|  |
| --- |
| P2.1 Specific Aims |

Monte Carlo (MC) simulation is a valuable tool for radiation therapy. Due to its well-known accuracy and flexibility to address a variety of problem geometries and physical processes, MC simulation plays an irreplaceable role in solving a wide spectrum of problems. Particularly for particle beam radiation therapy (PBRT), its remarkable value has been well recognized. Examples include, but not limited to, accurately calculating dose distributions that are highly sensitive to treatment geometry and anatomy, reducing range uncertainty, developing novel treatment verification techniques, capturing radiobiological effects from the microscopic level, and designing treatment facility. Hence, researchers are eager to have a fast, robust, and easy-to-use MC system for PBRT research.

Yet, two practical challenges impede the wide applications of MC. (1) Low efficiency. Being a statistic method, MC inherently requires simulations of a large number of particles to achieve a high precision, leading to, e.g., tens of hour computation time for a typical PBRT treatment plan dose calculation, on a conventional CPU platform. Most researchers do not have the facility to run large-scale MC simulations for PBRT problems. (2) Required expertise. Before running a simulation, a lengthy and cumbersome process is needed to set up problem geometry and physical processes, typically through cumbersome scripting languages. Fine-tuning of parameters for optimal code utilizations also require deep understandings of MC simulations and particle transport physics. These facts restrict the applications of MC to only a small group of researchers in PBRT, while many other researchers, especially radiobiologists and clinical physicists, do not have the expertise to use those novel MC simulation tools. As of today, the conflicts between the great desire of using MC and the difficulties of using it have impeded research and clinical activities in PBRT to a significant extent.

As part of the planning process for National Particle Therapy Research Center (NPTRC), we propose in this pilot project a highly accurate, efficient, yet user-friendly and centralized MC simulation system using novel graphics-processing unit (GPU) and cloud-computing technologies to promote the wide applications of MC to facilitate PBRT research. Different from conventional MC packages that run on the user’s end, our system remotely resides in a cloud inside NPTRC and provides MC simulation services to PBRT researchers though standard web browsers on desktop PCs, laptops, or even mobile devices. A user will be able to define the problem through friendly interactive tools with direct visualization capabilities. The backend GPUs will execute the simulations with an extremely high efficiency. While our long-term goal is to deliver a novel MC simulation platform to facilitate the establishments of NPTRC and its future research activities, as well as to service the entire PBRT community, the goal of this pilot project is to initially develop and validate a prototype system focusing on MC-based particle beam dose calculations to demonstrate its feasibility and impacts. The deliverability of this project has been clearly demonstrated by the mature technologies nowadays and our extensive preliminary studies. The strong team assembled for this project, particularly the integration of Dr. Parodi for particle physics modeling, also ensures our success. Our goal will be accomplished by pursuing the following **specific aims** (SAs):

1. System developments (Yr 1, Mon 1 - Yr2, Mon 3)
2. Develop web-interface and system-management components for user interface, simulation configuration, and server management.
3. Develop a comprehensive physics database for particle transport necessary to PBRT.
4. Develop core MC code including the general simulation structure, particle transport simulation modules and interaction sampling modules.
5. System validations (Yr 2, Mon 4-12)
6. Comprehensively validate the computational accuracy of our system by testing the results against theoretical predictions, results from golden standard MC packages, and experimental measurements.
7. Test computational efficiency
8. Perform end-to-end functionality test in a representative case of carbon therapy dose calculations and demonstrate its usability to non-MC experts.

This pilot project fits perfectly into the overall plan for the proposed NPTRC. (1) Being an integral component of NPTRC, it will play a critical role at the planning stage by offering virtual and realistic simulations of different clinical, physical, and technical scenarios. In the long run, the novel MC system will greatly expand the capacity of research activities at NPTRC and will hence significantly contribute to the establishments NPTRC’s leading role in PBRT research. (2) The high computation power, cloud nature, and user-friendly interface of our MC system make it possible to support a wide group of researchers inside/outside NPTRC to pursue novel PBRT research. Continuous developments of our system will add much more features to address needs in different research areas, such as microdosimetry for radiobiology modeling. Offering MC services is aligned with the NPTRC’s mission of providing resources to investigate important problems in PBRT.

|  |
| --- |
| P2.2 Research Strategy |

|  |
| --- |
| P2.2.1 Background and Significance |

P2.2.1.1 The importance of Monte Carlo methods for Particle Beam Radiation Therapy.

Monte Carlo (MC) simulation is a valuable tool for particle beam radiation therapy (PBRT). Due to flexibility to handle complex simulation geometry, physical interactions, and explicit production and transport of secondary radiations, MC simulations allow for highly accurate calculations as well as dedicated investigations for many problems in PBRT that are not accessible with conventional computation tools. Here we only list a few important examples. (1) Because of very conformal dose depositions achieved by advanced delivery techniques (e.g. beam scanning) and planning techniques (e.g. intensity or radiation quality modulations), MC methods are needed to capture dose distributions that are highly sensitive to treatment geometry and patient anatomy [[1-4](#_ENREF_1)]. (2) Because of the sharp distal dose fall-off, ranges of particle beams should be predicted as accurate as possible in the treatment planning and delivery process to fully utilize the potential advantage of particle therapy. MC methods allow to reduce range uncertainty margins and could thus add significant value to treatment precision [[5](#_ENREF_5)]. (3) MC calculations are at the heart of treatment verification techniques relying on the detection of secondary radiation [[6](#_ENREF_6), [7](#_ENREF_7)]. (4) Dramatically different from conventional photon therapy, particle therapy has higher relative biological effectiveness [[8](#_ENREF_8)] due to distinct physical interaction processes, which also varies along the beam path. Hardly do any methods other than MC capture this from a microscopic level. (5) With the flexibility to handle simulation geometry and scoring quantities, MC is also a valuable tool for many other studies pertaining to particle radiotherapy, such as design of treatment facilities and detectors.

P2.2.1.2 Conflicts between the desire for Monte Carlo methods and the difficulties of using it.

Over the years, the wide spectrum of research interests has generated a great desire of using MC for various purposes. Researchers are eager to have a fast, robust, and easy-to-use MC system. Nonetheless, this desire is contradicted by the difficulties encountered when using current MC tools. First, current MC methods are slow. Being a statistic method, MC inherently requires a huge number of particles to yield a satisfactory level of precision, causing a long computation time on a conventional CPU platform. For instance, it takes about 60 hours to compute the dose distribution for a typical proton plan on one CPU[[9](#_ENREF_9)]. While computer clusters have been utilized to accelerate calculations [[9-11](#_ENREF_9)], they are only available at national labs and major institutions but not to normal clinical users due to the high costs of facility deployment and maintenance. This prolonged computational time has limited uses of MC primarily in research projects, and only to those non-time-sensitive clinical applications, such as offline treatment quality assurance. Second, current MC methods require a high level of user expertise. General-purposed MC packages, such as Geant4[[12](#_ENREF_12), [13](#_ENREF_13)], FLUKA[[14](#_ENREF_14)], and MCNPX[[15](#_ENREF_15)], have been employed in particle therapy research. Before launching a simulation, there is a lengthy and tedious process to set up the problem physics and geometry, typically through scripting languages. This cumbersome process also lacks of effective visualization tools, making it error prone and difficult for non-MC experts. In addition, years of efforts are needed to master these packages in order to finely tune parameters to achieve an optimal utilization. These facts strongly restrict the applications of MC only to those experts.

P2.2.1.3 Our project overcomes challenges for the wide applications of Monte Carlo methods.

To resolve this significant conflict between the strong desire for MC and the difficulties of using it, we propose in this pilot project a centralized MC simulation system using novel graphics-processing unit (GPU) and cloud computing technologies. Different from conventional MC packages that run on the user’s end, our system remotely resides in a cloud at National Particle Therapy Research Center (NPTRC) planned in this proposal, and provides MC simulation services to users nationwide though standard web browsers on desktop PCs, laptops, or even mobile devices. A user will be able to easily define a problem via friendly interactive tools with direct visualization capabilities. The backend GPUs will execute the simulation with an extremely high efficiency. It is our ultimate objective to achieve this highly accurate, efficient, yet user-friendly MC system for the wide adoption of MC methods in a broad scope of PBRT problems, and therefore benefit the entire PBRT community. This pilot project will initiate efforts towards this long-term goal by developing and validating a prototype system for dose calculation purpose to demonstrate its feasibility and impacts.

P2.2.1.4 The novel MC system will support developments of NPTRC and its research projects

The GPU-cloud based MC system will also be an integral component for the planning of the NPTRC, as well as its future research activities. In fact, the novel MC system will play a critical role for the planning of NPTRC. The virtual simulations of different clinical, physical, and technical scenarios will offer realistic insights to many important problems, such as facility design. At the end of this pilot project, the MC system will also become an important tool for final-stage developments, e.g. commissioning of the proton beam at Texas Center for Advanced Radiation Therapy (TCART). The experience and technical achievements during that stage will be extremely valuable for the beamline developments at NPTRC. Down the road, the NPTRC is expected to be a national center pioneering in basic, translational, and clinical researches in PBRT regime. Example projects include, but not limited to (1) treatment plan optimization and adaptive PBRT, (2) treatment monitoring, verification, and quality assurance, and (3) particle beam radiobiology. Our MC system, which will be much improved at that stage compared to in this pilot project, will greatly expand the width (investigate a wide range of problems) and depth (answer questions from basic science level) of research activities at NPTRC and will therefore significantly contribute to the establishments NPTRC’s leading role.

P2.2.1.5 The developed MC system will become a valuable national PBRT research resource

As a national research center, NPTRC is also committed to provide resources to investigate important problems in PBRT. The high computation power and its cloud nature of our MC system make it possible to support numerous novel projects in PBRT across the nation. Its user-friendly interface maximally reduces difficulties of using MC simulations, allowing non-MC experts to benefit from this novel method. Under the umbrella of NPTRC, our system will offer MC services to the entire PBRT community. In a long run, this will become an integral platform for the entire PBRT field, on which researchers pursue solutions to their important problems using novel MC methods. Being a pilot project in this PBRT center-planning grant, we will conduct initial developments of our novel MC system. Yet, the end of this project is only a starting point of a new era. Our system will continuously grow with improved technical capacity to support research needs of the entire PBRT community.

|  |
| --- |
| P2.2.2 Rationale |

In addition to its significant impacts to the PBRT community, the rationale of this pilot project also lies in technical maturity of GPU and cloud-computing technologies, and our extensive research experience.

P2.2.2.1 GPU and cloud computing have laid a solid foundation to solve these problems

Recently, the rapid advances of GPU and cloud-computing technologies have shed lights to this situation. (1) Conventionally for graphics processing, a GPU card has thousands of processing threads integrated in a single chip offer a processing power comparable to a CPU cluster. Moreover, because of the driving force from computer game market, the cost of GPU-computing is low, such that with thousands of dollars, one can obtain a processing power similar to a modest CPU cluster. Nowadays, GPU has become an emerging source offering potent and affordable powers to tackle computationally demanding problems in medical physics [[16](#_ENREF_16), [17](#_ENREF_17)]. (2) Meanwhile, with novel computer virtualization framework and on-demand computational resources, cloud computing [[18-20](#_ENREF_18)] grants users with access to remotely-located computing powers conveniently through web interfaces on personal computers. Modern web technologies, e.g. HTML5[[21](#_ENREF_21)] and webGL[[22](#_ENREF_22)], are also mature enough to permit manipulating geometry and physics setups in MC simulations through web interface in a flexible fashion. The seamless combination of these two cutting-edge technologies leads to a novel solution to the aforementioned conflicts.

P2.2.2.2 We have accumulated extensive research experience.

The research team assembled for the proposed pilot project ensures its success. The project lead Dr. Jiang and co-investigator Dr. Jia from University of Texas Southwestern (UTSW) have pioneered in developing GPU applications in radiation therapy[[16](#_ENREF_16)], including MC simulations for particle transports[[23-30](#_ENREF_23)]. Dr. Parodi from Ludwig-Maximilians University (LMU) in Germany is a world-leading researcher in PBRT [[6](#_ENREF_6), [7](#_ENREF_7), [9](#_ENREF_9), [31-36](#_ENREF_31)], who will develop physics models for the current project. Drs. Timmerman and Story established researchers in radiotherapy and radiobiology, respectively. Their inputs will be valuable to shape the developed MC system suitable for problems of different types in PBRT. In the rest of this section, we will briefly describe some preliminary studies pertaining to the proposed pilot project.

|  |
| --- |
| Figure 1 a)~c) Dose distribution of a typical patient case. d) and e) Comparison of dose profiles from Geant4 and gPMC along the dash lines. |

**GPU**-based MC simulations Over the years, a set of GPU-based MC simulation packages have been developed by Drs. Jiang and Jia for various purposes, such as photon transport for x-ray imaging problems[[23](#_ENREF_23), [24](#_ENREF_24)], photon/electron transports for conventional photon radiotherapy[[25-28](#_ENREF_25)], and proton transport for proton therapy[[29](#_ENREF_29), [30](#_ENREF_30)]. Acceleration factors of a few hundred times have been achieved compared to CPU-based computations, and it takes only 36 sec to 40 sec to perform dose calculations for a clinical photon plan on a single NVIDIA GTX 580 card with ~1% relative uncertainty.

|  |
| --- |
| Figure 2. (a)-(d) TARGET interface for user log in, data upload, dose calculations, and result display. (3) Hardware architecture of TARGET system. |

Particular of relevance to the proposed pilot project is the package gPMC developed for proton beam dose calculation [[29](#_ENREF_29), [30](#_ENREF_30)]. In this package, we have carefully designed particle transport process to include necessary physics models for sufficiency accuracy, while avoiding overcomplicated processes to ensure efficiency. gPMC is implemented on GPU under NVIDIA CUDA platform [[37](#_ENREF_37)]. A large number of GPU threads are launched simultaneously to gain high efficiency. Novel techniques have been invented to tune the code suitable for GPU parallelization: separating transports of different particle types to avoid GPU thread divergence problem, and a multi-dose-counter method to mitigate a memory conflict issue.

The performance of gPMC has been evaluated extensively by comparing simulation results with those generated by TOPAS/Geant4, a golden-standard MC package for proton dose calculations[[12](#_ENREF_12), [13](#_ENREF_13), [38](#_ENREF_38)]. In all cases tested, the dose calculation results from gPMC and TOPAS/Geant4 are in a good agreement, such that ~97% of the voxels in the region above 10% of the maximum dose pass the gamma-index [[39](#_ENREF_39)] test with a 2mm/2% criterion. As an example, dose distribution for a typical patient case is shown in Figure 1. In terms of efficiency, it is found that the simulations of 107 source protons can be completed in 6~24 seconds on an NVIDIA GTX 580 card, yielding ~1% relative uncertainty, depending on the source energy and phantom materials. Compared to the 10~70 CPU hours for the TOPAS/Geant4 simulations, it is apparently that gPMC achieved an extremely high efficiency.

|  |
| --- |
| Figure 3. Example of FLUKA-based MC re-calculation (top) and original treatment plan (bottom) of RBE-weighted dose distribution for scanned carbon ion beam delivery at HIT. |

**Cloud**-based remote MC simulations We have also accumulated extensive experience in developing cloud-based applications to offer GPU-based MC simulations to remote users. One example is our TARGET (Treatment Assurance and Reporting using GPU Enhanced Tools) system. Currently being used clinically at UT Southwestern, TARGET is a web-based GPU MC tool for photon radiation therapy quality assurance (QA) purpose. Through a friendly web interface, users can upload treatment plan data computed by a commercial treatment planning system. The GPU-based MC dose engine is launched on a remote GPU cloud to verify the planned dose calculation accuracy. QA results are displayed to user and can be downloaded. Screen captures of TARGET interface are in Figure 2 (a)-(d).

Figure 2 (e) depicts hardware infrastructure to support TARGET. A web server serves as the middleman between remote users and the bulk of our system. It receives commands and data from user and sends them to a control server that further distributes requests to computation and data resources. A number of GPU workstations are dedicated to computations, which are equipped with multiple state-of-art NVIDIA GPUs and adequate system RAM. The data nodes with large capacity computer hard drives provide storage and fast access to massive data. InfiniBand connections link nodes of different kinds for high data throughput, low latency, and scalability. This mature and reliable hardware structure will also be the foundation of the proposed MC system in this pilot project.

Monte Carlo modeling of heavy ion beams Dr. Parodi’s group has also accumulated extensive experience in developing physics models for MC simulations in PBRT. Dr. Parodi was and still is strongly linked to the first clinical European heavy ion therapy facility, the Heidelberg Ion Therapy Center (HIT). Here, she has customized and used Monte Carlo methods to support the start-up and commissioning of the facility, as well as to generate the basic data of the treatment planning system which are still in clinical use [[33-35](#_ENREF_33)]. Moreover, she and her team developed a dedicated framework for automated calculation of dose and positron emitter distributions for treatment QA (Figure 3) [[7](#_ENREF_7), [36](#_ENREF_36)], which is still being used and extended in the framework of ongoing research projects under her supervision.

In the framework of the French Hadrontherapy network, Dr. Parodi’s group has been deeply involved in the improvement of hadronic physics models of Geant4 for the description of secondary radiation of relevance to in-vivo range verification techniques [[40](#_ENREF_40)]. In the framework of a European effort, they also actively participated to investigations aiming to identify reliable experimental data [[31](#_ENREF_31), [32](#_ENREF_32)] and to benchmark the performances of MC codes in terms of several physical quantities of relevance for correct computations in heavy ion beam therapy[[6](#_ENREF_6)].

**P2.2.2.3** Summary

The goal of the proposed pilot project is threefold: 1) solving the conflicts between the needs of MC in PBRT and its difficulty of use, 2) facilitating establishments of NPTRC and its continuous researches, 3) initiating continuous supports of a novel MC platform for the entire PBRT community. The mature technologies and our extensive experiences will be a solid foundation for the proposed project.

|  |
| --- |
| P2.2.3 Approach |

**P2.2.3.1 System developments (Yr 1, Mon 1 - Yr2, Mon 3)**

A complete MC system proposed in this project consists of hardware and software components. Due to the limited resources and time, we will use the hardware platform previously built for the TARGET project (Figure 2(e)) here and will only focus on the software developments. The resulting prototype system will demonstrate the feasibility of our system. The hardware will be expanded in scale depending on the user needs in future.

(a) Interface and management (UTSW)

|  |
| --- |
| Figure 4. Illustration of particle interaction network. Dash line circles a sub-network in a particular simulation. |

User interface and management A web browser will serve as the entry point to the system. The user interface will support user registration and management, data transfer, parameter setup, result visualization on virtually all modern browser clients on personal computers or touchpads. Based on the TARGET project, we plan to use the Django web framework [[41](#_ENREF_41)] to manage users, raw data, as well as issue job submissions. Visualization related tasks will be accomplished using HTML5 [[21](#_ENREF_21)] and WebGL [[22](#_ENREF_22)] technologies to support 3D rendering and blending of problem geometry and results. Data transfer will be through secure file transfer protocol.

MC simulation configuration To maximally reduce the burden of setting up a MC simulation job, we will employ a web-based graphics user interface (GUI) approach rather than conventional text/script based approach used in common MC packages [[14](#_ENREF_14), [15](#_ENREF_15), [42-44](#_ENREF_42)]. This is particularly important for non-MC expert. The user will navigate through the following steps to configure a job. (1) Problem geometry and material. Under the support of HTML5[[21](#_ENREF_21)] and WebGL[[22](#_ENREF_22)] technology, a user can define problem geometry and material by simply drawing them in a web browser and visualize it via interactive 3D rendering. Alternatively, the user can upload voxelized data, e.g. a CT image, to define the geometry and materials. (2) Particle source definition. A user defines particle source properties by inputting their positional, directional, and energy distributions using web-based equation editors. (3) Scoring quantity. A user defines quantities to be recorded during the simulation (dose, energy, fluence etc) by selecting from a drop-down list, as well as set associate parameters (region of interest etc.). (4) Physical process to be simulated (optional). A user explores the physics network (see Sec P2.2.3.1(b)**)** in an interactive way to select relevant interactions. Default interactions will be automatically set, if the user skips this step. (5) MC simulation parameters (optional). A user sets up relevant parameters, e.g. cutoff energies, by inputting them in the web. Default parameters will be used, if not specified.

Server and computation management After setting a simulation job, the user will submit it to our system. The backend GPU will execute the simulations. To manage a cluster of GPU servers and improve computation efficiency, parallel processing at the GPU level will be automatically implemented. MC simulations attain inherent parallelizability at the GPU level. We will simply divide number of source particles and parcel them to different GPUs. The inter-GPU parallelization will be implemented in a MapReduce-like model [[45](#_ENREF_45), [46](#_ENREF_46)]. A server status monitoring system and job scheduling system will be taken from the TARGET system and will be used in our system with necessary modifications.

(b) Physics database (LMU)

Extensive physics data are the ground for MC simulations. Since the reported literature on the application of MC methods to heavy ion beam therapy is still scarce, it is essential to rely on the expertise of LMU group who gathered several years of experience in the customization, benchmark and usage of general purpose MC packages such as FLUKA and Geant4.

The handling of ion transport simulation will be similar to what is in the gPMC package for proton transport. However, the main challenge here resides in the large amount of nuclear interactions and, more importantly, the possibility of projectile fragmentation for heavy ion beams. Given the fact that a particle GPU MC system will not be fully implementing the modeling of nuclear reactions, an extensive database that would provide meaningful assumptions and parameterizations will be necessary. To this aim, we will first generate a particle production network that includes all possible generations of different particles and fragments pertaining to ion therapy. As illustrated in Figure 4, nodes are particles and the links indicate the production from one particle to another under a certain physical process. On top of this network, two types of data will be included in our database. (1) Particle transport. Associated to each node (particle), relevant parameters such as cross section and stopping power, will be included that govern the transport properties. Parameterizations of total and partial nuclear reaction cross-sections that yield good agreement with measurements [[47](#_ENREF_47), [48](#_ENREF_48)] will be used. This will enable a fast analytical decision of the occurrence of a nuclear reaction and its final state products during MC simulations. (2) Interaction. Associate to each link (particle production), differential cross sections will be produced in this project using existing and largely benchmarked CPU MC tools. The database will provide angular, energy and Z spectra of secondary particle/fragment induced by nuclear reactions of specific projectile-target configurations.

|  |
| --- |
| Figure 5. The proposed universal code structure to handle the simulations. |

(c) Monte Carlo simulation core code (UTSW)

A versatile code structure There are two challenges to design our MC simulation code structure. (1) Efficiency. Due to GPU’s single instruction and multiple data scheme[[37](#_ENREF_37)], high efficiency is only achieved when GPU threads process the same instruction concurrently, e.g. all transporting particles of the same type. (2) Versatility. Users may simulate quite different physical processes depending on their specific applications. It is challenging to develop a universal yet robust code structure to handle all requests.

To overcome these challenges, we propose a simulation structure shown in Figure 5. Specifically, Step 2 determines all possible particle types that will be involved based on the user inputs at the configuration stage. The dashed line in Figure 4 illustrates a subgroup of particles involved in a particular job. Based on this, particle stacks will be allocated, one for each of the particle types. These stacks will hold any secondary particles generated during simulations. An active particle space (APS) will also be allocated, which holds the particles being simultaneously transported by the GPU. Step 3 populates the APS with particles either generated according to the source model (Step 3.3) or loaded from one of the particle stacks (Step 3.4). Once the APS is filled, depending on the particle type, a corresponding GPU module will be launched to transport them (Step 3.5). This iteration terminates when all stacks are empty and all source particles are simulated. The advantages of this scheme are twofold. (1) It reduces thread divergence by always simulating particles of the same type. Previous studies demonstrated the effectiveness of this approach [[49](#_ENREF_49), [50](#_ENREF_50)]. (2) This universal workflow can handle arbitrary physics process a user would like to simulate.

Particle transport simulations The particle transport modules will be written in OpenCL due to its portability across different GPU cards[[51-53](#_ENREF_51)]. (1) For charge-neutral particles, we will employ an analogous simulation scheme, with Woodcock transport method[[54](#_ENREF_54)] to eliminate cumbersome voxel boundary crossing and reduce memory access frequency, which is important for efficiency considerations. (2) For charged particles, Class II condensed history scheme [[55](#_ENREF_55)] with continuously slowing down approximation will be utilized with energy straggling and angular deflection[[56](#_ENREF_56)]. Catastrophic collision events are simulated explicitly.

Interaction sampling Sampling event consequences, usually productions of secondary particles and changes status of the current particle is specific to the event physics. (1) For events with known differential cross section, we will sample directly using the differential cross section as a probability function. An appropriate sampling method, e.g. rejection sampling will be selected for efficiency considerations [[42-44](#_ENREF_42)]. (2) When the event process is described by a physical model, e.g. intra-nuclear cascade model [[56](#_ENREF_56), [57](#_ENREF_57)], we will write GPU functions to simulate the consequences using the model.

**P2.2.3.2**  **System validations (Yr 2, Mon 4-12)**

After system developments, we will conduct a comprehensive set of validation studies to test computational accuracy, efficiency, and functionality of our system.

(a) Accuracy (UTSW+LMU)

Accuracy is the first priority issue, which will be evaluated through three stages:

(1) We plan to first validate our dose calculation accuracy by examining physical quantities predicated by theory, such as particle range in homogeneous phantom. Tests against established theories, e.g. Fano theory [[58](#_ENREF_58), [59](#_ENREF_59)], will also be conducted. Statistical t-tests will be used to compare the computed results. It is expected that the code will pass all tests with a significance level 0.05.

(2) We will also perform dose calculations in simulation studies and compare results with golden standard MC packages TOPAS/Geant4 [[9](#_ENREF_9), [13](#_ENREF_13)] and FLUKA[[14](#_ENREF_14)]. Phantoms composed of different materials and geometries, as well as those generated based on patient CT images will be used. We hypothesize that the dose distribution obtained from golden standard codes and obtained by our code will agree within 2%/2mm gamma index evaluation for all cases considered. We will also compare dose values using statistical two tail t-tests at each voxel to justify whether dose differences between two codes are statistically significant. Based on our preliminary studies, we expect that over 95% of the high dose region with dose > 20% of the maximum dose, pass the with a statistical significance of 0.05.

(3) We will further test against experiment measurements. Dr. Parodi is closely linked with HIT, which allows accessing experimental facilities and data to validate our calculations. The calculated quantities, such as depth dose profiles and lateral distributions, will be compared with measurements. Additional testing on the assumptions and modeling of the secondary particle creation due to nuclear reactions of heavy will be needed, in order to disentangle different physical phenomena that contribute to the overall macroscopic results. Validation will be done by benchmarking against high quality experimental data of nuclear fragmentation, such as in [[60-63](#_ENREF_60)]. Additional important experimental data (double differential cross sections of ion fragmentation using therapeutic beams on thin targets) will be soon available by international collaboration experiments such as the FIRST experiment in GSI and measurements at GANIL, and the European network ENLIGHT.

(b) Computation efficiency (UTSW)

Efficiency is a main factor to impact on the code practicality. We will evaluate the efficiency by timing all the aforementioned computations and characterize it by , where is computation time and is the relative uncertainty level for dose calculations. Based on our preliminary studies in gPMC, we expect the absolute efficiency will be ~103/sec (it takes ~10 sec to achieve 1% uncertainty) on a single GPU. We will also quantify the acceleration ratio relative to Geant4 and FLUKA.

(c) End-to-end functionality test in a clinical case (UTSW+LMU)

Finally, to demonstrate the usability, we will apply our system in a specific clinical problem of carbon ion therapy dose calculation. As such, we will select cancer patient cases treated at UTSW and design scanning beam carbon therapy plans on them. When doing dose calculations using our system, patient CT and source model will be defined at the code configuration stage. Calculation will be conducted and the results will be compared with the planned dose distribution. To demonstrate the convenience of using our system, non-MC experts in our group will be invited to conduct the tests as well. A successful end-to-end performance test will serve as the final validation for our system.

**P2.2.3.3 Outlook**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Task | Year 1 | | | | Year 2 | | | |
| SA1 | Develop web interface |  |  |  |  |  |  |  |  |
| Develop system management module. |  |  |  |  |  |  |  |  |
| Develop physics database. |  |  |  |  |  |  |  |  |
| Develop core MC codes |  |  |  |  |  |  |  |  |
| SA2 | Validate result accuracy |  |  |  |  |  |  |  |  |
| Test computational efficiency |  |  |  |  |  |  |  |  |
| End-to-end functionality test |  |  |  |  |  |  |  |  |

The timeline of the proposal is shown here. Upon completion, a prototype GPU-cloud-based MC system will be developed and tested. It will demonstrate the feasibility of offering accurate, efficient, and convenient MC simulations services to the entire PBRT community through cutting-edge GPU and cloud-computing technologies. Yet, this is only a starting point. The developed system will be a valuable component of NPTRC, providing supports internally to NPTRC, as well as externally to the entire PBRT community. Future expansion of this system will further broaden the scope of MC and enable its applications to a wide range of PBRT problems, e.g. micro MC for micro dosimetry, resulting in measurable and long-lasting impacts to PBRT research and clinic.

REFERENCE

1. Unkelbach, J., et al., *Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning.* Medical Physics, 2009. **36**(1): p. 149-163.

2. Lomax, A.J., *Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions.* Physics in Medicine and Biology, 2008. **53**(4): p. 1043.

3. Lomax, A.J., *Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties.* Physics in Medicine and Biology, 2008. **53**(4): p. 1027.

4. Soukup, M., et al., *Study of Robustness of IMPT and IMRT for Prostate Cancer Against Organ Movement.* International Journal of Radiation Oncology\*Biology\*Physics, 2009. **75**(3): p. 941-949.

5. Paganetti, H., *Range uncertainties in proton therapy and the role of Monte Carlo simulations.* Physics in Medicine and Biology, 2012. **57**(11): p. R99-R117.

6. Robert, C., et al., *Distributions of secondary particles in proton and carbon-ion therapy: a comparison between GATE/Geant4 and FLUKA Monte Carlo codes.* Physics in Medicine and Biology, 2013. **58**(9): p. 2879-2899.

7. Bauer, J., et al., *Implementation and initial clinical experience of offline PET/CT-based verification of scanned carbon ion treatment.* Radiotherapy and Oncology, 2013. **107**(2): p. 218-226.

8. Hall, E.J. and A.J. Giaccia, *Radiobiology for the Radiologist*. 2012: Wolters Kluwer Health.

9. Paganetti, H., et al., *Clinical implementation of full Monte Carlo dose calculation in proton beam therapy.* Physics in Medicine and Biology, 2008. **53**(17): p. 4825-4853.

10. Vadapalli, R., et al., *Grid-enabled treatment planning for proton therapy using monte carlo simulations.* Nuclear Technology, 2011. **175**(1): p. 16-21.

11. Jiang, H. and H. Paganetti, *Adaptation of GEANT4 to Monte Carlo dose calculations based on CT data.* Medical Physics, 2004. **31**(10): p. 2811-2818.

12. Geant4 Collaboration. *Geant4 Physics Referene Manual*. 2011; Available from: <http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/PhysicsReferenceManual/fo/PhysicsReferenceManual.pdf>.

13. Agostinelli, S., et al., *GEANT4-a simulation toolkit.* Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment, 2003. **506**(3): p. 250-303.

14. Ferrari, A., et al., *Fluka: a multi-particle transport code*. 2011, Geneva: CERN.

15. Briesmeister, J.F., *MCNP-a general Monte Carlo N-particle transport code.* Los Alamos National Laboratory Report 1993: p. LA-12625-M.

16. Jia, X., P. Ziegenhein, and S.B. Jiang, *GPU-based high-performance computing for radiation therapy.* Physics in Medicine and Biology, 2014. **59**(4): p. R151.

17. Pratx, G. and L. Xing, *GPU computing in medical physics: A review.* Medical Physics, 2011. **38**(5): p. 2685-2697.

18. Rangan, K., *The Cloud Wars: $100+ Billion at Stake. Tech. Rep.*, in *Merrill Lynch* 2008.

19. Armbrust, M., et al., *Above the clouds: A berkeley view of cloud computing.* Technical Report, EECS Department, University of California, Berkeley 2009. **UCB/EECS-2009**: p. 28.

20. Armbrust, M., et al., *A View of Cloud Computing.* Communications of the Acm, 2010. **53**(4): p. 50-58.

21. Hunt, L. *HTML5 Reference, The Syntax, Vocabulary and APIs of HTML5*. 2010; Available from: <http://dev.w3.org/html5/html-author/>.

22. WebGL group. *WebGL, OpenGL ES 2.0 for the Web*. 2013; Available from: <http://www.khronos.org/webgl/>.

23. Jia, X., et al., *Fast Monte Carlo simulation for patient-specific CT/CBCT imaging dose calculation.* Physics in Medicine and Biology, 2012. **57**(3): p. 577-590.

24. Jia, X., et al., *A GPU Tool for Efficient, Accurate, and Realistic Simulation of Cone Beam CT Projections.* Med. Phys., 2012. **39**(12): p. 7368-7378.

25. Jia, X., et al., *Development of a GPU-based Monte Carlo dose calculation code for coupled electron-photon transport.* Phys. Med. Biol., 2010. **55**: p. 3077.

26. Jia, X., et al., *GPU-based fast Monte Carlo simulation for radiotherapy dose calculation.* Phys. Med. Biol., 2011. **56**(22): p. 7017-31.

27. Townson, R.W., et al., *GPU-based Monte Carlo radiotherapy dose calculation using phase-space sources.* Physics in Medicine and Biology, 2013. **58**(12): p. 4341-56.

28. Tian, Z., et al., *Commissioning of a finite-size pencile beam dose calculation model via optimization*, 2014.

29. Jia, X., et al., *Development of gPMC v2.0, a GPU-based Monte Carlo dose calculation package for proton radiotherapy.* Medical Physics, 2013. **40**(6): p. 498.

30. Jia, X., et al., *GPU-based fast Monte Carlo dose calculation for proton therapy.* Physics in Medicine and Biology, 2012. **57**(23): p. 7783-7797.

31. Schwaab, J., et al., *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 and Biology, 2011. **56**(24): p. 7813-7827.

32. Sanchez-Parcerisa, D., et al., *Experimental study of the water-to-air stopping power ratio of monoenergetic carbon ion beams for particle therapy.* Physics in Medicine and Biology, 2012. **57**(11): p. 3629-3641.

33. Parodi, K., A. Mairani, and F. Sommerer, *Monte Carlo-based parametrization of the lateral dose spread for clinical treatment planning of scanned proton and carbon ion beams.* Journal of Radiation Research, 2013. **54**: p. 91-96.

34. Parodi, K., et al., *The influence of lateral beam profile modifications in scanned proton and carbon ion therapy: a Monte Carlo study.* Physics in Medicine and Biology, 2010. **55**(17): p. 5169-5187.

35. Parodi, K., et al., *Monte Carlo simulations to support start-up and treatment planning of scanned proton and carbon ion therapy at a synchrotron-based facility.* Physics in Medicine and Biology, 2012. **57**(12): p. 3759-3784.

36. Parodi, K., et al., *Monte Carlo Modeling and In-Vivo Imaging at the Heidelberg Ion Beam Therapy Center.* 2011 Ieee Nuclear Science Symposium and Medical Imaging Conference (Nss/Mic), 2011: p. 2795-2799.

37. NVIDIA, *NVIDIA CUDA Compute Unified Device Architecture, Programming Guide, 4.0*. 2011.

38. Perl, J., et al., *TOPAS: An innovative proton Monte Carlo platform for research and clinical applications.* Med. Phys., 2012. **39**(11): p. 6818-37.

39. Low, D.A., et al., *A technique for the quantitative evaluation of dose distributions.* Medical Physics, 1998. **25**(5): p. 656-661.

40. Dedes, G., et al., *Assessment and improvements of Geant4 hadronic2 models in the context of prompt-gamma3 hadrontherapy monitoring.* Phys. Med. Biol., 2014(accepted).

41. Django. *The Django framework*. 2014; Available from: https://[www.djangoproject.com/](http://www.djangoproject.com/).

42. Kawrakow, I., *Accurate condensed history Monte Carlo simulation of electron transport. I. EGSnrc, the new EGS4 version.* Medical Physics, 2000. **27**(3): p. 485-498.

43. Hirayama, H., et al., *The EGS5 code system.* Stanford Linear Accelerator Center Report, 2010: p. SLAC-R-730.

44. Salvat, F., J.M. Fernández-Varea, and J. Sempau, *PENELOPE-2008: A Code System for Monte Carlo Simulation of Electron and Photon Transport*. 2009, Issy-les-Moulineaux, France: OECD-NEA.

45. Farivar, R., et al., *MITHRA: Multiple data Independent Tasks on a Heterogeneous Resource Architecture*. 2009 Ieee International Conference on Cluster Computing and Workshops. 2009. 374-383.

46. He, B., et al., *Mars: A MapReduce Framework on Graphics Processors*. Pact'08: Proceedings of the Seventeenth International Conference on Parallel Architectures and Compilation Techniques. 2008. 260-269.

47. Silberbe.R and C.H. Tsao, *PARTIAL CROSS-SECTIONS IN HIGH-ENERGY NUCLEAR-REACTIONS, AND ASTROPHYSICAL APPLICATIONS .1. TARGETS WITH Z =28.* Astrophysical Journal Supplement Series, 1973. **25**(220): p. 315-333.

48. Sihver, L., et al., *TOTAL REACTION AND PARTIAL CROSS-SECTION CALCULATIONS IN PROTON-NUCLEUS (ZT-LESS-THAN-OR-EQUAL-TO-26) AND NUCLEUS-NUCLEUS REACTIONS (ZP AND ZT-LESS-THAN-OR-EQUAL-TO-26).* Physical Review C, 1993. **47**(3): p. 1225-1236.

49. Hissoiny, S., et al., *GPUMCD: A new GPU-oriented Monte Carlo dose calculation platform.* Medical Physics, 2011. **38**(2): p. 754-764.

50. Jia, X., P. Ziegenhein, and S.B. Jiang, *GPU-based High-Performance Computing for Radiation Therapy.* Accepted by Phys Med Biol, 2014.

51. Su, C.-L., et al., *Overview and comparison of OpenCL and CUDA technology for GPGPU.* 2012 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS 2012), 2012: p. 448-451.

52. Kakimoto, T., et al., *Performance comparison of GPU programming frameworks with the striped Smith-Waterman algorithm.* Computer Architecture News, 2012. **40**(5): p. 70-75.

53. Khronos OpenCL Working Group. *The open standard for parallel programming of heterogeneous systems*. 2013; Available from: https://[www.khronos.org/opencl/](http://www.khronos.org/opencl/).

54. Woodcock, E., et al., *Techniques used in the GEM code for Monte Carlo neutronics calculations in reactors and other systems of complex geometry*, in *Applications of Computing Methods to Reactor Problems: Argonne National Laboratories Report*1965. p. ANL-7050.

55. Berger, M.J., ed. *Monte Carlo calculations of the penetration and diffusion of fast charged particles*. Methods in computational physics, ed. B. Alder, S. Fernbach, and M. Rotenberg. Vol. 1. 1963, Academic: New York. 135-215.

56. Hagiwara, K., et al., *Review of particle physics.* Physical Review D, 2002. **66**(1).

57. Wan, H., C. Tseung, and C. Beltran, *A graphics processor-based intranuclear cascade and evaporation simulation.* Arxiv.org, 2013: p. 1309.7963.

58. Elles, S., et al., *Geant4 and Fano cavity test: where are we?* Journal of Physics: Conference Series, 2008. **102**(1): p. 012009.

59. Sterpin, E., et al., *A Fano cavity test for Monte Carlo proton transport algorithms.* Medical Physics, 2014. **41**(1): p. -.

60. Ben-Josef, E., et al., *Phase II trial of high-dose conformal radiation therapy with concurrent hepatic artery floxuridine for unresectable intrahepatic malignancies.* Journal of Clinical Oncology, 2005. **23**(34): p. 8739-8747.

61. Keall, P.J., et al., *The management of respiratory motion in radiation oncology report of AAPM Task Group 76.* Medical Physics, 2006. **33**(10): p. 3874-3900.

62. Gunzert-Marx, K., et al., *Secondary beam fragments produced by 200 MeVu(-1) (12)C ions in water and their dose contributions in carbon ion radiotherapy.* New Journal of Physics, 2008. **10**.

63. Dudouet, J., et al., *Double-differential fragmentation cross-section measurements of 95 MeV/nucleon C-12 beams on thin targets for hadron therapy.* Physical Review C, 2013. **88**(2).