The Astrophysical Journal, 928:141 (17pp), 2022 April 1 https://doi.org/10.3847/1538-4357/ac5626

A Multimission Method for the Reconstruction of Gamma-ray Events on Silicon Tracker Pair Telescopes

Alessio Aboudan[1,2], Andrea Bulgarelli[2], Valentina Fioretti[2], Andrea Giuliani[3], Marco Tavani[4], and Stefano Debei[1]
[1] CISAS G. Colombo University of Padova, via Venezia 15, I-35131 Padova, Italy; alessio.aboudan@unipd.it
[2] INAF-OAS Bologna, via Gobetti 93/3, I-40129 Bologna, Italy
[3] INAF-IASF Milano, via Alfonso Corti 12, I-20133 Milano, Italy
[4] INAF-IAPS Roma, via Fosso del Cavaliere, I-00133 Roma, Italy
Received 2021 February 13; revised 2022 January 18; accepted 2022 February 15; published 2022 April 4

Abstract

γ-ray astronomy in the energy range from MeV to GeV can provide a unique detection window for γ-ray bursts and other transient sources, fundamental information on particle acceleration mechanisms, MeV-blazar population studies up to $z \sim 4.5$, and a full overview of line emission from cosmic-ray interaction. Silicon-based pair tracking telescopes rely on γ-ray conversion into an electron–positron pair and its tracking using a stack of silicon strips. The method presented in this work is based on a Rauch–Tung–Striebel smoother. Its internal Kalman filter enables keeping multiple hypotheses about particle tracks and implementing statistically meaningful measurement selection among hits on different planes of the tracker. The algorithm can be easily configured to work with different tracker geometries and mass models. It can be used for the exploitation of data from past and current γ-ray missions as well as to assess the performances of new pair-tracking telescopes. The proposed method has been validated on Astrorivelatore Gamma a Immagini Leggero data and then used to investigate the performances of both e-ASTROGAM and All-Sky-ASTROGAM telescopes. The algorithm efficiency and its accuracy in estimating both the photon direction and energy were evaluated on γ-ray events simulated at different energies in the range between 30 MeV and 3 GeV. The point-spread function of each tracker was then compared with its angular resolution limit showing both the expected performances of the instrument and the margin of improvement that could be obtained by optimizing the reconstruction method.

Unified Astronomy Thesaurus concepts: Bayesian statistics (1900); Algorithms (1883); Gamma-ray telescopes (634); Gamma-ray detectors (630)

1. Introduction

γ-ray astronomy from space above tens of MeV is currently based on the data collected by Astrorivelatore Gamma a Immagini Leggero (AGILE; Tavani et al. 2009) and FERMI (Atwood et al. 2009). Both telescopes are made up of a stack of tungsten converter layers interleaved by silicon tracking layers (silicon trackers). The incoming γ-ray interacts with several tungsten layers producing an electron and a positron that are tracked by means of silicon strip detectors located on each layer of the instrument. Although AGILE and FERMI share a common design, the two instruments are different in terms of the number and thickness of tungsten layers, distance between silicon layers, number of silicon strips, and signal readout electronics. Nevertheless both instruments show a similar angular resolution in the energy range from about 100 MeV to some GeV (Sabatini et al. 2015). More recently, two γ-ray missions called e-ASTROGAM (De Angelis et al. 2018) and All-Sky-ASTROGAM (Tatischeff et al. 2019) have been proposed by a wide international community with a broad energy range (0.3 MeV–3 GeV) and improved performances (up to 1–2 orders of magnitude in continuum sensitivity) with respect to their predecessors. The full exploitation of the data collected by AGILE and the need to support the design and development of future missions such as e-ASTROGAM and All-Sky-ASTROGAM motivated the development of a multimission and easily configurable tool to reconstruct γ-rays in the pair domain. The method presented in this work is aimed at reconstructing the incoming direction of a γ-ray and its energy using as input the configuration of the tracker and a cloud of particle hits measured on the external faces of the tracker. Both the tracker model and the reconstruction algorithm can be configured to process different data sets, compare different tracker geometries, and investigate the overall system performances.

According to the physical process of pair production, the reconstruction of the γ-ray direction is based on the identification of both positron and electron paths plus an estimation of the distribution of the total energy between the two particles (Pittori & Tavani 2002). To this aim several track fitting approaches implementing both global (Billoir 1984) and recursive algorithms have been proposed (Billoir et al. 1985; Frühwirth 1987) and applied to FERMI (Jones 1998; Atwood et al. 2009), AGILE (Giuliani et al. 2006), and GAMMA400 (Cumani 2015). Recursive methods are based on Kalman filtering and smoothing algorithms implementing very similar models of the particles motion inside the tracker. Some methods start with the identification of an event *vertex*, i.e., where the e^+ and e^- paths became distinguishable; then the Kalman filter is used to further propagate and reconstruct the particles' paths (Billoir et al. 1985). The tracks are then smoothed and refined using outlier detection and removal methods (Frühwirth 1987). An effective method to search for the event vertex is to consider all the possible combinations of hits on the first two or three planes of the detector. In this case, to reduce complexity, the 2D projections of the tracks on the

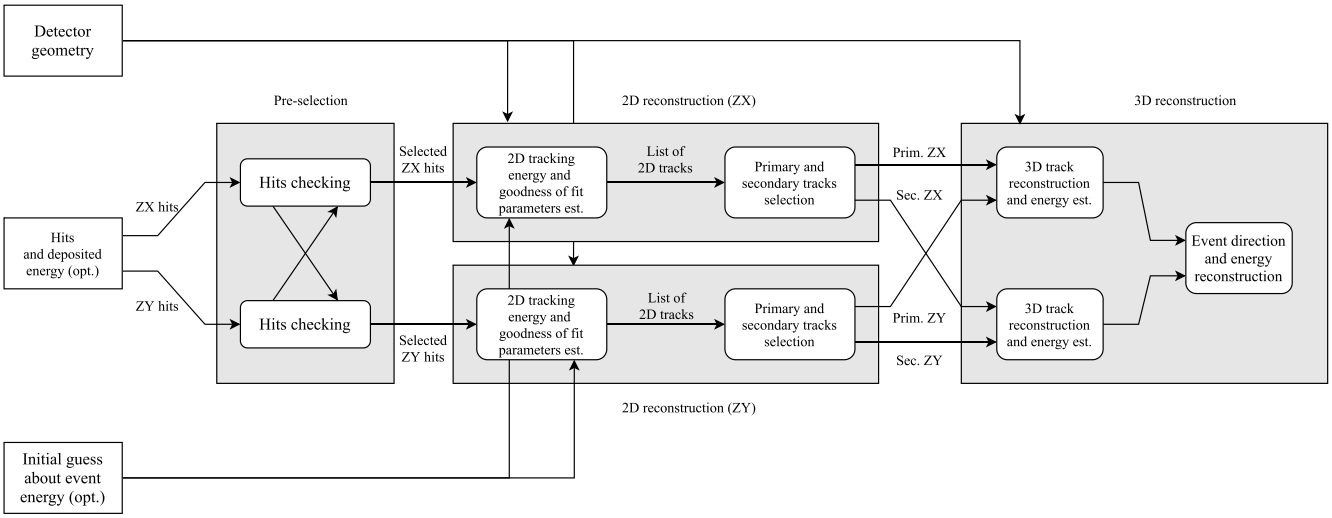

Figure 1. Tracker main blocks and processing flow diagram.

detector faces can be fitted separately before merging them in a full 3D reconstruction (Jones 1998; Giuliani et al. 2006). Additional measurements from a calorimeter are often used to have an initial estimate of the event energy and account for multiple scattering in the tracking (Jones 1998; Cumani 2015), or even to initialize the tracks to be reconstructed with the Kalman filter (Atwood et al. 2009). When the calorimeter is not available, other strategies to initialize the event energy have been developed (Giuliani et al. 2006; Atwood et al. 2009; Bernard & Frosini 2017; Frosini & Bernard 2017). The use of the calorimeter, the set of empirical rules proposed to perform vertex finding, track initialization, selection, and matching, mark the difference between the proposed methods limiting their applicability to the instrument for which they were designed.

The algorithm presented in this work is based on a Rauch–Tung–Striebel smoother. The first step of such an algorithm is a Kalman filter implementing a linear model of a particle moving inside the tracker and subject to Coulomb scattering; the same model can be found in Giuliani et al. (2006) and a similar one in Jones (1998). The algorithm begins reconstructing 2D tracks on the two faces of the detector; measured hits are processed sequentially from the top to the bottom of the instrument (forward pass). To get the method working with any detector configuration, it is assumed that each hit can be the starting point of a track and each track can be associated to more than one hit. As a consequence multiple hypotheses about the most probable tracks are generated and handled efficiently by the algorithm. After track smoothing (backward pass), one primary track and one secondary track are selected leading to the identification of the vertex. The 2D views of both the primary track and the secondary track are combined to build the corresponding 3D tracks; their energy is computed and used to reconstruct the incoming direction of the γ-ray and its energy. The measurement of the event energy and particle positions provided by a calorimeter can be included in the processing, if available. The validation of the tool presented in this work is based on Geant4 simulations (see Section 10) of the AGILE/GRID and e-ASTROGAM/All-Sky-ASTROGAM tracks. Simulated events were used to investigate the algorithm accuracy and its contribution to the overall instrument performance. The simulation setup has been first validated in Fioretti et al. (2020) by reproducing both the AGILE/GRID charge readout distribution obtained in laboratory measurements and the point-spread function (PSF) obtained in real observations.

The first part of this paper is aimed at providing an in-depth description of the developed tool. An overview of the whole algorithm with particular emphasis on tracks initialization, duplication, maintenance, and event reconstruction is given in Section 2. Section 3 describes the small-angle scattering model used to compute the uncertainty in the particle motion and to estimate the track energy. The model of the tracker and the state space model representing the particle motion are described in Sections 4 and 5. The algorithm used to identify the 2D tracks is detailed in Sections 6–8. Section 9 describes the method used to estimate the track energy. The second part of the paper deeply describes the simulation setup used to generate γ-ray events (see Section 10) and investigate the tracker performances (see Section 11). In particular, Section 11.3 reports the tracker validation with respect to the AGILE Reconstruction Pipeline (RECO) filter, while Section 11.4 presents the results of the analysis of both e-ASTROGAM and All-Sky-ASTROGAM data.

2. Description of the Method

2.1. Main Processing Steps

According to the physical process of pair production, the reconstruction of the γ-ray direction is based on the identification of both positron and electron paths plus an estimate of the distribution of the total energy between the two particles (Pittori & Tavani 2002). The inputs to the tracker are the detector geometry, the list of the hits observed on each tray, i.e., their position along the *X*- or *Y*-axis, the energy deposited on the sensitive strips, and an initial guess about the energy of the event; note that both the deposited energy and the event energy are optional. After event preselection, 2D tracks are identified independently on each face of the detector (either *ZX* or *ZY*); then 3D tracks are reconstructed, properly matching 2D tracks. The event processing shown in Figure 1 is divided in three steps: preselection, 2D reconstruction, and 3D reconstruction. In the preselection stage, it is possible to group nearby hits on the same layer to form a cluster. This step is optional, but it

Table 1
Tracker Parameters

Parameter	AGILE	e/a-ASTROGAM	References
Event preselection	$m = 3$, $n = 5$	$m = n = 10$	(Section 2.2)
Switch to confirmed	3	5	
Switch to terminated	$m = 3$, $n = 5$	$m = 3$, $n = 7$	
Duplication significance	75%	95%	(Sections 7, 2.3.2)
Update significance	99%	99%	
Initial position uncertainty 1σ	1 cm	1 cm	(Used to define $\boldsymbol{P}_0$ in Equation (14))
Initial direction uncertainty 1σ	3°	3°	
Measurement uncertainty 1σ	0.2 cm	0.08 cm	($\boldsymbol{R}_k$ in Equation (16))

is crucial to reduce the number of tracks to be reconstructed. Moreover, it is required that the event is detected at least on a minimum number of layers (e.g., three to five) to guarantee robust reconstruction of the γ-ray direction; otherwise the processing is stopped. The 2D reconstruction of the tracks is performed independently on both *ZX* and *ZY* faces of the detector. In this stage, the hits on one face of the tracker are processed starting from the top to the bottom of the detector (decreasing *Z*). After proper initialization a Bayesian estimator is used to process the measurements keeping multiple hypotheses about existing tracks and to get the optimal reconstruction of the most likely particle paths. The position and direction of a particle, moving inside the tracker and subject to Coulomb scattering, is modeled with a linear system of equations as in Giuliani et al. (2006). The filter is able to propagate each track down (Kalman filter part of the algorithm), along the *Z*-axis, and associate it with the best available hits in a statistically meaningful way. This reduces the number of multiple hypotheses maintained by the algorithm and improves both the performances and the efficiency of the tracker. All the identified tracks are then smoothed through a backward algorithm working from the bottom to the top of the detector. The complete algorithm (forward plus backward parts) is called Rauch–Tung–Striebel smoother and is thoroughly described in Särkkä (2013). The 3D reconstruction of e^+ and e^- tracks is made matching their 2D projections. Several criteria such as the track length, track continuity, number of matched hits, χ^2 measurement residuals, and estimated energy are used hierarchically to identify the primary and the secondary tracks and match them among *ZX* and *ZY* faces. Then the direction of the incoming photon is reconstructed weighting the primary and secondary track direction by their estimated energy (Pittori & Tavani 2002; Giuliani et al. 2006).

2.2. Event Preselection

To be properly reconstructed, an event should be visible on a minimum number of planes. In this work, an event is analyzed only if at least one detection is available on both *ZX* and *ZY* strips on at least *M* out of *N* consecutive tracker planes. As an example, setting $M = N = 5$ results in processing only events with at least five consecutive planes with observations on both *ZX* and *ZY* faces. In practice, tracks have their own consistency even if some gaps in the hits could be present; in this case, a less stringent constraint can be defined setting $M = 3$ and $N = 5$. This results in processing events with missing hits on one or two consecutive planes. Values used for AGILE and e-ASTROGAM are reported in Table 1.

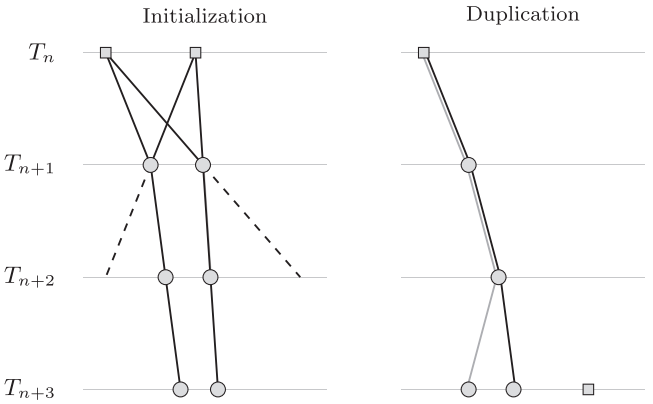

Figure 2. Examples of track initialization and duplication. Initialization occurs on T_n and T_{n+1} (left). On T_n initialization is partial (position only). Two hits on T_{n+1} are used to fully initialize four tracks (both position and direction). Then on T_{n+2} two tracks are associated with existing hits and updated accordingly, while the other two are not associated and hence pruned out (dashed). Track duplication occurs when a track is compatible with more than one hit (right). A fully initialized track is propagated to T_{n+3} and updated using the best hit (black line); then the track is duplicated to share all the history from T_n to T_{n+2} plus the propagated state on T_{n+3} and updated using the compatible hit (gray). Another hit on the same plane is not compatible and will be used to initialize a new track.

2.3. 2D Tracking

2.3.1. Track Initialization

The goal of the *track initialization* step is to compute the initial distribution of the particle state (both position and direction) on each face of the detector. The track is partially initialized using the measurement of a particle (hit) on one tray of the detector; in this case, only the position of the particle is set. The full initialization occurs when a second measurement of the particle position on another tray is used to compute the initial direction of the particle. The initial uncertainties on the particle position and direction are set as constants. The minimum requirement to initialize a track are two consecutive trays with at least one measurement per plane. When more measurements are available, then multiple hypotheses can be made about the existing tracks. In this case a *brute-force* approach is used, and all the possible tracks are initialized; an example is shown in Figure 2. The brute-force approach produces some wrong tracks that are not associated with hits on

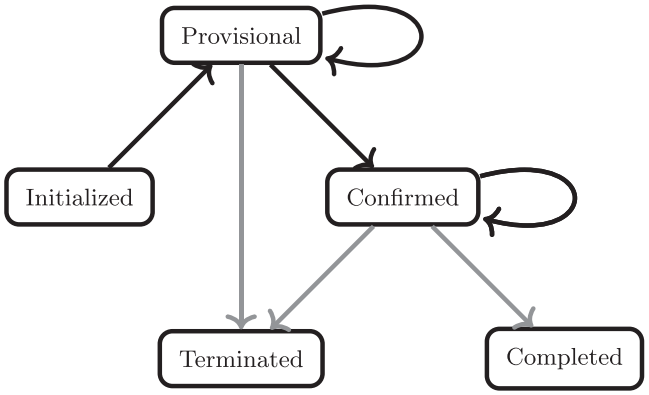

Figure 3. Particle tracker state machine: state transitions can be triggered by data association and update steps (black) or by the track maintenance algorithm (gray).

the successive planes of the detector and are set to *terminated* (see Section 2.3.2).

2.3.2. Track Propagation, Update, and Duplication

After *initialization*, the *forward iteration* of the tracking algorithm (Section 6.1) can take place. For each tray of the detector, the existing tracks are propagated to the next tray using the model in Section 5 and then paired with the available observations. If an observation is associated with the track, then the track state is properly updated according to the diagram in Figure 3. After the first update, each track is marked as *provisional* and becomes *confirmed* only after a given number of filter updates. This mechanism allows the identification of outliers that can be pruned out in an early stage of the algorithm. After the propagation of all the *provisional* and *confirmed* tracks, the Mahalanobis distance (Section 7) is computed for all the possible track-hit combinations. Then the track update step is performed on the track–hit pairs with highest association ranking. If a track is compatible with more than one hit, after the *update* with the best hit, track duplication can occur. In this case, the current track is duplicated and then updated using the remaining hits. Duplicated tracks share with the original one all the history (position, direction, and covariance) from the initialization plane to the last plane before duplication (see the example in Figure 2). The parent–child relations between tracks are kept in memory in a tree-like structure and used during the track *selection* phase. Two different thresholds on the significance level are used; one is used for the *update* step and another, less stringent one, is used for the *duplication* step. Hits that are not associated with a track are used to initialize new tracks from scratch.

2.3.3. Track Maintenance

Track maintenance is the process in which a decision is made about whether to end the life of a track. This decision is crucial to limit the number of active tracks, and it is based on empirical considerations. If the position or direction covariance tracked by the algorithm become too small, then the filter will collapse. If the track covariance becomes too big, this could signal a wrong data association or outliers. Moreover if the changing in particle direction from the first to the last traversed plane is too big, then it is very likely that the track is an outlier. To this aim the tracks are marked as *terminated* if one of the following conditions arise:

1. the track uncertainty has grown beyond a given threshold;
2. the track uncertainty has collapsed below a given threshold;
3. the track direction variation with respect to the first or the last value exceeds a given threshold.

If a track is not associated with a particle detection during the data association phase, then the track may no longer exist or the detector may have failed to identify the particle. In the latter case, there is the chance that the detector will find again the particle on the next plane. Hence when no measurements are associated for the past M out of N traversed detector trays, *provisional* tracks are *terminated* while *confirmed* tracks are marked as *completed*. Tracks marked as *completed* are smoothed by means of the *backward iteration* (Section 6.2). If a tray is made up of several layers, λ is approximated using the mean of the radiation lengths of the various materials weighted with the thickness of each layer. Then, using the smoothed trajectory, the track energy is estimated using the method in Section 9. Finally, the goodness of fit is computed with Equation (23).

2.4. Track Selection and Matching

The tracking algorithm described in Section 2.3.2 is used to reconstruct the set of possible 2D tracks on both *ZX* and *ZY* faces of the detector. For each face, a *primary* track and a *secondary* track are selected. Only the tracks that were marked as *completed* according to Section 2.3.3 are considered while the other tracks are discarded from the analysis. Optionally a threshold on the track length and/or on the goodness of fit can be defined to further reduce the number of candidates. The primary track is selected as the one with maximum length, maximum energy, and minimum χ^2 value. Then, the secondary track is defined as the remaining one with the maximum length, maximum energy, minimum χ^2 value, and event vertex located in the uppermost part of the detector; this corresponds to the secondary track with the minimum number of hits (if any) shared with the primary track. To efficiently evaluate the number of common planes between different tracks, the tree-like structure created during tracks propagation stage (Section 2.3.2) is used. In particular, if the primary and the secondary tracks have a common parent in the tree, then it corresponds to the vertex of the γ-ray event.

2.5. 3D Reconstruction

In the pair production process, the total energy of the incoming photon is divided between the e^+ and the e^- with one of the two particles carrying most of the total energy (Section 3). Particle energy is reflected on the rms scattering angle according to Equations (5) and (6) and can be used to associate the 2D tracks between the *ZX* and *ZY* planes, as shown in Figure 4. The 2D tracks are then merged to build the 3D profile of the two selected particles.

The energy of each 3D track is then computed with the method described in Section 9. The direction of the incoming photon is reconstructed combining the direction of the primary track and the one of the secondary track (on the first common plane) weighted with the track energy as suggested in Pittori & Tavani (2002), Giuliani et al. (2006). Let e_1 and $\boldsymbol{v}_1$ be the energy and the direction of the primary track, respectively. Here the direction is computed using the track position on the two uppermost planes of the detector. Additionally, let e_2 and

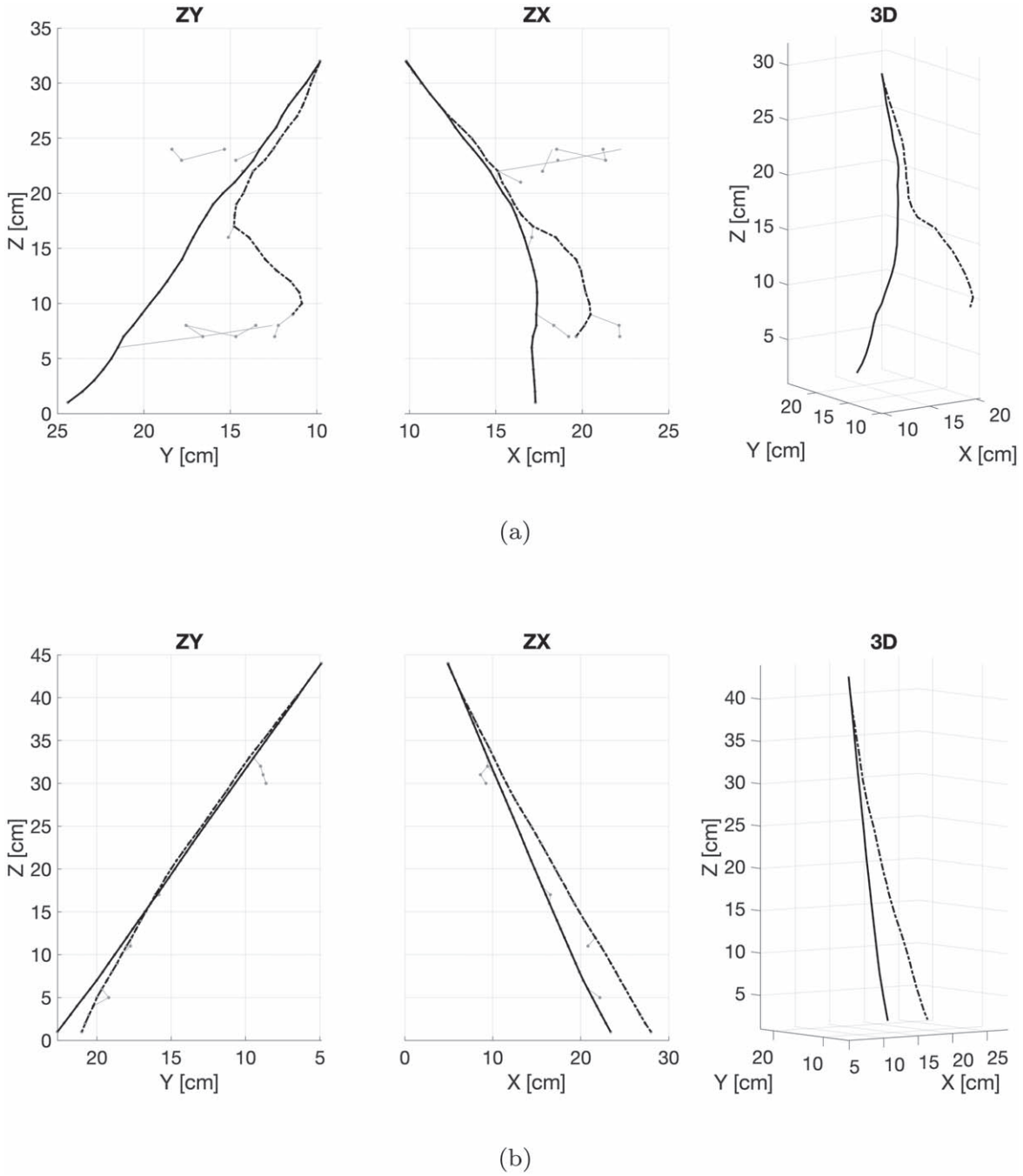

Figure 4. Input hits (gray) on *ZX* and *ZY* faces, selected primary (solid) and secondary (dashed) track, 3D reconstruction of charged particles paths. The examples are at (a) 30 MeV and (b) 500 MeV showing that as the energy increases the angular deviation decreases and in any case is reflected to the projections of the tracks on the detector faces; the estimated energy is the key both to correctly select primary and secondary 2D tracks as well as to match 2D tracks between the *ZX* and *ZY* faces and hence to retrieve the final 3D profiles.

$\boldsymbol{v}_2$ be the same parameters of the secondary track. Then the direction of the incoming photon in computed as

$$\boldsymbol{v} = \frac{e_1\,\boldsymbol{v}_1 + e_2\,\boldsymbol{v}_2}{e_1 + e_2}; \tag{1}$$

finally the incoming direction angles θ and ϕ are derived from $\boldsymbol{v}$ according to the detector reference frame.

3. Small-angle Scattering

γ-rays with energy E_γ exceeding 1.022 MeV (two times the rest mass of an electron) interact with matter producing electron–positron pairs. This *pair-production* effect takes place close to a nucleus that receives some fraction E_n of the total energy. The remaining energy is divided between the electron E_e and the positron E_p, and it is reflected in the scattering angles θ_e and θ_p (see Figure 5). The pair-production process obeys the following conservation laws for the energy E and momentum $\boldsymbol{p}$

$$E_\gamma + M\,c^2 = E_p + E_e + E_n, \tag{2}$$

$$\boldsymbol{p}_\gamma = \boldsymbol{p}_p + \boldsymbol{p}_e + \boldsymbol{p}_n, \tag{3}$$

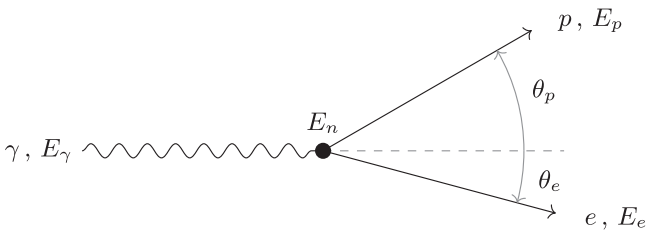

Figure 5. Schematic representation of the pair-production effect.

with M the mass of the nucleus of the interacting material. The left-hand side of Equation (2) is the photon energy plus the rest mass of the nucleus, while the right side contains the energy of the produced pair plus the kinetic energy related to the recoil of the nucleus. The amount of energy transferred to the nucleus E_n is very small, while almost all the original energy is divided between the e^+ and the e^-. The probability distribution of having a given E_p and E_e is almost flat for a total energy E_γ below 50 MeV while, for higher energetic photons, it is more likely that one of the two particles carries most of the total energy. Moreover, the angles θ_p and θ_e between the original photon and the direction of the e^+ and e^-, respectively, depend on the inverse of the photon energy (Rossi 1952). The electrons and positrons generated from the pair-production effect continue their paths undergoing several small-angle scattering and loosing energy, again with a scattering angle that depends on the particle energy.

The process of multiple small-angle scatterings that occur when a charged particle interacts with matter can be described by Moliere theory (Bethe 1953). Both the scattering angle and the scattering position distributions can be a approximated by Gaussians; hence the change in the particle flight direction can be represented by the variance of its angular direction σ_θ and position σ_x. The terms σ_θ and σ_x are obtained by interpolating the Moliere distribution and depend on the material radiation length X_0 and the material thickness z

$$\lambda = \frac{z}{X_0}, \tag{4}$$

$$\sigma_\theta^2 = \left[\frac{13.6\ \text{MeV}}{E_e}\sqrt{\lambda}(1 + 0.038\ln(\lambda))\right]^2, \tag{5}$$

$$\sigma_x^2 = \sigma_\theta^2 \frac{1}{3}\lambda^2, \tag{6}$$

where σ_θ is in rad, and σ_x is in cm. Equation (5) refers to the particle direction projected on a plane. The nonprojected variance (space) can be approximated as $\sigma^2_{\theta,\text{space}} = \sigma^2_{\theta,x} + \sigma^2_{\theta,y}$, where the x and y axes are orthogonal to the direction of motion, and the deflections on x and y directions are independent and identically distributed. This results in the following relation between the projected and nonprojected covariance

$$\sigma^2_{\theta,\text{space}} = \sqrt{2}\ \sigma^2_\theta. \tag{7}$$

Note that when a particle is traversing many different layers, the value λ of the combined scatterer should be used in Equations (5) and (6). Instead, adding the individual σ_θ in quadrature results in underestimating the angular deviation of the particle (Particle Data Group et al. 2020).

4. Tracker Modeling

The γ-ray tracking instrument is represented as a sequence of vertically stacked trays. Each tray is composed by one or more layers of different materials. Each layer is characterized by its thickness, density, and the radiation length of the corresponding material (Figure 6). It is assumed that the sensor strips used to detect the presence of charged particles lay on the top of the first layer of each tray and could be able to detect particles along the X direction, Y direction or both. The tracker reference frame is then defined with the X and Y axes parallel with the trays sides and the Z-axis pointing vertically. The origin of the frame is defined on the bottom of the tracker. By convention both the trays and the layers are numbered increasingly from the top to the bottom. Particles revealed by the sensor clusters on each sensitive plane are projected on the ZX and ZY faces of the detector, and their position in centimeters is computed; then ZX and ZY 2D measurements are processed independently using the algorithm described in the next sections.

5. State–space Model

Once projected on the side of the detector, the motion of a charged particle inside the tracker can be described by a linear discrete state–space model of the form

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{A}_k \boldsymbol{x}_k + \boldsymbol{q}_k \\ \boldsymbol{y}_k &= \boldsymbol{H}\boldsymbol{x}_k + \boldsymbol{r}_k, \end{aligned} \tag{8}$$

with state vector $\boldsymbol{x}_k$, state matrix $\boldsymbol{A}_k$, measurement vector $\boldsymbol{y}_k$, measurement matrix $\boldsymbol{H}$, state noise random process $\boldsymbol{q}_k \sim N(0, \boldsymbol{Q}_k)$, and measurement noise random process $\boldsymbol{r}_k \sim N(0, \boldsymbol{R}_k)$; here $N(\mu, \sigma^2)$ is the Gaussian distribution of the mean μ and covariance σ^2. As usual, the two random processes are assumed to be uncorrelated $E\boldsymbol{q}_i\boldsymbol{r}_j = 0 \forall\ i, j$. Consider the ZX face as shown in Figure 7 (the same holds for the other face). Let x_k be the position of the particle and θ_k its direction on layer k of the detector; then, the particle will reach the next layer $k+1$ at the position

$$x_{k+1} = x_k + d\tan(\theta_k),$$

with d the total distance between the two layers. Defining a new variable $\gamma_k = \tan(\theta_k)$, it is possible to obtain the linear relation $x_{k+1} = x_k + d\gamma_k$. Moreover, the angular direction of the particle is assumed to be constant between two trays, i.e., $\theta_{k+1} = \theta_k$ and $\gamma_{k+1} = \gamma_k$. Hence, the motion of a charged particle is represented by the following set of difference equations

$$\begin{bmatrix} x_{k+1} \\ \gamma_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_k \\ \gamma_k \end{bmatrix} + \begin{bmatrix} \nu_k \\ \eta_k \end{bmatrix}, \tag{9}$$

with ν_k representing the uncertainty in the particle position and η_k the uncertainty in the particle direction. The state vector of Equation (8) is defined as $\boldsymbol{x}_k = [x_k\ \gamma_k]^T$, and the state transition matrix becomes

$$\boldsymbol{A}_k = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \tag{10}$$

accounting for the thickness of the traversed material. The uncertainty terms are due to the interaction between the particle and the layers of the detector and are computed using the small-

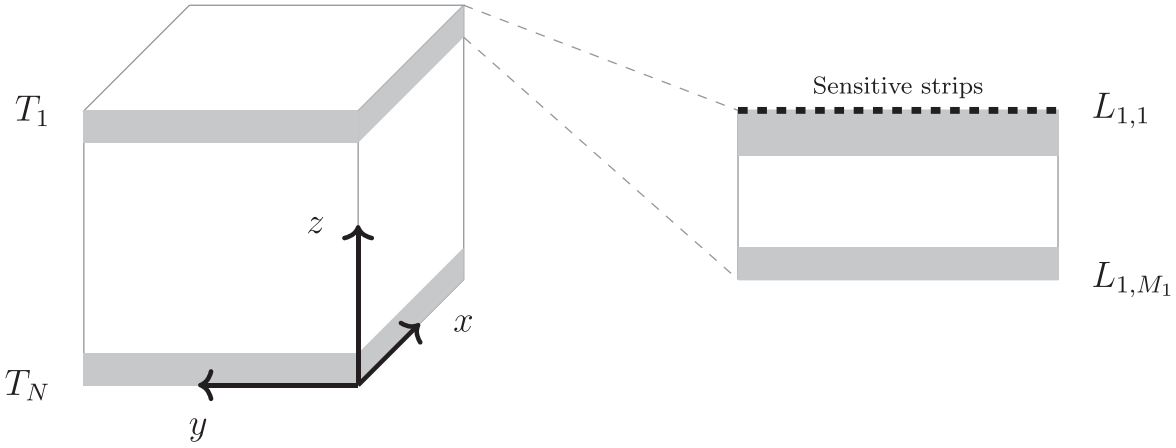

Figure 6. Model of a generic particle tracker made of N trays $T_1,\ldots,T_N$ stacked along Z-axis and related reference frame. Each tray is composed by one or more layers of different materials and thickness; for example, tray 1 has M_1 layers $L_{1,1},\ldots,L_{1,M_1}$.

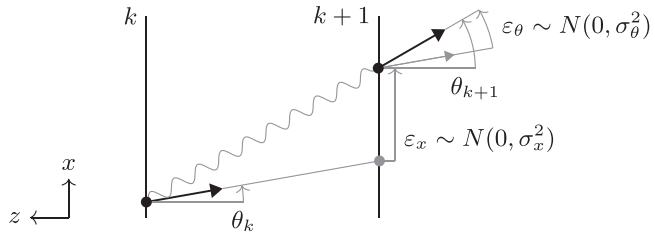

Figure 7. Charged particle motion inside the tracker between the sensitive layer of tray k and the one of tray $k+1$. Estimated particle position and direction are plotted in dark, while the predictions on the next plane are in gray. In particular ε_x and ε_θ are the errors between the predicted values and the best estimate provided by the filter. These terms are modeled in the filter using the Moliere distribution (Section 3) and, after the track reconstruction, are used to estimate the particle energy (Section 9).

angle scattering model described in Section 3. The position error is distributed as $\nu_k \sim N(0, \sigma_x^2)$, with the variance computed using Equation (6), while the direction error is distributed as $\eta_k \sim N(0, \sigma_\eta^2)$. This latter term has to be derived taking into account the nonlinear transformation $\gamma_k = \tan(\theta_k)$ and reads as

$$\sigma_\eta^2 = (1 + \tan^2(\theta))^2 \ \sigma_\theta^2 = (1 + \gamma^2)^2 \ \sigma_\theta^2 \tag{11}$$

with the variance of the particle direction σ_θ given by Equation (5). The position and direction uncertainties are correlated by a constant term $\sigma_{x\theta} = 0.87$ (Particle Data Group et al. 2020) and hence $\sigma_{\nu\eta} = (1+\gamma^2)\sigma_{x\theta}$; then the corresponding state covariance matrix reads as

$$\boldsymbol{Q}_k = \begin{bmatrix} \sigma_\nu^2 & \sigma_{\nu\eta} \\ \sigma_{\nu\eta} & \sigma_\eta^2 \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & (1+\gamma^2)0.87 \\ (1+\gamma^2)0.87 & (1+\gamma^2)^2\,\sigma_\theta^2 \end{bmatrix}. \tag{12}$$

The sensor strips placed on the top of each tray measure the position at which the particle interacts with the plane itself; if layer k is equipped with the strips, then the measurement of the particle position reads as

$$y_k = [\,1\ 0\,] \begin{bmatrix} x_k \\ \gamma_k \end{bmatrix} + r_k, \tag{13}$$

with constant measurement matrix $\boldsymbol{H} = [1\ 0]$ and r_k representing the uncertainty in the measurement of the particle position.

6. Bayesian Estimation of the Particle Track

The motion of the charged particle modeled in the previous section can be reconstructed using a Rauch–Tung–Striebel smoothing algorithm. The algorithm is made up of two steps. The *forward iteration* in which tracks are reconstructed starting form the top of the detector and moving toward the last plane on the bottom. During this phase, the observations available on each plane are matched to one or more tracks, as described in Section 7. Then *backward iteration* takes place to refine the reconstructed tracks using all the available data. A detailed description of this algorithm can be found in Särkkä (2013).

6.1. Forward Iteration

The forward iteration is the classic Kalman Filter algorithm; its goal is to compute the posterior distribution of the particle state $\boldsymbol{x}_k$ conditioned on the initial state and the measurements taken on the traversed planes of the detector

$$p(\boldsymbol{x}_k|\boldsymbol{x}_0, \boldsymbol{y}_1,\ldots,\boldsymbol{y}_k) = p(\boldsymbol{x}_k|\boldsymbol{x}_0, \boldsymbol{y}_{1:k}).$$

The state is distributed as a Gaussian random vector $\boldsymbol{x}_k \sim N(\boldsymbol{m}_k, \boldsymbol{P}_k)$ with $\boldsymbol{m}_k$ and $\boldsymbol{P}_k$ the mean and the covariance, respectively, that are estimated at tray k using all the observations from tray 1 to tray k. Then, the parameters of the state distribution at the next tray $k+1$ can be computed with the following *state prediction* step:

$$\begin{aligned} \boldsymbol{m}_{k+1}^- &= \boldsymbol{A}_k \boldsymbol{m}_k \\ \boldsymbol{P}_{k+1}^- &= \boldsymbol{A}_k \boldsymbol{P}_k \boldsymbol{A}_k^T + \boldsymbol{Q}_k. \end{aligned} \tag{14}$$

When tray k is made up of several layers $L_{k,1},\ldots,L_{k,M_k}$, as in Figure 6, the prediction step in Equation (14) shall be repeated for each layer computing $\boldsymbol{A}_{k,i}$ and $\boldsymbol{Q}_{k,i}$ with $i = 1,\ldots,M_k$ according with the specific material. The repeated application

of propagation steps results in the following equations

$$\begin{aligned} \boldsymbol{A}_k &= \prod_{i=1}^{M_k} \boldsymbol{A}_{k,i} \\ \boldsymbol{B}_i &= \prod_{j=M_k-i}^{M_k} \boldsymbol{A}_{k,j} \\ \boldsymbol{Q}_k &= \boldsymbol{Q}_{k,M_k} + \sum_{i=1}^{M_k-1} \boldsymbol{B}_i \boldsymbol{Q}_{k,i} \boldsymbol{B}_i^T. \end{aligned} \tag{15}$$

Then the parameters of the measurements distribution are computed with the following *measurements prediction* step:

$$\begin{aligned} \boldsymbol{v}_{k+1} &= \boldsymbol{H}_{k+1} \boldsymbol{m}_{k+1}^- \\ \boldsymbol{S}_{k+1} &= \boldsymbol{H}_{k+1} \boldsymbol{P}_{k+1}^- \boldsymbol{H}_{k+1}^T + \boldsymbol{R}_{k+1}. \end{aligned} \tag{16}$$

Finally the parameters of the state distribution given the new measurements at tray $k+1$ can be computed with the following *update* step:

$$\begin{aligned} \boldsymbol{e}_{k+1} &= \boldsymbol{y}_{k+1} - \boldsymbol{v}_{k+1} \\ \boldsymbol{U}_{k+1} &= \boldsymbol{P}_{k+1}^- \boldsymbol{H}_{k+1}^T \boldsymbol{S}_{k+1}^{-1} \\ \boldsymbol{m}_{k+1} &= \boldsymbol{m}_{k+1}^- + \boldsymbol{U}_{k+1} \boldsymbol{e}_{k+1} \\ \boldsymbol{P}_{k+1} &= \boldsymbol{P}_{k+1}^- - \boldsymbol{U}_{k+1} \boldsymbol{S}_{k+1} \boldsymbol{U}_{k+1}^T, \end{aligned} \tag{17}$$

where in this specific case matrix $\boldsymbol{A}_k$ can change on each step to reflect different thickness of the detector layers while $\boldsymbol{H}$ is constant.

6.2. Backward Iteration

The goal of the backward interaction is to compute the marginal posterior distribution of state $\boldsymbol{x}_k$ conditioned on the initial conditions and the whole measurements taken on all the planes of the detector

$$p(\boldsymbol{x}_k | \boldsymbol{x}_0, \boldsymbol{y}_{1:K}).$$

The computation of the smoothed state distribution starts at the last plane K and moves backward until the first traversed plane:

$$\begin{aligned} \boldsymbol{G}_k &= \boldsymbol{P}_k \boldsymbol{A}_k^T (\boldsymbol{P}_{k+1}^-)^{-1} \\ \boldsymbol{m}_k^s &= \boldsymbol{m}_k + \boldsymbol{G}_k (\boldsymbol{m}_{k+1}^s - \boldsymbol{m}_{k+1}^-) \\ \boldsymbol{P}_k^s &= \boldsymbol{P}_k + \boldsymbol{G}_k (\boldsymbol{P}_{k+1}^s - \boldsymbol{P}_{k+1}^-) \boldsymbol{G}_k^T. \end{aligned} \tag{18}$$

7. Data Association and Gating

Data association is performed only on the layers with sensitive strips. It takes place during the forward interaction of the estimator (Section 6.1) when each available particle hit is compared with the existing tracks and used to update or duplicate them, as described in Section 2.3.2. Having predicted the particle state at plane $k+1$ (Equation (14)), the distribution of the measurements computed by Equation (16) yields an area, in the tracker plane, where an observation is expected. This area is called *validation gate*. The validation gate is usually defined in terms of the Mahalanobis distance between vector $\boldsymbol{x}$ and a Gaussian distribution with mean $\boldsymbol{m}$ and covariance $\boldsymbol{P}$.

$$d(\boldsymbol{x}, \boldsymbol{m}, \boldsymbol{P}) = \sqrt{(\boldsymbol{x} - \boldsymbol{m})^T \boldsymbol{P}^{-1} (\boldsymbol{x} - \boldsymbol{m})}. \tag{19}$$

Considering the observation model of Equation (13), for each available measurement y_{k+1}, the Mahalanobis distance takes the simple form

$$d(y_{k+1}, v_{k+1}, S_{k+1}) = \frac{|y - v_{k+1}|}{\sqrt{S_{k+1}}}. \tag{20}$$

Thresholding over d distance results in an isoprobability contour. Hence, it is possible to quantify how likely measurements of the particle position fall in the validation gate specifying a desired significance level. Typical significance values are 0.95 or 0.99, and the actual value of d can be computed using the inverse cumulative distribution function of a χ^2 random variable with 1 degree of freedom. An observation is considered compatible with a particle track only if this quantity is below a given threshold d. If several measurements are compatible with a track, then the one with minimum distance is selected for the update.

8. Goodness of Fit

The smoothed track (Equation (18)) is then used to compute the goodness of fit. This enables the comparison of different tracks and the evaluation of the tracker performances. The goodness of fit is based on the distance between the particle position tracked by the smoother and the position measured by the detector. The distribution of the particle position can be derived as follows

$$\begin{aligned} \boldsymbol{v}_k^s &= \boldsymbol{H}_k \boldsymbol{m}_k^s \\ \boldsymbol{S}_k^s &= \boldsymbol{H}_k \boldsymbol{P}_k^s \boldsymbol{H}_k^T + \boldsymbol{R}_k: \end{aligned} \tag{21}$$

then, for each plane where a measurement has been associated to the track, the tracking error reads as

$$\boldsymbol{e}_k^s = \boldsymbol{y}_k - \boldsymbol{v}_k^s, \tag{22}$$

and the goodness of fit is represented by the following statistic

$$\chi^2 = \sum_{k=1}^{K} (\boldsymbol{e}_k^s)^T \boldsymbol{S}_k^s \boldsymbol{e}_k^s. \tag{23}$$

It can be proved that Equation (23) is distributed as a χ^2 random variable. It must be noted that the evaluation of Equation (23) is possible only on the planes where a measurement has been associated to the track (see Section 7).

9. Energy Reconstruction

Estimating the energy of a reconstructed track from its geometry is very important because energy is used to: (i) select the primary and secondary tracks, (ii) match 2D tracks across different detector faces to get the 3D profile, (iii) compute the photon incoming direction combining primary and secondary tracks directions.

The particle energy is correlated with the angular deviation of the track using the Moliere formula in Equation (5). Clearly the thickness of traversed material has to be corrected for the incidence angle writing

$$\lambda(\theta) = \frac{\lambda}{\cos(\theta)} = \frac{z}{X_0 \cos(\theta)}. \tag{24}$$

If the energy deposition on each layer is negligible with respect to the total particle energy, the angular deviation of the particle between two consecutive trays, as shown in Figure 7, is

approximated with the following Gaussian distribution

$$\theta_{n+1} - \theta_n \sim N\left(0, \left(\frac{13.6}{E_e}\right)^2 \frac{\lambda_n}{\cos(\theta_n)} \times \left(1 + 0.038 \ln\left(\frac{\lambda_n}{\cos(\theta_n)}\right)\right)^2\right); \quad (25)$$

the variance of this distribution can be expressed as the product of two factors σ_e^2 and ϕ_n^2, where $\sigma_e = 13.6/E_e$ while ϕ_n collects all the tray-dependent terms. Hence, we define the normalized angular deviation as follows

$$\Delta\theta_n = \frac{\theta_{n+1} - \theta_n}{\phi_n} \sim N(0, \sigma_e^2). \quad (26)$$

In the official AGILE pipeline (Giuliani et al. 2006), the left-hand side of Equation (26) is computed for each tray of the detector using the filtered or the smoothed-particle track. Then, the event energy is estimated taking the sample variance of these terms. Note that this holds both for 2D and 3D tracks up to a $\sqrt{2}$ scale factor (Particle Data Group et al. 2020). In this work we added a further step, and the raw estimate of the event energy is refined through a maximum-likelihood method. Angular deviations due to small-angle scattering can be considered statistically independent and identically distributed. Hence the corresponding total probability of having observed a sequence of angular deviations and the corresponding negative-log-likelihood function can be written as

$$p(\Delta\theta_1,\ldots,\Delta\theta_{n-1}|\sigma_e) = \prod_{i=1}^{n-1} p(\Delta\theta_i|\sigma_e) \quad (27)$$

$$-L(\sigma_e|\Delta\theta_1,\ldots,\Delta\theta_{n-1}) = -\sum_{i=1}^{n-1} \log(p(\Delta\theta_i)|\sigma_e) = \sum_{i=1}^{n-1}\left[\log(\sigma_e) + \frac{1}{2}\left(\frac{\Delta\theta_i}{\sigma_e}\right)^2\right]. \quad (28)$$

The refined value of σ_e can be obtained taking the minimum of the function in Equation (28) by means of a nonlinear minimization algorithm using the raw estimate from Equation (26) as the initial condition. In our experience, a fixed number (3 to 5) of simple Newton iterations is enough to get the final result.

10. Event Tracking Simulation

10.1. Simulation Framework

The interaction of γ-rays with the trackers under study is simulated with the BoGEMMS (Bulgarelli et al. 2012) framework, based on the latest version of the Geant4 (Agostinelli et al. 2003; Allison et al. 2006, 2016) particle transport code. The Geant4 toolkit, initially developed by CERN and then maintained by a large international collaboration, is a C++ Monte Carlo library for the simulation of particle interaction with matter. Built on top of the Geant4 library, BoGEMMS is an astronomy-oriented simulation tool for the evaluation of the instrument performance (e.g., background level, angular resolution, effective area) of X-ray and γ-ray space telescopes. Its modular architecture allows to use the same software to build any virtual model of the telescope and easily update its design. The BoGEMMS γ-ray module (Fioretti et al. 2014) is the base for AGILESim, the AGILE/GRID payload simulator recently validated against both laboratory and in-flight data (Fioretti et al. 2020). The same framework is used to build the mass model of the e-ASTROGAM and All-Sky-ASTROGAM electron–positron trackers. The Geant4 10.4 release and the *G4EmLivermorePolarizedPhysics* electromagnetic physics list are used throughout the work (Ivanchenko et al. 2019).

10.2. Tracker Mass Models

The AGILE-GRID tracker mass model is made up of 12 *X*–*Y* Si planes, consisting of two separate 410 μm thick layers of orthogonal single-sided strips. Each layer is divided into 3072 physical strips with a pitch of 121 μm, for a total side of $\sim$37 cm. Each tray of the tracker is composed by $\sim$1.5 cm of aluminum honeycomb enclosed by two layers of carbon fiber. The tungsten converter, with a thickness of 245 μm, is present in the first 10 trays (sky side; Bulgarelli et al. 2010).

In order to lower the minimum energy threshold, the e-ASTROGAM and All-Sky-ASTROGAM tracker design presents several changes with respect to AGILE. The lack of a tungsten converter and the minimization of the passive material in the supporting structure decrease the multiple scattering of e^-/e^+, which limits the tracking ability at low energies. The increase in the Si planes counterbalances the lack of tungsten to maintain the required interaction efficiency. A double-sided silicon strip detector (DSSD), in place of two single-sided X and Y silicon layers, enables the reconstruction of Compton events and low-energy pair production by increasing the spatial resolution.

The e-ASTROGAM tracker general design is described in Figure 8 (left panel). It is composed by 56 DSSD planes with a pitch of 1 cm. The silicon strips composing the planes have a pitch of 240 μm and a thickness of 500 μm. Since a minimum amount of passive material is foreseen in the trays, we assume that its impact in attenuating the particle path is negligible and no passive material is included in the present study. The All-Sky-ASTROGAM proposal reduces the mass budget of the tracker by including only 25 DSSD planes, each composed by 1152 strips in total.

10.3. Simulation, Post-processing, and Verification

In the AGILE-GRID simulation, the aim is comparing the performance of the present Bayesian method with the mission Kalman filter used by the on-board and on-ground official scientific pipeline. For this reason, we used in input parallel monoenergetic photon beams at 50, 100, 200, 400, and 1000 MeV as performed in Fioretti et al. 2020. For both e-ASTROGAM and All-Sky-ASTROGAM proposals the reconstruction algorithm is tested using parallel monoenergetic photon beams from the pair-production regime limit, from 10 MeV up to 3000 MeV. A minimum energy threshold of 15 keV is applied to both single- and double-sided strips in order to generate a trigger. The electronic cloud generated in the strip spreads during its travel in the electric field, and it can be measured by more than one electrode. This phenomenon, defined as charge sharing or capacitive coupling, is added in the post-processing of the simulation output by distributing the energy deposit following a Gaussian profile for the double-sided strips or the GRID measured distribution (Barbiellini et al. 2002). All sets of contiguous strips, generated by the

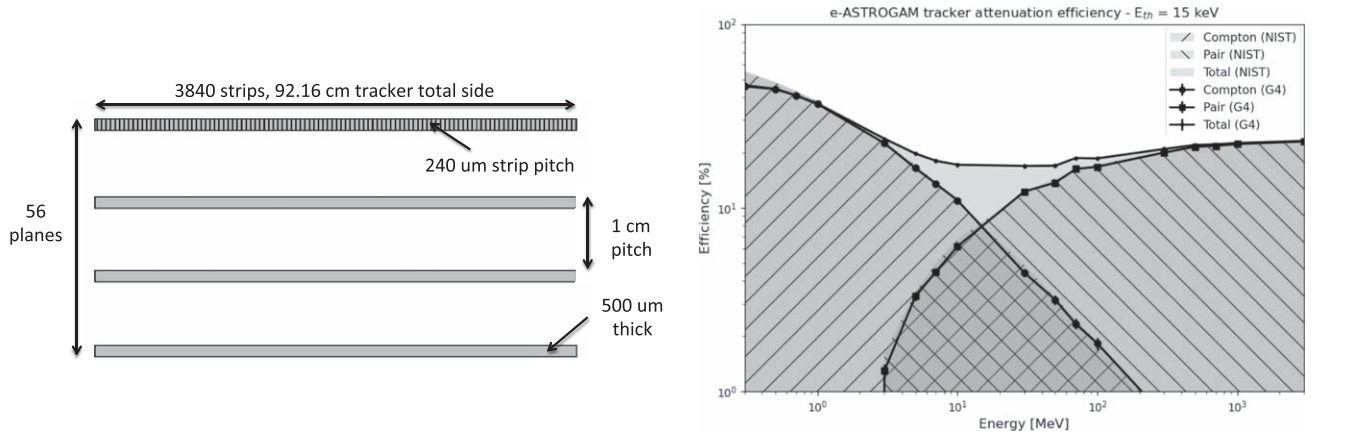

Figure 8. Left panel: simplified configuration of the tracker on board e-ASTROGAM. Right panel: comparison between the tracker simulated and tabulated attenuation efficiency from Compton scattering, pair production, and their combination.

same primary γ-ray, are then grouped into clusters, with the sum of the strip energy as the energy and the mean of the strip positions weighted by their energy as the barycenter.

The correct implementation of the Si tracker is verified by computing its attenuation efficiency toward monoenergetic beams. Figure 8 shows the verification applied to the e-ASTROGAM mass model in the 300 keV–3 GeV energy range by simulating a beam pointed at a 30° off-axis angle from the tracker normal. For each beam, we are able to discriminate simulated Compton scatterings from pair conversions (the few Rayleigh scatterings expected below 1 MeV are neglected). The number of detected interactions is divided by the input number of simulated particles to compute the attenuation efficiency shown by markers in Figure 8 (right panel). The Si attenuation coefficients μ tabulated by the National Institute of Standards and Technology (NIST; Berger & Hubbell 1987) are used in the Beer–Lambert formula to compute the expected attenuation efficiency provided by the total tracker thickness t, in the form $100 \times (1 - e^{-\mu t})$. The simulation well reproduces the analytical efficiency curves for both the selected processes and their sum. The lower efficiency obtained in the Compton regime below ~500 keV is due to a fraction of Compton scattering recoil energies falling below the energy threshold applied to the strips. This comparison validates the physics models implemented in the Geant4 simulation, the tracker mass model, and the post-processing algorithms applied to the simulation output. The list of cluster energies and positions, with an associated plane identification number, for each input energy and direction of the primary γ-ray beam is provided as input to the tracking algorithm.

All mass models include a scintillation calorimeter at the tracker bottom, electronics panels and anticoincidence scintillation panels surrounding the side and top of the tracker. The AGILE calorimeter consists of 30 Csi(Tl) bars, $15 \times 23 \times 375$ mm^3 each, organized in two orthogonal sets of 15 rows. Both the e-ASTROGAM and All-Sky-ASTROGAM calorimeters instead comprise one plane of vertical bars, 92 and 28 for each side, respectively, and a lateral size of 1 cm each. Silicon drift photodetectors (SDDs) reading out the signal at both ends of the CsI crystals are simulated by a 0.3 mm thick Si layer and a 1 mm thick circuit board.

11. Tracker Performance

11.1. Performance Definition

The performance of the algorithm described above is evaluated in terms of efficiency, angular resolution, and sensitivity. Also the event energy reconstructed using only the tracks geometry is considered. For AGILE, this the only mean of estimating the event energy, while for e-ASTROGAM and All-Sky-ASTROGAM this has to be considered as an additional figure of merit of the method that does not replace the information from the calorimeter.

The efficiency of the reconstruction method is simply defined as the ratio between the number of reconstructed events and the total number of processed events. The reconstruction can fail because of several issues. Some events are not analyzed because they do not satisfy the selection criteria described in Section 2.2; in other cases the tracker fail to reconstruct or to identify the primary or the secondary track (see Section 2.4). Sometimes the reconstructed track (both 2D or 3D) is too short to provide a reliable estimate of the energy, and hence the final event direction cannot be computed. The effective area is defined as the geometrical area of the tracker multiplied by its total efficiency given by the sum of the attenuation and reconstruction efficiencies.

The angular resolution represents the ability of a telescope to separate two γ-ray sources close to each other. The angular resolution of the telescope is defined considering the spread of the reconstructed γ-ray directions around the true one and then computing the radius containing the 68% of the events; the uncertainty on the containment radius is only statistical and derived from the Poisson fluctuation on the number of simulated events.

The sensitivity, or the minimum detectable flux, of a mission is a key performance descriptor because it depends on technological parameters such as the angular resolution, the effective area, and the instrumental background, linking them with the scientific goals of the mission (see Fioretti & Bulgarelli 2020 for a review).

11.2. Hyperparameters and Initial Event Energy Setup

The algorithm hyperparameters with major impact on the tracker performances, reported in Table 1, were derived by

manual tuning of the algorithm on a limited number of selected events. Event preselection and track maintenance parameters were defined to guarantee an efficiency above 80%. The efficiency depends on the number of detector sensitive planes and on the probability of associating a hit with a track during the forward pass of the algorithm (Section 2.3.2). In particular, the track duplication significance (Section 7) controls the number of alternative tracks that are generated by the tracker; higher significance means lower probability that a hit and a track are compatible. In this work, the threshold on the track duplication significance for AGILE is smaller than what applied to both All-Sky-ASTROGAM and e-ASTROGAM to increase the probability of identifying tracks on a limited number of planes. For all the detectors, the initial uncertainty in both the position and the direction of the particles is high with respect to the uncertainty in the localization of the hits on the sensitive planes; this is done to avoid the a priori exclusion of possible tracks.

The Moliere uncertainty model (Section 3) has proven to be adequate to guarantee correct particle tracking and was used to compute $\boldsymbol{Q_k}$ matrices in Equations (14) and (15). This requires an estimate of the event energy; for All-Sky-ASTROGAM and e-ASTROGAM, the simulated event energy was used because these two instruments are both equipped with a calorimeter. Instead, for AGILE a different approach was implemented because of the limitations of its mini-calorimeter. In the original work of Giuliani et al. (2006), the track reconstruction is repeated several times using different energy values; then the 3D tracks with lower χ^2 and the related energy level are selected as the output of the algorithm. In this work, when the event energy is not a priori known, the initial energy is fixed to 50 MeV to guarantee the best fitting performances when the tracks are more complex (low energy); at higher energy, this approach tends to increase the number of alternative tracks that are indeed pruned out during track selection and matching process. Note that in any case, after having reconstructed the 3D tracks, the final energy is estimated using only their geometry (Section 9).

Since the hit positions are determined by clustering raw measurements on neighboring strips, the measurement uncertainty $\boldsymbol{R}_k$ in Equation (16) was determined by Monte Carlo analysis on simulated data; the same event was generated several times, and then cluster positions were compared to true particle positions tacking their standard deviation as an estimate of the measurement uncertainty.

11.3. Algorithm Validation Using AGILE Data

The performances of the Bayesian Tracking Algorithm (BTA) presented in this work are evaluated on a set of AGILE/GRID γ-ray simulated events and compared with the results of the official AGILE RECO (Bulgarelli 2019; Bulgarelli et al. 2019). Simulation of the AGILE/GRID data has been performed with the AGILESim payload simulator framework, and in-flight data processing has been simulated with the on-board data handler simulator DHSim (Argan et al. 2004). The validation of the results has been performed by reproducing the real PSF of the instrument for different energy bands, as described in Fioretti et al. (2020). Simulation is carried out on five monoenergetic data sets with 50 MeV, 100 MeV, 200 MeV, 400 MeV, and 1000 MeV energy, selected within reconstructed energy intervals of 30–70 MeV, 70–140 MeV, 140–300 MeV, 300–700 MeV, and 700–1700 MeV, respectively. Each data set is composed by 10,000 γ-ray events and characterized by a fixed photon direction $\theta = 30°$ and $\phi = 225°$ corresponding to the worst case azimuth direction in terms of error projection on the detector axis.

Each simulated event has been processed with the BTA, and the output is the incoming γ-ray direction (θ and ϕ angles) and its energy. The distributions of estimated θ, ϕ reported in Figure 9 are approximately Gaussian and centered on the true values for all the energy levels. This shows the good accuracy of the reconstructed direction and that the BTA does not introduce systematic artifacts on both the θ and ϕ angles. As expected, the precision of the method increases as the energy increases (decreasing θ and ϕ std. deviation). Moreover, the estimated energy is distributed as a Landau probability density function with the most probable value close to the ground truth. The event energy is computed starting from the reconstructed 3D tracks using the maximum-likelihood algorithm in Section 9. In general, the energy is proportional to the discrete derivative of the tracks direction. Hence, as the energy increases, the angular deviations between neighboring planes become smaller and even small errors in the direction estimate result in huge errors in the computed energy. Furthermore, the likelihood function flattens out, and the standard deviation of the estimated energy increases with the energy itself.

Simulated events have also been processed by the AGILE RECO (Bulgarelli 2019), and the outputs are compared with the one of the BTA in terms of efficiency, PSF, and estimated energy. The AGILE RECO includes the track reconstruction module based on a recursive Kalman filter algorithm (Pittori & Tavani 2002; Giuliani et al. 2006) and the FM3.119 on-ground event classification filter (Bulgarelli et al. 2019). Such filter assigns a classification flag to each event depending on whether it is recognized as a γ-ray event, a charged particle, a "single-track" event, or an event of uncertain classification. The interaction between the track reconstruction module and the event classification filter must be taken into account to compare the performances of the two methods and to validate the BTA reproducing the performances of the AGILE RECO presented in Fioretti et al. (2020).

First, the comparison is carried out on all the simulated events without considering the contribution of the FM3.119 filter (see Table 2). The efficiency of the BTA, defined as the ratio between reconstructed and simulated events, is slightly lower than the efficiency of the AGILE RECO because of the event preselection step (not present in the official pipeline). Energy estimates are obtained as the most probable value from the Landau distribution fitting. The results are similar at 50 and 100 MeV; at higher energies, the estimates of the new tracker are better centered with respect to the true value but the standard deviation is higher as discussed above. Regarding the PSF, computed as the analytic 68% containment radius, the BTA outperforms the AGILE RECO for almost all the energy levels; only at 50 MeV the results are equivalent.

Then, the interaction of the BTA with the FM3.119 filter has been investigated (i) to evaluate its performance in a real case scenario and (ii) to compare the results with the mission reference values reported in Fioretti et al. (2020). To this aim, both the energy and PSF are computed considering only the events marked as genuine γ-rays by the FM3.119 filter, and the results are shown in Table 3; since the FM3.119 filter strictly interacts with the AGILE track reconstruction algorithm, the efficiency is not directly comparable and it is not reported in the

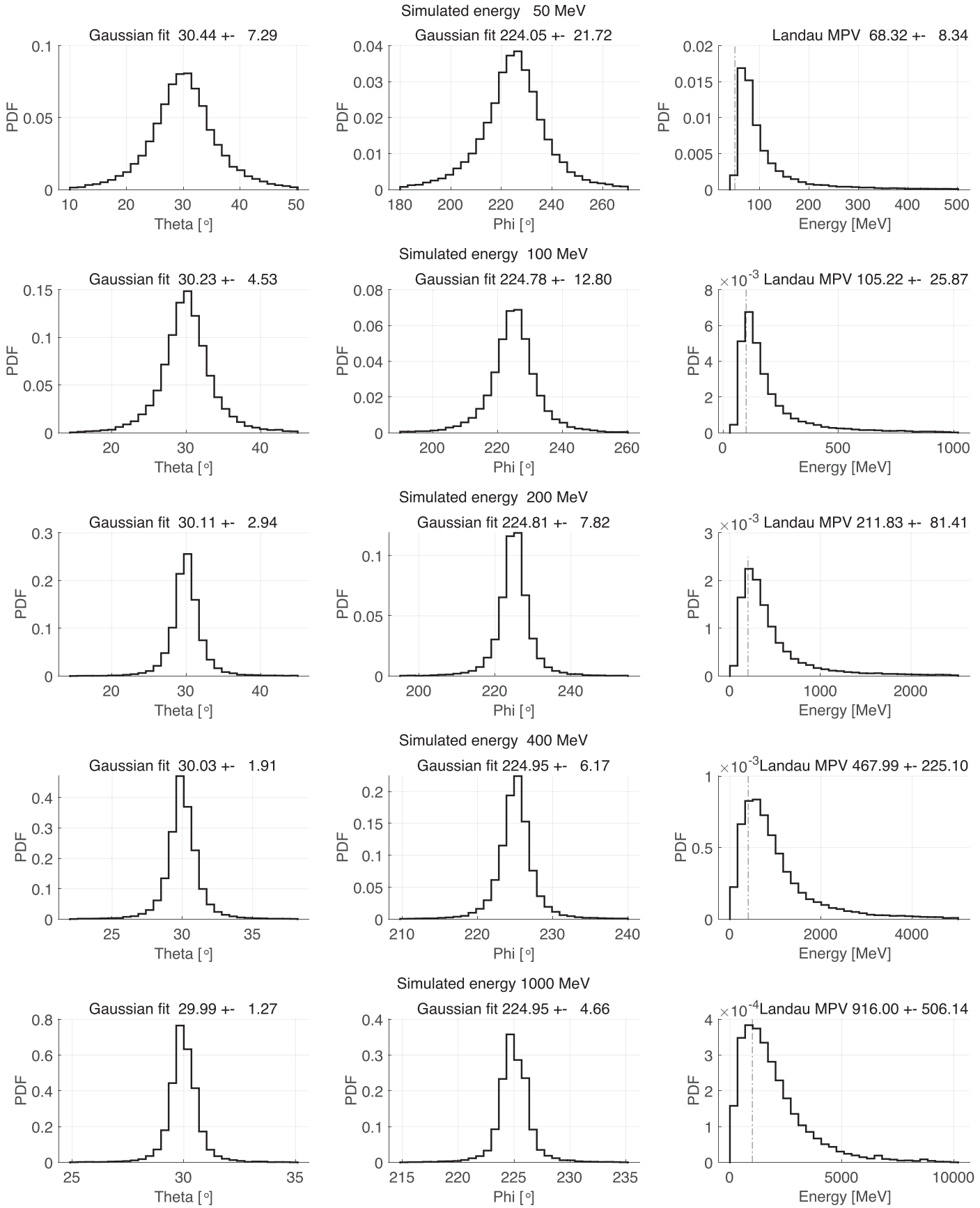

Figure 9. Incoming particle direction and energy distributions computed using the BTA on all the simulated events (see Table 2). Both the θ and ϕ angles are approximately Gaussian and centered on the true values for all the energy levels; the estimated energy is distributed as a Landau probability density function.

Table 2
Comparison of BTA and AGILE RECO Performances Considering all the Simulated γ-ray Events at Different Energy Levels

Energy (MeV)	AGILE RECO			BTA		
	Eff. (%)	Energy (MeV)	PSF (°)	Eff. (%)	Energy (MeV)	PSF (°)
50	98.75	61.00 ± 14.63	9.00 ± 0.06	98.67	68.32 ± 8.34	9.07 ± 0.08
100	99.37	89.00 ± 23.69	5.34 ± 0.04	98.15	105.22 ± 25.87	4.86 ± 0.04
200	99.58	160.00 ± 49.50	3.32 ± 0.04	97.70	211.83 ± 81.41	2.65 ± 0.02
400	99.65	292.00 ± 94.62	2.05 ± 0.03	97.32	467.99 ± 225.10	1.47 ± 0.01
1000	99.57	565.00 ± 147.83	1.14 ± 0.03	97.66	916.00 ± 506.14	0.82 ± 0.04

Note. The efficiency is defined as the ratio between the number of events reconstructed by the algorithms and the total number of simulated events; the PSF is defined as the analytic 68% containment radius, and the energy estimate is the most probable value from the Landau distribution fitting. All the errors are expressed as 1σ standard deviation.

Table 3
Comparison of BTA and AGILE RECO Performances Reported in Fioretti et al. (2020) Considering Only the Events Marked as Genuine γ-ray by the AGILE On-ground Event Classification Filter FM3.119

Energy (MeV)	AGILE RECO and FM3.119		BTA and FM3.119	
	Energy (MeV)	PSF (°)	Energy (MeV)	PSF (°)
50	57.12 ± 14.35	7.93 ± 0.13	50.55 ± 14.70	7.91 ± 0.11
100	99.00 ± 21.86	4.19 ± 0.06	102.08 ± 27.24	3.71 ± 0.07
200	166.50 ± 40.51	2.64 ± 0.04	208.18 ± 52.82	2.18 ± 0.03
400	290.93 ± 85.69	1.19 ± 0.02	426.02 ± 164.38	1.31 ± 0.02
1000	561.00 ± 140.24	0.58 ± 0.01	929.00 ± 383.14	0.65 ± 0.04

Note. The PSF is defined as the 68% containment radius, and the energy estimate is the most probable value from the Landau distribution fitting. All the errors are expressed as 1σ standard deviation.

table. The computed PSFs are similar, showing that the BTA is indeed able to reproduce the performances of the AGILE RECO. Moreover, these results suggest that dedicated fine tuning of the hyperparameters and proper integration with the event classification filter could improve the overall detector performances.

To investigate whether fine tuning of the hyperparameters could further improve the instrument PSF, the estimated PSF reported in Table 3 was compared with a lower bound on the angular resolution defined considering three terms: the detector characteristics, the small-angle scattering process, and the unknown recoil of the nucleus.

The PSF limits due to the detector characteristics depends on the ratio of the spatial resolution to the distance of consecutive trays (Fedel et al. 2000; Barbiellini et al. 2002; Sabatini et al. 2015) and reads as

$$\sigma_d = \tan^{-1}\left(\frac{d}{z}\right), \tag{29}$$

where d is the spatial resolution of the instrument, and z is the distance between trays.

The limit due to the unknown recoil of the nucleus is approximated as the ratio between the energy transfer to the nucleus E_n and the total energy of the γ-ray E_γ (Zoglauer 2005). The most probable energy transfer to the nucleus is close to the energy at rest of the electron $E_n \approx E_0$; hence the corresponding term is

$$\sigma_r = \tan^{-1}\left(\frac{E_0}{E_\gamma}\right). \tag{30}$$

The small-angle scattering term σ_θ is computed using Equation (5) and assuming that the pair production happens on the top interface of each layer and the energy is evenly shared between the e^- and the e^+. The overall angular resolution limit is derived combining quadratically the terms in Equations (29) and (30) with the the small-angle scattering of Equation (5)

$$\sigma_t = \sqrt{\sigma_d^2 + \sigma_r^2 + \sigma_\theta^2}. \tag{31}$$

The AGILE PSF considering all the simulated γ-ray events and the lower bound computed with Equation (31) are compared in Figure 10. For AGILE the limit on the angular resolution is dominated by the multiple Coulomb scattering (Moliere) in the whole energy range. Note that in Figure 10 the Moliere term coincides with the combined limit. In general, the PSF is greater than the lower bound for all the energy levels. This means that further optimization is possible even if it is beyond the scope of this work. The AGILE detector was designed to guarantee a spatial resolution of 40 μm for a broad energy range (Barbiellini et al. 2002; Bulgarelli et al. 2010). Nevertheless, the spatial resolution depends on the event energy and, for a small azimuth $\theta \leqslant 30°$, it could be greater than 110 μm, as reported in Barbiellini et al. (2002). Hence this latter value was used in the analysis; note that if 40 μm is considered then the contribution is negligible if compared with the other terms.

The AGILE point source sensitivity in the simulated energies is computed for extragalactic observations (1 Ms of exposure, 3σ detection significance) using as input the nominal and the current PSF and total tracker efficiencies. The background rate used for the sensitivity curves is obtained from the diffuse γ-ray emission spectrum of Abdo et al. (2010). An in-depth performance study of the tested missions is out of the scope of the present work, and we neglect any background contribution from cosmic rays, which would require a complete

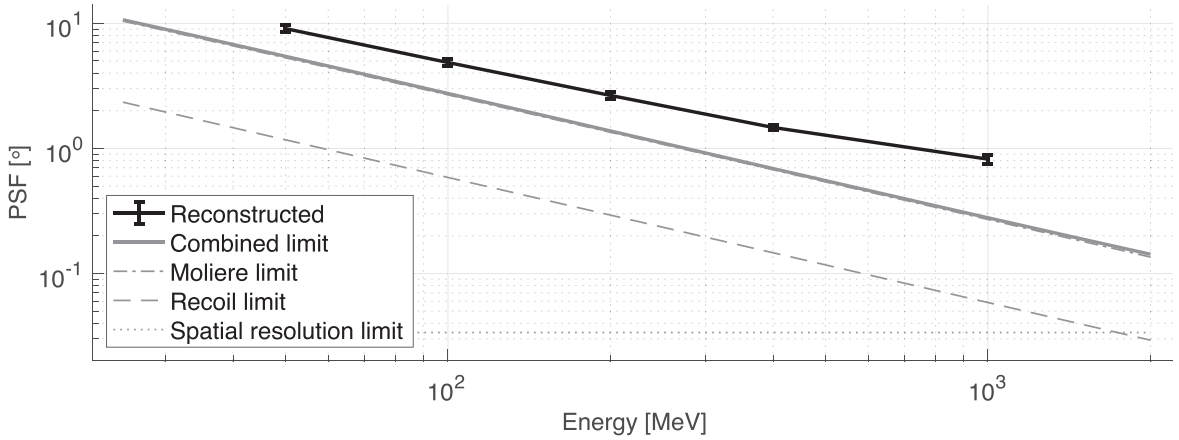

Figure 10. AGILE PSF (from Table 2) compared with the angular resolution lower bound computed considering the multiple Coulomb scattering, the unknown recoil of the nucleus, and the detector spatial resolution. In this plot, a spatial resolution limit of 110 μm has been considered. The main contribution to the total limit for AGILE is the multiple Coulomb scattering for all the considered energy levels; in the figure, the Moliere term and the combined limit are almost overlapping.

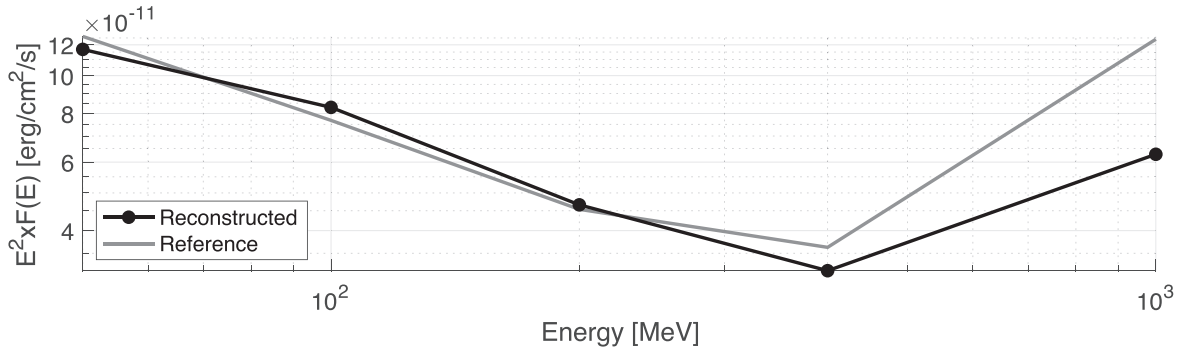

Figure 11. Sensitivity of AGILE detector computed with the algorithm presented in this work (reconstructed) compared with the nominal values. The plots show that the overall sensitivity at higher energy is improved.

Table 4
e-ASTROGAM Efficiency, PSF, and Energy Obtained on Simulated Data as a Function of the Event Energy

Energy (MeV)	Efficiency (%)	PSF (°)	Rec. Energy (MeV)
10	84.28	8.66 ± 1.13	19.35 ± 5.11
30	85.12	4.12 ± 0.54	48.21 ± 9.34
50	90.70	2.89 ± 0.43	54.32 ± 10.25
70	91.35	1.92 ± 0.18	73.86 ± 14.51
100	93.49	1.32 ± 0.15	118.99 ± 94.86
300	93.80	0.57 ± 0.09	327.11 ± 108.63
500	94.01	0.35 ± 0.07	662.64 ± 237.00
700	94.40	0.27 ± 0.03	1184.53 ± 628.15
1000	95.33	0.18 ± 0.04	1815.99 ± 780.19
3000	95.84	0.13 ± 0.03	2217.49 ± 1106.98

Table 5
All-Sky-ASTROGAM Efficiency, PSF, and Energy Obtained on Simulated Data as a Function of the Event Energy

Energy (MeV)	Efficiency (%)	PSF (°)	Rec. Energy (MeV)
10	83.10	14.15 ± 4.15	18.05 ± 2.24
30	84.35	7.22 ± 2.06	38.80 ± 8.43
50	87.31	3.18 ± 0.70	48.29 ± 12.13
70	88.84	2.36 ± 0.55	71.00 ± 15.09
100	90.03	1.44 ± 0.32	131.41 ± 99.70
300	90.94	0.91 ± 0.19	332.75 ± 124.10
500	91.00	0.60 ± 0.15	647.99 ± 219.02
700	83.09	0.56 ± 0.11	1270.02 ± 601.92
1000	91.71	0.38 ± 0.10	1641.39 ± 660.72
3000	91.69	0.28 ± 0.08	2358.60 ± 1024.00

simulation of the particle environment interaction with spacecraft and telescopes. The sensitivity of the AGILE detector when using the new filter is comparable to the reference one for energy below 400 MeV while it is improved at higher energy, as shown in Figure 11.

11.4. e-ASTROGAM and All-Sky-ASTROGAM Tracker Performances

After the validation, the tracker is used to process both the e-ASTROGAM and All-Sky-ASTROGAM simulated data and to investigate the expected performances of such detectors. The data set is composed by monoenergetic events in the range from 10 MeV to 3 GeV with fixed incoming direction $\theta = 30°$ and $\phi = 225°$, as was done for AGILE. The tracker configuration settings are reported in Table 1; the same parameters are used for both the detectors, and no fine tuning is performed. To carry out a worst-case analysis on the detector performance and to identify a set of parameters of general applicability (avoiding hyperparameters optimization), the preselection parameters are pretty stringent, requiring events to have at least 10 consecutive hits on each face of the detector. The algorithm efficiency is better than 84% for both detectors and for each energy. The simulated data at low energy contain a variable amount of

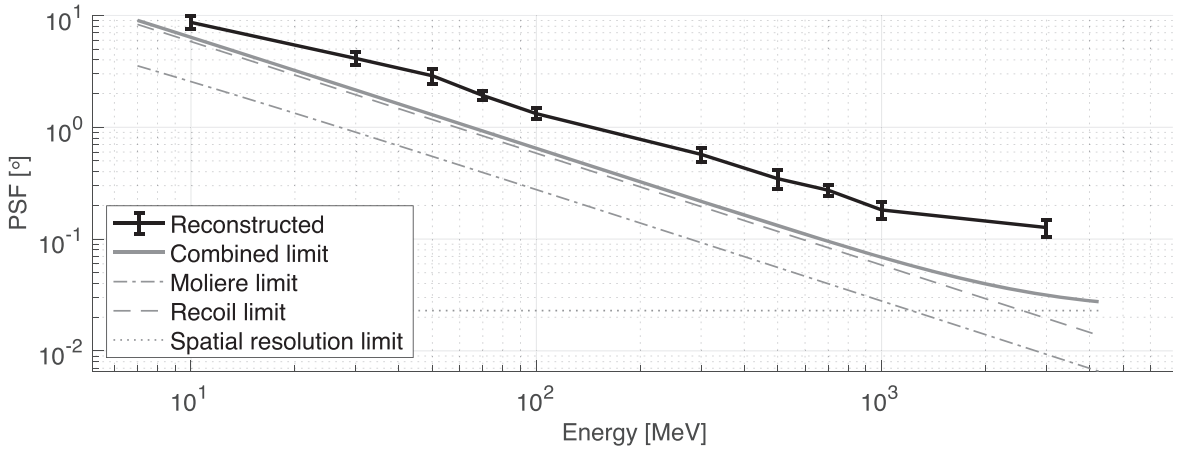

Figure 12. Estimated angular resolution of e-ASTROGAM compared with the theoretical limits computed assuming 40 μm spatial resolution.

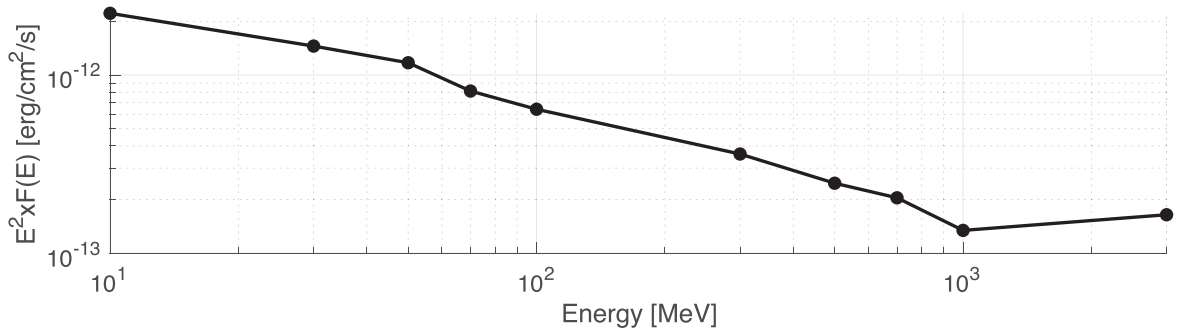

Figure 13. Sensitivity of e-ASTROGAM detector computed with the algorithm presented in this work.

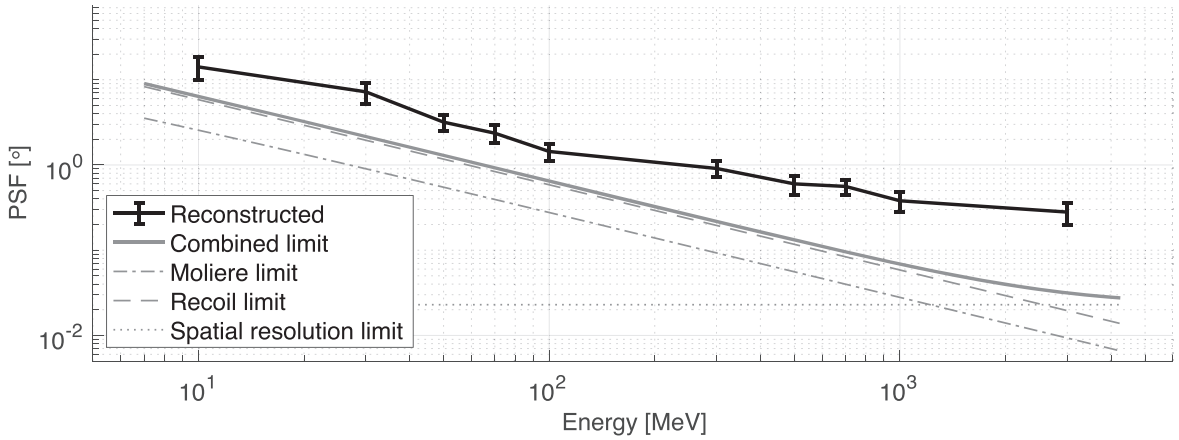

Figure 14. Estimated angular resolution of All-Sky-ASTROGAM compared with the theoretical limits computed assuming 40 μm spatial resolution.

Compton events; often the tracker handles such events correctly but sometimes it is not able to properly reconstruct the event and this results in a reduced efficiency even below 90% (see Tables 4 and 5). As expected, the efficiency tends to improve as the energy increases; nevertheless, at high energies it is limited by the event preselection algorithm. The PSF, effective area, and sensitivity in the pair regime for the e-ASTROGAM and All-Sky-ASTROGAM mission proposals (De Angelis et al. 2018; Tatischeff et al. 2019) were produced with a preliminary release of the present reconstruction algorithm. The e-ASTROGAM and All-Sky-ASTROGAM PSF values presented in Tables 4 and 5, obtained with the final code release, are in good agreement with the ones published in the mission proposal. To define the lower bound on the PSF a spatial resolution of 40 μm is used, corresponding to a $1/6$ strip pitch as reported in De Angelis et al. (2018). The reconstructed PSF in Figure 12 is close to the limit on the angular resolution even if the tracker is not optimized for a such detector, showing the overall goodness of the tracking algorithm. The accuracy of the estimated energy is good up to about 500 MeV; for higher energy levels, both the accuracy and the precision of the estimate are degraded, as already discussed in Section 11.3.

Similar conclusions about the efficiency, PSF, and estimated energy can also be drawn from the analysis of All-Sky-ASTROGAM events. The All-Sky-ASTROGAM detector and

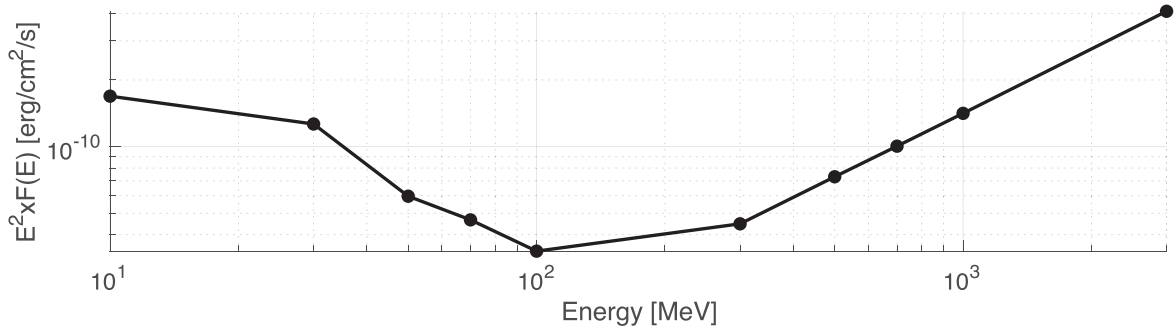

Figure 15. Sensitivity of the All-Sky-ASTROGAM detector computed with the algorithm presented in this work.

its scientific goals are presented in Tatischeff et al. (2019); the results in Table 5 provide further characterization of such a detector. Differences between the performance of All-Sky-ASTROGAM and e-ASTROGAM are due to the difference between the two detectors; in particular the reduced number of layers of All-Sky-ASTROGAM (25 instead of 56) resulted in lower efficiency and increased PSF.

The e-ASTROGAM and All-Sky-ASTROGAM point source sensitivity in the pair regime for extragalactic observations (1 Ms of exposure, 3σ detection significance) is shown in Figures 12, 13, 14, and 15 using the PSF and total efficiency values listed in Tables 4 and 5, respectively. As for the AGILE sensitivity, the background rate only includes the diffuse γ-ray emission, as modeled by Abdo et al. (2010).

12. Conclusions

The overall performance of a γ-ray telescope depends on both the physical characteristics of the instrument and the tracking algorithms used to reconstruct the observed event. In this work, a very general and flexible event reconstruction algorithm for the pair-production regime has been described. The proposed method is based on a Bayesian smoother embedding a simple model of the motion of the particles inside the tracker, and the Moliere formula is used to account for multiple Coulomb scattering effects.

The algorithm was validated using AGILE simulated events and the reconstruction accuracy in terms of the efficiency, angular resolution, energy estimation, and sensitivity was evaluated. The tracker alone was able to outperform the performance of the official AGILE algorithm and, if integrated into the in-flight data processing pipeline, it is able to reproduce the PSF of the AGILE telescope as derived using real mission data. Moreover, the comparison of the reconstructed PSF with the theoretical lower bound on the telescope angular resolution suggests that, with dedicated hyperparameter tuning, the overall performances may be improved.

After the validation, the algorithm was used to assess the expected performance of both the e-ASTROGAM and All-Sky-ASTROGAM γ-ray proposed telescopes with broad energy range and improved sensitivity with respect to its predecessors. For both telescopes, a single set of tracker configuration parameters was used. This work confirms the expected performances of the e-ASTROGAM telescope in the energy range from 30 MeV to 3 GeV, and the reconstructed PSF of All-Sky-ASTROGAM contributes to better define the performance of such instrument. For both e-ASTROGAM and All-Sky-ASTROGAM, the gap between the reconstructed PSF and the theoretical limit indicates that the overall performances of such instruments can be improved properly tuning the algorithm hyperparameters; this holds in particular in the GeV range.

The proposed algorithm and the described analysis method are very general and can be tailored to process already available data from existing instruments as well as to support the design of new instruments for γ-ray astronomy research.

In the near future, we plan to extended the algorithm working in the following directions: (i) reduce the gap between the estimated PSF and the theoretical limit testing different strategies for the tracks selection and matching step, (ii) increase the accuracy of the energy reconstruction embedding the energy estimation in the Bayesian filter by means of expectation-maximization algorithm, (iii) further investigate the topology of the secondary scattering pairs within the same event to extend the reconstruction in the energy range from few to tens of MeV, where both pair production and Compton effect overlap.

ORCID iDs

Alessio Aboudan https://orcid.org/0000-0002-8290-2184
Andrea Bulgarelli https://orcid.org/0000-0001-6347-0649
Valentina Fioretti https://orcid.org/0000-0002-6082-5384
Marco Tavani https://orcid.org/0000-0003-2893-1459

References

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, PhRvL, 104, 101101
Agostinelli, S., Allison, J., Amako, K., et al. 2003, NIM A, 506, 250
Allison, J., Amako, K., Apostolakis, J., et al. 2006, ITNS, 53, 270
Allison, J., Amako, K., Apostolakis, J., et al. 2016, NIMPA, 835, 186
Argan, A., Tavani, M., Giuliani, A., et al. 2004, in IEEE Symp. Conf. Record Nuclear Science 2004 (Piscataway, NJ: IEEE), 371
Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, ApJ, 697, 1071
Barbiellini, G., Fedel, G., Liello, F., et al. 2002, NIMPA, 490, 146
Berger, M., & Hubbell, J. 1987, XCOM: Photon cross sections on a personal computer NBSIR 87-3597, National Bureau of Standards
Bernard, D., & Frosini, M. 2017, arXiv:1710.10886
Bethe, H. A. 1953, PhRv, 89, 1256
Billoir, P. 1984, NIMPR, 225, 352
Billoir, P., Frühwirth, R., & Regler, M. 1985, NIMPA, 241, 115
Bulgarelli, A. 2019, ExA, 48, 199
Bulgarelli, A., Argan, A., Barbiellini, G., et al. 2010, NIMPA, 614, 213
Bulgarelli, A., Fioretti, V., Malaguti, P., et al. 2012, Proc. SPIE, 8453, 845335
Bulgarelli, A., Fioretti, V., Parmiggiani, N., et al. 2019, A&A, 627, A13
Cumani, P. 2015, PhD thesis, Università degli studi di Trieste
De Angelis, A., Tatischeff, V., Grenier, I., et al. 2018, JHEAp, 19, 1
Fedel, G., Barbiellini, G., Liello, F., et al. 2000, Proc. SPIE, 4140, 274
Fioretti, V., & Bulgarelli, A. 2020, in Fundamental Concepts; Data Reduction and Analysis, ed. C. Bambi (Berlin: Springer), 55
Fioretti, V., Bulgarelli, A., Tavani, M., et al. 2020, ApJ, 896, 61
Fioretti, V., et al. 2014, Proc. SPIE, 9144, 91443N
Frosini, M., & Bernard, D. 2017, NIMPA, 867, 182

Frühwirth, R. 1987, NIMPA, 262, 444
Giuliani, A., Cocco, V., Mereghetti, S., Pittori, C., & Tavani, M. 2006, NIMPA, 568, 692
Ivanchenko, V., Bagulya, A., Bakr, S., et al. 2019, EPJWC, 214, 02046
Jones, B. B. 1998, PhD thesis, Stanford Univ.
Particle Data Group, Zyla, P., Barnett, R. M., et al. 2020, PTEP, 2020, 083C01
Pittori, C., & Tavani, M. 2002, NIMPA, 488, 295
Rossi, B. 1952, High-energy Particles (Englewood Cliffs, NJ: Prentice-Hall), https://books.google.it/books?id=wLQmAAAAMAAJ
Sabatini, S., Donnarumma, I., Tavani, M., et al. 2015, ApJ, 809, 60
Särkkä, S. 2013, Bayesian Filtering and Smoothing (Cambridge: Cambridge Univ. Press), https://books.google.it/books?id=5VlsAAAAQBAJ
Tatischeff, V., Angelis, A. D., Tavani, M., et al. 2019, arXiv:1905.07806
Tavani, M., Barbiellini, G., Argan, A., et al. 2009, A&A, 502, 995
Zoglauer, A. 2005, PhD thesis, Technische Universität München