrt{3}}, -\frac{b}{2}, 0), \tag{8}$$

and the location of the four alpha particles are:

$$\begin{aligned} \boldsymbol{r}_1^{O16} &= \frac{b}{2\sqrt{2}}(1,1,1), & \boldsymbol{r}_2^{O16} &= \frac{b}{2\sqrt{2}}(-1,-1,1), \\ \boldsymbol{r}_3^{O16} &= \frac{b}{2\sqrt{2}}(1,-1,-1), & \boldsymbol{r}_4^{O16} &= \frac{b}{2\sqrt{2}}(-1,1,-1), \end{aligned} \tag{9}$$

where $\boldsymbol{r}_j^{C12}$ and $\boldsymbol{r}_j^{O16}$ are also randomly rotated with an Euler angle. The quantity b is a parameter to control the distance between the α particles, and it was sampled from the range of [2.1,2.9] fm. Thus, the nucleons of the projectile (^{12}C) and the target (^{16}O in water) are set in the initialization of the hadronic collision as,

$$\boldsymbol{r}_{ij} = R(\theta_\alpha^j\theta_\beta^j\theta_\gamma^j)\boldsymbol{R}_i^\alpha + R(\theta'_\alpha\theta'_\beta\theta'_\gamma)\boldsymbol{r}_j^{C12},\ i \in (1,2,3,4), j \in (1,2,3), \tag{10}$$

$$\boldsymbol{r}_{ij} = R(\theta_\alpha^j\theta_\beta^j\theta_\gamma^j)\boldsymbol{R}_i^\alpha + R(\theta'_\alpha\theta'_\beta\theta'_\gamma)\boldsymbol{r}_j^{O16},\ i \in (1,2,3,4), j \in (1,2,3,4), \tag{11}$$

respectively. Note that the Euler angles here and the b distances in (8) and (9) are not the same for carbon projectile and oxygen target (although the same symbols are used). The distribution of the initial nucleons by this method may not reproduce the experimental binding energy of the ground state nucleus. After the α-cluster configuration, therefore, the nucleon locations are finely tuned by solving the damped EOM to agree with the total binding energy within 1 keV. Those prepared nuclei are expected partly to keep

an α-cluster structure with good ground state properties and less computational effort than the AMD and the inclusion of the Pauli potential.

2.2. Simulation configuration

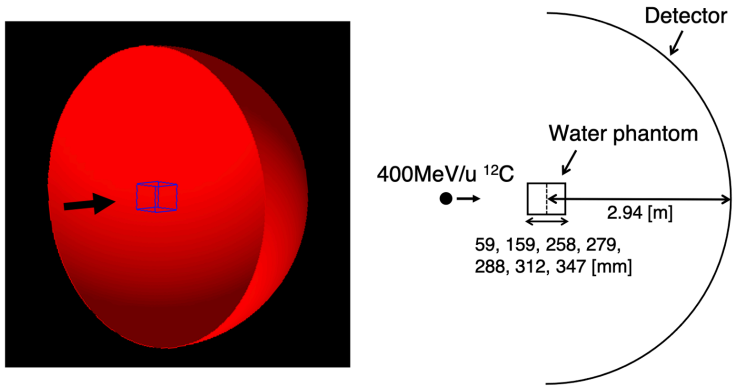

Figure 1. Simulation setup. (Left) three-dimensional graphic in the current simulation, where a box located in the center is the water phantom, whereas the red hemisphere shows the detector area, and (Right) its projection illustration. Seven different thicknesses of the box (width $\times$ height of 50 cm $\times$ 50 cm) in the direction of the beam were tried: 59, 159, 258, 279, 288, 312, and 347 mm. The images are not in scale.

The geometrical simulation configuration is shown in Fig. 1, where the experimental setup adopted by Haettner et al. (2013) is reproduced. This configuration is the same as that used by Bolst et al. (2017). In this section, the simulation setup is described briefly. More details can be found elsewhere (Haettner et al. 2013). A mono-energetic 400-MeV/u ^{12}C pencil beam with a full width at half maximum (FWHM) of 5 mm and energy sigma of 0.15% representing an FWHM of 1.4 MeV/u was incident on a box water phantom with a lateral size of 50 by 50 cm. Seven independent simulations were run, with box thickness of 59, 159, 258, 279, 288, 312 and 347 mm. The fragments produced in the water phantom were tallied when traversing a hemisphere with a radius of 2.94 m corresponding to the distance from the mid target to the detector, imitating the experimental setup. Fragment yields of $Z = 1 - 5$ emerging from the phantom in

the forward direction were scored within a 10-degree cone. The angular distribution was calculated as the number of fragments at a ring-shaped solid angle to compare with that in experimental data. The experimental data for the fragment yields and the angular distributions were used as references in the QMD model parameter optimization. The fitting parameters and the method for the model comparison will be described in the next subsection. Although the experimental data includes the kinetic energy distribution of fragments at each solid angle, it was not used in the fitting process but rather in the verification of the optimized model parameters.

2.3. Parameters for QMD model optimization

In spite of a large number of parameters, the parameter tuning in the QMD model has rarely been performed in hadron therapy simulations. In this context, a study was done in order to achieve a better description of prompt-gamma yields for hadron therapy monitoring (Dedes et al. 2014), where two quantities, the square of the width of the Gaussian wave packet denoted as L in Eq. (1) and the maximum distance criterion in the fragment cluster judgement denoted as R,

$$\tilde{\boldsymbol{R}}_{ij}^2 = (\boldsymbol{r}_i - \boldsymbol{r}_j)^2 + \gamma_{ij}^2((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot \boldsymbol{\beta}_{ij})^2 \leq R^2 \tag{12}$$

between the ith and jth particles using

$$\boldsymbol{\beta}_{ij} = \frac{\boldsymbol{p}_i + \boldsymbol{p}_j}{E_i + E_j}, \quad \gamma_{ij} = \frac{1}{\sqrt{1 - \boldsymbol{\beta_{ij}}^2}}, \tag{13}$$

were independently optimized.

As described by Dedes et al. (2014), $\sqrt{L}$ in fm, which may be realistic in the range of [0.8, 3.0] fm, controls the effective interaction range of the nuclear potential, and thus can influence the final states of the collision sensitively. Ideally, L should be dynamically changed depending on the nucleon state during the collision, or, at least, it should be optimized for certain configurations for each projectile/target nucleus, as proposed by Wang et al. (2002). The default value of L in the G4QMD model (JQMD as well) is 2 fm^2. This value seems to be too large for light nuclei, such as ^{12}C and ^{16}O, as verified

in a later section. Furthermore, no study has been conducted to optimize the L and R parameters simultaneously; therefore, L and R were chosen as the parameters for the QMD model optimization. These two parameters were determined for each Skyrme parameter set using the optimization process described in the next subsection.

2.4. Optimization criteria and model comparison

First, parameters L and R in each Skyrme parameter set were optimized by fitting the experimental data of fragment yields and their angular distributions in water phantoms, as denoted in Section 2.2. The loss function used in the fitting was defined by the squared Mahalanobis's distance (Bishop 2006) for each fragment ($Z = 1 - 5$) as,

$$F_Z^Y(L,R) = \sum_i^{N_thicknesses} \frac{(\mathrm{Sim}_{i,Z}^Y(L,R) - \mathrm{Exp}_{i,Z}^Y)^2}{2(\sigma_{Exp,i,Z}^Y)^2}, \tag{14}$$

where the superscript Y means the total yields and $\mathrm{Sim}_{i,Z}^Y$ and $\mathrm{Exp}_{i,Z}^Y$ are the yields of the atomic number Z in the ith phantom thickness considered in the MC simulation study and experiment, respectively. Here, $N_{thicknesses}$ is the number of phantom thicknesses under study, which depends on the utilized experiments (in this study, $N_{thicknesses} = 7$). Moreover, $\sigma_{Exp,i,Z}^Y$ is the standard deviation of the nuclear fragment yield experimentally observed for a specific Z and phantom thickness with index i. For the angular distributions, the function

$$F_{i,Z}^A(L,R) = \sum_j^{N_{angle}} \frac{(\mathrm{Sim}_{i,j,Z}^A(L,R) - \mathrm{Exp}_{i,j,Z}^A)^2}{2(\sigma_{Exp,i,j,Z}^A)^2}, \tag{15}$$

was defined for each thickness case i, and its mean value over all the phantom thicknesses is

$$F_Z^A(L,R) = \sum_i^{N_{thicknesses}} \frac{F_{i,Z}^A(L,R)}{N_{thicknesses}}, \tag{16}$$

where the superscript A means the angular distribution and $\mathrm{Sim}_{i,j,Z}^A$ and $\mathrm{Exp}_{i,j,Z}^A$ are the yields at angle with index j for a fragment with atomic number Z, obtained with phantom thickness with index i, for MC simulation and the experiment, respectively. Here, N_{angle} is the total number of angle bins considered.

Parameter fitting could be done by only using Eq. (16) for the objective function, because the yield can be calculated by integrating over the angular distribution multiplied by $2\pi \sin\theta$. Due to this multiplication factor, however, the contribution to the total yield of larger angles to the loss function Eq. (16) is smaller. Consequently, the yield is not optimized as intended, if only Eq. (16) were employed. To reproduce the yield and angular distributions, Eqs. (14) and (16) are combined:

$$F_Z(L, R) = F_Z^Y(L, R) + F_Z^A(L, R). \tag{17}$$

For this combination, one must determine the relative weights of the two terms. In this study, $F_Z^Y(L, R)$ and $F_Z^A(L, R)$ had the same order of magnitude for some parameter values. Consequently, Eq. (17) was used in the objective function of the parameter optimization. The objective function planes for L and R were determined by linear interpolation after the polynomial fitting for the data obtained from grid searching in the ranges of $0.5 \leq L \leq 2.25$ fm^2 with a constant spacing of 0.25 fm^2 and of $2.0 \leq R \leq 8.0$ fm with a constant spacing of 1.0 fm. In each grid search, 10^4 carbon ion histories were simulated.

The comparison among the Skyrme models listed in Table 1 was performed after optimizing the parameters L and R. In general, it is difficult to judge the superiority in a model-to-model (or version-to-version also) comparison, particularly, when some objectives are in a trade-off relation, because the result depends sensitively on the definition of the objective function. In the present case, the fragment yields and angular distributions for each element under study emerges from a trade-off relationship; for example, while the yields of $Z = 1$ particles are overestimated, the yields of $Z = 2$ particles are underestimated and vice versa, depending on the specific parameters adopted. To overcome this difficulty, a MOO scheme was adopted to compare quantitatively the different models. Many solutions, called "Pareto optimal solutions", are a result of MOO, and a set of Pareto optimal solutions can form the so-called "Pareto optimal front" in the objective function space (Deb 2011, von Lücken et al. 2014). The Pareto optimal front can be formed by MOO, in particular, with trade-off relations. By

visualizing the Pareto front, called "Pareto optimal", the best model can be determined easily. In this work, three objective functions in MOO were employed:

$$G_1 = F_{Z=1}, \quad G_2 = F_{Z=2}, \quad G_3 = F_{Z=3} + F_{Z=4} + F_{Z=5}, \tag{18}$$

where three objective functions for fragments H (G_1), He (G_2), and the others (G_3) are defined. Reducing the objective functions to three makes it possible to visualize the Pareto front in the objective function space. Total number of fragments is comparable in each objective function, which is also another reason for the choice of Eq. (18). That is, L and R were optimized in each Skyrme model using MOO criteria, where the two-dimensional parameter space was discretized, and three objective values from Eq. (18) were evaluated. This method was used to identify the best model for describing the fragment production in the water phantom experiment used as a reference, based on the formation of the Pareto front. Platypus (Hadka 2017), a framework for evolutionary computing in Python with a focus on multi-objective evolutionary algorithms for MOO, was used, and OMOPSO (Sierra & Coello Coello 2005) was selected as an algorithm for parameter optimization. For this problem, it showed excellent performance in forming a Pareto front efficiently compared with other algorithms available in Platypus. The parameters of OMOPSO were set as follows; population size, 100; swarm size, 100; leader size, 100; mutation probability, 0.1; mutation perturbation, 0.5; and number of function evaluations, 10^4. In the present study, the best solution was also defined as that providing the minimum value of the $L1$ norm of the objective function vector. Owing to the nature of the sampling algorithm, variations in the solution can be observed. Therefore, 10 independent optimization trials were performed, and the mean value of solutions was adopted as the final best solution.

2.5. Evaluation of model parameters

In this study, the experiment is used for two purposes: a) the extraction of model parameters and b) the validation of results. For a), the experimental data of the total

fragment yield and of the angular distribution were employed, and the pareto front of three objective functions in Eq. (18) as well as its $L1$ norm were evaluated in the model comparison. For b), the experimental data of the energy distribution and the charge radius of the ground state nucleus were employed. For energy distributions, the normalized difference (ND) defined in Appendix B was used with the evaluation,

$$F^{E,A}_{i,j,Z} = \sum_{k}^{N_{energy}} \frac{(\mathrm{Sim}^{E,A}_{i,j,k,Z} - \mathrm{Exp}^{E,A}_{i,j,k,Z})^2}{2(\sigma^{E,A}_{Exp,i,j,k,Z})^2}, \tag{19}$$

where the superscript E, A means the energy distribution with a specific angle and N_{energy} is the number of energy bins used to plot the energy distribution (equal to the number of data in the bottom row of Fig. 6). $\mathrm{Sim}^{E,A}_{i,j,k,Z}$ and $\mathrm{Exp}^{E,A}_{i,j,k,Z}$ are the fragment yields for phantom-thickness index i, angle index j, and energy bin k, for atomic number Z, for simulation and experiment, respectively. $\sigma^{E,A}_{Exp,i,j,k,Z}$ is the respective experimental standard deviation. For the charge radius, the absolute difference between simulation and experiment is used to judge which models are reproducible.

3. Results

3.1. Comparison among Skyrme models

Figure 2 shows a comparison in the objective function space using Eq. (18) among the Skyrme models with the JQMD, SLy4, SkM*, and SIII parameter sets, where the (a) G_1-G_2 plane, (b) G_2-G_3 plane, (c) G_3-G_1 plane, and (d) the three-dimensional (3D) plot are depicted. The scattered data (in magenta, cyan, green, and yellow) represent the Pareto fronts obtained using MOO. Then, the parameters L and R giving the minimum value of the $L1$ norm of the objective functions ($G_1 + G_2 + G_3$) with 10 trials of MOO are chosen, which are indicated by circular markers with error bars for each Skyrme model in Figure 2. The corresponding optimized R vs. L plots are shown in Fig. 3, where the circular markers indicate the best parameters in each model that minimize the $L1$ norm. Note that the circular marker in SLy4 overlaps with that in SkM*.

Figure 2 reveals that the Pareto solutions in SLy4 (cyan) and SkM* (magenta) are

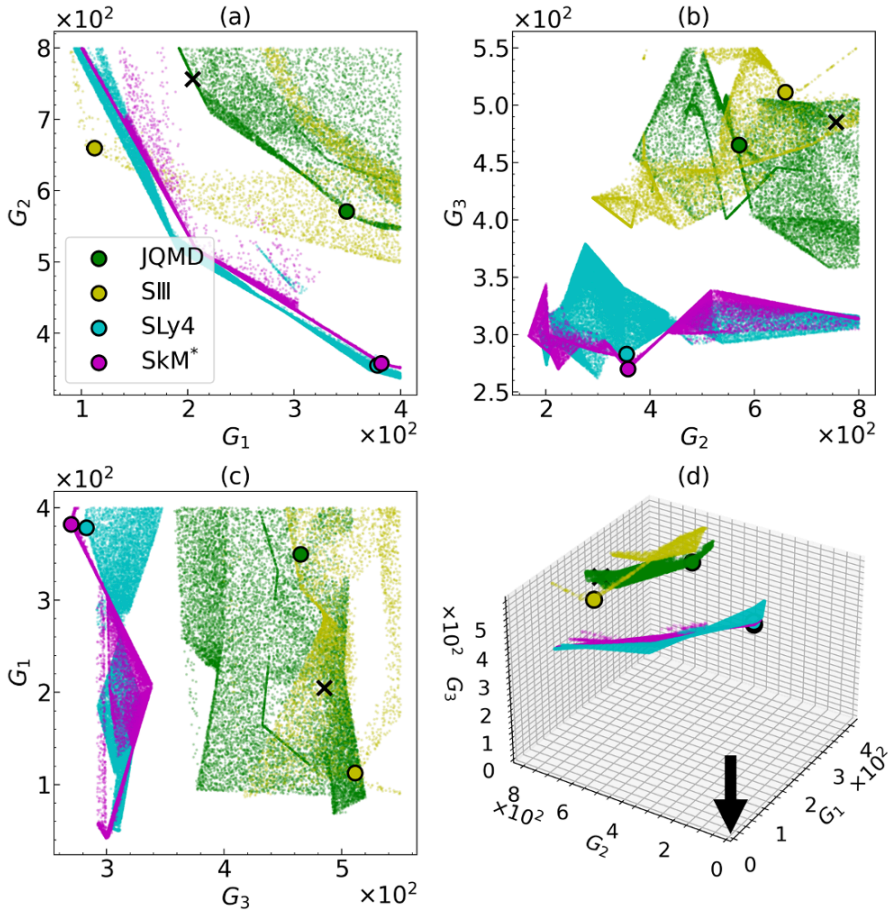

Figure 2. Pareto front in the objective function space: (a) G_1-G_2 plane, (b) G_2-G_3 plane, (c) G_3-G_1 plane, and (d) the three-dimensional plot of the Pareto fronts, for each Skyrme parameter listed in Table 1. The circular markers and the error bars show the mean and the standard deviation of 10 independent multi-objective optimization trials to give the minimum $L1$ norm ($G_1 + G_2 + G_3$) for each Skyrme parameter. The "×" indicates the G4QMD result (JQMD with default parameters). If not visible, the error bars are within the markers.

closer to the origin (black arrow) in the objective function space than those in JQMD (green) and SIII (yellow), suggesting that the SLy4 and SkM* parameter sets could reproduce the experimental data for water phantoms better than those JQMD and SIII. The optimized L and R in SLy4 and SkM* were similar, whereas those in JQMD and SIII were significantly different from the other two. Reasonably, the best L in JQMD was almost the default value (2 fm^2, see Table 2).

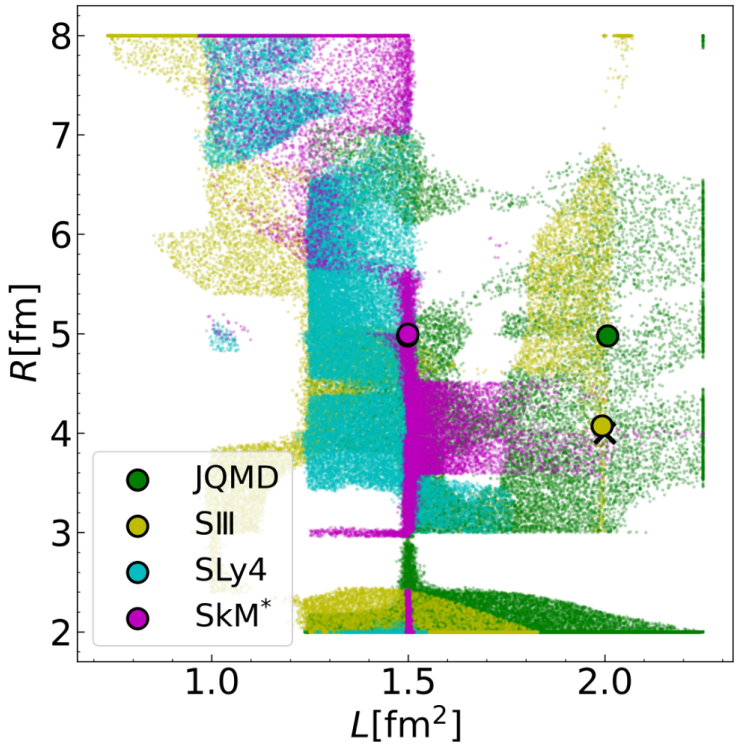

Figure 3. The R vs. L space corresponding to Fig. 2. The colors in marker are the same as in Fig. 2. The circular marker of SLy4 overlaps with that of SkM*. The "×", which overlaps with the circular marker of SIII, indicates G4QMD (JQMD with default parameters). If not visible, the error bars are within the markers.

3.2. Effect of inclusion of the α-cluster model

The effect of the inclusion of the α-cluster model described in Section 2.1.2 in the fragmentation model was investigated using the SLy4 and SkM* parameter sets as shown in Fig. 4. The inclusion of the α-cluster model produces better solutions in the objective function space. In other words, the solutions with the α-cluster model (blue for SLy4α and red for SkM*α) are distributed more closely to the origin than those without the α-cluster model (cyan for SLy4 and magenta for SkM*). The SkM*α performed better than the SLy4α, although the models were almost comparable. The parameters L and R which giving the minimum value of the $L1$ norm in each model are shown in Fig. 5 and Table 2. The optimized values with and without the α-cluster model are similar to each other in SkM* ($(L, R) = (1.50, 5.00)$ for SkM*, and $(1.26, 3.99)$ for

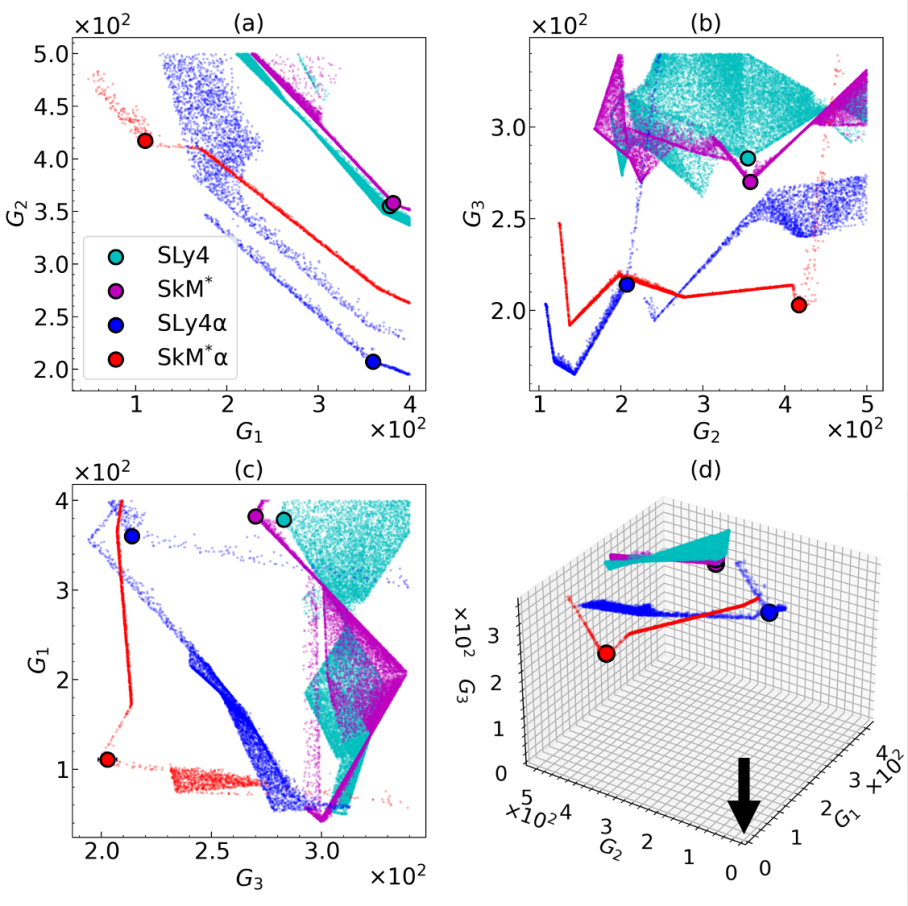

Figure 4. Pareto front in the objective function space for SLy4 and SkM* parameter sets with α-cluster model: (a) G_1-G_2 plane, (b) G_2-G_3 plane, (c) G_3-G_1 plane, and (d) the three-dimensional plot. The circular markers and the error bars show the mean and the standard deviation of 10 independent multi-objective optimization trials to give the minimum $L1$ norm $(G_1 + G_2 + G_3)$. If not visible, the error bars are within the markers.

SkM*α), whereas they are slightly different in SLy4 ($(L, R) = (1.50, 4.98)$ for SLy4, and $(1.00, 6.03)$ for SLy4α), although both Skyrme parameter sets with the α-cluster model can drastically reduce the objective values from the original G4QMD (JQMD with the default parameter setting).

With the estimated best parameters in L and R, fragment simulation was performed by increasing the number of incident carbon ions to 10^6 to reduce the statistical error in the simulation. The representative distributions of the fragment yield with $Z = 1-5$ for some specific models are shown in Fig. 6, where the top, middle, and bottom rows

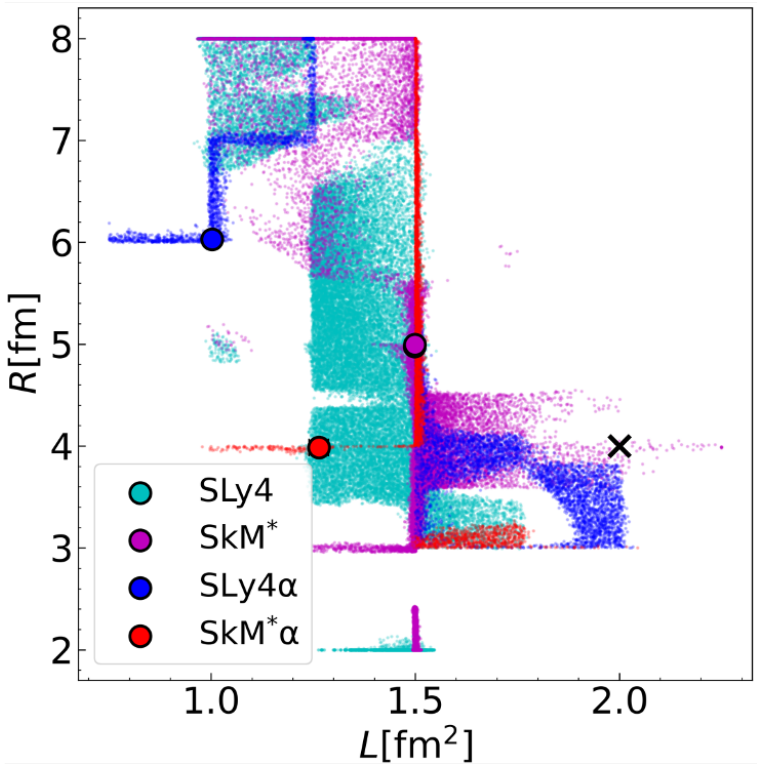

Figure 5. The R vs. L space corresponding to Fig. 4. The colors in marker are the same as in Fig. 4. The circular marker of SLy4 overlaps with that of SkM*. The "×" indicates G4QMD (JQMD with default parameters). If not visible, the error bars are within the markers.

Table 2. Best parameter set and the corresponding objective value with $L1$ norm ($G1 + G2 + G3$) for each model.

Model	L[fm^2]	R[fm]	$L1$ norm
SkM*	1.50	5.00	1009.9
SLy4	1.50	4.98	1016.1
SkM*α	1.26	3.99	730.7
SLy4α	1.00	6.03	781.0
JQMD	2.01	4.98	1385.4
G4QMD†	2.0	4.0	1428.9

†: JQMD with default parameter setting.

indicate the total yield, the angular distribution in the water phantom with a thickness of 258 mm, and energy distribution in the forward direction (angle = 0 degrees) for

the same water phantom, respectively. The simulation results for G4QMD (JQMD with default parameters $(L, R) = (2.0, 4.0)$), the SkM*α(default) (the SkM*α model with default parameters $(L, R) = (2.0, 4.0)$), and the SkM*α(Best) (the SkM*α model with the parameters $(L, R) = (1.26, 3.99)$) are compared with the experimental data. It should be noted here that the two different experimental results are plotted in the total fragment yield distribution of $Z = 5$. In the experiment using a time-of-flight measurement by Haettner et al. (2013), the incident carbons contaminated the signals of the B-fragments, resulting in the black circle data with larger errors in the plot. Haettner et al. conducted the direct measurement of B-fragments independently, which produced the results represented by the open-circled markers in the plot. Data from a time-of-flight measurement were used for parameter optimizations. Nevertheless, the direct measurements are more appropriate for comparison with the simulation results. These distributions show that the nuclear model (orange) is somewhat effective in improving the fragment production in water phantom and for further improvement, the tuning of the QMD parameters (L and R) is required.

3.3. Validation in energy distribution

In this study, the energy distribution was excluded in the objective functions, because the statistical error was relatively large in the simulation for the grid searching of the objective plane, when 10^4 histories were used. Nevertheless, the energy distribution was reproducible using the SkM*α model well. All comparison results between JQMD with default parameters (G4QMD result) and SkM*α with the best parameters are shown in Figs. 8-12 in Appendix B, where the total yields, angular distributions, and energy distributions for each Z are compared. Although the comparison shows that the different models reproduce some results better in some specific cases and less well in others, overall the SkM*α seems to reduce the discrepancy with the experimental data; i.e., the number of positive (negative) values in ND is 82 (70), meaning that SkM*α(Best) is better than G4QMD to reproduce the experimental energy distributions. All the results

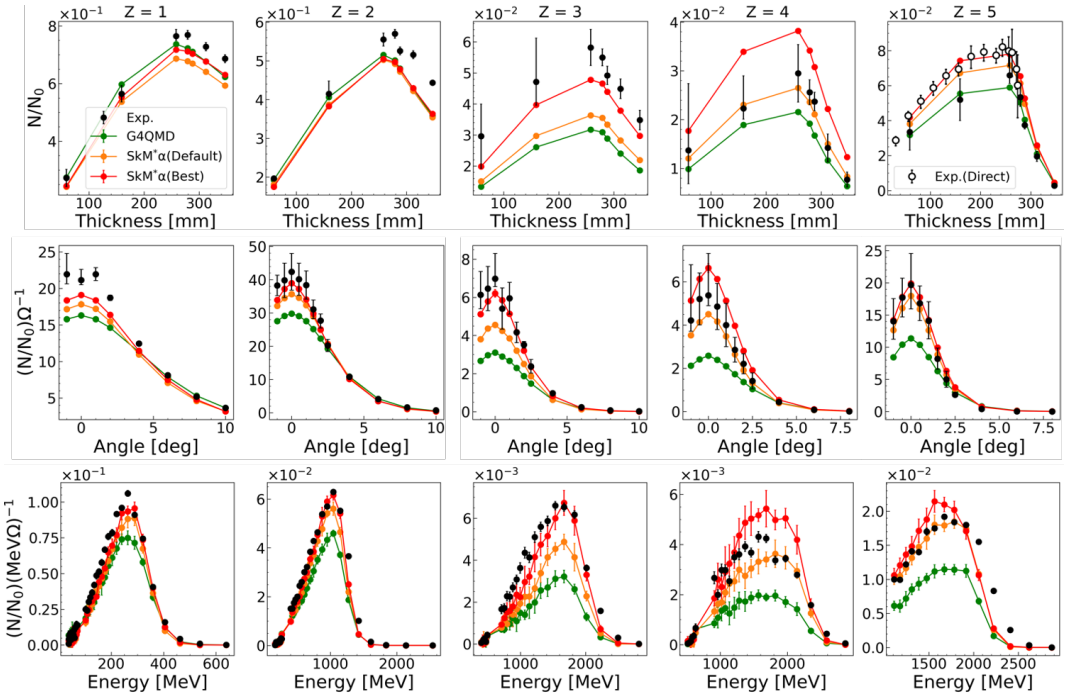

Figure 6. Fragment distributions for simulation and experimental data. The top, middle, and bottom rows indicate the total yield, the angular distribution in the water phantom with a thickness of 258 mm, and energy distribution in the forward direction (angle = 0 degrees) in the same water phantom, respectively. The results for each atomic number are shown in the columns. For the total yield in $Z = 5$ (B-fragments), the experimental data of the time-of-flight measurement (the black circles) and the direct measurement (the open-circles) are plotted (Haettner et al. 2013).

of the fragment distributions, as well as the comparison of the depth dose curves, are given in the Supplementary Information, where the results obtained with both Skyrme QMD model and the BIC model are included and again, it is found included, and it is shown that the experimental data, including energy distributions, can be reproduced by the SkM*α very well.

3.4. Validation of optimized L

The nuclear charge radius is one of the most obvious and important observables that provides information about the nuclear model and the influence of effective interactions on the nuclear structure (Angeli & Marinova 2013). Because the charge radii in

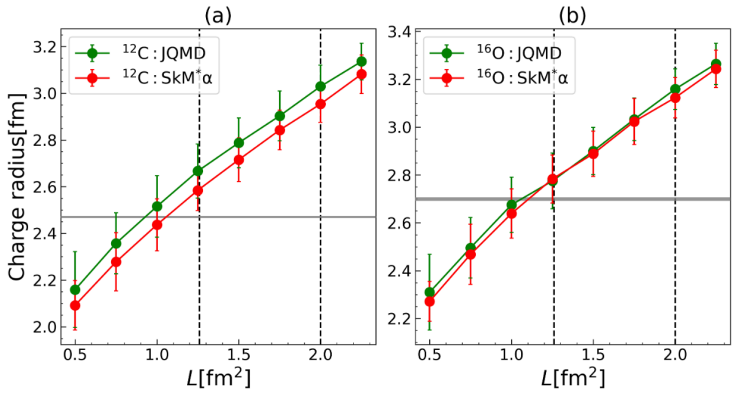

Figure 7. Nuclear charge radius of ^{12}C (a) and ^{16}O (b) as a function of L. The gray horizontal line is the experimental value with the standard deviation as a vertical width. The dashed vertical lines show a default value ($L = 2$) and an optimized value for SkM*α ($L = 1.26$).

the ground state nucleus can be varied by sensitively changing the squared standard deviation of the Gaussian distribution L in the QMD model, here, the optimized value of L in SkM*α as compared with JQMD is validated. Figures 7(a) and (b) show the nuclear charge radius for ^{12}C and ^{16}O, respectively, simulated 100 times with the QMD model as a function of L, where the experimental data were obtained from Angeli & Marinova (2013). With both models, it has been resulted that L should be approximately 1.0 to reproduce the charge radii. The optimized L in SkM*α ($L = 1.26$) seems to be a reasonable parameter value compared with the default one ($L = 2.0$) and the optimized one in JQMD ($L = 2.01$), also implying that the SkM*α is superior to JQMD.

4. Discussion

In this study, SkM* and SLy4 parameter sets of the improved QMD produced physical distributions of secondary fragments closer to the reference experimental results than JQMD and SIII. In addition, SkM* was slightly superior to SLy4 in the $L1$ norm as seen in Table 2. The SkM* has a refined model that includes the nuclear surface tension, introduced to improve the incompressibility modulus of nuclear matter, whereas SIII gives an incompressibility that is too high ($K_0 \sim 365$ MeV) (Bartel et al. 1982, Beiner

et al. 1975). The advantage of using the ImQMD model is that the model parameters are obtained directly from the well-established Skyrme interaction. In addition, the effects emerging from the terms with g_0, g_τ, and κ_s in Eq. (4) were introduced. Because these parameters are related to the surface energy density (g_0 and κ_s) and the kinetic energy density (g_τ), including the derivatives of density, the improvement obtained by using SkM* suggests that the surface-related terms are significant in the fragment simulation for the light projectile/target system. Although it was not the best model in the present analysis, SLy4 is also one of the refined parameter sets in the Skyrme model, characterised by a greater emphasis on the isospin degrees of freedom. A series of sets of SLy parameterizations was proposed to reproduce the properties of the nuclei from the β stability line to the drip lines (Chabanat et al. 1998). The results of the present analysis reflects the experimental situation in which the fragment products were not distinguished by isotope but rather by the atomic number, so that the isospin properties in the fragmentation would be weakly observed. In addition to the importance of the nuclear surface potential, the fact that the Pareto front obtained by MOO using SkM*α (the SkM parameter with the α-cluster model), was closer to the origin than without using it, implies that the α-cluster structure in the ground state of the projectile and the target is non-negligible in the fragmentation. Thus, we can say that the modern Skyrme parameters combined with the α-cluster structure can significantly improve the performance in describing the fragmentation process in carbon ion therapy. However, there remains a trade-off among objective functions for fragments with $Z = 1, 2$ and for those with $Z \geq 3$. This indicates that it remains difficult to reproduce the experimental results in all fragments simultaneously, although this difficulty might be solved by additional tuning of the parameters in Skyrme model shown in Table 1.

The QMD parameters, including the squared standard deviation L of the Gaussian distribution that describes the nucleons and the threshold of the cluster radius R, were optimized for each nuclear model. Whether the optimized values are reasonable for the analysis of other physical properties should be verified. For L, in the present work,

the charge radii for the ground state in ^{12}C and ^{16}O formed by QMD simulation with each nuclear model were checked. It was confirmed again that the charge radii using optimized parameters in SkM* and SLy4 were more realistic than those in JQMD and SIII. It should be noted that this result is aligned with the modifications made by Dedes et al. (2014), who found that $L = 0.8$ fm^2 gave a significant improvement to prompt gamma yields. A confirmation of this will require our future study. In contrast, all nuclear models yielded the similar optimized values, implying that the cluster radius is model independent. This was expected, because R can directly affect to the fragment formation at the end of the time evolution in inelastic collisions, without the knowledge of the nuclear interaction. However, R could be isospin-dependent, as proposed for neutron-rich nuclei such as the neutron skin or neutron halo effect and a Coulomb effect, for example, $R_{nn} = R_{np} = 6$ fm and $R_{pp} = 3$ fm (Zhang et al. 2012). It is difficult to optimize the iso-R in the current study, where only fragments with different Z were distinguished. Isotope analysis with a heavier target nucleus provides the isospin-dependent radius for the cluster recognition.

The present work showed that the appropriate nuclear model with optimized QMD parameters reproduces not only the properties used in the optimization (yield and its angular distribution) but also the properties in the verification (energy distribution and charge radii). Therefore, the best model can be selected through verification analysis. Even so, it would still be difficult to determine with certainty whether the selected model is superior to the others without any indicator. In this study, using MOO to identify the best QMD models under study was proposed. From the visualization in the multi-objective space, as shown in Figs. 2 and 4, one can confidently conclude that SkM* and SLy4 are better models than JQMD and SIII. In addition, MOO showed that SkM* with the α-cluster structure model is superior to the others. One can also conclude that MOO can be useful for further model tuning in G4QMD.

In this study, three methods to improve G4QMD were considered. As seen in Table 2, the introduction of the modern Skyrme-type interaction reduced the $L1$ norm

($G_1 + G_2 + G_3$ in Eq. (18)) to be 1009.9 for SkM* and 1016.1 for SLy4 from the default G4QMD (1428.9). The introduction of the α-cluster model reduced it further, 1009.9 $\rightarrow$ 730.7 for SkM*α and 1016.1 $\rightarrow$ 781.0 for SLy4α. The effect of the parameter optimization was obvious; The SkM*α model with the best parameters reproduced the experimental data better than that with the default parameters, as shown in Fig. 6. Therefore it can be said that all of three methods are effective to improve the discrepancy between the simulation and the experiment.

The proposed improved QMD model was developed in Geant4, and it improved the fragmentation simulation in water phantoms as described. However, other tasks remain; in this study, only carbon ions incident at 400 MeV/u on a water target were investigated. For application in hadron therapy, the developed model should be verified in target materials of interest, e.g. calcium and other tissues. For this purpose, the comparison with experimental fragmentation data would be useful. This study should also be repeated for the entire energy range of the carbon ions of clinical interest. Other parameters of the QMD model (and not only L and R that were chosen here) should also be tuned. These include the maximum time to simulate the inelastic collision by QMD, the threshold for the violation of energy conservation during nucleon-nucleon and/or nucleus-nucleus collision, the sampling radius of the b-value, and new parameters such as the isospin dependence of cluster judgement. For the nuclear model, the relativistic mean field approach has been developed recently in the QMD model (Nara & Stoecker 2019, Nara et al. 2020). Because the nuclear force is essentially different from that of Skyrme model, drastic changes can be expected when tuning these extra parameters.

5. Conclusion

Geant4 QMD model was expanded to make it possible to use various Skyrme parameter sets. In addition, an *ad hoc* α-cluster model was implemented to describe the formation of the ground state nucleus. After the parameters L and R were optimized, the developed models considerably reduced the discrepancy with the reference experimental data. This

work highlighted that MOO can be a powerful method for improving physical models.

The potential of the QMD model to describe the nuclear fragmentation satisfactorily was demonstrated. To improve this further, optimizing the remaining parameters of the QMD model using additional experimental data as a reference is planned, as well as further developing the model within the MOO framework.

Acknowledgments

This study has been supported by JSPS KAKENHI (Grant No. 19K08201). The simulation has been partly done on a computer server at Research Center for Nuclear Physics, Osaka University.

Appendix A

Geant4 models used in the current work are listed in Table. 3. It should be noted here that we adopted the QMD model in hadron inelastic process even for the proton projectile, which is treated as a secondary particle in the present simulation.

Table 3. Geant4 models used for the simulation of hadronic interactions.

Interaction	Projectile	Energy range	G4Model
Radioactive Decay	GenericIon	–	G4RadioactiveDecayPhysics
Hadron Elastic	GenericIon	0 – 100 TeV/u	G4HadronElasticPhysics
	d, t, ^{3}He, α		
	p, n	0 – 100 TeV	G4ChipsElasticModel
Hadron Inelastic	GenericIon	0 – 100 MeV/n	BIC*
		100 MeV/u – 10 GeV/u	QMD**
	p, d, t, ^{3}He, α	0 – 100 MeV/n	BIC*
		100 MeV/u – 10 GeV/u	QMD**
	n	0–20 MeV	G4NeutronHPInelastic
		20 MeV – 6 GeV	BIC*
Fission	n	0 – 20 MeV	G4NeutronHPFission
		20 MeV – 100 TeV	G4LFission
Capture	n	0 – 20 MeV	G4NeutronHPCapture
		20 MeV – 100 TeV	G4nRadCapture

$*$: "BIC" means binary cascade, binary ion cascade, and binary light ion cascade, for neutron, proton, and the other particles, respectively.

$**$: The Geant4 built-in or the developed QMD in this study is used.

Appendix B

In this section, instead of plotting all the fragment distributions (these are provided in Fig. 6 and in the Supplementary Information), we show the results of quantitative comparisons between G4QMD and SkM*α(Best) using the normalized difference (ND):

$$ND = \frac{F_{\mathrm{G4QMD}} - F_{\mathrm{SkM^*\alpha(Best)}}}{F_{\mathrm{G4QMD}} + F_{\mathrm{SkM^*\alpha(Best)}}}, \tag{20}$$

where F_{model} is given by Eq. (14) and Eq. (15) for fragment yield and angular distribution analysis respectively. For energy distribution analysis, F_{model} is evaluated in each panel of the distributions represented in Fig. 6 by using Eq. (19).

The results are shown in Figs. 8-12, where the labels attached to the bar charts are used to identify the corresponding figures (Fig. 6 and the ones provided in the Supplementary Information). For example, "Y-1" means a total fragment yield for $Z = 1$, "A-1-59" means an angular distribution of $Z = 1$ with phantom thickness of 59mm, "E-1-258-0" means an energy distribution of $Z = 1$ with phantom thickness of 258mm at 0-degrees emission angle, etc. The positive (negative) ND means that SkM*α(Best) is better (worse) than G4QMD. For total yields and angular distributions, which are used in the parameter optimization, the number of positive (negative) values is 26 (14), whereas for energy distributions, which are used in the validation, the number of positive (negative) values is 82 (70), meaning that SkM*α(Best) is better than G4QMD to reproduce comprehensively the experimental data. Note that above counts are given with the time-of-flight measurement for $Z = 5$ fragment, and it is expected that the number of positive values are increased if the energy and angular distributions in the direct measurement are applicable.

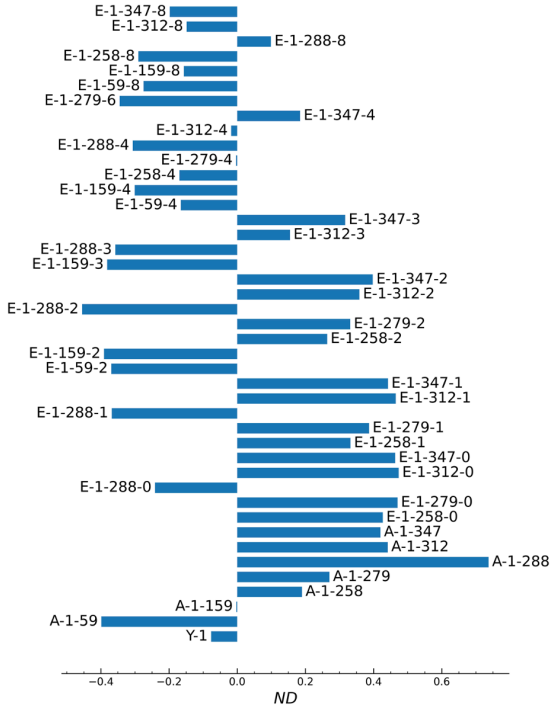

Figure 8. Normalized difference (ND) in total yield, angular distribution, and energy distribution for $Z = 1$. G4QMD and SkM*α(Best) are compared.

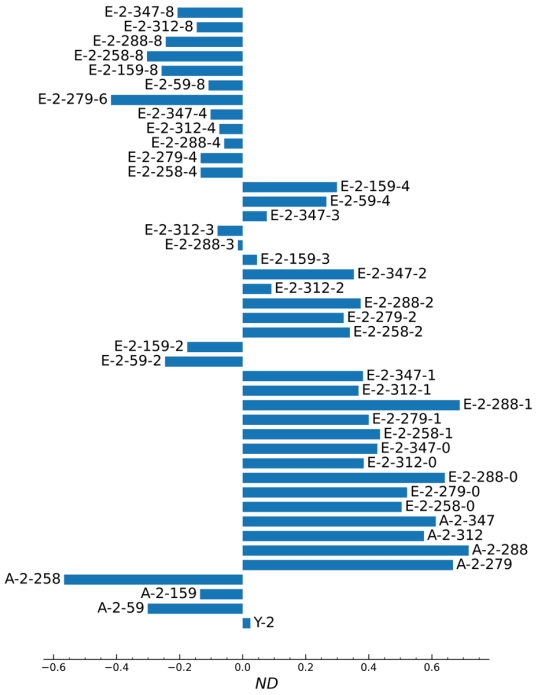

Figure 9. Same as in Fig. 8 except for $Z = 2$.

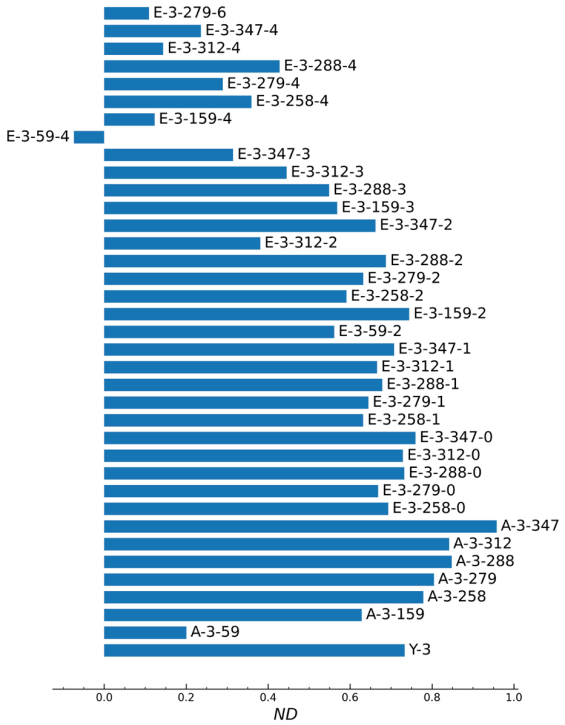

Figure 10. Same as in Fig. 8 except for $Z = 3$.

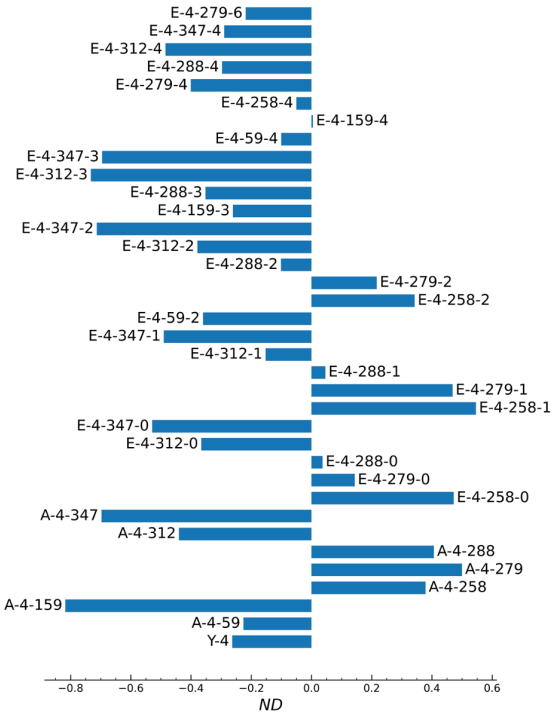

Figure 11. Same as in Fig. 8 except for $Z = 4$.

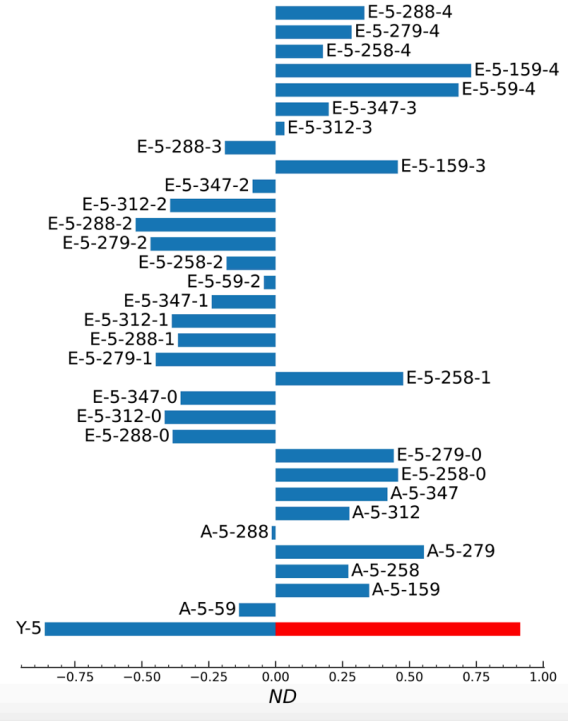

Figure 12. Same as in Fig. 8 except for $Z = 5$, where the different color bars indicate the different sets of experimental data used as a reference (Haettner et al. 2013)

.

References

Agostinelli, S., Allison, J., Amako, K., Apostolakis, J., Araujo, H., Arce, P., Asai, M., Axen, D., Banerjee, S., Barrand, G., Behner, F., Bellagamba, L., Boudreau, J., Broglia, L., Brunengo, A., Burkhardt, H., Chauvie, S., Chuma, J., Chytracek, R., Cooperman, G., Cosmo, G., Degtyarenko, P., Dell'Acqua, A., Depaola, G., Dietrich, D., Enami, R., Feliciello, A., Ferguson, C., Fesefeldt, H., Folger, G., Foppiano, F., Forti, A., Garelli, S., Giani, S., Giannitrapani, R., Gibin, D., Gómez Cadenas, J., González, I., Gracia Abril, G., Greeniaus, G., Greiner, W., Grichine, V., Grossheim, A., Guatelli, S., Gumplinger, P., Hamatsu, R., Hashimoto, K., Hasui, H., Heikkinen, A., Howard, A., Ivanchenko, V., Johnson, A., Jones, F., Kallenbach, J., Kanaya, N., Kawabata, M., Kawabata, Y., Kawaguti, M., Kelner, S., Kent, P., Kimura, A., Kodama, T., Kokoulin, R., Kossov, M., Kurashige, H., Lamanna, E., Lampén, T., Lara, V., Lefebure, V., Lei, F., Liendl, M., Lockman, W., Longo, F., Magni, S., Maire, M., Medernach, E., Minamimoto, K., Mora de Freitas, P., Morita, Y., Murakami, K., Nagamatu, M., Nartallo, R., Nieminen, P., Nishimura, T., Ohtsubo, K., Okamura, M., O'Neale, S., Oohata, Y., Paech, K., Perl, J., Pfeiffer, A., Pia, M., Ranjard, F., Rybin, A., Sadilov, S., Di Salvo, E., Santin, G., Sasaki, T., Savvas, N., Sawada, Y., Scherer, S., Sei, S., Sirotenko, V., Smith, D., Starkov, N., Stoecker, H., Sulkimo, J., Takahata, M., Tanaka, S., Tcherniaev, E., Safai Tehrani, E., Tropeano, M., Truscott, P., Uno, H., Urban, L., Urban, P., Verderi, M., Walkden, A., Wander, W., Weber, H., Wellisch, J., Wenaus, T., Williams, D., Wright, D., Yamada, T., Yoshida, H. & Zschiesche, D. (2003), 'Geant4—a simulation toolkit', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506**, 250–303.

Allison, J., Amako, K., Apostolakis, J., Araujo, H., Arce Dubois, P., Asai, M., Barrand, G., Capra, R., Chauvie, S., Chytracek, R., Cirrone, G., Cooperman, G., Cosmo, G., Cuttone, G., Daquino, G., Donszelmann, M., Dressel, M., Folger, G.,

Foppiano, F., Generowicz, J., Grichine, V., Guatelli, S., Gumplinger, P., Heikkinen, A., Hrivnacova, I., Howard, A., Incerti, S., Ivanchenko, V., Johnson, T., Jones, F., Koi, T., Kokoulin, R., Kossov, M., Kurashige, H., Lara, V., Larsson, S., Lei, F., Link, O., Longo, F., Maire, M., Mantero, A., Mascialino, B., McLaren, I., Mendez Lorenzo, P., Minamimoto, K., Murakami, K., Nieminen, P., Pandola, L., Parlati, S., Peralta, L., Perl, J., Pfeiffer, A., Pia, M., Ribon, A., Rodrigues, P., Russo, G., Sadilov, S., Santin, G., Sasaki, T., Smith, D., Starkov, N., Tanaka, S., Tcherniaev, E., Tome, B., Trindade, A., Truscott, P., Urban, L., Verderi, M., Walkden, A., Wellisch, J., Williams, D., Wright, D. & Yoshida, H. (2006), 'Geant4 developments and applications', *IEEE Transactions on Nuclear Science* **53**, 270–278.

Allison, J., Amako, K., Apostolakis, J., Arce, P., Asai, M., Aso, T., Bagli, E., Bagulya, A., Banerjee, S., Barrand, G., Beck, B., Bogdanov, A., Brandt, D., Brown, J., Burkhardt, H., Canal, P., Cano-Ott, D., Chauvie, S., Cho, K., Cirrone, G., Cooperman, G., Cortés-Giraldo, M., Cosmo, G., Cuttone, G., Depaola, G., Desorgher, L., Dong, X., Dotti, A., Elvira, V., Folger, G., Francis, Z., Galoyan, A., Garnier, L., Gayer, M., Genser, K., Grichine, V., Guatelli, S., Guèye, P., Gumplinger, P., Howard, A., Hřivnáčová, I., Hwang, S., Incerti, S., Ivanchenko, A., Ivanchenko, V., Jones, F., Jun, S., Kaitaniemi, P., Karakatsanis, N., Karamitros, M., Kelsey, M., Kimura, A., Koi, T., Kurashige, H., Lechner, A., Lee, S., Longo, F., Maire, M., Mancusi, D., Mantero, A., Mendoza, E., Morgan, B., Murakami, K., Nikitina, T., Pandola, L., Paprocki, P., Perl, J., Petrović, I., Pia, M., Pokorski, W., Quesada, J., Raine, M., Reis, M., Ribon, A., Ristić Fira, A., Romano, F., Russo, G., Santin, G., Sasaki, T., Sawkey, D., Shin, J., Strakovsky, I., Taborda, A., Tanaka, S., Tomé, B., Toshito, T., Tran, H., Truscott, P., Urban, L., Uzhinsky, V., Verbeke, J., Verderi, M., Wendt, B., Wenzel, H., Wright, D., Wright, D., Yamashita, T., Yarba, J. & Yoshida, H. (2016), 'Recent developments in Geant4', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **835**, 186–225.

Ando, K. & Kase, Y. (2009), 'Biological characteristics of carbon-ion therapy', *International journal of radiation biology* **85**(9), 715–728.

Angeli, I. & Marinova, K. P. (2013), 'Table of experimental nuclear ground state charge radii: An update', *Atomic Data and Nuclear Data Tables* **99**(1), 69–95.

Arce, P., Bolst, D., Bordage, M.-C., Brown, J., Cirrone, P., Cortés-Giraldo, M. A., Cutajar, D., Cuttone, G., Desorgher, L., Dondero, P. et al. (2021), 'Report on G4-Med, a Geant4 benchmarking system for medical physics applications developed by the Geant4 Medical Simulation Benchmarking Group', *Medical physics* **48**(1), 19–56.

Bartel, J., Quentin, P., Brack, M., Guet, C. & Håkansson, H.-B. (1982), 'Towards a better parametrisation of Skyrme-like effective forces: A critical study of the SkM force', *Nuclear Physics A* **386**(1), 79–100.

Bayram, T., Akkoyun, S. & Kara, S. O. (2014), 'A study on ground-state energies of nuclei by using neural networks', *Annals of Nuclear Energy* **63**, 172–175.

Beiner, M., Flocard, H., Van Giai, N. & Quentin, P. (1975), 'Nuclear ground-state properties and self-consistent calculations with the Skyrme interaction:(I). Spherical description', *Nuclear Physics A* **238**(1), 29–69.

Bertolli, O., Eleftheriou, A., Cecchetti, M., Camarlinghi, N., Belcari, N. & Tsoumpas, C. (2016), 'PET iterative reconstruction incorporating an efficient positron range correction method', *Physica medica* **32**(2), 323–330.

Bishop, C. M. (2006), *Pattern Recognition and Machine Learning*, Springer, Berlin.

Bolst, D., Cirrone, G. A., Cuttone, G., Folger, G., Incerti, S., Ivanchenko, V., Koi, T., Mancusi, D., Pandola, L., Romano, F. et al. (2017), 'Validation of Geant4 fragmentation for heavy ion therapy', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **869**, 68–75.

Chabanat, E., Bonche, P., Haensel, P., Meyer, J. & Schaeffer, R. (1998), 'A Skyrme

parametrization from subnuclear to neutron star densities Part II. Nuclei far from stabilities', *Nuclear Physics A* **635**(1-2), 231–256.

Chacon, A., Guatelli, S., Rutherford, H., Bolst, D., Mohammadi, A., Ahmed, A., Nitta, M., Nishikido, F., Iwao, Y., Tashima, H. et al. (2019), 'Comparative study of alternative Geant4 hadronic ion inelastic physics models for prediction of positron-emitting radionuclide production in carbon and oxygen ion therapy', *Physics in Medicine & Biology* **64**(15), 155014.

Chen, J., Yun, S., Dong, T., Ren, Z. & Zhang, X. (2018), 'Validation of Geant4 physics models for nuclear beams in extended media', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **434**, 113–119.

Deb, K. (2011), Multi-objective optimisation using evolutionary algorithms: an introduction, *in* 'Multi-objective evolutionary optimisation for product design and manufacturing', Springer, London, pp. 3–34.

Dedes, G., Pinto, M., Dauvergne, D., Freud, N., Krimmer, J., Létang, J., Ray, C. & Testa, E. (2014), 'Assessment and improvements of Geant4 hadronic models in the context of prompt-gamma hadrontherapy monitoring', *Physics in Medicine & Biology* **59**(7), 1747–1772.

Dobaczewski, J., Nazarewicz, W. & Stoitsov, M. (2003), Nuclear ground-state properties from mean-field calculations, *in* 'Exotic Nuclei and Atomic Masses', Springer, Berlin, Heidelberg, pp. 55–60.

Dudouet, J., Cussol, D., Durand, D. & Labalme, M. (2014), 'Benchmarking Geant4 nuclear models for hadron therapy with 95 MeV/nucleon carbon ions', *Physical Review C* **89**(5), 054616.

Dudouet, J., Juliani, D., Labalme, M., Cussol, D., Angélique, J., Braunn, B., Colin, J., Finck, C., Fontbonne, J., Guérin, H. et al. (2013), 'Double-differential fragmentation

cross-section measurements of 95 MeV/nucleon ^{12}C beams on thin targets for hadron therapy', *Physical Review C* **88**(2), 024606.

Dutra, M., Lourenço, O., Martins, J. S., Delfino, A., Stone, J. R. & Stevenson, P. (2012), 'Skyrme interaction and nuclear matter constraints', *Physical Review C* **85**(3), 035201.

Elsässer, T., Weyrather, W. K., Friedrich, T., Durante, M., Iancu, G., Krämer, M., Kragl, G., Brons, S., Winter, M., Weber, K.-J. et al. (2010), 'Quantification of the relative biological effectiveness for ion beam radiotherapy: direct experimental comparison of proton and carbon ion beams and a novel approach for treatment planning', *International Journal of Radiation Oncology* Biology* Physics* **78**(4), 1177–1183.

Folger, G., Ivanchenko, V. N. & Wellisch, J. P. (2004), 'The binary cascade', *The European Physical Journal A-Hadrons and Nuclei* **21**(3), 407–417.

Geant4 Collaboration (2020), 'Physics reference manual documentation release 10.7, https://geant4-userdoc.web.cern.ch/usersguides/physicsreferencemanual/backupversions/v10.7/fo/physicsreferencemanual.pdf', p. 374.

Habermehl, D., Debus, J., Ganten, T., Ganten, M.-K., Bauer, J., Brecht, I. C., Brons, S., Haberer, T., Haertig, M., Jäkel, O. et al. (2013), 'Hypofractionated carbon ion therapy delivered with scanned ion beams for patients with hepatocellular carcinoma–feasibility and clinical response', *Radiation oncology* **8**(1), 1–8.

Hadka, D. (2017), 'Platypus documentation'.
URL: *https://readthedocs.org/projects/platypus/downloads/pdf/latest/*

Haettner, E., Iwase, H., Krämer, M., Kraft, G. & Schardt, D. (2013), 'Experimental study of nuclear fragmentation of 200 and 400 MeV/u ^{12}C ions in water for applications in particle therapy', *Physics in Medicine & Biology* **58**(23), 8265–8279.

Haettner, E., Iwase, H. & Schardt, D. (2006), 'Experimental fragmentation studies with ^{12}C therapy beams', *Radiation protection dosimetry* **122**(1-4), 485–487.

Huang, B.-S. & Ma, Y.-G. (2021), 'Dipole excitation of ^{6}Li and ^{9}Be studied with an extended quantum molecular dynamics model', *Physical Review C* **103**(5), 054318.

Inaniwa, T., Kanematsu, N., Hara, Y. & Furukawa, T. (2014), 'Nuclear-interaction correction of integrated depth dose in carbon-ion radiotherapy treatment planning', *Physics in Medicine & Biology* **60**(1), 421–435.

Inaniwa, T., Lee, S. H., Mizushima, K., Sakata, D., Iwata, Y., Kanematsu, N. & Shirai, T. (2020), 'Nuclear-interaction correction for patient dose calculations in treatment planning of helium-, carbon-, oxygen-, and neon-ion beams', *Physics in Medicine & Biology* **65**(2), 025004.

Kanada-En'yo, Y., Kimura, M. & Ono, A. (2012), 'Antisymmetrized molecular dynamics and its applications to cluster phenomena', *Progress of Theoretical and Experimental Physics* **2012**(1).

Karger, C. P. & Peschke, P. (2017), 'RBE and related modeling in carbon-ion therapy', *Physics in Medicine & Biology* **63**(1), 01TR02.

Kawashiro, S., Yamada, S., Okamoto, M., Ohno, T., Nakano, T., Shinoto, M., Shioyama, Y., Nemoto, K., Isozaki, Y., Tsuji, H. & Kamada, T. (2018), 'Multi-institutional study of carbon-ion radiotherapy for locally advanced pancreatic cancer: Japan Carbon-ion Radiation Oncology Study Group (j-cros) Study 1403 Pancreas', *International Journal of Radiation Oncology* Biology* Physics* **101**(5), 1212–1221.

Kean, K. R., Nishikawa, T. & Iwata, Y. (2020), Perturbation scheme for the effective nuclear force, *in* 'Proceedings of 13th International Conference on Nucleus-Nucleus Collisions', p. 010018.

Komatsu, S., Fukumoto, T., Demizu, Y., Miyawaki, D., Terashima, K., Sasaki, R., Hori, Y., Hishikawa, Y., Ku, Y. & Murakami, M. (2011), 'Clinical results and risk factors of proton and carbon ion therapy for hepatocellular carcinoma', *Cancer* **117**(21), 4890–4904.

Krämer, M. & Durante, M. (2010), 'Ion beam transport calculations and treatment plans in particle therapy', *The European Physical Journal D* **60**(1), 195–202.

Mancusi, D., Niita, K., Maruyama, T. & Sihver, L. (2009), 'Stability of nuclei in peripheral collisions in the JAERI quantum molecular dynamics model', *Physical Review C* **79**(1), 014614.

Matsufuji, N., Fukumura, A., Komori, M., Kanai, T. & Kohno, T. (2003), 'Influence of fragment reaction of relativistic heavy charged particles on heavy-ion radiotherapy', *Physics in Medicine & Biology* **48**(11), 1605–1623.

Napoli, M. D., Agodi, C., Battistoni, G., Blancato, A. A., Cirrone, G. A. P., Cuttone, G., Giacoppo, F., Morone, M. C., Nicolosi, D., Pandola, L., Patera, V., Raciti, G., Rapisarda, E., Romano, F., Sardina, D., Sarti, A., Sciubba, A., Scuderi, V., Sfienti, C. & Tropea, S. (2012), 'Carbon fragmentation measurements and validation of the Geant4 nuclear reaction models for hadrontherapy', *Physics in Medicine and Biology* **57**(22), 7651–7671.

URL: *https://doi.org/10.1088/0031-9155/57/22/7651*

Nara, Y., Maruyama, T. & Stoecker, H. (2020), 'Momentum-dependent potential and collective flows within the relativistic quantum molecular dynamics approach based on relativistic mean-field theory', *Physical Review C* **102**(2), 024913.

Nara, Y. & Stoecker, H. (2019), 'Sensitivity of the excitation functions of collective flow to relativistic scalar and vector meson interactions in the relativistic quantum molecular dynamics model RQMD. RMF', *Physical Review C* **100**(5), 054902.

Niita, K., Chiba, S., Maruyama, T., Maruyama, T., Takada, H., Fukahori, T., Nakahara, Y. & Iwamoto, A. (1995), 'Analysis of the (N, xN′) reactions by quantum molecular dynamics plus statistical decay model', *Physical Review C* **52**(5), 2620–2635.

Niita, K., Maruyama, T., Nara, Y., Chiba, S. & Iwamoto, A. (1999), Development of JQMD (JAERI quantum molecular dynamics) code, Technical report, Japan Atomic Energy Research Inst.

Ohnishi, A., Maruyama, T. & Horiuchi, H. (1992), 'Near- and far-side contributions to ejectile polarization', *Progress of Theoretical Physics* **87**(2), 417–432.

Ono, A. (2019), 'Dynamics of clusters and fragments in heavy-ion collisions', *Progress in Particle and Nuclear Physics* **105**, 139–179.

Rahmim, A., Qi, J. & Sossi, V. (2013), 'Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls', *Medical physics* **40**(6Part1), 064301.

Sakata, D., Guatelli, S. & Simpson, E. C. (2020), 'Improved integrated nucleus-nucleus inelastic cross sections for light nuclides in Geant4', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **463**, 27–29.

Schardt, D., Elsässer, T. & Schulz-Ertner, D. (2010), 'Heavy-ion tumor therapy: Physical and radiobiological benefits', *Reviews of Modern Physics* **82**(1), 383.

Schardt, D., Schall, I., Geissel, H., Irnich, H., Kraft, G., Magel, A., Mohar, M., Münzenberg, G., Nickel, F., Scheidenberger, C., Schwab, W. & Sihver, L. (1996), 'Nuclear fragmentation of high-energy heavy-ion beams in water', *Advances in Space Research* **17**(2), 87–94.

Schulz-Ertner, D., Nikoghosyan, A., Thilmann, C., Haberer, T., Jäkel, O., Karger, C., Kraft, G., Wannenmacher, M. & Debus, J. (2004), 'Results of carbon ion radiotherapy in 152 patients', *International journal of radiation oncology* biology* physics* **58**(2), 631–640.

Schwaab, J., Brons, S., Fieres, J. & Parodi, K. (2011), 'Experimental characterization of lateral profiles of scanned proton and carbon ion pencil beams for improved beam models in ion therapy treatment planning', *Physics in Medicine & Biology* **56**(24), 7813–7827.

Sierra, M. R. & Coello Coello, C. A. (2005), Improving PSO-based multi-objective optimization using crowding, mutation and ϵ-dominance, *in* C. A. Coello Coello,

A. Hernández Aguirre & E. Zitzler, eds, 'Evolutionary Multi-Criterion Optimization', Springer, Berlin, Heidelberg, Berlin, Heidelberg, pp. 505–519.

Torikoshi, M., Minohara, S., Kanematsu, N., Komori, M., Kanazawa, M., Noda, K., Miyahara, N., Itoh, H., Endo, M. & Kanai, T. (2007), 'Irradiation system for HIMAC', *Journal of Radiation Research* **48**(Suppl_A), A15–A25.

Tran, L. T., Bolst, D., Guatelli, S., Pogossov, A., Petasecca, M., Lerch, M. L., Chartier, L., Prokopovich, D. A., Reinhard, M. I., Povoli, M. et al. (2018), 'The relative biological effectiveness for carbon, nitrogen, and oxygen ion beams using passive and scanning techniques evaluated with fully 3D silicon microdosimeters', *Medical physics* **45**(5), 2299–2308.

von Lücken, C., Barán, B. & Brizuela, C. (2014), 'A survey on multi-objective evolutionary algorithms for many-objective problems', *Computational Optimization and Applications* **58**, 707—756.

Wang, N., Li, Z. & Wu, X. (2002), 'Improved quantum molecular dynamics model and its applications to fusion reaction near barrier', *Physical Review C* **65**(6), 064608.

Wang, N., Li, Z., Wu, X., Tian, J., Zhang, Y. X. & Liu, M. (2004), 'Further development of the improved quantum molecular dynamics model and its application to fusion reactions near the barrier', *Physical Review C* **69**(3), 034608.

Zhang, Y. & Li, Z. (2006), 'Elliptic flow and system size dependence of transition energies at intermediate energies', *Physical Review C* **74**(1), 014602.

Zhang, Y., Li, Z., Zhou, C. & Tsang, M. (2012), 'Effect of isospin-dependent cluster recognition on the observables in heavy ion collisions', *Physical Review C* **85**(5), 051602.

Zhang, Y., Tsang, M., Li, Z. & Liu, H. (2014), 'Constraints on nucleon effective mass splitting with heavy ion collisions', *Physics Letters B* **732**, 186–190.

Zhang, Y.-X., Wang, N., Li, Q.-F., Ou, L., Tian, J.-L., Liu, M., Zhao, K., Wu, X.-Z. &

Li, Z.-X. (2020), 'Progress of quantum molecular dynamics model and its applications in heavy ion collisions', *Frontiers of Physics* **15**(5), 1–64.