PAPER • OPEN ACCESS

Near-field coded-mask technique and its potential for proton therapy monitoring

To cite this article: Ronja Hetzel *et al* 2023 *Phys. Med. Biol.* **68** 245028

View the article online for updates and enhancements.

Physics in Medicine & Biology

IPEM
Institute of Physics and
Engineering in Medicine

CrossMark

PAPER

Near-field coded-mask technique and its potential for proton therapy monitoring

Ronja Hetzel[1,*] , Vitalii Urbanevych[2,*] , Andreas Bolke[3] , Jonas Kasper[1], Monika Kercz[2,4] , Magdalena Kołodziej[2,4] , Andrzej Magiera[2] , Florian Mueller[5] , Sara Müller[1] , Magdalena Rafecas[3] , Katarzyna Rusiecka[2] , David Schug[5,6] , Volkmar Schulz[1,5,6] , Achim Stahl[1] , Bjoern Weissler[5,6] , Ming-Liang Wong[2] and Aleksandra Wrońska[2]

1   III. Physikalisches Institut B, RWTH Aachen University, Aachen, Germany
2   Marian Smoluchowski Institute of Physics, Jagiellonian University, Kraków, Poland
3   Institute of Medical Engineering, University of Lübeck, Lübeck, Germany
4   Doctoral School of Exact and Natural Sciences, Jagiellonian University, Kraków, Poland
5   Physics of Molecular Imaging Systems, RWTH Aachen University, Aachen, Germany
6   Hyperion Hybrid Imaging Systems GmbH, Aachen, Germany
*   Authors to whom any correspondence should be addressed.

E-mail: ronja.hetzel@rwth-aachen.de and vitalii.urbanevych@doctoral.uj.edu.pl

Abstract

Objective. Prompt-gamma imaging encompasses several approaches to the online monitoring of the beam range or deposited dose distribution in proton therapy. We test one of the imaging techniques - a coded mask approach - both experimentally and via simulations. *Approach.* Two imaging setups have been investigated experimentally. Each of them comprised a structured tungsten collimator in the form of a modified uniformly redundant array mask and a LYSO:Ce scintillation detector of fine granularity. The setups differed in detector dimensions and operation mode (1D or 2D imaging). A series of measurements with radioactive sources have been conducted, testing the performance of the setups for near-field gamma imaging. Additionally, Monte Carlo simulations of a larger setup of the same type were conducted, investigating its performance with a realistic gamma source distribution occurring during proton therapy. *Main results.* The images of point-like sources reconstructed from two small-scale prototypes' data using the maximum-likelihood expectation maximisation algorithm constitute the experimental proof of principle for the near-field coded-mask imaging modality, both in the 1D and the 2D mode. Their precision allowed us to calibrate out certain systematic offsets appearing due to the limited alignment accuracy of setup elements. The simulation of the full-scale setup yielded a mean distal falloff retrieval precision of 0.72 mm in the studies for beam energy range 89.5–107.9 MeV and with 1×10^8 protons (a typical number for distal spots). The implemented algorithm of image reconstruction is relatively fast—a typical procedure needs several seconds. *Significance.* Coded-mask imaging appears a valid option for proton therapy monitoring. The results of simulations let us conclude that the proposed full-scale setup is competitive with the knife-edge-shaped and the multi-parallel slit cameras investigated by other groups.

1. Introduction

There is a consensus in the proton therapy community that the implementation of methods that enable real-time monitoring of proton therapy would allow to better exploit the potential of this treatment modality and thus offer better and safer therapy to patients (NuPECC 2014). What requires verification is the spatial distribution of the dose delivered during therapy and its compliance with the one resulting from the treatment plan, preferably in a continuous manner during irradiation, on a single-spot (or at most a few-spot) basis. It is also of interest for the development of an online adaptive proton therapy, which is already used for photon

irradiations (Albertini *et al* 2020). Many of the proposed approaches rely on the detection of prompt-gamma (PG) radiation, where the spectral and spatial characteristics are correlated with the beam range (Min *et al* 2006, Verburg and Seco 2014, Pinto *et al* 2015, Kelleter *et al* 2017). An overview of PG-based monitoring methods for proton therapy can be found in (Wrońska and Dauvergne 2021).

Several groups have been developing imaging setups featuring gamma cameras with passive collimation. The most mature projects involve the use of a camera combined with a collimator with a knife-edge-shape slit. Such a setup has been tested in clinical conditions of pencil-beam scanning and passively-scattered beam and proven to provide control of inter-fractional range changes with a precision of about 2 mm (Richter *et al* 2016, Xie *et al* 2017). However, in view of the fact that the number of registered PG quanta is one of the main limiting factors in PG imaging (PGI), multi-slit systems have been investigated too. Not only do they register more gamma quanta than single-slit cameras, but they offer also a larger field of view (FOV). The performance of multi-slit setups was studied extensively via GEANT4 simulations for different geometrical configurations in Pinto *et al* (2014), and experimentally in Park *et al* (2019). However, no superiority with respect to the knife-edge-shaped camera could be shown. An extension of the latter was proposed in Ready *et al* (2016), Ready (2016), where a collimator with many knife-edge-shaped slits was studied. Although the obtained results were quite impressive ($2\sigma = 1$ mm range retrieval precision), the studies were conducted at a beam energy of 50 MeV, i.e. below the lower limit of the clinically applied energy range. Unfortunately, the studies of that group were discontinued. The concept, however, was picked up and extended to a dual-head setup, enabling 3D-imaging (Lu *et al* 2022). Simulation results indicate the feasibility of using the setup for online range monitoring in proton therapy, with a position resolution better than 2 mm across the whole FOV. The group is currently developing a prototype setup to verify their simulation results. Yet a different approach has been presented in Sun *et al* (2020a), where a setup consisting of a pixelated detector and a coded-mask collimator with a modified uniformly redundant array (MURA) pattern (Gottesman and Fenimore 1989) has been studied via simulations. This kind of gamma collimation is widely used in astronomy and proven to work well also in reconstructing the positions of gamma sources (see, e.g. Anderson *et al* 2006, Braga 2019), though one needs to stress that those applications deal with far-field imaging and mainly point-like sources, which is in general a less demanding imaging scenario. A MURA collimator is undoubtedly easier to manufacture than one with multiple knife-edge shaped slits. The setup like the one proposed in Sun *et al* (2020a) offers 2D-imaging with a much larger detection efficiency than the solutions discussed above, since half of the collimator pixels are open. The authors report an accuracy of range determination better than 0.8 mm, but this number holds for 10^{10} impinging protons, which exceeds by 2 orders of magnitude the number typically applied in a single spot.

Coded-mask (CM) imaging is an extension of the well-known pinhole camera concept, widely used in various areas: from photography to space imaging (Young 1989). In a pinhole camera, the detector is fully covered with impenetrable material, except for a small hole that decodes a source while projecting it onto the detector. Although it may provide a good image resolution when there are no constraints on irradiation time, it is not applicable in PGI in clinical conditions, when the fixed number of impinged protons per irradiated spot specified by the prescribed dose to the patient limits the number of emitted gammas. An almost completely opaque collimator further reduces the number of registered gamma quanta, leading to a situation in which the reconstructed image is strongly affected by statistical fluctuations.

In the CM approach, a detector is covered not with a single-hole shield but with a mask consisting of a number of such holes forming a specific pattern. Usually, the number of open pixels is similar to the number of filled ones, so approximately 50% of the mask is opaque. In comparison to the single-slit camera, such a setup registers more photons which allows to increase the detector efficiency. The optimisation of mask patterns is an interesting problem that has been widely explored in the last few decades (Fenimore and Cannon 1978, Gottesman and Fenimore 1989). In this work, we are using a MURA pattern (Gottesman and Fenimore 1989) as it is beneficial in performance metrics such as signal-to-noise ratio and image resolution. A specific MURA mask is characterised by a prime number (called the rank or order of the mask) which defines the number of pixels per dimension and the construction of the pattern of opaque and empty pixels.

In this work, we present the results of our experimental studies with two small-scale prototypes of detection setups exploiting the coded-mask technique, conducted using point-like radioactive sources. The two setups featured different detectors and mask patterns, allowing 1D- and 2D-imaging. The images are reconstructed using the maximum-likelihood expectation maximisation (MLEM) algorithm (Rockmore and macovski 1976). We report the obtained point-spread function for different source positions within FOV and the detection efficiency of the system. Using the experimental results to benchmark the simulation application, we furthermore simulate the performance of a full-scale prototype, currently under development as a scatterer module within the **Si**licon Photomultiplier and Scintillating **Fi**bre-based **C**ompton **C**amera (SiFi-CC) project (Kasper *et al* 2020), including its capability to reconstruct a continuous linear gamma source present during proton therapy.
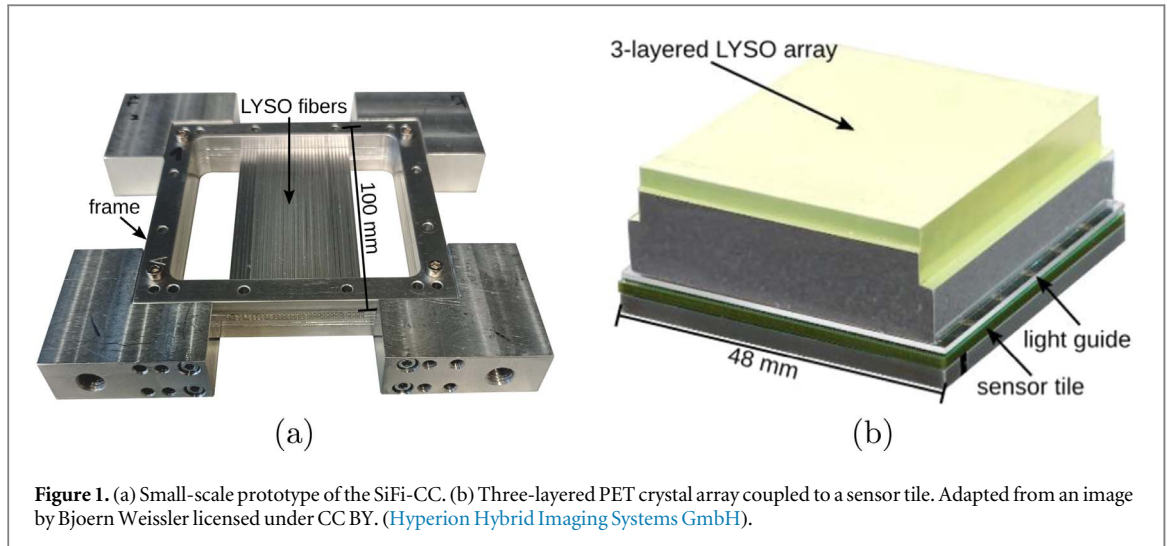
Figure 1. (a) Small-scale prototype of the SiFi-CC. (b) Three-layered PET crystal array coupled to a sensor tile. Adapted from an image by Bjoern Weissler licensed under CC BY. (Hyperion Hybrid Imaging Systems GmbH).

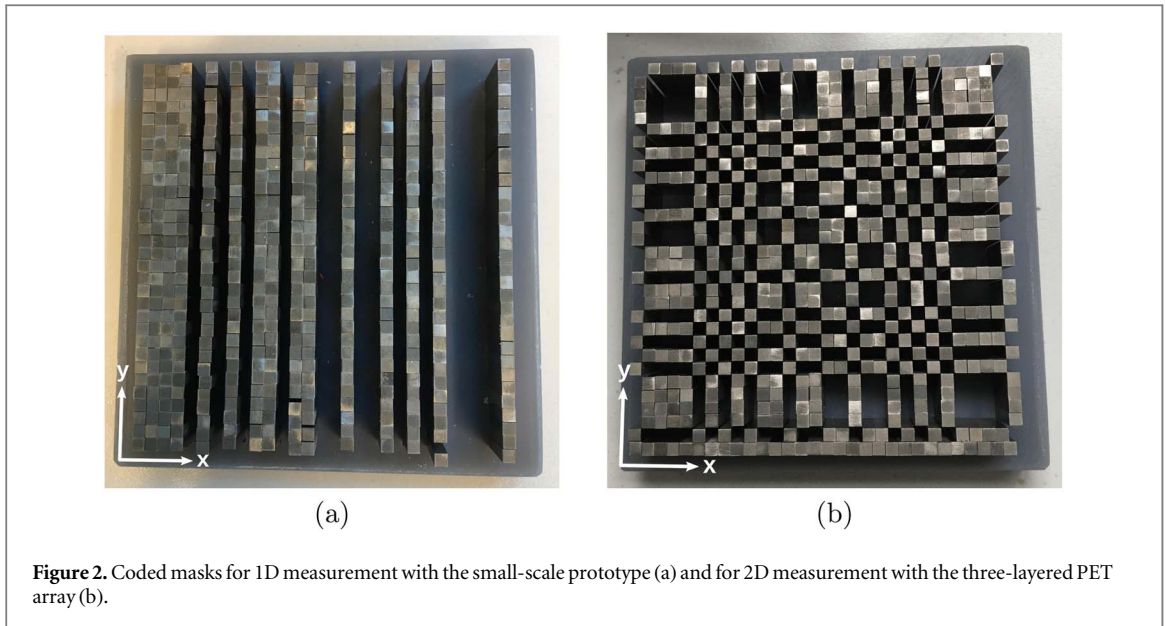## 2. Materials and methods

2.1. Detector components

2.1.1. Scintillators

The measurements are conducted with two different scintillation detectors. In both cases, the Ce:LYSO scintillator was used as a sensitive material because of its large effective atomic number and large density, resulting in high efficiency of gamma detection. Good availability and moderate price were the additional arguments supporting this choice (Rusiecka *et al* 2021). We use the small-scale prototype (SSP) of a SiFi-CC module consisting of 64 Ce:LYSO fibers (see figure 1(a)). The fibers have a squared cross section of 1 mm × 1 mm and are 100 mm long. For this measurement, they are arranged in two layers of 32 fibers each. Every fiber is wrapped individually in aluminum foil and they are held together in an aluminum frame. The pitch between both two fibers and the two layers is 1.36 mm.

The second scintillator used is a three-layered Ce:LYSO array developed for positron emission tomography (PET) imaging simultaneously to magnetic resonance imaging (MRI). The array has a base area of 45 mm × 48 mm, a total height of 15 mm and three layers (see figure 1(b)). Each layer consists of individual needles with a pitch of 1.33 mm resulting in 3425 needles in total. The needles within a layer are optically separated by $BaSO_4$. Each layer is shifted by half a needle pitch with respect to the layer below to enable the identification of the layer and thus the depth-of-interaction in the array by the footprint of the detected light. The height of the individual layers is optimised for uniform absorption of 511 keV-photons.

2.1.2. Readout platform

As a readout system we use the Hyperion III platform (Hyperion Hybrid Imaging Systems GmbH, Weissler *et al* 2015, 2022). It is developed by Hyperion Hybrid Imaging Systems as MRI-compatible detector platform for PET systems and includes hardware, firmware and software for data acquisition, processing and analysis. We use sensor tiles equipped with digital silicon photomultipliers by Philips Digital Photon Counting (PDPC). The dimensions of one sensor tile are (48×48) mm^2 and the whole tile holds 6×6 DPC-3200-22 (Frach *et al* 2009, 2010). Each digital photon counter (DPC) has four readout channels for the detected number of photons and delivers one common timestamp. Each of the four readout channels contains 3200 single-photon avalanche diodes (SPADs). The DPCs are self-triggering if one of the four readout channels passes a given trigger and validation threshold based on sub-regions on the sensor. During these measurements, a DPC is triggered if on average 3.0 ± 1.4 photons are registered in one channel and the average validation threshold for an event to be recorded to disk is 53 ± 1.5 photons (Tabacchini *et al* 2014). We use a validation time to accept a trigger of 40 ns and an integration time of 325 ns to collect photons on the sensor tile. The overvoltage of the silicon photomultipliers (SiPMs) is set to 3 V. As this is a digital tile it is possible to disable the SPADs which produce a high number of dark counts. This inhibit fraction is set to 10 %. The surface of each sensor tile is covered with a glass plate of 1.1 mm. The sensor tiles are connected to a singles processing unit which manages their voltage supply and feeds their data to the data acquisition and processing server. During the measurement, the tiles are cooled by a 15 °C liquid cooling system.

Figure 2. Coded masks for 1D measurement with the small-scale prototype (a) and for 2D measurement with the three-layered PET array (b).

2.1.3. Masks

We perform measurements in one and two dimensions, i.e. we reconstruct an image along one axis or on a plane. For these two tasks, we use a one- and a two-dimensional versions of a MURA mask of rank 476, clipped to 31×31 central pixels (see figure 2). The mask rank as well as the setup geometry have been optimised via Monte Carlo simulations before the experiment, which is explained further in section 2.3.1. To construct the physical masks we use tungsten rods of $(2.26 \times 2.26 \times 20)$ mm^3 which are inserted into 3D printed rasters made from Pro Grey Resin. The rod manufacturing reaches a precision of 0.1 mm. The resulting masks have a dimension of (73.6×73.6) mm^2. The rasters have a total thickness of 13 mm and the holes to insert the rods are 10 mm deep. To prevent the rods from falling out, the assembled masks are wrapped in cling film.

2.2. Radioactive sources

For image reconstruction, the experimental data were obtained with a radioactive ^{22}Na source with an activity of 2.89 MBq . The active material in that source covers an area of 1 mm $\times$ 1 mm. As a β^+-emitter, ^{22}Na provides two photons of 511 keV emitted back-to-back, which can be used for electronic collimation. For calibration of the detectors we additionally used the 1275 keV gamma line of ^{22}Na and two more radioactive sources: a ^{137}Cs source with a gamma line at 662 keV with an activity of 1.73 MBq and a ^{133}Ba source with a prominent line at 356 keV with an activity of 1.52 MBq. During the measurement the sources were placed in front of the detector on a grid allowing for easy repositioning of the source for different measurement configurations in 1 cm-steps in vertical and horizontal direction.

2.3. Experiment: setup 1D

2.3.1. Experimental setup for imaging

In the 1D setup of our experiment, we aimed for the reconstruction of a source position along one axis, so only in one dimension. We used the small-scale prototype as scintillation detector and the 1D coded mask (see figure 3(a)). In the current experiment, the distance between the source plane and the front part of the detector is 236.5 mm while the distance from the centre of the mask to the front part of the detector is 66.5 mm. Just like the mask patterns, the distances were optimised via simulations. For this purpose, for a number of positions of a point-like source across the whole FOV, the setup response was simulated using Geant4 and the corresponding images were reconstructed for a fixed number of MLEM iterations (200). This was repeated for many setup geometries and mask patterns. The varied parameters were the mask pixel size and the mask-detector distance. The distance between the source plane and the front surface of the mask was fixed at 160 mm. The metrics used in the comparison of different setup geometries were the mean resolution of the reconstructed position, as well as the mean residual of this quantity and its variation over the FOV.

The orientation of the bars forming the mask pattern is vertical, the same as of the fibers. Both ends of the fibers are coupled to one silicon photomultiplier (SiPM) tile each with an optical silicon pad of 0.4 mm thickness made of Elastosil RT 604. The pad has a size of (8×48) mm^2 so it covers one row of DPCs and light-sharing between different readout channels is enabled. The other five DPC rows of the tile are not directly exposed to light from the fibers.
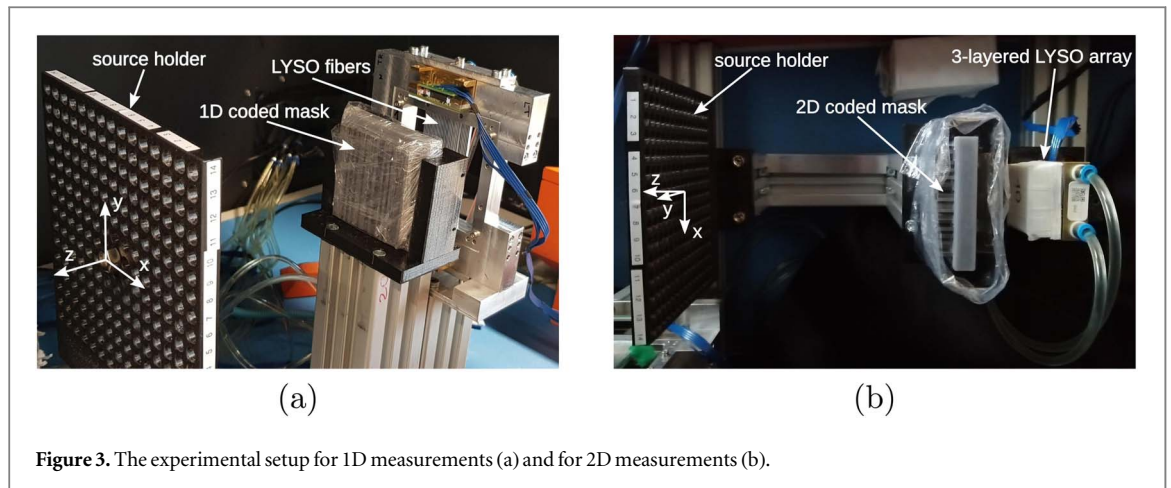
Figure 3. The experimental setup for 1D measurements (a) and for 2D measurements (b).

2.3.2. Setup for energy and y-position calibration

As the light yield on the SiPMs heavily depends on the position of the interaction along the fiber, a position-dependent energy calibration is needed. For this, we use a fan beam collimator which is described in detail in (Müller *et al* 2018, 2019, Hetzel *et al* 2020). The collimator consists of lead blocks with adjustable slits to two sides so that particles emitted by a radioactive source placed in the centre leave the collimator in thin elongated beams. To calibrate the small scale prototype with this setup, we use a ^{22}Na source and employ the three-layered PET crystal as coincidence detector on the other side of the collimator to enable electronic collimation. The used slit width of 1.9 mm leads to a beam width on the fibers of 2.5 mm (FWHM). The coincidence window is set to 10 ns. The fibers of the small scale prototype are irradiated at nine different positions in 10 mm steps.

2.4. Experiment: setup 2D

2.4.1. Experimental setup for imaging

In the two-dimensional setup we used the three-layered PET crystal as a detector with the 2D coded mask (see figure 2(b)) placed in front of it. The distance between the source plane and the front plane of the detector was 220 mm and the distance from the centre of the mask to the front plane of the detector was 50 mm (see figure 3(b)). The scintillator was coupled to the sensor tile over a 1.1 mm thick light guide glued to it with SCIONIX RTV 481.

2.4.2. Setup for energy calibration

For calibration of the detector in this configuration, we use frontal irradiations of the scintillator with radioactive ^{22}Na, ^{137}Cs and ^{133}Ba sources.

2.5. Analysis chain for experimental data

2.5.1. Preprocessing

As all DPCs are triggered independently, a clustering algorithm is needed to form events. We use a cluster window of 40 ns to combine triggered DPCs on one tile. These combined signals, we refer to as one hit in the following. The first timestamp recorded in one of the DPCs is used as the timestamp of the hit. If we use several tiles in one measurement, we apply a coincidence window of 10 ns to combine hits from the different tiles. This we call an event.

2.5.2. Fiber identification in 1D setup

With its base area of (1×1) mm^2, one fiber is smaller than the area covered by one readout channel which is approximately (4×4) mm^2. To identify single fibers in which impinging gammas deposited energy, we used the light-sharing through the optical pad. Depending on the position of the fiber, one or two horizontally neighbouring DPCs are triggered, so there are four or eight signals which can be used to calculate a centre of gravity (CoG). Plotting all CoG in one histogram yields a so-called floodmap. When separating events by the number of triggered DPCs, one can identify light accumulation points on the floodmaps corresponding to interactions in individual fibers. These peaks are determined in projections of the floodmaps. To tag the hit fibers, the floodmaps are first divided into layers by horizontal lines and afterwards the fibers are separated within the layers by the centreline between two peaks so that each fiber corresponds to a rectangular region on the floodmap. To consider an event valid, we demand that the same hit fiber is identified on the two sensor

boards. This is the case for 84.8 % of all recorded events. More details on the fiber identification are presented in Rusiecka (2023).

2.5.3. Energy and y-position calibration in 1D setup

The calibration of the energy and the interaction position within a fiber is performed individually for each fiber. For this purpose, we used the nine measurements with a ^{22}Na source obtained with the fan beam collimator and employed the ELAR model described in detail in Rusiecka *et al* (2021). This provides a position-dependent energy calibration based on the measurement of the 511 keV-peak of ^{22}Na on both ends of the fibers. The description of the calibration of this data set can be found in Rusiecka (2023).

2.5.4. Needle identification in 2D setup

To identify the hit needle in the three-layered PET array, again floodmaps of the CoG positions are used. Due to the shift of the three layers with respect to each other, all needles yield different light accumulation points on the floodmap and are thus distinguishable. Here, separate floodmaps for four different regions of interest (ROIs) dependent on the main channel, which is the channel with the highest photon count in this event, are employed: (1) the main DPC containing the main channel, (2) the main DPC and the vertically neighbouring DPC to the main channel triggered, (3) the main DPC and the horizontally neighbouring DPC to the main channel triggered, (4) the main DPC, the vertically neighbouring DPC, the horizontally neighboring DPC and the diagonally neighbouring DPC triggered. If more than one region of interest (ROI) is valid for one event, all are evaluated. For each of the four floodmaps the light accumulation points are identified on a two-dimensional grid and assigned to a needle ID. A predecessor of this algorithm is described in Schug *et al* (2015). During the needle identification process, the event is assigned to the needle ID of the closest light accumulation point on the floodmap. If different ROIs yield different needle IDs, for each ROI a quality factor $QF = d_1/(d_1 + d_2)$ is calculated, where d_1 is the distance to the closest light accumulation point and d_2 the distance to the second closest light accumulation point on the respective floodmap. The final needle ID is then taken from the ROI with the smallest quality factor. The three top rows and three bottom rows of needles, i.e. with the highest and lowest y-coordinates, could not be resolved with this approach and were therefore not taken into account in the further steps of the analysis. An in-depth description of the procedure can be found in Müller (2022).

2.5.5. Energy calibration in 2D setup

The energy calibration is performed separately for each needle and for each of the four ROIs explained in section 2.5.4. For each of these cases, we take the energy spectra of the calibration measurements with radioactive sources and fit the 511 keV peak in the ^{22}Na spectrum, the 662 keV peak in the ^{137}Cs spectrum and the 356 keV peak in the ^{133}Ba spectrum with a Gaussian function plus a linear function as background approximation. The three peak positions were used to obtain a first linear calibration. After the needles responses were individually calibrated, the energy spectra for all needles were added up. Then the statistics were sufficient to also take the 1274.5 keV peak in the ^{22}Na spectrum into account. Its position was used to apply a global saturation correction to all data:

$$E(Q) = p_2 \cdot \exp\left(-\frac{Q}{p_1}\right) \cdot \left[\exp\left(-\frac{Q}{p_1}\right) - \exp\left(-\frac{p_0}{p_1}\right)\right], \tag{1}$$

where Q represents the pre-calibrated energy and fitted p_i are the parameters of the saturation correction. All the intermediate steps are reported in Müller (2022).

2.6. Implementation of MLEM for image reconstruction

In order to reconstruct the image decoded with CM, we are using the MLEM algorithm (Rockmore and Macovski 1976) while different approaches (like FISTA (Sun *et al* 2020a) or OSEM (Mu *et al* 2016)) are also possible. MLEM utilises prior information about the probability of a photon being emitted in a particular position of the source plane to be registered in each detector pixel.

MLEM is a widely used iterative algorithm, examples of its application in PET can be found e.g. in Shepp and Vardi (1982), Lange and Carson (1984). It serves for the reconstruction of Poissonian data:

$$\mathbf{f}^{[k]} = \frac{\mathbf{f}^{[k-1]}}{\mathbf{S}}\left(\mathbf{H}^T \cdot \frac{\mathbf{l}}{\mathbf{H} \cdot \mathbf{f}^{[k-1]}}\right), \quad \text{for } k = 1, 2, \ldots \tag{2}$$

where $\mathbf{l}$ is a vector of measured data, $\mathbf{f}^{[k]}$ is the image estimate after k-th iteration ($\mathbf{f}^{[0]} = \mathbb{1}$), $\mathbf{H}$ is a system matrix and $\mathbf{S}$ a normalisation term (or sensitivity map). Equation (2) is written in vector form and all multiplications indicated by dots represent the vector (matrix) multiplications, while all other operations are element-wise.

An element of a system matrix $\mathbf{H}_{ij}$ is a probability that a particle originated from the j-th voxel of the source plane will be detected in a i-th voxel of the detector:

$$\mathbf{H}_{ij} = p(V_i|O_j), \tag{3}$$

where V_i is an event that particle was detected in the detector voxel i, O_j is an event in which a particle originated from the source plane voxel j.

In order to obtain such a set of probabilities, we used a Monte Carlo simulation. The chosen FOV is divided into pixels and simulations with a single point-like source placed in the centre of each pixel were performed. The detector response from each of these simulations gives a column in a system matrix, since the number of registered particles in each detector pixel is proportional to the probability from equation (3) (given that the number of simulated events in each vertex j is the same). The detailed explanation of the system matrix generation procedure is in section 2.8.1.

After normalising each column of a system matrix by the sum of its elements, the elements $\mathbf{H}_{ij}$ can be interpreted as a conditional probability to register the particle in the detector voxel i given that the particle originated from the voxel j and was detected somewhere.

Note that both the image $\mathbf{l}$ and the reconstructed object $\mathbf{f}$ are one-dimensional vectors, even when we want to reconstruct 2D (or even 3D) objects. It allows us to utilise one- and two-dimensional setups with the same reconstruction algorithm.

2.7. Efficiency and background corrections

Having a calibrated detector response, before proceeding with image reconstruction, we perform a data correction step. In order to do that, we make use of two auxiliary elements: a background measurement $\mathbf{l}_{bg}$ performed without a radioactive source and without a mask, and a detection efficiency map. For the latter, we register the detector response $\mathbf{l}_{no-mask}$ when irradiating the detector without a mask, and with a source placed far enough to consider the gamma flux uniform over the surface of the detector. Having those data we can construct an efficiency map which will reflect the relative detection efficiency of individual detector elements (fibers or needles) ϵ:

$$\epsilon = \frac{\mathbf{l}_{no-mask} - \mathbf{l}_{bg}}{\max\{\mathbf{l}_{no-mask}^{(i)} - \mathbf{l}_{bg}^{(i)}\}}, \tag{4}$$

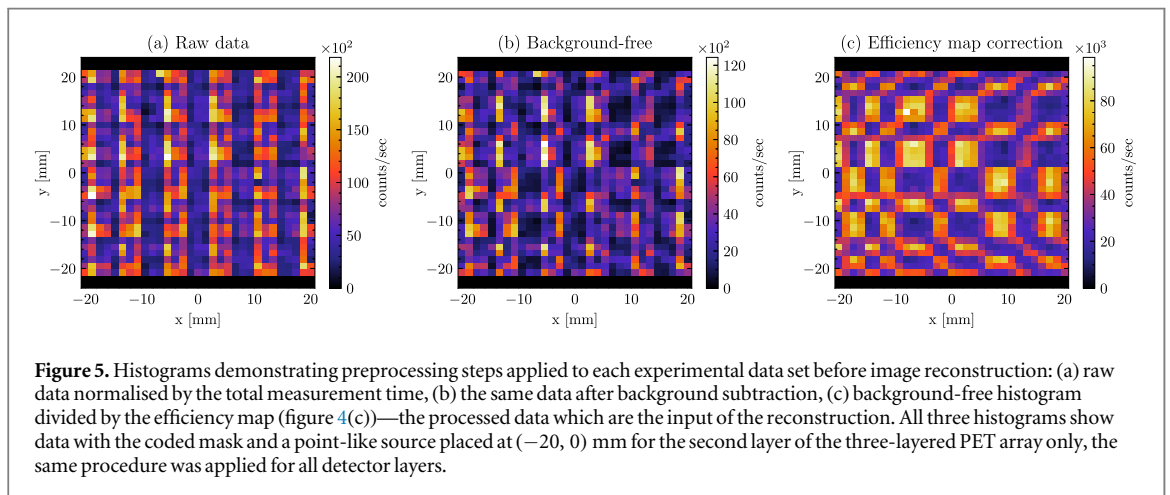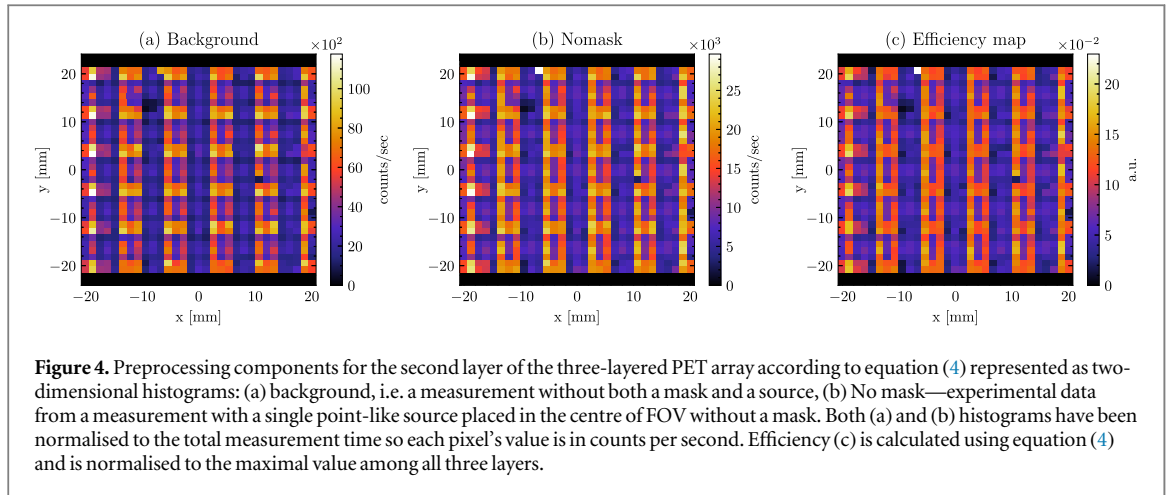where the index i runs over all detector elements.

This step is performed once for each detection setup, the resulting efficiency map and $\mathbf{l}_{bg}$ are used to correct each registered data set that serves as input for image reconstruction. In the correction, we subtract the background from each raw data vector $\mathbf{l}_{raw}$ and divide it by the efficiency map:

$$\mathbf{l} = \frac{\mathbf{l}_{raw} - \mathbf{l}_{bg}}{\epsilon}. \tag{5}$$

The resulting vector $\mathbf{l}$ is ready to be used with the MLEM algorithm according to equation (2) together with the system matrix prepared beforehand (see section 2.8.1).

The preparation and application of the efficiency correction we demonstrate in figure 4 taking the second layer of the three-layered PET array as an example. The background data $\mathbf{l}_{bg}$ is presented in histogram figure 4(a), where each bin value is a number of counts per second in the corresponding crystal. Figure 4(b) presents results of measurement with a single point-like source but without a mask ($\mathbf{l}_{no-mask}$ in equation (4)) with the same units as the $\mathbf{l}_{bg}$ histogram. By evaluating equation (4) for these two histograms, we obtain an efficiency map presented in figure 4(c). One can notice that those three histograms have some similarities in their patterns, representing the DPC structure of the sensor tile. With other algorithms, a more homogeneous efficiency distribution can be achieved. In our approach, this is cancelled out when the image is corrected with the efficiency map and so does not compromise our image reconstruction. In order to avoid singularities resulting from a division by zero, we add a small value (10^{-6}) to all elements of the efficiency map.

In figure 5 we demonstrate all intermediate steps from equation (5), applied to the same second layer of the three-layered PET array: the raw detector response $\mathbf{l}_{raw}$ in panel (a) is followed by the background-free histogram $\mathbf{l}_{raw}$–$\mathbf{l}_{bg}$ and finally the background-subtracted and efficiency-corrected detector response $\mathbf{l}$ is shown in panel (c) which serves as an input for the MLEM algorithm. This data is taken in one of the full-fledged measurements, with a coded mask and a point-like source placed at $(-20, 0)$ mm. Figure 5 shows how this procedure transforms the seemingly messy raw detector response into the histogram with a recognisable mask pattern (see figure 2(b)) projected onto it.
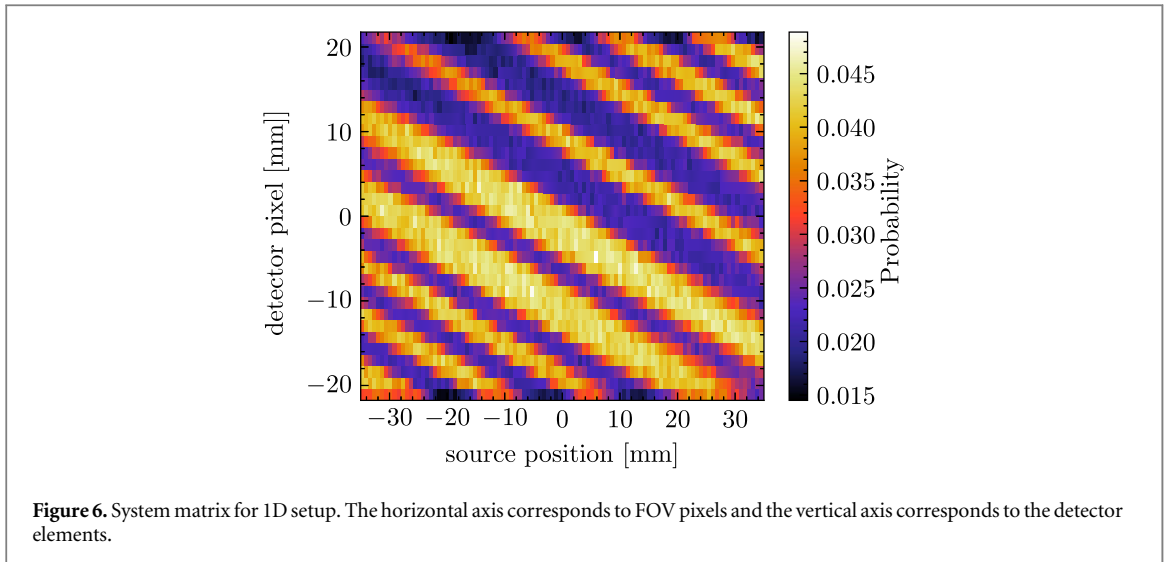
Figure 4. Preprocessing components for the second layer of the three-layered PET array according to equation (4) represented as two-dimensional histograms: (a) background, i.e. a measurement without both a mask and a source, (b) No mask—experimental data from a measurement with a single point-like source placed in the centre of FOV without a mask. Both (a) and (b) histograms have been normalised to the total measurement time so each pixel's value is in counts per second. Efficiency (c) is calculated using equation (4) and is normalised to the maximal value among all three layers.

Figure 5. Histograms demonstrating preprocessing steps applied to each experimental data set before image reconstruction: (a) raw data normalised by the total measurement time, (b) the same data after background subtraction, (c) background-free histogram divided by the efficiency map (figure 4(c))—the processed data which are the input of the reconstruction. All three histograms show data with the coded mask and a point-like source placed at $(-20, 0)$ mm for the second layer of the three-layered PET array only, the same procedure was applied for all detector layers.

2.8. Simulations

2.8.1. Generation of system matrix

MLEM requires a system matrix (SM) which, in this work, is calculated prior to the reconstruction. For both 1D and 2D setups we utilise Monte Carlo simulations in order to obtain corresponding probabilities which are the elements of the system matrix. The FOV was chosen to be of the same size as the mask and it is 70 mm (in each dimension for 2D) and is divided into 100 pixels (in each dimension). By performing simulations of 10^6 γ-particles with a point source placed subsequently in the centre of each pixel of the FOV we obtain the system matrix elements column by column. Such simulated statistics result in the statistical uncertainty of matrix elements being below 1.5 %. All simulated particles have the same energy 4.4 MeV as our investigations revealed that our approach to SM calculation is not sensitive to energy. During the simulation, each element of the system matrix is incremented by the energy deposit obtained from the corresponding detector bin. As a result, we are working with accumulated energy values that are then transformed into probabilities through a normalisation process.

A system matrix for a one-dimensional coded-mask setup (described in section 2.3) has a size of 32×100 elements and is shown in figure 6. We see that its elements reflect the pattern (or its central part) of the coded mask (see figure 2(a)). With shifting the source position (horizontal axis in figure 6), the pattern moves vertically in the histogram, which corresponds to the translation of the mask shadow along the detector plane.

For our simulations, we use GEANT4 V10.6 with PhysicsList QGSP_BIC_HP and electromagnetic option 4. The generation of each matrix column (one simulation) takes around 25 s on INTEL(R) CORE(TM) I7-9700 CPU @ 3.00GHZ and utilizing 6 CPU cores in parallel, which means a full matrix for the 1D setup (100 simulations) can be generated within 10 min. In the case of a 2D system matrix with similar conditions, the number of auxiliary simulations will be 10000 (100 in each direction). Consequently, the total time will increase proportionally to about 16 h.

Figure 6. System matrix for 1D setup. The horizontal axis corresponds to FOV pixels and the vertical axis corresponds to the detector elements.

2.8.2. Setup for 1D full-scale prototype

In this paper, we show experimental results for point-like sources only. Those laboratory experiments allowed us to estimate the performance of the prototype detectors (energy- and spatial resolutions, capabilities, limitations, etc) and conclude about the near-field imaging capabilities of the setups based on them. However, the ultimate goal of the whole project is to come up with an approach applicable in proton therapy. This requires imaging of continuous source distributions, with a particular focus on the distal part of the gamma production depth profile with its falloff occurring close to the Bragg peak position (Min *et al* 2006). This scenario is clearly more demanding than the case of point-like sources. In fact, the small-scale prototypes presented in this paper (figure 3(a)) are not suitable for that task; their small surface areas limit the spatial information that is used as input for image reconstruction (i.e. limited views, incomplete sampling). However, this does not necessarily imply that the technology is not well-suited for range monitoring. Encouraged by the results presented in this study, we started the construction of a larger detector within the SiFi-CC project, with its design being very similar to the small-scale prototype used in the 1D setup. It consists of 7 layers with 55 fibers each; the fiber size is $(1.94 \times 1.94 \times 100)$ mm^3, and the pitch is 2.01 mm. We tested via Monte Carlo simulations configured similarly as in section 2.8.1 how the detector—called henceforth the full-scale prototype—performs in a coded mask setup, i.e. combined with a structured collimator. In GEANT4, we coupled that detector with a larger section of the 467-th order MURA mask, with the same pixel size as the one presented in figure 2(a), but dimensions extended to 51 central bins from the full array (instead of 31) in the horizontal direction and 45 pixels in vertical. The detector-to-mask and mask-to-source distances remained the same as those used for the 1D setup of the small-scale prototype.

In the Monte Carlo simulations, a poly(methyl methacrylate) (PMMA) phantom of the dimensions $(60 \times 60 \times 90)$ mm^3 and density 1.19 g cm^{-3} was irradiated with a proton beam. The phantom was centred with respect to our FOV and the beam impinged from the positive direction of the *x*-axis, along the longest axis of the phantom. A schematic of the setup geometry is presented in figure 7. In the simulations, we recorded energy deposits in individual fibers. The processes of production and detection of scintillation photons are not simulated, but their effect is taken into account by smearing the obtained energy deposits with a resolution function, found via more realistic simulations and benchmarked with laboratory tests with single fibers (Kasper *et al* 2020, Rusiecka *et al* 2021):

$$\frac{\sigma_E(E)}{E} = a + b \cdot E^{-1/2} + c \cdot E^{-3/2}, \tag{6}$$

where $a = 0.0322$, $b = 0.6730$ and $c = -0.0013$.

The simulations were conducted for the following beam energies: 85.9, 90.7, 95.1 and 107.9 MeV, corresponding to the Bragg peak depth in the phantom of 5.0, 5.5, 6.0 and 7.5 cm, respectively, where the latter were determined using SRIM (Ziegler *et al* 2010). For each beam energy, a set of 1000 simulations, each with 10^7 protons, was performed. It is worth noting, that in proton therapy the typical spot strengths range from a few times 10^7 for proximal spots up to about 2×10^8 for distal ones (Pausch *et al* 2020). Having separate simulations with 10^7 protons, we were able to investigate the effect of statistics in image reconstruction, the resulting resolutions and uncertainties, and conclude about the feasibility of beam range shift detection on the basis of data from a single spot or an iso-energy layer. For each file simulated for a beam with a kinetic energy of $T_p = 85.9$ MeV, we have around 840 000 PGs generated, out of which our detector registers about 15 000
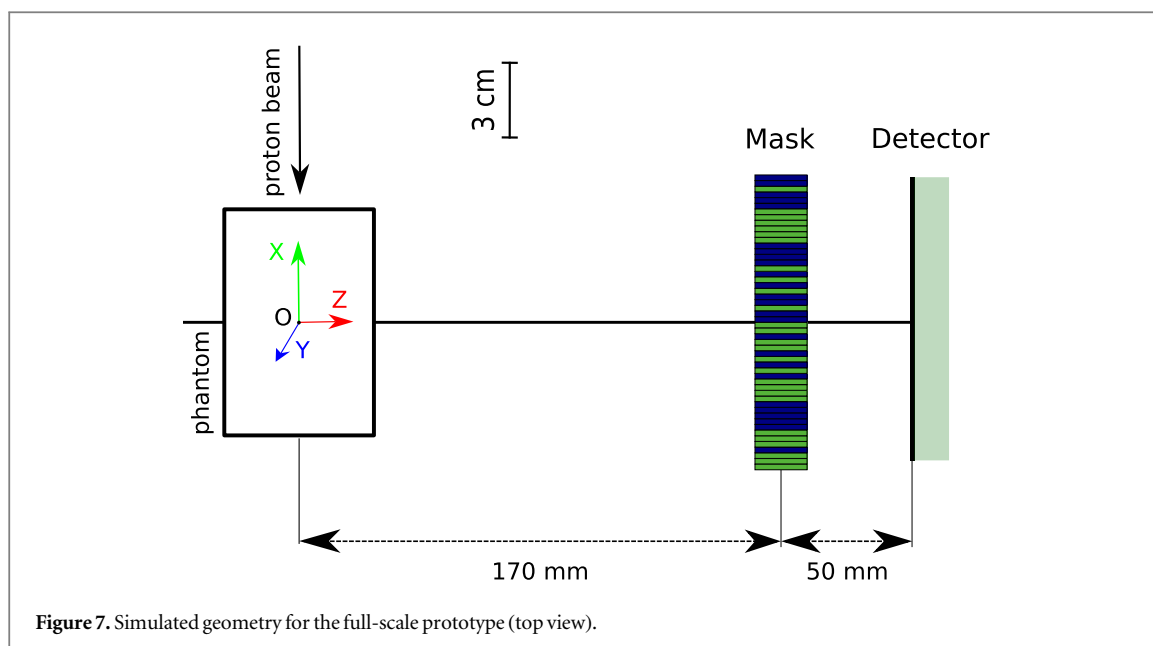
Figure 7. Simulated geometry for the full-scale prototype (top view).

gammas which is about 1.8 %. The overall detection efficiency (including geometrical acceptance) is thus 1.5×10^{-3}. Those numbers remain similar for other studied beam energies.

A system matrix for this setup was generated assuming one-dimensional reconstruction. A FOV of 130 mm along the *x*-direction was assumed and divided into 200 pixels. In this case, unlike in the studies with radioactive sources, the gamma source is not monoenergetic. Nevertheless, when generating the system matrix, we shot 10^6 4.4 MeV gamma particles from each FOV bin, since our earlier studies showed that the resulting system matrix is not very strongly energy-dependent. To test that, we in a preliminary work we investigated if the reconstruction of a point source is influenced by the energy used to generate the system matrix. Namely, we generated one-dimensional system matrices with various energies ranging from 0.5 to 5 MeV, in increments of 0.5 MeV. We then reconstructed the same simulation data obtained from a point-like source with a sample energy value of 4.4 MeV and a position $(-10, 0)$ mm. The results showed that the reconstructed source position varied by only 0.03 mm, and the width of the fitted Gaussian curve exhibited a variation of just 3 %. These findings indicate that the reconstruction process is nearly independent of the energy used for the system matrix generation. The observed differences can be attributed to statistical fluctuations.

We evaluated the image reconstruction performance by inspecting the uncertainty on the distal fall-off position determination (DFPD), which is calculated in a similar way as defined in Gueth *et al* (2013) and the procedure is demonstrated in figure 8. Namely, we take the reconstructed profile (blue dots in the figure), and interpolate it with a cubic spline, obtaining a smooth function (green line). Subsequently, we subtract from the function its minimum value and then normalise it by the resulting function maximum. Like this, we obtain a normalised reconstructed profile described by a smooth function of the values between 0 and 1. We find all the points where this function equals 0.5 (red cross and orange stars) and take the left-most one (red cross) as our estimation of the distal falloff position. The same is being done with the MC source distribution so that we can compare these two values.

2.9. Performance evaluation of MLEM

In order to examine the performance of MLEM, we have prepared sample reconstruction results based on experimental data from the 2D coded-mask setup with the three-layered PET detector. The very same data as shown in figure 4(c) (but for all three detector layers) were loaded into a single vector and used as an input for the image reconstruction. Having a system matrix of size 10 000 × 3425 and an image vector with size 3425, 100 iterations take 6.51 sec on a INTEL(R) CORE(TM) I7-9700 CPU @ 3.00GHZ to be performed.

Results after 100 iterations are shown in figure 9 as a 2D histogram (a) and its projections on the *x*- and *y*-axes (panels (b) and (c), respectively). In the 2D histogram, we observe a clear image without artefacts that exhibits a single peak. The peak is close to the designed source position (marked as a green cross) with a small shift to the upper left. Numerical evaluation of the reconstruction quality was performed via a Gaussian fit—its results are visible in figures 9(b) and (c) as a green solid line. The peak width in the reconstructed image (here 1.7 and 1.5 mm for *x*- and *y*-direction) depends on the number of MLEM iterations and will further decrease when using a higher number of iterations.
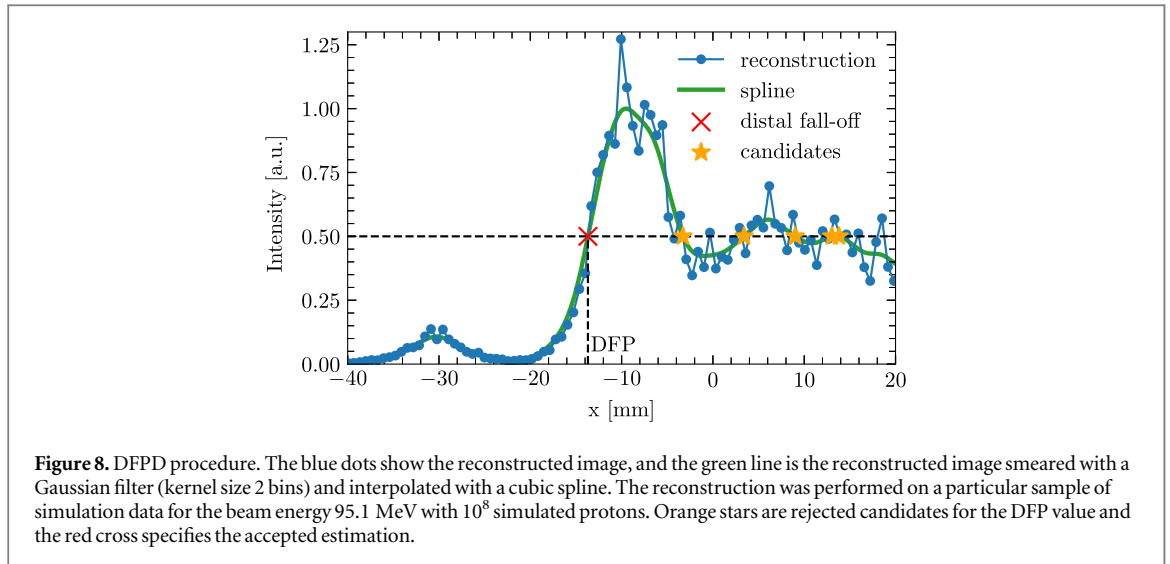
Figure 8. DFPD procedure. The blue dots show the reconstructed image, and the green line is the reconstructed image smeared with a Gaussian filter (kernel size 2 bins) and interpolated with a cubic spline. The reconstruction was performed on a particular sample of simulation data for the beam energy 95.1 MeV with 10^8 simulated protons. Orange stars are rejected candidates for the DFP value and the red cross specifies the accepted estimation.
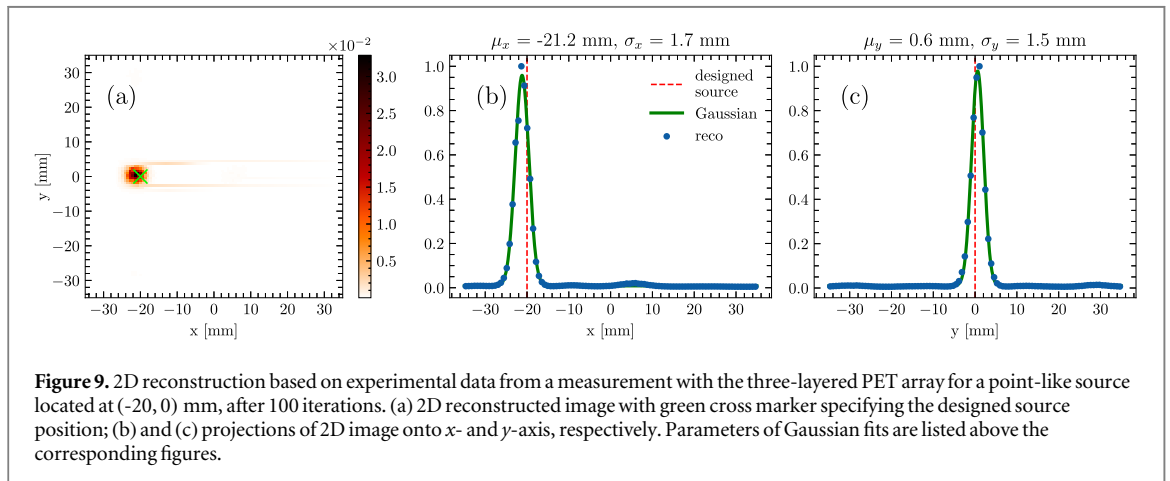

Figure 9. 2D reconstruction based on experimental data from a measurement with the three-layered PET array for a point-like source located at (-20, 0) mm, after 100 iterations. (a) 2D reconstructed image with green cross marker specifying the designed source position; (b) and (c) projections of 2D image onto *x*- and *y*-axis, respectively. Parameters of Gaussian fits are listed above the corresponding figures.
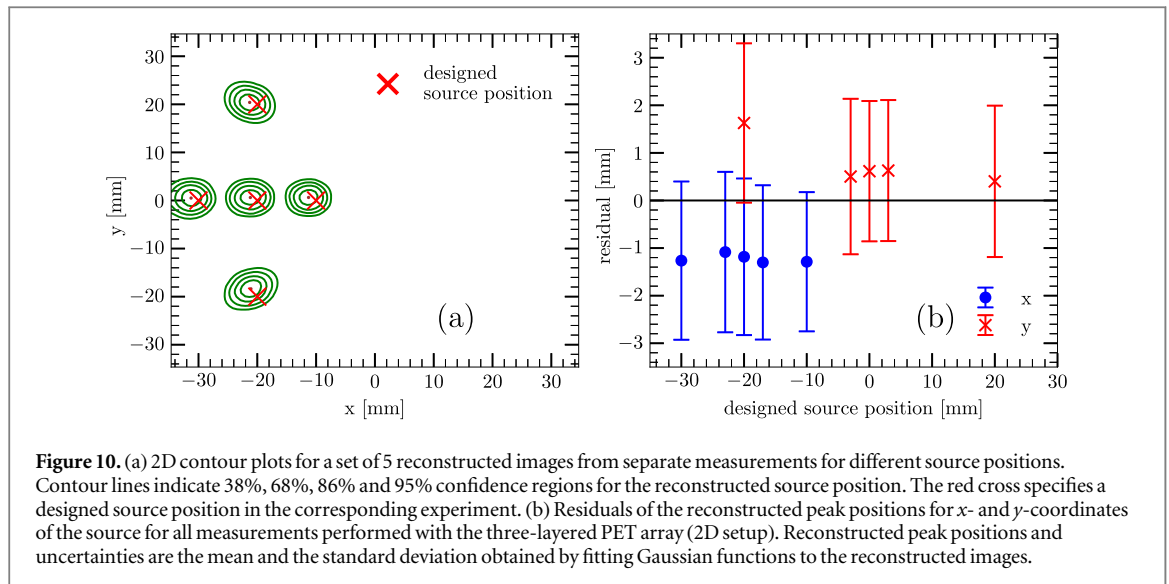
Evaluation of the reconstruction performance is straightforward in the case of point-like sources: comparing the reconstructed peak position with the designed one and, using Gaussian fitting, determining the peak width and position. In the case of continuous source distributions, we are using a universal image quality index (UQI) (Wang and Bovik 2002) that allows comparing two vectors **X** and **Y**, representing in our case the reconstructed image and the expected source distribution. Its application in the near-field coded-aperture imaging was demonstrated in Sun *et al* (2020b). The index is defined as follows:

$$\text{UQI} = \frac{2\text{cov}(\mathbf{X}, \mathbf{Y})}{\text{var}(\mathbf{X}) + \text{var}(\mathbf{Y})} \frac{2\bar{\mathbf{X}}\bar{\mathbf{Y}}}{\bar{\mathbf{X}}^2 + \bar{\mathbf{Y}}^2}, \tag{7}$$

where cov and var are covariance and variance functions, and $\bar{\mathbf{X}}$, $\bar{\mathbf{Y}}$ are the means of the vector components. UQI takes values in the range $[-1; 1]$ where 1 corresponds to identical images and -1 to inverted ones. Vectors $\bar{\mathbf{X}}$ and $\bar{\mathbf{Y}}$ are standardised before calculating UQI, namely we apply:

$$\mathbf{X}'_n = \frac{\mathbf{X}_n - \mathbf{X}_{\min}}{\mathbf{X}_{\max} - \mathbf{X}_{\min}} \tag{8}$$

to each vector, where *n* loops over all vector elements. UQI is beneficial in the case of simulations because one knows the true source distribution, unlike in the case of experimental data. When investigating how the UQI depends on the number of performed iterations, we observe first a steep rise up to about 800 iterations; in that range, subsequent iterations significantly improve the image quality. For larger numbers of iterations, a slow decrease is observed due to the amplification of statistical fluctuations inherent to MLEM. The maximum corresponds to the optimal number of iterations at which the reconstructed image is as close to the true one as possible, given UQI is used as a metric. However, the maximum is rather broad, for the presented example the UQI does not drop below 0.95 of its maximum value for the numbers of iterations from the range 361–1628. In this range, the reconstructed picture does not change significantly, thus it is not critical to run MLEM iterations exactly the number of times corresponding to the maximum UQI, but close to it. This has important

Figure 10. (a) 2D contour plots for a set of 5 reconstructed images from separate measurements for different source positions. Contour lines indicate 38%, 68%, 86% and 95% confidence regions for the reconstructed source position. The red cross specifies a designed source position in the corresponding experiment. (b) Residuals of the reconstructed peak positions for x- and y-coordinates of the source for all measurements performed with the three-layered PET array (2D setup). Reconstructed peak positions and uncertainties are the mean and the standard deviation obtained by fitting Gaussian functions to the reconstructed images.

implications for image reconstruction using experimental data, where we have no reference image - even if the number of iterations deduced based on the simulations will not be identical with the *a priori* unknown optimal number of iterations for experimental data, the latter is likely to be close enough, i.e. the reconstructed image will be close to the best possible reconstruction.

3. Results

3.1. Calibration results

3.1.1. Energy and y-position calibration in 1D setup

A detailed description of the calibration results including the intermediate steps is presented in Rusiecka (2023). The average energy resolution of the fibers is $\sigma_E/E = 7.7(4)\%$ and the average position resolution along the fibers is 34(3) mm (FWHM) at 511 keV.
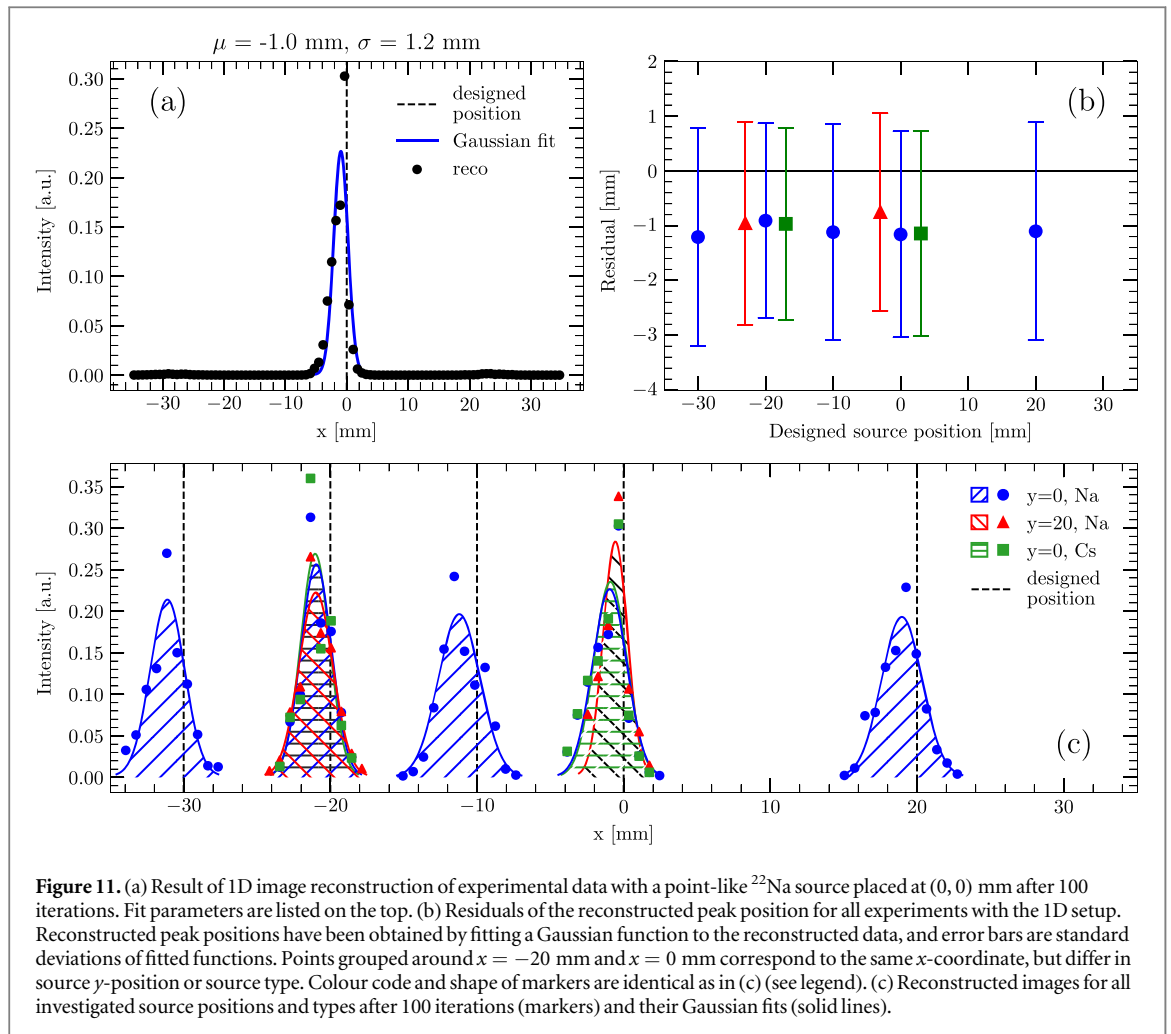
3.1.2. Energy calibration in 2D setup

The intermediate results of the calibration procedure of the 2D setup can be found in (Müller 2022). The overall energy resolution obtained at 511 keV is $\sigma_E/E = 5.35\%$. For 14 additional needles, the calibration process failed because their spectra could not be fitted satisfactorily. Events with hits in those needles were excluded from the further analysis.

3.2. Results from imaging experiments

We performed a series of five measurements with the three-layered PET array, the 2D coded mask and a point-like source placed in different positions. In all cases, the measurement time was 1201.2 ± 0.2 s and the number of registered hits was about 2×10^8. Combined results of all image reconstructions, based on the collected data, after 100 iterations, are presented in figure 10(a) as contour plots. Each contour plot consists of ellipses bounding 38%, 68%, 86% and 95% confidence regions for a particular reconstructed peak. In addition, the green point inside the smallest ellipse (for each reconstruction) points to the reconstructed peak position, while the red cross is the designed source position. Similarly to the results from figure 9(a) (which are also presented here), all reconstructions exhibit offsets of the reconstructed peak positions with respect to the designed source positions and all of them are in the same direction (upper left). For the measurements where the source was off *x*-axis (source coordinates $(-20, -20)$ mm and $(-20, 20)$ mm) we observe that the major axis of both groups of ellipses is rotated and the rotation is symmetrical with respect to the *x*-axis. Nevertheless, the offset of the reconstructed peak position is not symmetrical with reference to the origin but always into the same direction. This hints towards a systematic effect rather than an artefact of the reconstruction.

The latter conclusion is further supported by figure 10(b) where we present the residuals of the reconstructed source positions as a function of the designed source position for *x*- and *y*-axes (blue circles and red crosses, respectively). Error bars represent standard deviations of the fitted Gaussian functions, peak positions were determined in the same fit. For $x = -20$ mm and for $y = 0$ mm we have three measurements so in order to demonstrate all results in one plot, the points have been slightly shifted around the design values. In the figure it is clearly visible that indeed all residuals for the *x* position have the same sign and very similar values; the same

Figure 11. (a) Result of 1D image reconstruction of experimental data with a point-like ^{22}Na source placed at (0, 0) mm after 100 iterations. Fit parameters are listed on the top. (b) Residuals of the reconstructed peak position for all experiments with the 1D setup. Reconstructed peak positions have been obtained by fitting a Gaussian function to the reconstructed data, and error bars are standard deviations of fitted functions. Points grouped around $x = -20$ mm and $x = 0$ mm correspond to the same x-coordinate, but differ in source y-position or source type. Colour code and shape of markers are identical as in (c) (see legend). (c) Reconstructed images for all investigated source positions and types after 100 iterations (markers) and their Gaussian fits (solid lines).

applies to the residuals of the reconstructed y coordinate: the average residual value for x-coordinate is -1.23 (8) mm and for y-coordinate it is 0.73(45) mm (stated uncertainties are standard deviations).

We performed also nine measurements with the one-dimensional setup and the point-like sources. First, data for the ^{22}Na source placed on the x axis in the following positions -30, -20, -10, 0, and 20 mm were taken. Additional two points with the same source off x-axis were examined: $(-20, 20)$ mm and $(0, 20)$ mm. In addition, two experiments with the ^{137}Cs source were performed for $x = -20$ and 0 mm. A sample reconstructed image is presented in figure 11(a) along with a Gaussian fit (solid line) and the designed source position (vertical dashed line). Fit parameters are listed at the top of the figure. In figure 11(b) we demonstrate the residuals for all 1D reconstructions. As previously, the reconstructed source position is taken from the Gaussian fit and the error bars represent standard deviations of the fitted functions. All reconstructed source positions are consistently shifted towards negative values, i.e. the residuals have very similar, negative values with an average of $-1.03(14)$ mm. The observed offset of the reconstructed source position is independent of that position and is constant throughout FOV. The offset is consistent with the 2D measurement within 1σ. We present the whole set of reconstructed images in figure 11(c). The mean standard deviation of the Gaussian fits for all source positions is 1.14(18) mm.

3.3. Results of simulations

In this section we discuss the results of simulations with the full-scale prototype for 1D imaging. Figure 12(a) shows a reconstructed image (blue filled line) of the gamma vertex distribution resulting from the simulation of 10^8 protons of 90.7 MeV interacting with the PMMA phantom. The image was smeared with a Gaussian filter with a kernel of 2 pixels. The shown reconstructed image is a result of 795 MLEM iterations which corresponds to a maximum of UQI. The image is smooth, without significant noise artefacts. There are small peaks in the tail, behind the main one, but each of them is less than half as high as the main peak so their presence does not affect the DFPD. The orange line represents a depth profile constructed from Monte Carlo true information, in which entries are weighted by gamma initial energies, and the profile is subsequently interpolated to create a smooth line. In general, the reconstructed main peak position is very close to the true one and the shape of the gamma
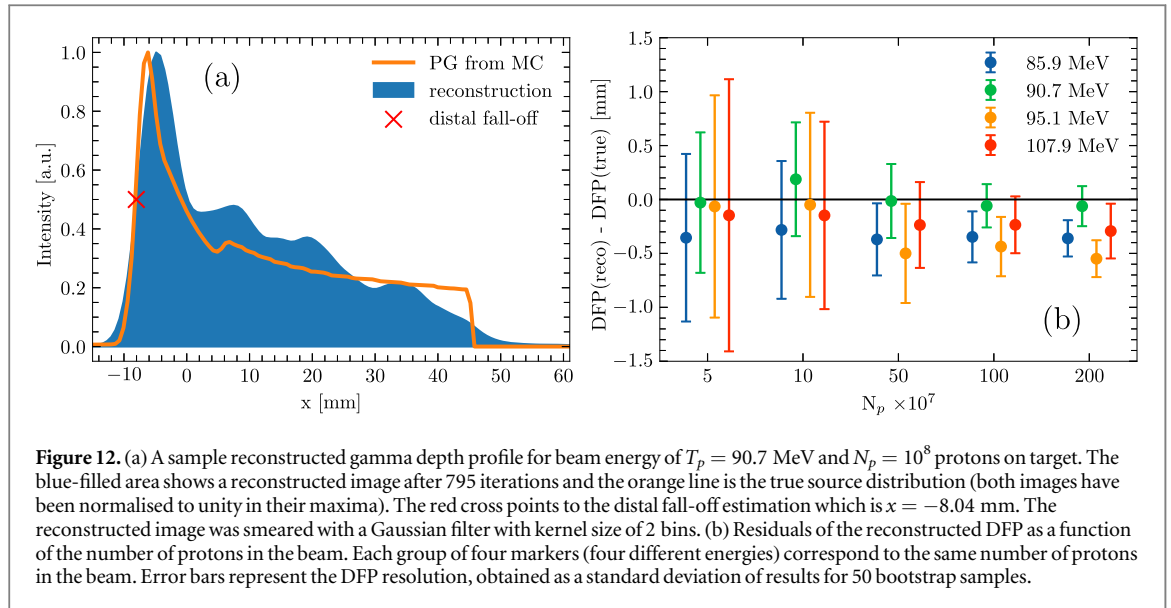
Figure 12. (a) A sample reconstructed gamma depth profile for beam energy of $T_p = 90.7$ MeV and $N_p = 10^8$ protons on target. The blue-filled area shows a reconstructed image after 795 iterations and the orange line is the true source distribution (both images have been normalised to unity in their maxima). The red cross points to the distal fall-off estimation which is $x = -8.04$ mm. The reconstructed image was smeared with a Gaussian filter with kernel size of 2 bins. (b) Residuals of the reconstructed DFP as a function of the number of protons in the beam. Each group of four markers (four different energies) correspond to the same number of protons in the beam. Error bars represent the DFP resolution, obtained as a standard deviation of results for 50 bootstrap samples.

Table 1. DFPs for different energies of the proton beam and different numbers of protons N_p. Values in the table are the means and the standard deviations calculated from 50 bootstrap samples, expressed in mm. 'MC' represents the DFP of the Monte Carlo gamma profile for $N_p = 10^{10}$ protons. The statistical uncertainty for the 'MC' estimation is so small (below 0.04 %) that it can be neglected.

Proton energy	85.9 MeV	90.7 MeV	95.1 MeV	107.9 MeV
N_p		DFP in mm		
1×10^8	-3.64 ± 0.64	-8.02 ± 0.53	-13.56 ± 0.85	-29.03 ± 0.87
5×10^8	-3.73 ± 0.33	-8.22 ± 0.34	-14.01 ± 0.46	-29.12 ± 0.40
1×10^9	-3.70 ± 0.24	-8.26 ± 0.20	-13.94 ± 0.28	-29.12 ± 0.26
2×10^9	-3.72 ± 0.17	-8.27 ± 0.19	-14.06 ± 0.17	-29.18 ± 0.25
MC	-3.36	-8.20	-13.51	-28.89
Bragg peak	-5.00	-10.00	-15.00	-30.00

depth profile is reflected properly. The distal fall-off position (DFP), determined for this particular sample reconstructed image, is -7.64 mm and for the MC truth it is -8.20 mm, which means that the offset is below 0.6 mm.

Offsets of the determined DFP for all regarded energies are shown in figure 12(b) as a function of the initial number of protons impinged on the target. Presented means and resolutions of the DFP were obtained based on 50 bootstrap samples taken out of 1000 samples corresponding to 10^7 protons each. We see that starting from 50×10^7 protons, the DFP mean values are very stable and only error bars are being reduced. The same data are presented numerically in table 1. Starting from the statistics 50×10^7, the DFP resolutions, calculated as a standard deviation of 50 bootstrap samples, are less than 0.6 mm for all beam energies.

4. Discussion

We performed a set of measurements varying the source position for two setups: one- and two-dimensional. In the two-dimensional setup, we observe that after 100 iterations the reconstructed image has an evident single peak without significant noise artefacts (see figure 9). The peak position and its σ-value were calculated via a Gaussian fit to both *x*- and *y*-projections of the image. The obtained reconstructed peak position is slightly shifted with respect to the designed position in both horizontal and vertical directions. Having reconstructed the images for all measurements at different source positions (figure 10(a)), we see that all images have similar offsets in the same direction. The different components of our setup were mechanically aligned with accuracy on the millimetre level. We did not employ a prior calibration to correct for systematic effects that might arise from uncertainties in the alignment. Therefore, systematic shifts can occur when we reconstruct positions. As long as the relative distances of the reconstructed positions match the expected distances, a systematic shift can easily be calibrated out once the setup is finally mounted. The measured offsets in the 2D setup show a small standard
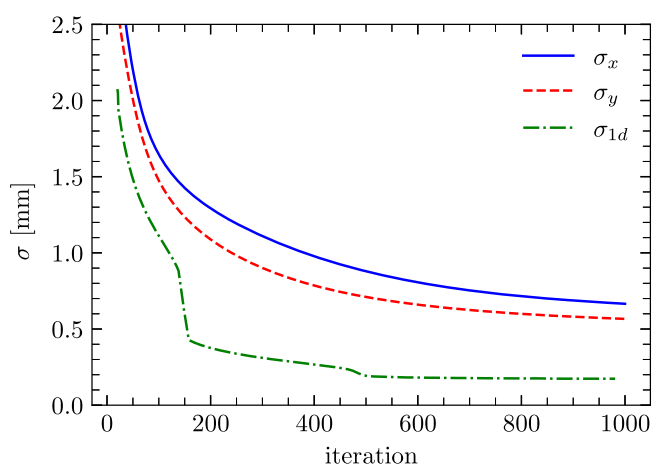
Figure 13. σ-value of the fitted Gaussian as a function of the number of iterations for: x- and y- projections of 2D reconstruction for a source at $(-20, 0)$mm (solid blue and dashed red lines, respectively) and 1D reconstruction for a source at $(0, 0)$mm (dashed–dotted green line).

deviation: $\Delta x = -1.23(8)$ mm and $\Delta y = 0.73(45)$ mm, which hints towards their systematic origin rather than a method artefact. Results from the single-dimensional setup are similar as for the 2D case, except that there is no sensitivity to the source y-position. In this particular case, we have a peak position shifted on average by -1.03 (14) mm with respect to the designed position which is still within one standard deviation ($\sigma = 1.2$ mm) (see figure 11(b)). It is consistent with the shift in the x-direction determined from 2D measurements. This shows that in fact, a set of measurements with radioactive sources like the one performed can be used to calibrate out such offsets resulting from e.g. setup misalignments.

Furthermore, for determining a proton range in a clinical setup, relative shifts are often more relevant than absolute positions. For instance, in Richter *et al* (2016) inter-fractional range variations are investigated and in Ferrero *et al* (2018) the evolution of a PET image over time is analysed to monitor the dose delivery process. Thus, a constant shift does not hinder a precise range shift determination.

While the three-layered PET array coupled with a structured collimator provides good-quality two-dimensional images of a radioactive source, we show that we also achieve a good reconstruction in one dimension with our 1D setup. A 1D gamma depth profile is already sufficient for the determination of a range shift in a clinical scenario and a 1D setup requires less readout channels (and therefore is more cost-effective) for its operation than a full 2D setup, if the system is scaled up to extend its FOV.

All reconstructions performed for experimental data were stopped after 100 iterations. For a point-like source, the number of iterations (after a certain point) does not change the reconstructed picture much in the sense that the peak position remains the same and only its width is being reduced with each subsequent iteration. In figure 13 we show a dependence of the peak width on the number of iterations for a sample 2D reconstruction (solid blue and dashed red lines) and 1D reconstruction (dashed–dotted green line). It is clearly visible that the peak widths are constantly decreasing with the number of iterations. In the 1D reconstruction, the large number of iterations reduces the peak to a few bins, which hinders a Gaussian fit and causes the steps in the figure resulting from fit instabilities rather than the real change of the peak width.

In order to test the method with source distributions relevant for proton therapy monitoring via PGI, we investigated a setup featuring a larger detector: the full-scale prototype and its performance via simulations of its response to PGs originating from a PMMA phantom irradiated with a proton beam. The simulations were performed for four proton energies. The example simulated PG profile in figure 12(a) resembles shape-wise the profiles presented in other works (see e.g. Pinto *et al* 2016, Krimmer *et al* 2018, Livingstone *et al* 2021), though the main peak is enhanced due to weighting the entries by their energy deposits. Additionally, it shows a small peak at a distance of 15 mm to 18 mm to the Bragg peak. This corresponds to a proton energy between 43 and 49 MeV. Our earlier experimental data of Wrońska *et al* (2021) also do reveal the presence of such a peak, similar structures have been reported by other groups as well (see Min *et al* 2006). We evaluate our results by comparing the DFP values obtained from the reconstructed images with the true ones (obtained from Monte Carlo truth distribution). Among other possibilities, such as fitting a sigmoid function to the distal edge of the profile, we choose to use the 50 % location to define DFP. Due to the smoothness and steepness of our reconstructed profiles, this is an on the one hand a very simple and on the other hand a very robust method to assign a range value to the profile. The dependence of the precision of the reconstructed DFP on the number of impinged protons is presented graphically in figure 12(b) and numerically in table 1. For 10^8 protons, the error of the distal

Table 2. Precision (one standard deviation) of beam range estimation by different groups and different PGI approaches. Question marks (?) represent the missing information for reports from clinical tests (Richter *et al* 2016, Xie *et al* 2017), where the numbers of impinged protons and/or beam energy range were not stated.

Source	Energy [MeV]	N_p	Precision [mm]
CM simulation (this work)	85.9–107.9	10^8	0.72
CM simulation (Sun *et al* 2020a)	122.7	10^8	2.1
MPS experiment (Park *et al* 2019)	95.09	3.8×10^8	1.2
MPS simulation (Pinto *et al* 2014)	160	10^8	1.30–1.66
KES clinical (Xie *et al* 2017)	100–160	?	0.7–1.3
KES clinical (Richter *et al* 2016)	?	?	2.0

fall-off estimation is within 0.6 mm for all regarded energies and an average precision of the DFP estimation is 0.72 mm. This value is confronted with precision values reported by other groups working with different setups (both simulation and experimental) in table 2. In the comparison, we include besides our results also those obtained with a two-dimensional coded mask setup (Sun *et al* 2020a), multi-parallel slit (MPS) simulation (Pinto *et al* 2014) and experimental (Park *et al* 2019) results as well as clinical results obtained with knife-edge slit (KES) design (Richter *et al* 2016, Xie *et al* 2017). Although precision-wise our results outperform most of the works under comparison, we admit that our simulation model does not take into account certain effects which could deteriorate the result, e.g. the neutron background. However, the authors of Sun *et al* (2020a) showed that this contribution can be efficiently eliminated by employing a Discrete Cosine Transform. We also do not take into account other effects, e.g. fully realistic resolutions, the time structure of a clinical proton beam leading to high prompt gamma rates, as well as statistics reduction due to the dead time of the detector and finite throughput of the data acquisition system, which are obviously included in the experimental results of Richter *et al* (2016), Xie *et al* (2017), Park *et al* (2019). We suppose that these effects will not deteriorate our results a lot as the energy resolution is validated against experiments and we could show in Kasper *et al* (2020) that our system should be able to deal with clinical rates.

5. Conclusions

So far, the use of coded mask systems for proton therapy monitoring via PGI has been considered only theoretically, i.e. via Monte Carlo simulations. In this paper, we show the experimental results of its practical implementation and evaluation of performance using the MLEM algorithm for image reconstruction. In the first step, we use small-scale prototypes of the detector and test the image reconstruction framework with point-like sources. Experimental results confirmed that the near-field coded-mask imaging is feasible with gamma sources, and our implemented image reconstruction framework is able to reconstruct source positions quite well with both 1D and 2D approaches: a clear peak is always visible, and the reconstructed images are free from artefacts. The reconstructed source positions show a constant systematic offset to the designed source position presumably due to the limited accuracy of the setup alignment. A calibration with radioactive sources prior to the use of the setup for beam range verification is a straightforward method to correct for it.

The second step comprised Monte Carlo simulations with a realistic source distribution, i.e. obtained from a PMMA phantom irradiated with a proton beam. Here, the simulated detector had a larger size of 110.6×100 mm^2 and was coupled with a larger mask compared to the small-scale prototypes. This not only led to a larger FOV, but also increased the setup sensitivity to the details of the imaged source distribution. Our investigation conducted for different beam energies from the range 85.9–107.9 MeV and varying statistics shows promising results. The reconstructed images resembled the Monte Carlo truth PG depth profiles. At the statistics of 1×10^8 impinging protons, the mean precision of beam range estimation in the investigated beam energy range was 0.72 mm (1σ), which makes the setup competitive to other PGI approaches with passive collimation, such as KES or MPS investigated by other groups. Due to the promising results of the simulation of the full-scale prototype, we proceed with the construction of a detector with exactly this design.

Acknowledgments

electronics were provided by the Department of Physics of Molecular Imaging Systems, Institute for Experimental Molecular Imaging, RWTH Aachen University. They were developed within the European Union's Horizon 2020 research and innovation programme under grant agreement No 667211. When working on the project, Ronja Hetzel was supported from the German Research Foundation (DFG) project number 288267690, and Andreas Bolke from the DFG Grant COMMA, project number 383681334.

The authors are responsible for the content of this publication.

Data availability statement

The data cannot be made publicly available upon publication because they are not available in a format that is sufficiently accessible or reusable by other researchers. The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Ronja Hetzel ● https://orcid.org/0000-0002-6142-6609
Vitalii Urbanevych ● https://orcid.org/0000-0002-1858-4708
Andreas Bolke ● https://orcid.org/0000-0001-7772-163X
Monika Kercz ● https://orcid.org/0009-0007-6792-8945
Magdalena Kołodziej ● https://orcid.org/0000-0001-8016-9495
Andrzej Magiera ● https://orcid.org/0000-0002-7561-6366
Florian Mueller ● https://orcid.org/0000-0002-9496-4710
Sara Müller ● https://orcid.org/0000-0003-1916-354X
Magdalena Rafecas ● https://orcid.org/0000-0001-5691-7756
Katarzyna Rusiecka ● https://orcid.org/0000-0003-2554-0612
David Schug ● https://orcid.org/0000-0002-5154-8303
Volkmar Schulz ● https://orcid.org/0000-0003-1341-9356
Achim Stahl ● https://orcid.org/0000-0002-8369-7506
Bjoern Weissler ● https://orcid.org/0000-0003-1119-785X
Ming-Liang Wong ● https://orcid.org/0000-0002-9530-1222
Aleksandra Wrońska ● https://orcid.org/0000-0003-0126-3315

References

Albertini F *et al* 2020 Online daily adaptive proton therapy *Br. J. Radiol.* **93**
Anderson D N *et al* 2006 Detection and location of gamma-ray sources with a modulating coded mask *Technometrics* **48** 252
Braga J 2019 Coded aperture imaging in high-energy astrophysics *Publ. Astron. Soc. Pac.* **132** 012001
Fenimore E E and Cannon T M 1978 Coded aperture imaging with uniformly redundant arrays *Appl. Opt.* **17** 337
Ferrero V *et al* 2018 Online proton therapy monitoring: clinical test of a silicon-photodetector-based in-beam PET *Sci. Rep.* **8** 4100
Frach T *et al* 2009 The digital silicon photomultiplier—principle of operation and intrinsic detector performance *IEEE Nucl. Sci. Symp. Conf. Rec.* pp 1959–65
Frach T *et al* 2010 The digital silicon photomultiplier—system architecture and performance evaluation *IEEE Nucl. Sci. Symp. Med. Imaging Conf.* (IEEE) pp 1722–7
Gottesman S R and Fenimore E E 1989 New family of binary arrays for coded aperture imaging *Appl. Opt.* **28** 4344
Gueth P *et al* 2013 Machine learning-based patient specific prompt-gamma dose monitoring in proton therapy *Phys. Med. Biol.* **58** 4563
Hetzel R *et al* 2020 Characterization and simulation of an adaptable fan-beam collimator for fast calibration of radiation detectors for PET *IEEE Trans. Radiat. Plasma Med. Sci.* **4** 538
Hyperion Hybrid Imaging Systems GmbH. https://hyperion-his.com/
Kasper J *et al* 2020 The SiFi-CC project—feasibility study of a scintillation-fiber-based Compton camera for proton therapy monitoring *Phys. Med.* **76** 317
Kelleter L *et al* 2017 Spectroscopic study of prompt-gamma emission for range verification in proton therapy *Phys. Med.* **34** 7
Krimmer J *et al* 2018 Prompt-gamma monitoring in hadrontherapy: a review *Nucl. Instrum. Methods Phys. Res.* A **878** 58
Lange K and Carson R 1984 EM reconstruction algorithms for emission and transmission tomography *J. Comput. Assist. Tomogr.* **8** 306
Livingstone J *et al* 2021 Influence of sub-nanosecond time of flight resolution for online range verification in proton therapy using the line-cone reconstruction in Compton imaging *Phys. Med. Biol.* **66** 125012
Lu W *et al* 2022 Simulation study of a 3D multislit prompt gamma imaging system for proton therapy monitoring *IEEE Trans. Radiat. Plasma Med. Sci.* **6** 355
Min C-H *et al* 2006 Prompt gamma measurements for locating the dose falloff region in the proton therapy *Appl. Phys. Lett.* **89**
Mu Z, Dobrucki L W and Liu Y-H 2016 SPECT imaging of 2D and 3D distributed sources with near-field coded aperture collimation: Computer simulation and real data validation *J. Med. Biol. Eng.* **36** 32
Müller F *et al* 2018 Gradient tree boosting-based positioning method for monolithic scintillator crystals in positron emission tomography *IEEE Trans. Radiat. Plasma Med. Sci.* **2** 411
Müller F *et al* 2019 A novel DOI positioning algorithm for monolithic scintillator crystals in PET based on gradient tree boosting *IEEE Trans. Radiat. Plasma Med. Sci.* **3** 465

Müller S 2022 *Use of PET Readout Electronics for a Scintillating Fibre based Compton Camera* RWTH Aachen University, Aachen, Germany Master's Thesis (https://www.institut3b.physik.rwth-aachen.de/global/show_document.asp?id=aaaaaaaabtuxbew)

NuPECC report 2014: Nuclear Physics for Medicine. (http://www.nupecc.org/pub/npmed2014.pdf). Accessed: 2021-07-13.

Park J H *et al* 2019 Multi-slit prompt-gamma camera for locating of distal dose falloff in proton therapy *Nucl. Eng. Technol.* **51** 1406

Pausch G *et al* 2020 Detection systems for range monitoring in proton therapy: needs and challenges *Nucl. Instrum. Methods Phys. Res.* A **954** 161227

Pinto M *et al* 2014 Design optimisation of a TOF-based collimated camera prototype for online hadrontherapy monitoring *Phys. Med. Biol.* **59** 7653

Pinto M *et al* 2015 Absolute prompt-gamma yield measurements for ion beam therapy monitoring *Phys. Med. Biol.* **60** 565

Pinto M *et al* 2016 Assessment of Geant4 prompt-gamma emission yields in the context of proton therapy monitoring *Front. Oncol.* **6** 1–7

Ready J F 2016 Development of a multi-knife-edge slit collimator for prompt gamma ray imaging during proton beam cancer therapy *PhD Thesis* University of California, Berkeley (http://adsabs.harvard.edu/abs/ PhDT.......151R)

Ready J *et al* 2016 MO-FG-CAMPUS-JeP1-01: Prompt gamma imaging with a multi-knife-edge slit collimator: evaluation for use in proton beam range verification *Med. Phys.* **43** 3717

Richter C *et al* 2016 First clinical application of a prompt gamma based in vivo proton range verification system *Radiother. Oncol.* **118** 232

Rockmore A J and Macovski A 1976 A maximum likelihood approach to emission image reconstruction from projections *IEEE Trans. Nucl. Sci.* **23** 1428

Rusiecka K 2023 The SiFi-CC detector for beam range monitoring in proton therapy—characterization of components and a prototype detector module *PhD Thesis* Jagiellonian University, Kraków (https://doi.org/10.48550/arXiv.2306.10820)

Rusiecka K *et al* 2021 A systematic study of LYSO:Ce, LuAG:Ce and GAGG:Ce scintillating fibers properties *J. Instrum.* **16** P11006

Schug D *et al* 2015 Data processing for a high resolution preclinical PET detector based on Philips DPC Digital SiPMs *IEEE Trans. Nucl. Sci.* **62** 669

Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **1** 113

Sun S, Liu Y and Ouyang X 2020a Design and performance evaluation of a coded aperture imaging system for real-time prompt gamma-ray monitoring during proton therapy *Radiat. Phys. Chem.* **174**

Sun S, Liu Y and Ouyang X 2020b Near-field high-resolution coded aperture gamma-ray imaging with separable masks *Nucl. Instrum. Methods Phys. Res.* A **951** 163001

Tabacchini V *et al* 2014 Probabilities of triggering and validation in a digital silicon photomultiplier *J. Instrum.* **9** P06016

Verburg J M and Seco J 2014 Proton range verification through prompt gamma-ray spectroscopy *Phys. Med. Biol.* **59** 7089

Wang Z and Bovik A 2002 A universal image quality index *IEEE Signal Process Lett.* **9** 81

Weissler B *et al* 2015 A digital preclinical PET/MRI insert and initial results *IEEE Trans. Med. Imaging* **34** 2258–70

Weissler B *et al* 2022 Hyperion III-MRI-Compatible PET Detector Platform *PSMR-TBP 2022 9th Conf. on PET/MR and SPECT/MR & Total-Body PET Workshop* https://agenda.infn.it/event/28667/book-of-abstracts.pdf

Wrońska A and Dauvergne D 2021 Range verification by means of prompt-gamma detection in particle therapy *Radiation Detection Systems* ed K Iniewski and J Iwańczyk (Boca Raton: CRC Press/Routledge) vol 2, pp 139–60

Wrońska A *et al* 2021 Prompt-gamma emission in Geant4 revisited and confronted with experiment *Phys. Med.* **88** 250

Xie Y *et al* 2017 Prompt gamma imaging for in vivo range verification of pencil beam scanning proton therapy *Int. J. Radiat. Oncol. Biol. Phys.* **99**

Young M 1989 The pinhole camera: imaging without lenses or mirrors *Phys. Teach.* **27** 648

Ziegler J F, Ziegler M D and Biersack J P 2010 SRIM—the stopping and range of ions in matter *Nucl. Instrum. Methods Phys. Res.* B **268** 1818