

A High-Granularity Proton CT Enhanced by Track Discrimination

Authors: Dr. Huangchao Shi, Miss Danqi Wang, Ms. Meiqi Liu, Yang, Yidong, Baohua Qi, Liu, Dr. Yong, Qin, Mr. Yanghui, Feng, Dr. Changqing (Nuclear electronics), Liu, Dr. Yi, Dr. Hongbo Zhu, Zhu, Hongbo

Date: 2025-04-30T00:00:00+00:00

Abstract

Proton Computed Tomography (pCT) provides a promising solution to enhance the accuracy of Relative Stopping Power (RSP) required for proton therapy planning. This research introduces a novel high-granularity pCT architecture that incorporates a silicon pixel tracking system and a calorimetric range telescope, which uniquely integrates range telescope functionality with track discrimination capabilities. The Bortfeld function fitting and Convolutional Neural Network (CNN) classifier algorithms are developed and applied for discrimination. In simulation studies, both approaches demonstrate the capability to reduce uncertainty in Water Equivalent Path Length (WEPL) determination for individual proton tracks to below 3 mm. The standard imaging protocol (3.2 mGy, 4×10^8 protons) achieves sub-millimeter spatial resolution (~ 0.5 mm) with sub-1% RSP accuracy and 1.1 mm spatial resolution in simulation. This low-dose performance significantly expands clinical applicability, particularly for pediatric imaging or frequent imaging scenarios. Furthermore, the target 10 MHz proton detection rate suggests potential for real-time image guidance during radiotherapy. By circumventing the need for ultra-precise energy measurements, this design minimizes hardware complexity and provides a scalable foundation for future pCT systems.

Full Text

Preamble

A High-Granularity Proton CT Enhanced by Track Discrimination*

Huang-Chao Shi,¹ Dan-Qi Wang,¹ Mei-Qi Liu,² Yi-Dong Yang,² Bao-Hua Qi,^{3,4}

Yong Liu,^{3,4} Yang-Hui Qin,^{5,6} Chang-Qing Feng,^{6,7} Yi Liu,⁸ and Hong-Bo Zhu^{1,†}

¹School of Physics, Zhejiang University, Hangzhou 310058, China

²School of Physical Sciences, University of Science and Technology of China, Hefei 230026, China

³Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

⁴University of Chinese Academy of Sciences, Beijing 100049, China

⁵School of Nuclear Science and Technology, University of Science and Technology of China, Hefei 230026, China

⁶State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei 230026, China

⁷Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China

⁸School of Physics, Zhengzhou University, Zhengzhou 450001, China

Proton Computed Tomography (pCT) provides a promising solution to enhance the accuracy of Relative Stopping Power (RSP) required for proton therapy planning. This research introduces a novel high-granularity pCT architecture that incorporates a silicon pixel tracking system and a calorimetric range telescope, which uniquely integrates range telescope functionality with track discrimination capabilities. The Bortfeld function fitting and Convolutional Neural Network (CNN) classifier algorithms are developed and applied for discrimination. In simulation studies, both approaches demonstrate the capability to reduce uncertainty in Water Equivalent Path Length (WEPL) determination for individual proton tracks to below 3 mm. The standard imaging protocol (3.2 mGy, 4×10^8 protons) achieves sub-millimeter spatial resolution (0.5 mm) with sub-1% RSP accuracy. With proton count requirements reduced by track discrimination, an ultra-low-dose protocol (0.16 mGy, 2×10^7 protons) is proposed with achieved sub-1% RSP accuracy and 1.1 mm spatial resolution in simulation.

This low-dose performance significantly expands clinical applicability, particularly for pediatric imaging or frequent imaging scenarios. Furthermore, the target 10 MHz proton detection rate suggests potential for real-time image guidance during radiotherapy. By circumventing the need for ultra-precise energy measurements, this design minimizes hardware complexity and provides a scalable foundation for future pCT systems.

Keywords: Proton computed tomography, Range telescope, Track Discrimination, Monte Carlo simulation

Introduction

Proton and heavy ion therapies have gained significant recognition for their efficacy in tumor treatment due to the unique Bragg peak feature [?]. This physical characteristic enables concentrated radiation deposition within tumor volumes

while sparing surrounding healthy tissues. Precise beam positioning, angular alignment, and energy modulation during treatment planning are crucial for ensuring accurate dose delivery. The Relative Stopping Power (RSP) map serves as the fundamental bridge connecting these phases, traditionally derived from Hounsfield Units (HU) in X-ray CT through empirical HU-to-RSP calibration [?]. However, inherent uncertainties of 2.4-5% in this conversion [?] have motivated active exploration of advanced modalities, including proton CT (pCT) and dual-energy CT (DECT), to improve RSP accuracy.

Current pCT implementations employ two distinct methodologies for RSP mapping: proton-integrating and proton-tracking. The integrating approach measures aggregate energy deposition from proton ensembles, while the tracking method reconstructs individual proton tracks and residual energies. Although proton-integrating systems benefit from simplified detector architectures and rapid imaging capabilities [?], their spatial resolution is fundamentally limited by Multiple Coulomb Scattering (MCS) [?], with additional constraints in dose efficiency due to proton pileup [?]. In contrast, proton-tracking systems achieve superior spatial resolution (~ 1 mm) and dose efficiency through single-proton event reconstruction. Due to advancements in detector technologies and high-speed electronics [?], the clinical viability of proton tracking systems has significantly enhanced in recent years [?, ?]. It should be noted that clinical pCT requires 3D RSP mapping with a density resolution better than 1% and a spatial resolution better than 1 mm. Such goals are empirically achieved through detection of more than 10^8 protons within a 100×100 mm² field [?]. To ensure clinical practicality, these systems must maintain sustained proton detection rates exceeding 10 MHz to reduce scanning time to the tens of seconds range.

Proton tracking systems typically consist of tracking detectors that measure the positions and directions of both incoming and outgoing protons, followed by range or energy detectors that measure the residual energies of the outgoing protons. Silicon-based detectors are extensively used in tracking applications due to their excellent spatial resolution, implemented either as strip detectors [10-13] or pixel detectors [?, ?]. The strategies for range or energy detection, however, vary significantly across different system designs. The Phase-II scanner [?] utilized a 5-stage scintillator-photomultiplier tube (PMT) stack to measure proton energy [?]. The PRaVDA [?], Hi-CT [?] and Bergen [?] systems all employed alternating silicon-absorber layers to determine proton range. Meanwhile, iMPACT [?] and ASTRA [?] used high-granularity calorimeters to simultaneously measure proton ranges and resolve multi-proton events, aided by fine longitudinal and lateral segmentation in the detectors. Although both architectures rely on energy deposition data, their implementations differ fundamentally. In the iMPACT calorimeter design, dual-threshold energy measurements were employed to identify the buildup region and sub-peak position of the proton energy deposition profiles, thus serving as a dedicated range telescope. In contrast, the ASTRA design combined calorimetric energy measurement with range estimation to reconstruct the residual proton energy, thereby functioning as a dual-purpose range telescope and calorimeter.

This study presents an innovative pCT design that integrates a silicon pixel tracking system and a high-granularity calorimetric range telescope. Following the iMPACT paradigm, our architecture fundamentally redefines the calorimeter as a range telescope rather than a precision energy spectrometer. This approach significantly alleviates the stringent demands for ultra-precise energy measurement. However, the range telescope exhibits compromised energy-range correlation when protons undergo nuclear interaction or large-angle scattering. To address this limitation, the acquired energy deposition profile is novelly transformed into a track classifier through systematic analysis of the proton interaction dynamics and assisted with Bortfeld function fitting and Convolutional Neural Network (CNN) techniques. Dedicated algorithms have been developed for proton track reconstruction and discrimination. The performance of the pCT system has been thoroughly verified through Monte Carlo simulations, demonstrating effective integration of range telescope with energy measurement. In addition, a simplified prototype is under development to demonstrate the viability of the detector and electronics components.

II. Materials and Methods

A. pCT System Design

The pCT system, as illustrated in Fig. 1 [Figure 1: see original paper], consists of a tracking system and a high-granularity calorimetric range telescope. It is designed to handle 200 MeV protons at a maximum rate of 10 MHz across a $100 \times 100 \text{ mm}^2$ detection area. To reduce lateral proton leakage, the detector coverage is extended to $128 \times 128 \text{ mm}^2$. The target imaging performance aims for a RSP accuracy of $<1\%$ and a spatial resolution of 1 mm. Within this detector coordinate system, the z-axis corresponds to the direction of the incoming proton beam. In the transverse plane, the x-axis is oriented horizontally perpendicular to the beam (z-axis), while the y-axis points vertically. For tomographic imaging, the phantom rotates around the y-axis during CT scans.

The tracking system includes four tracking stations that are instrumented with Complementary Metal Oxide Semiconductor (CMOS) pixel sensors. The two front stations measure positions and directions of protons entering the phantom, while the two back stations measure exiting protons. The utilized CMOS pixel sensor, named TaichuPix-3 [?, ?], is a full-reticle-sized sensor featuring small pixels of $25 \times 25 \text{ }\mu\text{m}^2$. Its high granularity and short dead time ($< 500 \text{ ns}$) enable it to achieve high counting efficiency under extreme beam conditions, with the capacity to handle up to $3.6 \times 10^7 \text{ hits}/(\text{cm}^2 \cdot \text{s})$ [?]. These exceptional performance characteristics ensure that the tracking system can process the target proton fluxes without saturation. The TaichuPix-3 sensor features a pixel array of 1024 columns and 512 rows, forming an active area of $12.8 \times 25.6 \text{ mm}^2$.

Through modular assembly, five sensors are mounted sequentially on a flexible printed circuit (FPC) to form a ladder measuring $12.8 \times 128 \text{ mm}^2$. Ten such

ladders are attached to both sides of a 400 μm -thick carbon fiber plate, with five ladders on each side. They are interleaved with appropriate overlap to minimize the dead area occupied by the peripheral readout electronics. This arrangement provides each tracking station with a full detection area coverage of $128 \times 128 \text{ mm}^2$, while maintaining a material budget of approximately 0.37% radiation length (X_0).

The calorimetric range telescope comprises 100 stacked layers, each with a thickness of 3 mm, aligned along the beam axis to determine proton stopping positions. Each layer is divided into 16 independent scintillator bars. The orthogonal orientation of scintillator bars in adjacent layers enables distinct reconstruction of proton tracks in the x-y plane through cross-layer position correlation. This configuration also allows the estimation of proton entry points into the scintillator, facilitating the matching of proton tracks reconstructed from the downstream stations of the tracking system. Each scintillator bar measures $128 \times 8 \times 3 \text{ mm}^3$, resulting in a total coverage area of $128 \times 128 \text{ mm}^2$ and a total thickness of 300 mm. BC-408 plastic scintillators are adopted to detect proton energy deposition due to its advantages of short decay time ($< 2.1 \text{ ns}$), high light yield and cost-effectiveness. The calculated penetration depth of proton beam in BC-408 is 261.2 mm, which is well within the longitudinal containment capacity of the range telescope. The scintillator bars are encapsulated with ESR reflector film to increase the light yield and aluminum foil to minimize the inter-channel crosstalk.

Each scintillator bar is coupled to a Hamamatsu SiPM and read out with ASIC MPT2321, which features 32-channels with integrated 12-bit ADC and 20-bit TDC for each individual channel [?]. The range telescope attains an expected detection rate beyond 10 MHz per channel, enabled by SiPM decay time below 100 ns and MPT2321 integration time of 50 ns. Data from MPT2321 readout chips and TaichuPix-3 pixel sensors are routed through a FELIX-based DAQ system [?], achieving sustained throughput up to 10 Gb/s for real-time event packaging and PC transmission.

B. pCT Simulation

A Geant4 simulation with 200 MeV parallel proton beam was developed to evaluate the system performance and compare designed configurations. The physics list QGSP_{BIC} was employed for proton and secondary particle interactions, while G4EmStandardPhysics_{option4} governed scintillation photon generation and transport. Each tracking station comprised TaichuPix-3 chips, FPCs, support, and interconnects, resulting in a material budget of 0.37% X_0 . Noise and crosstalk were intentionally excluded in this preliminary modeling, allowing direct recording of all proton and secondary particle hits without energy thresholds.

In the range telescope, scintillator bars were encapsulated with an 80 μm -thick composite reflective layer (ESR + aluminum) as insensitive region. Scintillation

light transport was decoupled from the main simulation through a standalone single-bar modeling. The photo-electron yield (nPE) was determined from the simulated photon count at the SiPM, then applying a 40% quantum efficiency and introducing a 10% smearing to account for coupling uncertainties. The nPE exhibited strong correlation with proton hit position along scintillator length (128 mm), varying by 50% from one side to the other. The position dependence was modeled by nPE distributions according to the hit position. During full-system simulation, light yields were determined by sampling from these position-dependent distributions.

Two phantoms were built in Geant4 to systematically evaluate the pCT system performance and reconstruction algorithm: one dedicated to RSP imaging, and the other for Water Equivalent Path Length (WEPL) calibration and validation, where WEPL represents the integral of RSP along the proton path.

- **Phantom 1 (RSP Characterization):** A cylindrical water phantom (30 mm radius and 40 mm height) containing four cylindrical inserts (5 mm radius and 40 mm height) of polypropylene (PP), teflon, air, and bone-equivalent material (bone-100). The CT scan employed 1° rotational increments over $0-180^\circ$ for RSP imaging.
- **Phantom 2 (WEPL Calibration):** A water cuboid ($128 \times 128 \text{ mm}^2$ base area) with adjustable thickness (0–140 mm in 5 mm steps). This configuration directly calibrates WEPL from proton residual range measurement by matching the thickness to WEPL.

The simulation captured both detector outputs and ground truth data. The detector outputs comprised triggered cells in tracking system and nPE map in range telescope. The ground truth data includes proton energy loss in the phantom and interaction features, whose impact on proton track concluded as follows [?]:

- **Ionization:** Induces quasi-continuous energy loss.
- **Elastic Coulomb scattering:** Causes track deviations through scattering.
- **Non-elastic nuclear reactions:** Generates secondary particles (protons, neutrons, heavy ions and gamma rays) accompanied by abrupt energy loss in secondary protons.
- **Bremsstrahlung:** Simultaneously alters track and induces radiative energy loss, negligible for 200 MeV protons.

The range telescope configuration prioritizes a strong proton energy-range correlation. Thus, only ionization and small-angle elastic Coulomb scattering are considered beneficial for reconstruction. To facilitate the development of noise suppression algorithms, the simulation recorded binary flags for occurrences of nuclear reactions and maximum scattering angles in both phantom and range telescope, respectively.

C. pCT Reconstruction

The RSP map is derived from detector outputs via a three-step workflow:

1. **Proton track reconstruction:** Proton tracks through the tracking system and range telescope are reconstructed to determine both positions and WEPL values, which are derived from residual range measurement with calibrated WEPL-range conversion.
2. **WEPL imaging:** WEPL maps are generated through statistical accumulation of sufficient proton events.
3. **RSP imaging:** The 3D RSP map is reconstructed from multi-angle WEPL projections using the filtered back-projection (FBP) algorithm.

The current reconstruction framework predominantly addresses single-proton events with an emphasis on noise suppression, whereas the extension to multi-proton reconstruction remains under development.

1. Position Measurement and Filter on Scattering Angle

The fired pixels are matched between the front and back tracking stations to reconstruct the incoming and outgoing proton tracks, respectively. A matching algorithm similar to that presented in Ref. [?] is used for future multi-proton reconstruction. In the current single-proton scenario, this algorithm demonstrates sufficient robustness to effectively exclude false matching from secondary particles.

Once matched tracks are established, the Most Likely Path (MLP) algorithm is applied to estimate the track's spatial coordinates in both the x- and y-directions. The MLP framework leverages a matrix-based formalism [?], which combines Bayesian statistical inference with a Gaussian approximation model of MCS. Inputs to the scattering geometry algorithm include predefined phantom boundaries, though these may be replaced in clinical implementations by fixed patient position or a priori imaging data. To benchmark the MLP algorithm's localization improvement, the midpoint between the incoming and outgoing track intersections on the mid-plane of the phantom is computed as another position estimator for comparative analysis.

The proton scattering angle (θ) constitutes a critical parameter for background suppression in tracking system. Fig. 2 [Figure 2: see original paper] presents the energy loss versus θ correlation for 200 MeV protons traversing a 100 mm water-equivalent phantom. While the majority of protons show an expected energy loss near 51 MeV, the subset with higher energy loss exhibits non-negligible angular dispersion. This correlation enables effective event filtering based on θ .

2. Residual Range Measurement and Filter on Deposit Energy

The nPE is measured by SiPMs, producing maps in alternating projections in the range telescope: X-Z planes for even-numbered layers and Y-Z planes for odd-numbered layers. Proton track reconstruction is performed using an iterative algorithm:

1. **Seed initialization:** Utilize extrapolated tracks from back tracking stations as initial seeds.
2. **Candidate selection and iteration:**
 - Select candidate hits from scintillator bars aligned with or adjacent to current seed.
 - Prioritize aligned hit as next-iteration seed. If none exist, choose adjacent hit inclined to the track direction.
 - Continuously update the seed and track direction, which are initially derived from extrapolated tracks and refined through iterations.
3. **Termination criteria:** Halt iteration when two consecutive layers within the same projection (X-Z or Y-Z) lack candidate hit. A single-layer skip tolerance accommodates insensitive regions between scintillator bars.

The reconstructed tracks are primarily labeled as “perturbed” tracks if nuclear interaction and large angle scattering ($>10^\circ$) occurs, otherwise as “unperturbed” tracks. The distribution of deposited energy (Edep) is crucial for track discrimination, thereby reducing the uncertainty in reconstructed range. The conversion from nPE to Edep is calibrated through standalone single-bar simulations, accounting for position-dependent light attenuation along the 128-mm scintillator bar length. The hit position along the bar length is determined by analyzing the reconstructed track across adjacent orthogonal layers.

The Edep profile is modeled using the Bortfeld function [?], which provides an analytical representation expressed as:

$$D(z) = 0.65 \cdot D_{100} \cdot \left[\left(\frac{z - R_0}{z - R_0} \right) + \sigma \cdot k \cdot PCF \right]$$

where $D(z)$ is Edep at depth z , D_{100} is the peak of Edep, p is range-energy exponent (material-dependent), R_0 is proton range, σ is composite Gaussian width accounting for both range straggling and detector resolution, k is fluence reduction coefficient, PCF is parabolic cylinder function, implemented from Ref. [?]. The model-dependent parameters (D_{100} , p , σ and k) are determined by MC study on unperturbed tracks and fixed during track discrimination as shown in Fig. 3 [Figure 3: see original paper] (a, b), while R_0 remains free. To mitigate effects of energy leakage from insensitive regions, energy deposits below 80% of the first layer energy (empirically determined) receive increased uncertainty weighting, thereby improving fitting robustness for unperturbed tracks.

The metric $\phi = \log_{10}(\chi^2/NDF)$ is introduced to quantify goodness of the fit, where $\chi^2 = \sum(y_i - f(x_i))^2$ represents the discrepancy between the observed values y_i and the fitted values $f(x_i)$, $NDF = N_{layers} - N_{parameters}$ is the number of degrees of freedom, N_{layers} is the number of data points in the fit and $N_{parameters}$ is number of free parameters. Fig. 3 (c, d) shows the ϕ -range correlation for 200 MeV protons without a phantom. Unperturbed tracks are centered around the target range with $\phi \approx 1$, while other tracks exhibit increasing ϕ as the range decreases. This result provides a basis for selecting tracks based on ϕ .

As an alternative track discrimination method, convolutional neural network (CNN) is implemented in the range telescope inspired by the work in Ref. [?]. The CNN is constructed using TensorFlow [?] with the following specifications:

- **Input:** 2D projections of the scintillator light yield (nPE) in the X-Z and Y-Z planes.
- **Architecture:** Four convolutional layers with ReLU activation, kernel sizes 3×3 , and max-pooling. Two fully connected layers with dropout regularization ($p = 0.5$) to mitigate overfitting.
- **Training:** Adam optimizer [?] with a learning rate of 10^{-3} , 128 batch size and 10 training epochs, loss function is binary cross-entropy for probabilistic discrimination.
- **Output:** A scalar probability (0-1) indicating the likelihood of a track being classified as “good”.
- **Sample:** Mixed sample similar to the Phantom 2 group, where protons pass a water phantom with thickness ranging from 0 to 150 mm, the thickness step is 1 mm instead of 5 mm, 80% of the sample is used for training and the rest for testing.

Proton tracks are labeled using two complementary methods:

- **Process tags:** Derived from Geant4 ground truth data. Tracks are labeled as “good” if they are not perturbed by nuclear interaction and large angle scattering ($>10^\circ$).
- **Range tags:** Reconstructed ranges for protons passing a same phantom thickness are fitted with a Gaussian distribution. Tracks within $\pm 3\sigma$ are labeled as “good”.

Process tags are more fundamental but can only be accessed through simulation, so it requires MC-data agreement in nPE distribution. Range tags can be realized in practical experiment but lack physical interpretability.

3. WEPL Calibration

The reconstructed residual range is converted to WEPL through a calibration process. The cuboid phantom group (Phantom 2) is utilized to allow protons to pass through varying thicknesses of the phantom. The reconstructed range distributions for 0 mm (no phantom) and 100 mm thicknesses are fitted with

Gaussian functions as shown in Fig. 4 [Figure 4: see original paper] (a) and (b). The means and widths of the Gaussian fits are plotted against the thicknesses, i.e., the WEPL values, and are fitted with a linear function, as illustrated in Fig. 4 (c).

4. pCT Imaging

The protons are grouped into $0.5 \text{ mm} \times 0.5 \text{ mm}$ pixels to construct the WEPL maps. Each reconstructed proton track provides a WEPL measurement at a spatial position determined by the MLP method, where the WEPL value is determined from the reconstructed proton range and the position is calculated via the MLP method. Following previous pCT studies [?], a “pixel-wise” filtering is applied under the assumption of uniform WEPL within each pixel. The mean and standard deviation (σ) of WEPL values within each pixel are calculated, and WEPL measurements deviating by more than 2σ from the mean are discarded.

The WEPL maps acquired at different angular projections are systematically organized into sinograms. Subsequently, the three-dimensional RSP map is reconstructed through filtered back projection (FBP) utilizing a Hann filter for frequency-domain noise suppression. In this preliminary investigation, advanced CT reconstruction algorithms are intentionally avoided to specifically isolate and evaluate the fundamental performance characteristics of both the detector system and reconstruction methodology.

III. Results

A. Performance of Track Reconstruction

The σ and Edep thresholds are determined by balancing the reconstructed WEPL standard deviation (STD) and selection rate—two critical metrics for imaging noise reduction. Optimized thresholds yield selection rates of 97% for the σ filter and 73% for the Edep filter.

Fig. 5 [Figure 5: see original paper] compares the reconstructed WEPL distributions under Edep filters after applying the σ filter. The WEPL profile is segmented into three regions: the peak, proximal tail and distal tail. While the σ filter marginally reduces background in the tails, all Edep filters achieve significant suppression (1-2 order of magnitude) in both tails. Proximal tail events consist of proton tracks undergoing large-angle scattering or nuclear interaction, so CNN filters outperform Bortfeld fitting by leveraging additional scattering discrimination from nPE maps. Conversely, distal tail events are governed by nuclear interaction with anomalous Edep distributions, leading to comparable performance between CNN filters and the Bortfeld fitting. The range-tagged CNN exhibits superior performance in the $\text{WEPL} > 150 \text{ mm}$ region, likely due to the training dataset limitations (0-150 mm WEPL range).

To quantify pixel-wise filter performance, WEPL measurements from same thicknesses are virtually grouped with Poisson-sampled group sizes ($N=56$, see Sec.

III B). Fig. 5 (b) displays the STDs of track-level WEPL after applying the pixel-wise filter. The STDs are suppressed to 3 mm-level, showing the effectiveness of track discrimination. While pixel-wise filters effectively suppress distal tail events, their limited efficacy in proximal tails results in similar STD performance for process-tagged and range-tagged CNNs, whereas Bortfeld fitting exhibits an 8% higher STD. The range-tagged CNN shows anomalies at 0 mm and 140 mm thicknesses, attributable to training set truncation (0-150 mm range).

Although proton energy is not directly reconstructed due to the WEPL-range calibration, its resolution can be derived from the range resolution via:

$$\sigma_E/E = p \cdot \sigma_R/R$$

where $\sigma_R=3$ mm (estimated from WEPL STD), $R=260$ mm (200 MeV proton range in water), and $p=1.77$ (Geiger exponent). This yields a relative energy resolution of 0.6% for 200 MeV protons.

Fig. 5 (c) presents the deviation between reconstructed and true WEPL values under combined Edep and pixel-wise filter. Systematic biases are observed: 0.2 mm (process-tagged CNN), 0.3 mm (range-tagged CNN) and 0.5 mm (Bortfeld fitting). This bias originates from differences in range determination during calibration versus reconstruction. During calibration, the range for a given WEPL is defined as the peak position in the Gaussian-fitted range distribution. In contrast, reconstructed ranges represent filtered mean values, where the residual tail slightly reduces the mean of reconstructed range, thereby overestimating WEPL. Crucially, the bias magnitude remains consistent across phantom thicknesses, confirming its independence from phantom geometry. This thickness-independent bias confirms the filter's stability. Thus, a fixed calibration offset (equal to the negative bias) is applied for each method.

In evaluating Edep filter methods, the process-tagged CNN achieves the lowest STD/bias with negligible phantom thickness dependence, serving as a theoretical benchmark for detector performance limits due to its geometry independence. However, its experimental infeasibility restricts practical use unless nPE consistency between data and MC is proved. The range-tagged CNN, while experimentally viable, exhibits significant sample-dependent variations (e.g., WEPL range and step). To mitigate this instability, Bortfeld fitting provides a physics-constrained parameterized alternative with enhanced robustness. For real-world applications, the latter two methods will be experimentally validated to balance accuracy and feasibility.

B. Performance of Benchmark RSP Imaging

Benchmark pCT imaging employs 4×10^8 protons with 180 projections, achieving $0.5 \times 0.5 \times 0.5$ mm³ pixel resolution over a 100×100 mm² field. The simulated radiation dose is approximately 3.2 mGy. Assuming uniform proton

distribution, this configuration yields an average of $\bar{n} = 56$ protons per pixel before filtering and 40 after filtering. Pixel-wise WEPL standard deviations below 0.5 mm are achieved for phantom thicknesses of 0–140 mm, exhibiting agreement with the theoretical relationship $\sigma\{pixel\} = \sigma\{proton\}/\sqrt{N_{\{proton\}}}$.

The pCT scan is performed on Phantom 1, which includes material inserts comprising PP, Teflon, air and bone-100. The reconstructed images for a slice are illustrated in Fig. 6 [Figure 6: see original paper]. To quantify the accuracy of the RSP map, cylindrical regions of interest (ROIs) are defined for each insert. Each ROI is dimensioned to half the radius and height of its corresponding insert. For volumetric analysis, each ROI spanned 40 consecutive slices, with 80 pixels per slice, yielding a total of 3200 pixels (denoted as $n_{\{pixel\}}$). The mean RSP ($RSP_{\{mean\}}$) and width ($\sigma_{\{pixel\},RSP}$) within each ROI are obtained by Gaussian fitting. The RSP accuracy, representing the systematic deviation from the reference value, is calculated in a same way with Ref. [?]:

$$RSP_{acc} = 100 \cdot \frac{RSP_{mean} - RSP_{ref}}{RSP_{ref}}$$

where $RSP_{\{ref\}}$ denotes the ground-true RSP value of each material. Additionally, $\sigma_{\{pixel\},RSP}$ characterizes the noise of spatial heterogeneity values at pixel level. To estimate the statistical precision of the RSP measurement, the noise is derived by normalizing $\sigma\{RSP\}$ with the square root of the total sampled pixels:

$$RSP_{noise} = \frac{\sigma_{pixel,RSP}}{\sqrt{n_{pixel}}}$$

The spatial resolutions are quantified by extracting RSP values near insert-water boundaries, following fits with a step function convolved with a Gaussian function. All materials demonstrate spatial resolution of ≈ 0.5 mm, with MLP implementation shown a marginal improvement of 0.03 mm compared to non-MLP methods. Notably, PP exhibits measurement instability due to significant pixel-level RSP fluctuations, stemming from its near-water equivalence with only 4% RSP difference from water. When excluding the unstable PP results, MLP demonstrates approximately 10% performance enhancement over conventional methods. This limited improvement can be attributed to the fundamental spatial resolution limit imposed by the 0.5 mm pixel size inherently constrains potential improvements.

The pCT imaging results are summarized in Table 1. RSP accuracies remain below 0.5% for all materials except air, where the small absolute RSP value leads to an amplified relative error. Pixel-level RSP fluctuations show consistency across all inserts, maintaining a STD of approximately 0.013.

C. Performance of Low-Dose RSP Imaging

Track discrimination demonstrates a significant advantage in statistical efficiency by implementing preliminary suppression of the WEPL tail distribution. This preprocessing step significantly reduces the proton density requirements for subsequent pixel-wise filtering. From study of virtually grouping of WEPL measurements, a sub-3 mm track-level WEPL STD is achieved with a remarkably low proton density of 11 protons per pixel. So a low-dose proton imaging protocol is proposed with only 2×10^7 protons distributed across 180 projection angles, corresponding to a total absorbed dose of 0.16 mGy. The imaging configuration adopts an enlarged pixel size of $1.0 \times 1.0 \times 1.0 \text{ mm}^3$, with each ROI containing 400 pixels.

The primary reconstructed sinogram and pCT image in Fig. 7 [Figure 7: see original paper] (a) exhibit linear artifacts caused by noisy pixels in the sinogram where insufficient proton counts compromise the effectiveness of pixel-wise filtering. To address this issue, a noise correction protocol is implemented in which pixels with <5 proton counts are corrected using linear interpolation of adjacent valid pixels (≥ 5 counts). This count-driven correction method effectively preserves WEPL accuracy while mitigating artifacts, as demonstrated by the refined sinogram patterns and enhanced pCT reconstruction quality shown in Fig. 7 (b). In addition, significant potential exists for developing advanced algorithms leveraging both WEPL data and proton count distributions across neighboring pixels.

The quantified RSP accuracy and spatial resolution are presented in Table 2. While the RSP uncertainty increases by 5–6 times compared to conventional pCT protocols, the RSP accuracy remains $<1\%$ (except for air) while spatial resolution increased to 1.1 mm. These results demonstrate the feasibility of clinical pCT implementation at ultralow radiation doses (0.2 mGy level), achieving the dual objectives of $<1\%$ RSP accuracy and millimeter-scale resolution—specifications sufficient for pretreatment positioning in fractionated proton therapy. Though exhibiting elevated noise levels, this dose-efficient reconstruction paradigm maintains sufficient imaging fidelity for repeated beam alignment applications where cumulative radiation exposure must be strictly controlled. The scanning time is reduced to 2 seconds at a 10 MHz proton detection rate, suggesting the potential of real-time image-guided radiotherapy. Imaging quality can be further enhanced through integration with pre-acquired X-ray reference images [?, ?].

IV. Discussion and Prospect

A. Study on Detector Configurations

The detector configurations are investigated through systematic simulations, with key design parameters compared in Table 3. The process-tagged CNN is used as Edep filter due to its model-independent tagging. The WEPL STDs at the track level serve as the primary evaluation metric. Furthermore, recon-

structed RSP images demonstrate that $\sigma_{\{\text{pixel}\},\text{RSP}}$ increases proportionally with WEPL variations. Notably, all configurations maintain RSP accuracy within 1% (excluding air) through individualized calibration implemented for each detector setup.

The range telescope paradigm is designed to relax the requirements for energy resolution of individual channels. Reducing the bit depth from 12-bit to 4-bit results in WEPL STDs with robust performance (less than 2% variance). These findings also suggest faster readout speeds by adopting lower-bit ADCs and reducing data size.

The scintillator bar thickness is a critical parameter in range telescope design, as it directly determines the WEPL uncertainty, quantified by:

$$\sigma_{WEPL}^2 = k^2 \cdot (\sigma_{proton}^2 + \sigma_{det}^2)$$

where k denotes the WEPL-range calibration coefficient from Fig. 4 (c). The intrinsic proton range straggling in plastic scintillator contributes $\sigma_{\{proton\}} = 2.5 \text{ mm}$, while the detector uncertainty $\sigma_{\{det\}} = \text{Thickness}/\sqrt{12}$ arises from geometrical resolution limitations. Theoretical calculations demonstrate a baseline uncertainty of 2.75 mm at 3 mm thickness. Progressive thickness increases to 4, 5, and 6 mm elevate the uncertainty by 4.1%, 9.1%, and 15.0% respectively. The simulated STDs in Table 3 show agreement with theoretical predictions, thereby confirming the analytical model. These results establish 3 mm as the optimal scintillator thickness for precision range measurement applications.

The lateral segmentation is primarily intended to reconstruct multi-proton tracks. However, the STD of single-proton discrimination with CNN increases when the scintillator width is enlarged, becoming comparable to that of the Bortfeld fitting method. This result indicates an additional advantage from lateral segmentation by enabling the distinction of large-angle scattering events.

B. Proof-of-Concept Prototype

At the preliminary stage, a simplified prototype is designed to validate the core principles. The tracking system configuration is reduced to a single TaichuPix-3 chip per detection layer, while the range telescope is implemented as a $50 \times 50 \times 3 \text{ mm}^3$ scintillator plane, as shown in Fig. 8 [Figure 8: see original paper]. This configuration provides a detection area of $25.6 \times 12.8 \text{ mm}^2$, allowing the reconstruction of only one proton per event. The range telescope architecture comprises 64 layers covering the range of 100 MeV protons, with each layer incorporating one SiPM.

A 100 MeV proton beam evaluation is planned for 2025, with this prototype serving for initial system verification and experimental data acquisition. Subsequent enhancements will focus on: (1) Implementation of nonideal simulation and mitigation strategies in reconstruction addressing noise, dead area and

crosstalk in both tracking system and range telescope. (2) Discrimination verification of Bortfeld fitting and range-based CNN with realistic nPE distributions. (3) Testing of the real system performance and establishment of the imaging protocols, especially the statistical requirement per pixel.

C. Further Improvement

Further improvements in proton detection rates primarily depend on multi-proton reconstruction techniques. The multi-proton reconstruction algorithm under development for the range telescope aims to discriminate overlapping proton tracks while maintaining the existing single-track discrimination performance. By addressing challenges in accurately iterating hits and properly allocating light yield at proton track intersections, this algorithm targets the simultaneous reconstruction of 3–5 protons while maintaining a WEPL STD of 3 mm. Additionally, the tolerance for energy measurement precision (validated via nPE digitization studies) enables potential detection rate improvements by adopting faster SiPMs and readout chips.

V. Conclusion

This study presents an innovative pCT system that addresses traditional limitations in energy measurement precision and energy-range correlation through a multi-stage filtering strategy integrated into the imaging workflow. By sequentially applying scattering angle filtering in the tracking system, energy deposition filtering in the range telescope, and pixel-wise filtering during WEPL image reconstruction, the system achieves remarkable performance: 0.5 mm spatial resolution and $<1\%$ RSP accuracy in simulations using 4×10^8 protons. The key innovation lies in this multi-filter architecture's statistics-easing effect, which substantially reduces proton count requirements compared to conventional pixel-wise filtering alone. This enables ultra-low-dose imaging feasibility, demonstrating sub-1% RSP accuracy with just 2×10^7 protons (0.16 mGy dose) while maintaining 1.1 mm resolution. Such capabilities directly address two critical clinical needs: repeatable patient positioning for fractionated therapy and potential real-time imaging applications.

Systematic design analyses reveal fundamental principles for system optimization: scintillator geometry studies establish thickness and width tolerances for range telescope performance, and digitization investigations demonstrate robustness against energy resolution compromises. Future work will focus on prototype validation and enhancement strategies, particularly crosstalk and noise mitigation. This architecture, combining physical filtering with advanced algorithms, provides a scalable framework for pCT systems that simultaneously achieve high RSP resolution, ultra-low radiation dose, and clinical practicality.

VI. Bibliography

[1] R. Mohan., A review of proton therapy –current status and future directions.

- Prec. Rad. Onc. 6(2), 164-176 (2022). <https://doi.org/10.1002/pro6.1149>
- [2] B. Schaffner and E. Pedroni., The precision of proton range calculations in proton radiotherapy treatment planning: experimental verification of the relation between CT-HU and proton stopping power. Phys. Med. Biol. 43(6), 1579 (1998). <https://dx.doi.org/10.1088/0031-9155/43/6/016>
- [3] M. Yang, X.R. Zhu, P.C. Park et al., Comprehensive analysis of proton range uncertainties related to patient stopping-power-ratio estimation using the stoichiometric calibration. Phys. Med. Biol. 57(13), 4095 (2012). <https://dx.doi.org/10.1088/0031-9155/57/13/4095>
- [4] P. Pemler, J. Besserer, J. de Boer et al., A detector system for proton radiography on the gantry of the Paul-Scherrer-Institute. Nucl. Inst. Meth. A 432(2), 483-495 (1999). [https://doi.org/10.1016/S0168-9002\(99\)00284-3](https://doi.org/10.1016/S0168-9002(99)00284-3)
- [5] U. Schneider and E. Pedroni, Multiple coulomb scattering and spatial resolution in proton radiography. Med. Phys. 21(11), 1657-1663 (1994). <https://doi.org/10.1118/1.597212>
- [6] M. Testa, J.M. Verburg, M. Rose et al., Proton radiography and proton computed tomography based on time-resolved dose measurements. Phys. Med. Biol. 58(22), 8215 (2013). <https://dx.doi.org/10.1088/0031-9155/58/22/8215>
- [7] R. He, X.Y. Niu, Y. Wang et al., Advances in nuclear detection and readout techniques. Nucl. Sci. Tech. 34(12), 205 (2023). <https://doi.org/10.1007/s41365-023-01359-0>
- [8] G. Poludniowski, N.M. Allinson and P.M. Evans, Proton radiography and tomography with application to proton therapy. Brit. J. Radiol. 88(1053), 20150134 (2015). <https://doi.org/10.1259/bjr.20150134>
- [9] R.P. Johnson, Review of medical radiography and tomography with proton beams. Rep. Prog. Phys. 81(1), 016701 (2017). <https://dx.doi.org/10.1088/1361-6633/aa8b1d>
- [10] H.F.W. Sadrozinski, T. Geoghegan, E. Harvey et al., Operation of the preclinical head scanner for proton CT. Nucl. Inst. Meth. A 831(21), 394-399 (2016). <https://doi.org/10.1016/j.nima.2016.02.001>
- [11] M. Esposito, C. Waltham, J.T. Taylor et al., PRAVDA: The first solid-state system for proton computed tomography. Phys. Med. 55, 149-154 (2018). <https://doi.org/10.1016/j.ejmp.2018>
- [12] F. Baruffaldi, R. Iuppa, E. Ricci et al., iMPACT: An innovative tracker and calorimeter for proton computed tomography. IEEE Trans. Radiat. Plasma Med. Sci. 2(4), 345-352 (2018). <https://doi.org/10.1109/TRPMS.2018.2825499>
- [13] J. Alme, G.G. Barnaföldi, R. Barthel et al., A high-granularity digital tracking calorimeter optimized for proton CT. Front. Phys. 8 (2020). <https://www.frontiersin.org/journals/physics/articles/10.3389/fphy.2020.568243>
- [14] V.A. Bashkirov, R.W. Schulte, R.F. Hurley et al., Novel scintillation detector design and performance for proton radiography and computed tomography. Med. Phys. 43(2), 664-674 (2016). <https://doi.org/10.1118/1.4939255>
- [15] R.P. Johnson, V. Bashkirov, L. DeWitt et al., A fast experimental scanner for proton CT: Technical performance and first experience with phantom scans. IEEE Trans. Nucl. Sci. 63(1), 52-60 (2016). <https://doi.org/10.1109/TNS.2015.2491918>

- [16] Y.L. Chen, H.K. Wang, S.Y. Zhang et al., Hi' CT: a pixel sensor-based device for ion tomography. Nucl. Sci. and Tech. 34(7), 111 (2023). <https://doi.org/10.1007/s41365-023-01280-6>
- [17] N. Pozzobon, F. Baruffaldi, D. Bisello et al., Calorimeter prototyping for the iMPACT project pCT scanner. Nucl. Inst. Meth. A 936, 1-4 (2019). <https://doi.org/10.1016/j.nima.2018.10.155>
- [18] M. Granado-González et al., A novel range telescope concept for proton CT. Phys. Med. Biol. 67(035013), 2022. <https://iopscience.iop.org/article/10.1088/1361-6560/ac4b39>
- [19] T. Wu, W. Wei, S. Grinstein et al., The TaichuPix1: a monolithic active pixel sensor with fast in-pixel readout electronics for the CEPC vertex detector. J. Inst. 16(09), 09020 (2021). <https://dx.doi.org/10.1088/1748-0221/16/09/P09020>
- [20] Y. Zhang, W. Wei, X.T. Li et al., Development of a CMOS pixel sensor prototype for the high hit rate CEPC vertex detector. Nucl. Inst. Meth. A 1042, 167442 (2022). <https://doi.org/10.1016/j.nima.2022.167442>
- [21] T.Y. Wu, S.Q. Li, J. Zhou et al., Beam test of a 180nm CMOS pixel sensor for the CEPC vertex detector. Nucl. Inst. Meth. A 1059, 168945 (2024). <https://doi.org/10.1016/j.nima.2023>
- [22] Y. Liu, D.J. Du, B. Lu et al., Characterisation studies of two front-end electronics chips designed for SiPM readout. Nucl. Inst. Meth. A 1069, 169808 (2024). <https://doi.org/10.1016/j.nima.2024.169808>
- [23] W. Wu et al.[ATLAS TDAQ Collaboration], FELIX: the New Detector Interface for the ATLAS Experiment. IEEE Trans. Nucl. Sci. 66(7), 986-992 (2019). <https://doi.org/10.1109/TNS.2019.2913617>
- [24] W.D. Newhauser and R. Zhang, The physics of proton therapy. Phys. Med. Biol. 60(8), 155 (2015). <https://dx.doi.org/10.1088/0031-9155/60/8/R155>
- [25] R.W. Schulte, S.N. Penfold, J.T. Tafas et al., A maximum likelihood proton path formalism for application in proton computed tomography. Med. Phys. 35(11), 4849-4856 (2008). <https://doi.org/10.1118/1.2986139>
- [26] T. Bortfeld, An analytical approximation of the Bragg curve for therapeutic proton beams. Med. Phys. 24(12), 2024-2033 (1997). <https://doi.org/10.1118/1.598116>
- [27] S. Zhou, Parabolic-Cylinder-Functions-C-(2022), GitHub. <https://github.com/civerjia/Parabolic-Cylinder-Functions-C->
- [28] H.E.S. Pettersen, M. Aehle, J. Alme et al., Investigating particle track topology for range telescopes in particle radiography using convolutional neural networks. Acta Oncol. 60(11), 1413-1418 (2021). <https://doi.org/10.1080/0284186x.2021>
- [29] M. Abadi, A. Agarwal, P. Barham et al., Tensorflow: A system for large-scale machine learning (2015). <https://www.tensorflow.org/>
- [30] D.P. Kingma and J. Ba, Adam: A method for stochastic optimization. CoRR, abs/1412.6980 (2014). <https://api.semanticscholar.org/CorpusID:6628106>
- [31] G. Dedes, J. Dickmann, K. Niepel et al., Experimental comparison of proton CT and dual energy x-ray CT for relative stopping power estimation in proton therapy. Phys. Med. Biol. 64(16), 165002 (2019).

<https://doi.org/10.1088/1361-6560/ab2b72>

[32] Y.Q. Yang, W.C. Fang, X.X. Huang et al., A new imaging mode based on X-ray CT as prior image and sparsely sampled projections for rapid clinical proton CT. Nucl. Sci. Tech. 34(8), 126 (2023). <https://doi.org/10.1007/s41365-023-01280-6>

[33] Y.Q. Yang, W.C. Fang, X.X. Huang et al., Static superconducting gantry-based proton CT combined with X-ray CT as prior image for FLASH proton therapy. Nucl. Sci. Tech. 34(1), 11 (2023). <https://doi.org/10.1007/s41365-022-01163-2>

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv –Machine translation. Verify with original.