VRIJE
UNIVERSITEIT
BRUSSEL

Graduation thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Physics and Astronomy

A SIMULATION STUDY OF MUOGRAPHIC IMAGING OF THE EGYPTIAN PYRAMIDS

Dora Geeraerts

June 2025

Promotor: Prof. Dr. Michael Tytgat
Supervisor: Zhe Wang

Sciences and Bioengineering Sciences

Abstract

More than a century after the discovery of cosmic rays, cosmic ray muons have become a powerful tool in non-invasive imaging, with applications ranging from ancient structures and volcanoes to nuclear casks and beyond. Their ability to penetrate dense materials enables the imaging of internal structures by measuring the muon transmission through an object, a technique better known as transmission muography. The ScIDEP project applies this method to investigate the Pyramid of Khafre, aiming to generate two- and three-dimensional density maps using scintillator-based detectors stationed at multiple viewpoints. A key challenge lies in the low probability of muons traversing the detectors after passing through the pyramid, resulting in the need for a computationally optimized simulation. Using Geant4 and EcoMug, a simulation was developed to constrain the muon generation to the detector's field of view, including the extrapolation of muon trajectories and energies before entering the pyramid. A custom tracking algorithm based on the Hough transform and DBSCAN clustering was developed to reconstruct the muon trajectories from a hit collection, showing sensitivity to parameter tuning and track smearing. Simulations for the free sky, the current pyramid structure, and the pyramid with a hypothetical void enabled the creation of transmission maps, revealing the known pyramid features and highlighting the system's sensitivity to potential hidden chambers. While statistical limitations prevent definitive conclusions, results demonstrate the value of muography for the investigation of archaeological structures. Future work involves extended simulations, dual-detector configurations for 3D reconstruction, improved modeling of the detector response, and refinement of the tracking algorithm. The ScIDEP project illustrates the promise of muon imaging in uncovering ancient construction techniques, through interdisciplinary and international collaboration.

Acknowledgements

First and foremost, I would like to thank my promotor, Prof. Dr. Michael Tytgat, without whom I might have never discovered the fascinating research topic of muography. In spite of his demanding schedule, he was always willing to make time for a meeting, offering valuable insights. His support and clear direction helped me regain focus whenever I encountered obstacles. The weekly meetings not only kept me motivated but also provided an opportunity to reflect on my progress and gain insights from other research topics. His dedication and enthusiasm are deeply appreciated. Furthermore, I would like to thank my supervisor, Zhe Wang - better known as Kevin - for his continuous support. He was always approachable and willing to help with any question or simulation challenge, and my progress would not have been the same without his guidance.

Beyond academia, I wish to express my heartfelt appreciation to my friends, housemates, fellow students, and colleagues. Their genuine interest in my research and sympathy during the more challenging moments provided much-needed emotional support. Your encouragement never went unnoticed.

Finally, I am intensely grateful to my family, especially my mom, dad, and brother. Thank you for your unwavering support, for encouraging me to explore new paths, and for always believing in me - while also giving me the freedom to grow and move at my own pace. Your inexhaustible love and trust have been instrumental in shaping the person I am today.

Contents

General Introduction

Cosmic rays are high-energetic particles that continuously penetrate the Earth's atmosphere, originating from both galactic and extragalactic sources. Their existence was first discovered in 1912 by Austrian physicist Hess, who, during a balloon flight, observed an increase in ionization rates at higher altitudes compared to sea level. Primary cosmic rays consist predominantly of protons, with a small portion of heavier nuclei. Upon colliding with atmospheric nuclei, they initiate complex cascades that produce a variety of secondary particles—called secondary cosmic rays. At sea level, muons make up approximately 63% of the secondary cosmic rays. Due to their high velocity and the effect of time dilation predicted by special relativity, these muons are able to reach the surface of the Earth, despite their short intrinsic lifetime. This property, combined with their high energy, makes cosmic ray muons highly penetrative, capable of traversing dense and massive structures on Earth, rendering them ideal candidates for an imaging technique called muography—a non-invasive technique that exploits cosmic ray muons to investigate the internal composition of an object. As muons travel through matter, they will lose energy through a range of interactions. The mean energy loss of a muon per distance traveled is minimal in a certain energy range comprising the average energy of cosmic ray muons. As a result, the mean energy loss rate is primarily dependent on the density of the traversed matter. Consequently, by measuring the cosmic ray muon flux traversing an object, it is possible to acquire information about the density of this object. Various models, ranging from approximate to highly detailed, are used to characterize this flux, and accurate modeling is crucial for muographic applications. The imaging technique relies on either the absorption or scattering of muons, the former known as transmission muography and the latter as muon scattering tomography; where the focus of this thesis lies on the former. Transmission muography is based on measurements of the opacity, defined as the integrated density along a certain line of sight, by comparing the number of muons transmitted through a target with the expected number in the absence of that target. This technique offers a powerful, non-invasive way to probe the unknown, ranging from geological applications to archaeological structures. One such application is the Pyramids of Giza in Egypt, monuments that continue to fascinate both researchers and the public, as the construction of these pyramids is not fully understood yet. The ScIDEP project aims to contribute to solving this puzzle by using transmission muography to investigate the internal structure of the Pyramid of Khafre, the second-largest Pyramid of Giza. Using scintillator-based tracking detectors placed both inside and outside the pyramid, the trajectories of incoming muons are reconstructed from recorded hit coordinates across multiple detector planes. Through the reconstructed muon directions, a transmission map can be generated by comparing the measured muon flux through the pyramid with that from an unobstructed view of the sky. Comparing this measured map with simulations based on a model of the pyramid's structure as known today, anomalies may be revealed that deviate from the expected transmission, indicating the presence of hidden chambers or corridors. Such discoveries have the potential to offer new insights into ancient construction techniques and illustrate

the adaptability of muography to a wide range of research fields.

Chapter 1 covers the production of cosmic ray muons, as well as their interaction with matter, and reviews several models of the cosmic ray muon flux. Additionally, the principle of transmission muography is presented by introducing its core quantity—the muon transmission—and outlining the different types of muon telescopes. This chapter concludes by illustrating some archaeological applications that demonstrate the power and versatility of transmission muography. Next, Chapter 2 provides an overview of the Pyramid of Khafre and highlights some pioneering efforts in applying muography to the Egyptian pyramids. It also outlines the objectives of the ScIDEP project, including details of the detector configuration and the general simulation pipeline, from muon generation to image reconstruction. The optimized simulation developed in this thesis is introduced in Chapter 3, where techniques were explored to constrain cosmic ray muon generation to an inverted cone comprising the detector's field of view, significantly improving computational efficiency. Chapter 4 dives into the development of a novel track reconstruction algorithm, based on existing line detection techniques. The performance of this algorithm was evaluated to determine the optimal reconstruction parameters and to study the influence of noise. Finally, in Chapter 5, the components of the previous chapters are integrated to reconstruct two-dimensional images of the Pyramid of Khafre. Simulations were run for both the unobstructed sky and the pyramid model as known today. The resulting muon tracks were reconstructed to generate a transmission map that highlights the features of the pyramid. An additional simulation incorporating a hypothetical hidden chamber was conducted to explore the detector's sensitivity to muon excesses, indicating the presence of hidden voids. At last, Chapter 6 summarizes the objectives and results of this thesis, and discusses future steps for continued analysis and development.

The following list outlines the specific contributions of the author:

- In Chapter 3, I developed the optimized simulation framework, building upon an existing Geant4-based simulation of the ScIDEP project. This included performing subsidiary simulations to acquire data for the energy and traversed length histograms to enable muon extrapolation to a hemisphere covering the pyramid, implementing sensitive detectors with associated hit classes, and data recording improvements.
- In Chapter 4, I developed a custom track reconstruction algorithm. The algorithm was based on a line detection algorithm developed by Dalitz *et al.*, tailored to the needs of the ScIDEP project. I evaluated its performance and analyzed the impact of smearing and additional noise.
- In Chapter 5, I generated the muon hit datasets for all three simulation configurations, assessed the quality of the track reconstruction, produced the resulting transmission images, and proposed a runtime-based approximation to estimate muon rates.

The generative AI tool ChatGPT was used for the installation of Geant4, and for debugging purposes of the Geant4 simulation, the Python code, and the Latex code [1].

Chapter 1

Muon Radiography

Muography is an imaging technique based on the attenuation and scattering of cosmic ray muons as they pass through matter. The origin of primary cosmic rays—and consequently, the production of cosmic ray muons —is discussed in Section 1.1. This section also covers models of the cosmic ray muon flux and their interactions with matter, both of which are fundamental to muon radiography. The technique can be divided into two categories: transmission muography and muon scattering tomography[1]. These methods are explained in Subsection 1.2.1, where the concept of muon transmission—a key measure for transmission muography—is introduced. To detect cosmic ray muons, specialized muon detectors are employed. Their types and operating principles are detailed in Subsection 1.2.2. Finally, within the context of this thesis, which focuses on imaging of the Egyptian pyramids, relevant archaeological applications of muon radiography are presented in Subsection 1.2.3.

1.1 Cosmic ray muons

1.1.1 Cosmic ray muon production

Cosmic rays (CR) were first discovered by Austrian physicist V. Hess in 1912 while noticing during a balloon flight that the rate of ionization was almost three times higher at 5300 m than at sea level [2]. Primary cosmic rays are highly energetic particles originating from high-energetic sources in the Universe, ranging from the Sun to galactic and extra-galactic sources [3, 4]. Supernova remnants are the most probable galactic sources, while active galactic nuclei and pulsars are additionally speculated to be extragalactic sources [5, 6, 7]. These primary CRs consists mainly of protons (90%), with a smaller fraction of heavier nuclei [8]. When entering the Earth's atmosphere, primary CRs will interact with the atmospheric nuclei via the strong force. These high-energy collisions will cause the nuclei to break up into their constituents, which will reorganize into hadrons through a process called hadronization. The products of these hadronic showers are called secondary cosmic rays. The cascade of these secondary particles resulting from the primary CR interaction with the atmosphere is called a cosmic ray air shower. A general overview of a cosmic ray air shower is shown in Figure 1.1. Mesons are hadrons composed of an equal number of quarks and antiquarks. The most common mesons in secondary cosmic

[1]Muon scattering tomography is often abbreviated as muon tomography, although this term is sometimes used more broadly as an umbrella term. Transmission muography is also referred to as muon radiography or absorption muography. In this text, the general term muography refers collectively to both techniques.

rays are charged pions (π^+ and π^-) and kaons (K^+ and K^-). Neutral pions (π^0) and kaons (K^0) are also produced in cosmic ray air showers, but these have a significantly smaller lifetime compared to their charged counterparts [9]. These secondary CRs will either interact with atmospheric atoms, causing more hadronic cascades, or will decay without interaction. Secondary cosmic rays consist mainly of electrons, positrons, muons, and protons [10]. At sea level, muons make up most of the secondary cosmic rays, with a fraction of 63% [11].

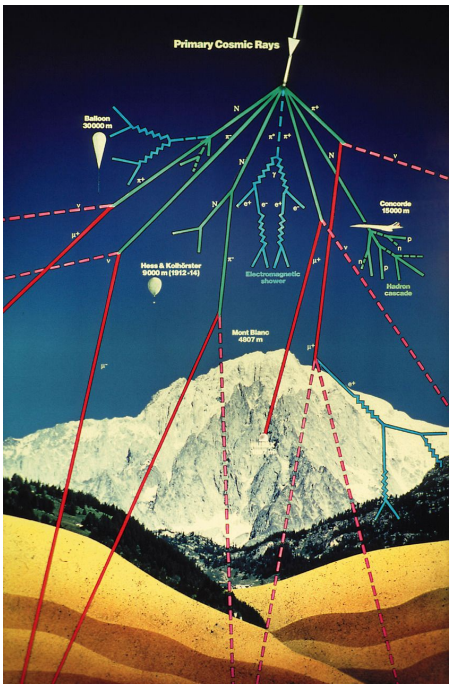


Figure 1.1: Cosmic ray air shower. Figure taken from [12].

A muon μ is an elementary particle of the Standard Model, which is classified as a lepton with a charge convention of -1 and spin of $1/2$ [13]. The muon can be considered the heavier "brother" of the electron: with a mass of 105.7 MeV, it is almost 208 times heavier than the electron. It is an unstable particle with an average lifetime of 2.2×10^{-6} s for a muon at rest. However, since cosmic ray muons travel at high speed, it is possible to observe these muons at sea level on Earth, as a result of Einstein's theory of special relativity. If a muon moves at a velocity v relative to a stationary reference frame, it will perceive the passage of time t in that reference frame as [14]:

$$t' = \frac{t}{\sqrt{1 - \frac{v^2}{c^2}}} \,. \tag{1.1}$$

Here the speed of light is denoted as c with a constant value of $c = 3 \times 10^8$ m s^{-1} [15]. Due to time dilation (Equation 1.1), cosmic ray muons are able to travel much larger distances compared to muons at rest. Their high mass and velocity allow cosmic-ray muons to penetrate hundreds of meters through matter without being significantly scattered or absorbed [10, 16]. These properties make it a highly penetrating particle capable of traversing large and dense objects along a straight-line trajectory, and therefore also the ideal probe for radiographic imaging. Muons eventually decay into an electron, an electron antineutrino, and a muon neutrino via the

weak interaction [8].

Cosmic ray muons—also called atmospheric muons—are produced in the decay of charged cosmic ray pions and kaons. For a charged pion, the decay channel with the highest branching ratio ($\sim 100\%$) is [10]

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \quad \text{and} \quad \pi^- \rightarrow \mu^- + \bar{\nu}_\mu \,, \tag{1.2}$$

with a pion mean lifetime of approximately 26.03×10^{-9} s. Charged kaons produce muons in a similar manner as in Equation 1.2, but here the decay channels have a smaller branching ratio (63.5%), with a kaon mean lifetime of 12×10^{-9} s. Neutrinos only interact faintly with matter, and can in general pass through large amounts of material without being observed. When traveling through matter, muons will lose energy through electromagnetic interactions, as will be explained more profoundly in Subsection 1.1.3. The energy loss—more precisely, the mean energy loss per distance travelled—of charged particles in matter is described by the Bethe-Bloch formula [17]. For atmospheric muons, the mean energy at sea level is 4 GeV near zenith angle 0° [11]. As shown on Figure 1.2, between 100 MeV and 100 GeV, the Bethe-Bloch formula falls to a minimum, making muons in this energy range minimum ionization particles [3, 10]. As a result, the mean energy loss rate is nearly independent of other properties besides the density of the traversed matter. Consequently, by measuring the cosmic ray muon flux of muons traversing an object, it is possible to acquire information about the density of this object.

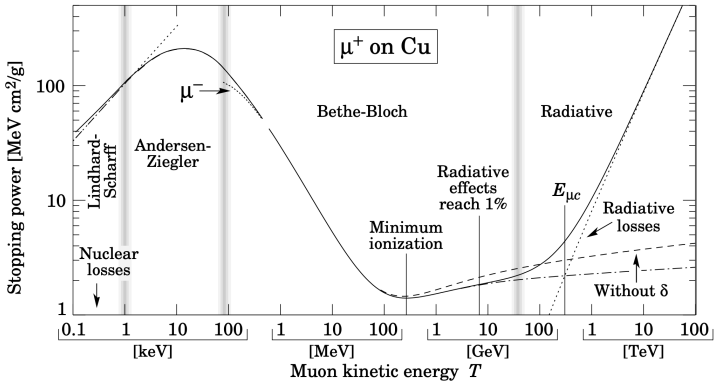


Figure 1.2: Stopping power for positive muons in copper as a function of their kinetic energy T. Figure taken from [17].

1.1.2 Cosmic ray muon flux models

Due to the production of secondary cosmic rays, as described in the previous section, the Earth's atmosphere can be regarded as a natural source of muons. These atmospheric muons are generated at the top of the atmosphere with different energies and inclinations. Their traveled distance s can be calculated using their inclination θ as

$$s(\theta) = \frac{d}{\cos\theta} \tag{1.3}$$

for a flat atmosphere with a thickness d [14]. This means that muons with a larger inclination will travel a longer distance, thus generally taking a longer time to reach the surface of the

Earth. Due to the finite lifetime of the muon, a longer travel time corresponds to a lower survival probability; hence, for muons with a larger inclination, a lower arrival rate is expected. Moreover, a muon that travels longer distances will have a higher survival rate if its energy is higher. Thus, the horizontal muon flux is more energetic than the vertical one, but the vertical flux has a higher average intensity [10]. This is illustrated in Figure 1.3, where the CR muon flux is shown for different measurements for an inclination of $\theta = 0°$ and $\theta = 75°$. As a consequence of these two relations, the muon flux measured at Earth's surface is modeled as a function of the inclination θ and energy E of the muon.

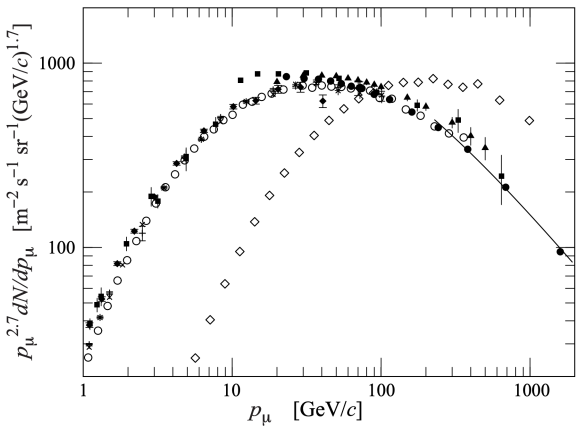


Figure 1.3: Cosmic ray muon spectrum at $\theta = 0°$ ($\circ, \bullet, \blacksquare, \blacktriangledown, \times, +$) and $\theta = 75°$ ($\Diamond$). The line plots the result from Equation 1.8 for vertical muons. Figure taken from [5].

Typically, the rate of muons at sea level is around 70 m^{-2} s^{-1} sr^{-1} [5]. Most muons are generated in the atmosphere at altitudes between 26 km and 15 km above sea level with an average muon energy of around 6 GeV, which drops to an average energy of 4 GeV at sea level [10]. As mentioned in the previous subsection, muons are observed on Earth as a consequence of time dilation. Suppose a muon with rest mass $m_0 = 105.7$ MeV is produced in the atmosphere with energy $E = 6$ GeV, then the relativistic gamma factor is

$$\gamma = \frac{E}{m_0 c^2} = \frac{6 \text{ GeV}}{105.7 \text{ MeV}} = 56.8 \,, \tag{1.4}$$

corresponding to a time dilation of

$$t' = \gamma \times \tau = 56.8 \times 2.2 \ \mu\text{s} = 125.0 \ \mu\text{s} \,, \tag{1.5}$$

where τ is the mean lifetime of a muon at rest. Using the expression of gamma in terms of the speed of the particle, it can be deduced that the muon has a speed of $v = 0,9998$c. Thus, the distance this muon will travel is $s = vt' = 37$ km. Taking into account energy losses caused by interactions with the atmosphere, this decay length decreases to 25 km [10]. As a result, it is clear that atmospheric muons can reach the Earth's surface.

The cosmic ray muon flux is usually characterized by the symbol $\Phi(\theta, E)$ and expressed in units of cm^{-2} s^{-1} sr^{-1} GeV^{-1} [11]. For various applications in fundamental physics, as well as in muon radiography, it is important to model this muon flux thoroughly. Different models

have been proposed and can generally be split into three categories: empirical models, analytical models, and Monte Carlo models [18].

The most popular empirical model is the cosine-power model [5]. The angular distribution of muons at sea level in function of the zenith angle θ is proportional to $\cos^2\theta$, for muons with an energy of $E \sim 3$ GeV. The flux can thus be modeled as

$$\Phi(\theta) = \Phi_0 \cos^2\theta \,. \tag{1.6}$$

The front term Φ_0 is a scaling factor, which corresponds to the vertical muon flux, i.e. for inclination $\theta = 0°$. The first factor $\cos\theta$ can be explained by the probability rate declining when the muon travels a longer distance, as seen in Equation 1.3. The second factor comes from the fact that a muon traveling with inclination angle θ sees the Earth's surface area scaled by a factor $\cos\theta$ approximately [14]. This model omits the dependence of the flux on the muon energy. At higher energies, the angular distribution will flatten, while at lower energies it will become steeper.

Analytical models are based on an empirical model and are usually derived from physical phenomena [18]. These models take the energy dependence into account, resulting in a more complete model of the cosmic ray muon flux. This class of models is usually obtained by fitting empirical parametric curves to the measured muon flux at sea level. The parametrization can be motivated by physics related to the production of the CR muons or can be suggested by seeking a tight fit [11]. A famous analytical flux model is the Gaisser model, proposed by physicist T. K. Gaisser in 1990 [8]. This model has the following form:

$$(\Phi(\theta, E))_{\text{Gaisser}} = A_G E^{-\gamma} \left(\frac{1}{1 + \frac{E\cos\theta}{E_\pi^{\text{CR}}}} + \frac{B_G}{1 + \frac{1.1E\cos\theta}{E_K^{\text{CR}}}} \right) . \tag{1.7}$$

The first term in the parentheses takes into account the atmospheric muon flux caused by pion decay, the second term the muon flux caused by kaon decay. The adjustable parameters are the scale factor A_G, the power index γ, and the balance factor B_G—which depends on the ratio of muons produced by kaons with respect to pions. Parameters E_π^{CR} and E_K^{CR} are the critical energies[2] of pions and kaons for a vertical inclination. In theory, these parameters are physical quantities that should be determined through quantum mechanical calculus. However, some authors consider E_π^{CR} and E_K^{CR} as adjustable parameters when fitting the model to measured data. Replacing these factors by their calculated values yields:

$$(\Phi(\theta, E))_{\text{Gaisser}} = \frac{0.14E^{-2.7}}{\text{cm}^2\ \text{s}\ \text{sr}\ \text{GeV}} \left(\frac{1}{1 + \frac{1.1E\cos\theta}{115\ \text{GeV}}} + \frac{0.054}{1 + \frac{1.1E\cos\theta}{850\ \text{GeV}}} \right) . \tag{1.8}$$

The power index $\gamma = -2.7$ relates the atmospheric muon flux to the energy spectrum of their parents, the primary cosmic rays [5]. An important remark is that this model does not take the curvature of the Earth into account, and moreover, it overestimates the flux for muon energies below 10 GeV due to muon decay and energy loss becoming more prominent [8]. This overestimation is taken into account in other models, like the Gaisser-Tang model [19]. Other analytical models have been suggested as well. These are usually modifications of the Gaisser model, taking into account certain assumptions about the curvature of the Earth or muon decay. Examples

[2]The critical energies are the energy thresholds above which pions and kaons interact with the atmosphere before decaying [10].

of these are the Guan model, Bugaev model and Reyna-Hebbeker model [20, 21, 22].

The last category of flux models is based on Monte Carlo simulations. Monte Carlo (MC) algorithms are computational algorithms performed to compute numerical calculations of quantities that are too difficult to solve analytically. These MC algorithms are based on repeated random sampling of quantities that follow known probability distributions. Examples of these for the cosmic ray muon flux are CORSIKA, CRY and EcoMug [23, 24, 25]. Depending on how thorough the model is, it starts by generating the primary cosmic rays and then simulates their interactions with the atmospheric atoms. Next, the process of the hadronic shower is simulated, with as a result the decay of secondary cosmic ray particles to atmospheric muons. Lastly, the procedure of how these muons travel to and reach Earth's surface is simulated. A downside of the Monte Carlo method is that MC algorithms usually require a longer computational time and more computational power. However, the approximations of empirical and analytical methods might result in a less accurate result for the atmospheric muon flux. None of these models is universally suitable. Thus, the choice for a certain model should be made based on the conditions of the experiment, i.e., ranges of θ and E, and required accuracy.

1.1.3 Muon interactions

Muography relies on the interaction of cosmic ray muons with matter. The probability of a muon traversing a certain volume depends not only on the muon's decay, but also on its interaction with the surrounding material. Muons will lose energy through electromagnetic interactions like ionization, atomic excitation, and radiative processes. There are three main radiative processes for a muon [17]:

- Bremsstrahlung: the deflection of a muon caused by the electric field of another charged particle, resulting in the emission of a photon,
- Pair production: the creation of an electron-positron pair,
- Photonuclear interactions: interaction of a high-energy muon with the electric field of a nucleus, causing the emission of a virtual photon that consequently is absorbed by the nucleus.

Each of these processes has a different energy curve, so the dominating process depends on the energy of the muon. At energies up to $E \approx 500$ GeV, the ionization and excitation processes will dominate [26]. Bremsstrahlung is dependent on the mass m of a particle through $1/m^2$ [3]. Since muons are approximately 200 times heavier than electrons or positrons, they will lose significantly less energy through Bremsstrahlung.

As mentioned before, the mean energy loss per distance traveled of charged particles in matter is described by the Bethe-Bloch formula. The contribution of the ionization and radiative processes to the loss of energy of a muon as it traverses matter can be represented as [17]:

$$-\frac{dE}{dx} = a(E) + b(E)E\,. \tag{1.9}$$

Here $a(E)$ is the energy lost through ionization and atomic excitation, and $b(E)$ is the fraction of energy lost through all radiative processes together. These parameters $a(E)$ and $b(E)$ are a function of the energy and will vary with the material properties. A table with standard values for a certain material composition can be consulted from the Particle Data Group [5].

For muons with an energy ranging from 100 MeV up to 100 GeV, the energy loss rate is constant at first order because these muons are minimum ionization particles. The value for the stopping power in dense materials for a muon energy of a few GeV is around [10]

$$-\frac{dE}{dx} = 2.1 \text{ MeV g}^{-1} \text{ cm}^2 \tag{1.10}$$

and less than this in air. This explains the approximately 2 GeV energy loss of muons when traveling from the atmosphere to sea level. In the energy range where the muon is minimum ionizing, the mean energy loss rate is proportional to the density of the traversed matter and is nearly independent of other properties [3]. As a result, the cosmic ray muon flux can be related to the density of the traversed material. Thus, by measuring the flux of cosmic-ray muons traversing an object, it is possible to acquire information about the density of this object. The opacity σ of an object is defined as the density of that object integrated along a certain line of sight and can be calculated as follows [27]:

$$\sigma(L) = \int_L \rho(x)dx \,, \tag{1.11}$$

where L corresponds to the thickness of the object along the muon's trajectory, ρ is the density of the object and x is the length variable along the muon's trajectory. The measurement of opacity is at the core of muon radiography, with a schematic overview presented in Figure 1.4.

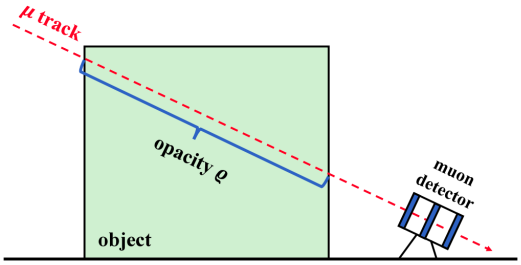


Figure 1.4: Illustration of the opacity measured through muon radiography. Figure taken from [27].

1.2 Muography

Muon imaging relies on the detection of cosmic ray muons using a muon telescope, with the purpose of retrieving a density image of the object of interest. There are two methods of muon imaging: transmission muography and muon scattering tomography. The first uses the electromagnetic interactions causing absorption of muons through energy loss, while the latter relies on the principle of Coulomb scattering of muons in the traversed object [3]. Figure 1.5 illustrates the difference between transmission muography and muon scattering tomography.

1.2.1 Operating principle

Muon radiography was first used in the last century, with the earliest known application being in 1955, when the overburden of a tunnel in Australia was measured by implementing a

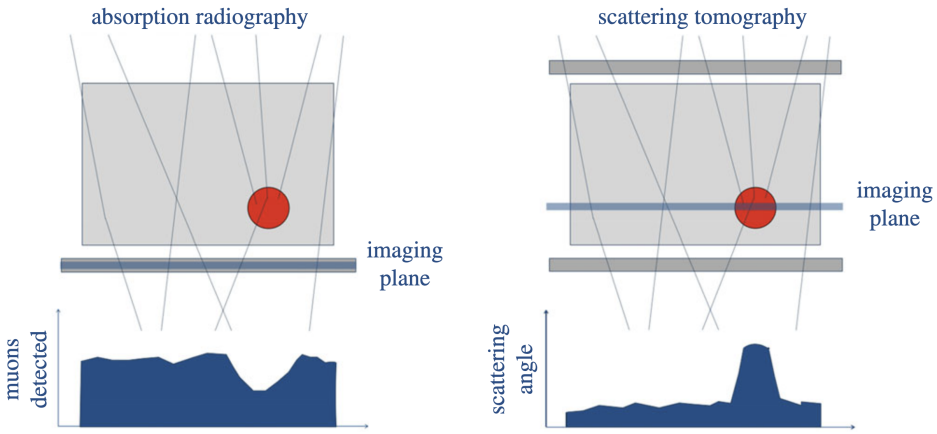


Figure 1.5: Illustration showcasing the difference between transmission muography and scattering tomography. Figure taken from [28].

Geiger-Müller counter [29]. With the cosmic-ray muon flux being maximal in the vertical direction and decreasing for larger zenith angles, the muon telescope is typically placed at a lower altitude looking at the target in an upward direction [3]. An illustration of a simplified setup for transmission muography is shown in Figure 1.4. The telescope is used to track the trajectories of the incoming muons after traversing the target. For a certain amount of time, muons are detected and the muon absorption in the object is measured. The total time of data acquisition is such that a sufficiently high number of muons is detected, to ensure a qualitative image reconstruction. Due to the possibility to reconstruct the muon's trajectory—and thus its direction—the flux of the muons passing through the object is measured for each direction in the field of view. Transmission muography relies on the same principle as X-ray imaging: a quantity called the transmission is measured when the source particle (a photon in the case of X-rays) passes through the matter [14]. In function of the zenith angle θ and azimuth angle ϕ, the transmission can be defined as

$$T(\theta, \phi) = \frac{N(\theta, \phi)}{N_0(\theta, \phi)} , \tag{1.12}$$

where $N_0(\theta, \phi)$ is the number of muons that are expected without target and $N(\theta, \phi)$ is the number of muons that are transmitted through the target [10]. The expected number of muons is calculated with Monte Carlo simulations or can be acquired by performing free sky data measurements. The latter can either be done by rotating the detector to a point without obstructed view or—if the former is impossible—by putting the same detector in a different location to observe free sky [3]. Free sky measurements have the advantage of taking into account the detector efficiency [10]. Including the detector efficiency and the acquisition time of the measurements, the measured transmission, according to Equation 1.12, becomes:

$$T_m(\theta, \phi) = \frac{\epsilon_{FS}}{\epsilon} \frac{t_{FS}}{t} \frac{N_m(\theta, \phi)}{N_{FS}(\theta, \phi)} , \tag{1.13}$$

where ϵ_{FS} is the detector efficiency for the free sky data acquisition with acquisition time t_{FS}, ϵ is the detector efficiency for the target data acquisition with acquisition time t, $N_{FS}(\theta, \phi)$ is

the number of muons that were counted without target, and $N_m(\theta, \phi)$ is the number of muons that were transmitted through the target. As a result, the detector acquires a 2D image of the transmission of the object. This transmission can be related to the density and thickness of the object.

The interaction processes of a muon traveling through an object are stochastic; for an individual muon, it cannot be predicted deterministically if this muon will escape [14]. However, the average behavior of muons can be described. Suppose the muons have an average energy E and traverse an object with a uniform density ρ and thickness d, then the muon transmission is empirically described as [30]

$$T_\mu(E, \rho) = e^{-\mu(E,\rho)d} \,. \tag{1.14}$$

Here, $\mu(E, \rho)$ is the linear attenuation coefficient. This coefficient can be redefined as $\mu' = \frac{\mu}{\rho}$, which is independent of the material, reforming Equation 1.14 as follows:

$$T_\mu(E, \rho) = e^{-\mu'(E)\rho d} \,. \tag{1.15}$$

When the object has a non-uniform density, the density has to be integrated over:

$$T_\mu(E, \rho) = e^{-\mu'(E)\int_L \rho(x)dx} \,. \tag{1.16}$$

Here, x is a coordinate variable moving along the trajectory of the muon. The integral in this equation is precisely the opacity as defined in Equation 1.11, such that the non-uniform transmission can be rewritten as:

$$T_\mu(E, \rho) = e^{-\mu(E)\sigma} \tag{1.17}$$

According to Equation 1.17, the transmission is related to the density ρ and size d of the object through the opacity σ: for a larger opacity—corresponding to a denser material or larger volume—the transmission will be less. Thus, by measuring the muon transmission of an object, a 2D mapping of the amount of traversed matter of the object can be reconstructed. An important remark is that this results in a map of the average density in a certain direction. This means that a low density region without cavities can result in the same image as a high density region with cavities.

To examine the internal structure of a target object, the measured transmission $T_{measured}$ is typically compared to the expected transmission $T_{expected}$, where the latter is usually predicted through simulations. This comparison can be quantified by the relative transmission, which is defined as

$$R(\theta, \phi) = \frac{T_{measured}(\theta, \phi)}{T_{expected}(\theta, \phi)} \,, \tag{1.18}$$

and is equal to unity when no unknown internal structures are present, assuming that the expected transmission was accurately reproduced [10].

The second technique of muon imaging, muon scattering tomography (MST), makes use of Coulomb scattering of muons. When a muon traverses an object and passes sufficiently close to an atomic nucleus to feel the electromagnetic influence, it will be scattered in a different direction [3]. The atomic nuclei have a different atomic number Z depending on the kind of material of the object; thus, the electromagnetic interaction will differ in strength for different materials. The larger Z, the stronger the interaction. The principle is to correlate this deflection of the muon inside the target to the atomic number Z. Domains where this is of interest are

nuclear waste, nuclear security, and material control, where it is essential to distinguish higher Z materials from lower Z materials [31]. To detect the deflection of the muons, a detector should be placed before and after the muon traverses an object. This technique was not investigated in this thesis, so it will not be elaborated upon. From now on, the general term *muography* will be used to denote transmission muography.

1.2.2 Detectors in muon telescopes

Muon detectors are devices to indirectly detect the presence of a muon: they detect the secondary particles that are created when a muon passes the detector and interacts with the detector material. The type of detector that is used for a muography experiment depends on several factors, like the necessary resolution and the budget [3]. There are three main types of detectors commonly applied in muography: plastic scintillators, gasses, and emulsion films. The core mechanism for each of these is the same: the ionization of atoms due to a muon passing through.

- Organic scintillator-based detectors rely on the principle of luminescence: when a charged particle passes through the scintillating material, the atoms get excited, and when the excited electron falls back to a lower energy state, light is emitted in the form of photons. These photons are collected and converted into an electronic signal, which is proportional to the energy deposit of the muon in the scintillator. Organic scintillator-based detectors can be plastic, crystal, or even liquid. The plastic ones are popular as they are robust, reliable, and simple to operate [26]. They are most optimal for situations that do not require a high resolution and are low budget. Another advantage is their flexibility to be produced into all kinds of shapes.

- Gaseous detectors consist of a narrow gas volume. When a muon passes through the gas, the gas will be ionized, creating electron-ion pairs. To collect these charges, an electric field is applied across the gas volume. Considering that some types of gaseous detectors include an internal amplification mechanism, a gas detector generally performs at a better resolution for a larger area compared to scintillator detectors [3]. On the downside, the gas will age, resulting in a decrease in the efficiency, and some gases used nowadays are flammable and/or polluting for the environment [26]. For gaseous detectors, multiple types of detectors exist: the Multi-Wire Proportional Chamber (MWPC), drift chambers, Migromegas detectors, Gas Electron Multipliers (GEM), and Resistive Plate Chambers (RPC) [16].

- Emulsion films are photographic plates with grains of silver halide crystals. When a muon traverses the plate, it ionizes the crystals along its path, resulting in a migration of electrons from the valence band to the conduction band. This leads to the transformation of some silver ions (Ag^+) into silver atoms (Ag), triggering a chemical change [32]. After recording the muons for a certain amount of time, the films are developed. The angular resolution is defined by the grains in the film, which is typically in the order of milliradians, resulting in better performance compared to gaseous and scintillating detectors [3]. Emulsion films don't require a power supply, making them perfect for remote areas. However, there is no information about the timing, making it impractical to study dynamical systems, and the equipment necessary to develop them is specific and very complex, resulting in a limited accessibility.

Other types of detectors are currently being developed, such as semiconductor detectors or detectors relying on the principle of the Cherenkov effect [26, 33]. There will always be a trade-

off between the resolution, the cost of construction, maintenance, the dependency on power sources, accessibility, etc. Since a lot of cosmic-ray muons will never reach the detectors—either they will be completely absorbed by the target or they will miss the detector—it generally takes a long time to acquire enough data, ranging from hours to months [3]. This means that the material needs to remain stable during a long period.

To be able to retrieve information about the trajectory—and thus the direction—of the cosmic-ray muons, multiple muon detectors have to be combined together. This ensemble of stacked detectors is also called a muon telescope or hodoscope [3]. A detector part of this ensemble is called a tracking station if it is able to retrieve a hit coordinate. When a muon passes through a muon telescope, each tracking station where the muon passes through will record a hit spatial coordinate. To retrieve the trajectory of the muon, it is usually assumed that the muon follows a straight trajectory. This assumption is plausible since cosmic-ray muons are highly penetrable, as already stated in the previous section. Hence, with a minimum of 2 tracking stations, the muon's trajectory can be determined. The direction of a trajectory can be characterized by the azimuth angle ϕ and zenith angle θ.

1.2.3 Applications

Muography can be applied in numerous fields. One particular field where it is being implemented is archaeology, considering that muography is a non-invasive technique to image a structure. This makes it an appealing method to investigate the internal structure of tombs, tunnels, or other archaeological formations. In this subsection, a few archaeological applications of muography, ranging from ancient urban settlements in Italy to a Chinese defense wall, will be described.

In Naples, Italy, muography has become a valuable tool to study the structure of the remains of ancient urban settlements [34]. Mount Echia, an overburden of soft volcanic rock with a maximum altitude of 60 m in the historical center of Naples, marks the location of the earliest urban settlement in the area. Founded by the Greek colony Parthenope in 8th century BCE, it harbors a network of tunnels and chambers, among them the Bourbon Tunnel, excavated in 1853. In 2017, a scientific team installed an upgraded version of a MU-RAY (scintillator-based) telescope [35]. After acquiring data for 26 days, preliminary results indicated the presence of a void. For further investigation, the MU-RAY telescope was relocated to a different location for an additional data acquisition of 8 days. Furthermore, an additional telescope, a portable MIMA (scintillator-based) telescope, was installed at a third location, for a total data collection time of 50 days [36]. After 3D reconstruction of the cavity, it was deduced that the cavity is around 4 m wide, 3-4 m high and 7 m long.

In 5th century BCE, the development of Parthenope led to the nearby foundation of Neapolis, the ancient city of Naples [34]. Located 10 m below current street level, it contains remains of buildings, streets, aqueducts and necropolis [37]. Due to the dense population in Naples, archaeological excavations are not always safe or practical, raising the need for non-invasive methods like muography. In 2023, two modules of nuclear emulsions were installed in an underground corridor and analyzed independently in Naples and Nagoya, Japan. After comparing measured data with simulations, a muon flux excess was revealed that could not be explained by the known architecture of the underground corridors, as seen in the dashed pink circled area on Figure 1.6. One of the observed anomalies appeared to be compatible with the existence of a currently inaccessible burial chamber, with a size of 2-3.5 m.

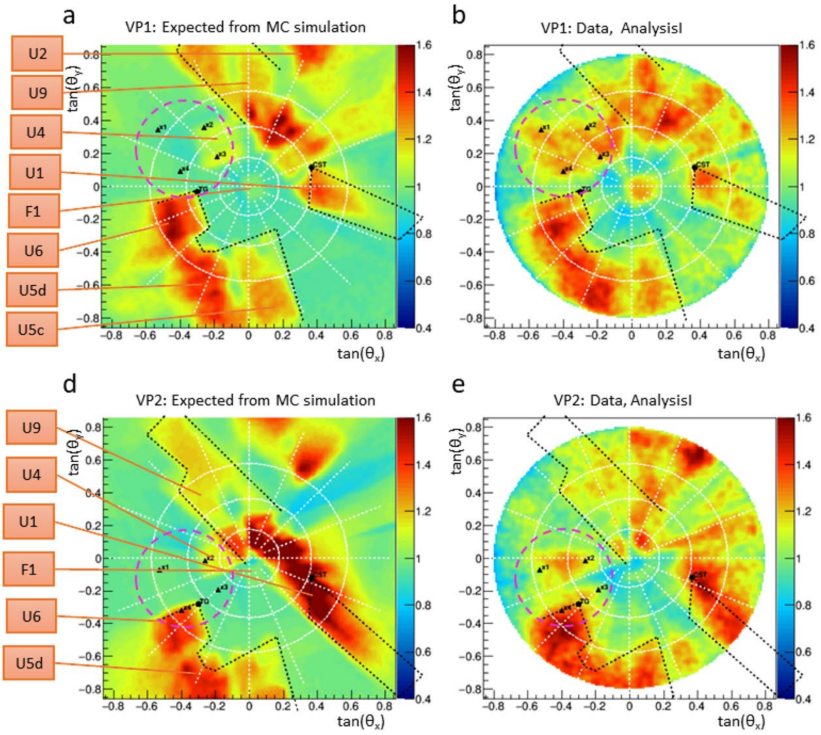


Figure 1.6: Reconstructed images of the muon flux for simulations (a,d) and measurements (b,e) from the two detector viewpoints (VP1 and VP2). Black dashed lines indicate the shapes of some known structures. The pink dashed circle indicates a muon excess which cannot be explained by any known structure. Figure taken from [37].

Another notable application of muography in archaeology is a study done on the defensive wall in Xi'an, China [38]. This wall is one of the best-preserved ancient city walls in the country, stretching roughly 13.74 km. Built more than 600 years ago during the Ming Dynasty, it holds significant cultural and historical value. However, the structure has suffered damage over time due to adverse weather conditions, such as heavy rainfall in 2006 and a monsoon in 2020. Research identified 214 cracks over 1 cm, as well as other damage like spalling and holes. Documentation of past repairs of the wall is scarce, meaning that current conservation mostly relies on external observations, which is insufficient to understand the internal damage. Muography offers a non-destructive approach to explore the wall's internal structure, by analyzing the density distributions. A study focused on a 12 meter high rampart of the defensive wall. A portable muon imaging system, CORMIS, was developed, and data was collected over three months total, from six different detector positions. Using this data from multiple viewpoints to reconstruct a 3D image, the system achieved a spatial resolution of $1 \times 1 \times 1$ m^3. This unprecedented level of precision demonstrates the research's ability to address several limitations of earlier muography techniques, highlighting its potential for future conservation projects for archeological structures.

These examples illustrate how muography enriches our understanding of ancient structures while showcasing the precision and adaptable nature of transmission muography in uncovering otherwise possibly inaccessible archeological features. The technique has considerably evolved, both from instrumental point of view and data analysis methodologies, and is likely to become a widely used method in the future, providing further insight in the field of archeology, as well as volcanology, civil engineering, geology, etc. [39]. In the case of muon scattering tomography, this technique is predominantly used to monitor nuclear waste or nuclear reactors, although recent studies have suggested a promising potential for applications related to the conservation of cultural heritage [31, 40].

Chapter 2

The ScIDEP Project

The construction of the pyramids on the Giza Plateau in Egypt remains a subject of fascination and mystery. The ScIDEP project is an international collaboration that aims to reveal the internal structure of the Pyramid of Khafre using transmission muography. Despite only being slightly smaller, the known internal layout of the Pyramid of Khafre is simpler than that of the Great Pyramid of Giza. Details about the Pyramid of Khafre and the ScIDEP project are provided in Section 2.1, which also touches upon several pioneering studies previously conducted on the Pyramids of Giza. The collaboration intends to install muon telescopes both inside and outside the pyramid to search for hidden cavities, with the ultimate goal of deepening our understanding of the pyramid construction techniques. The muon telescope system designed for this purpose is described in Section 2.2. Finally, the overall simulation pipeline—from muon generation to image reconstruction—is outlined in Section 2.3.

2.1 Pyramid of Khafre

The pyramids at the Giza pyramid complex in Egypt are among the oldest monuments on Earth, being built during the Fourth Dynasty of the Old Kingdom of ancient Egypt, around $2560 - 2510$ BCE [41]. Despite their historical and architectural significance, no documents about the construction process written during the Fourth Dynasty have been discovered so far, such that there is no definite consensus on how they were built [42]. King Khafre was the son of Khufu, who ordered the construction of the Great Pyramid of Khufu, the largest pyramid of Giza [43]. The grandiosity of the Great Pyramid set a high standard for King Khafre to follow. The pyramid of Khafre is built on higher ground, giving the illusion that it is taller than Khufu's, and has a pink granite bottom. Today, the Pyramid of Khafre stands as the second tallest pyramid on the Giza Plateau, with a height of 136.4 m and a base length of 215.3 m [44]. It has a slope of $53°10'$ and is constructed out of limestone blocks, each weighing over 2 tons. The majority of the outer limestone casing was stripped over time. Despite only being slightly smaller than the Great Pyramid of Khufu, the presently known internal structure of the Pyramid of Khafre appears considerably simpler. As seen in Figure 2.1, its interior contains two entrance corridors leading to a small chamber and a larger—slightly off-center w.r.t. the vertical axis—chamber. The central chamber, the King's burial chamber (also called Belzoni chamber), has dimensions of $14 \times 5 \times 6.8$ m^3, while the subsidiary chamber has dimensions of $10 \times 3 \times 2.6$ m^3. The simplicity of the interior might have been influenced by the experience of the builders completing the Great Pyramid, but also raises the suggestion that the pyramid might contain chambers or

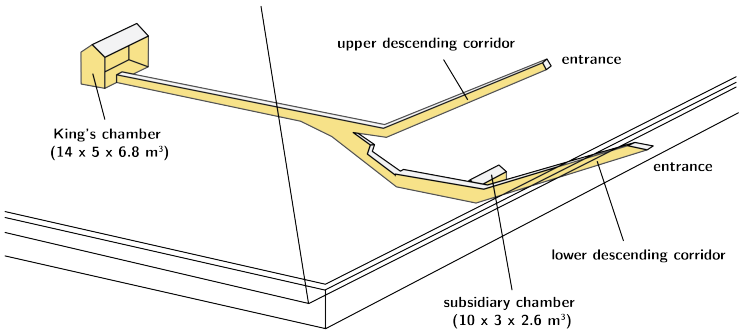


Figure 2.1: Interior layout of the Pyramid of Khafre. Figure adapted from [44].

corridors that have yet to be revealed [45].

A pioneering example of muography applied to the pyramids is a study conducted by Physics Nobel Laureate L.W. Alvarez *et al.* in 1970, who investigated the internal structure of the Pyramid of Khafre, in search of hidden chambers [46]. Two 1.8-meter spark chambers, separated by a distance of 0.3 m, were placed in the King's chamber at the bottom of the pyramid, with scintillation counters to trigger the spark chambers upon the passage of a muon through all three counters. Data was acquired over several months to construct a two-dimensional map of 19 % of the volume of the pyramid. While their research successfully visualized the known features of the pyramid—the diagonal ridges and the cap—as shown in Figure 2.2, the presence of a hidden room or corridor could not be confirmed.

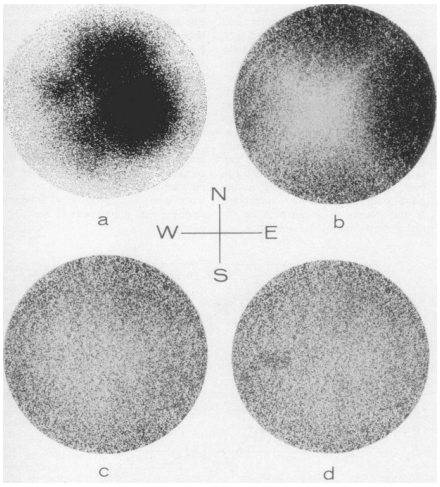


Figure 2.2: Scatter plots illustrating the muographic image reconstruction of the Pyramid of Khafre for the (a) uncorrected data, (b) data corrected for the geometrical acceptance of the detector, (c) data additionally corrected for the features of the pyramid, and (d) corrected data including a simulated hidden chamber. Figure taken from [46].

Half a century later, in 2017, the ScanPyramids collaboration claimed to have discovered a large new void above the Grand Gallery in the Great Pyramid of Khufu [42]. This effort was part of a broader project to scan four Old Kingdom Egyptian Pyramids. A combination of scintillators and nuclear emulsions placed inside the pyramid, along with gaseous detectors outside the pyramid, was employed. As seen in Figure 2.3, an excess of muons with over 5σ statistical evidence was observed, corresponding to a void dubbed ScanPyramids Big Void. This void was estimated to be at least 30 m long, comparable to the length of the Grand Gallery. In 2023, the collaboration provided a detailed analysis of another discovered void [47], called the ScanPyramid North Face Corridor (NFC). Using high-resolution multipoint observations, the inclination, location, and dimensions could be revealed. Combining two independent analyses from CEA and Nagoya University, the void was found to be more or less horizontally oriented with a cross-section of approximately 2×2 m and a length around 9 m. Furthermore, a study was done combining three non-destructive testing techniques (ground penetrating radar, ultrasonic testing, and electrical resistivity tomography) to support these findings and enable a more detailed examination [48].

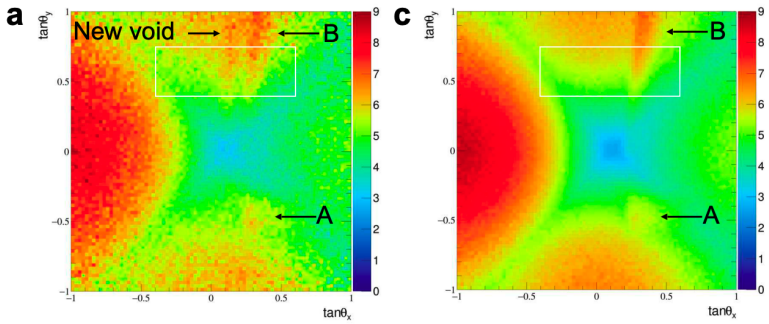


Figure 2.3: Detected (a) and simulated (c) muon flux from within the Pyramid of Khufu, viewed from one of the nuclear emulsion detectors. Region A denotes the King's chamber, B the Grand Gallery and "New void" the region of unexpected muon excess. Figure taken from [42].

These studies highlight the potential of muography as a powerful tool for analysing the structure of the Pyramids of Giza and other archaeological sites.

The ScIDEP (Scintillator Imaging Detector for the Egyptian Pyramids) project is an international scientific collaboration using transmission muography to reveal the internal structure of the Pyramid of Khafre at the Giza Plateau [49]. It is a collaboration between institutes from Belgium (VUB), Egypt (E-JUST), the USA (U. New Mexico and PNLL) and Romania (IFIN-HH). By positioning muon telescopes at multiple viewpoints in and around the pyramid, the project aims to reconstruct a three-dimensional image of its interior and search for hidden cavities. The potential discovery of hidden internal features could provide valuable insight into the construction techniques of the pyramids. Beyond its scientific goals, the ScIDEP project emphasizes the value of interdisciplinary and international collaboration to enrich our knowledge about the pyramids and other historical monuments.

2.2 Detector setup

The ScIDEP project aspires to install muon telescopes, with improved technology over Alvarez's experiment, in the Pyramid of Khafre to continue the search for hidden structures using the principle of transmission muography. The objective is to install muon telescopes at various viewpoints in and outside the pyramid, with the intention of reconstructing a three-dimensional image of the pyramid. Two kinds of plastic scintillator-based muon telescopes are being developed, one made from plain polyvinyl toluene (PVT) plates, the other from polystyrene bars [45]. For both detectors, the scintillation photons which pass a wavelength shifting (WLS) fiber will be absorbed by these fibers and re-emitted at a longer wavelength, after which they are guided through the fibers via internal reflection to silicon photomultipliers (SiPM), from where the signal will be acquired using a readout system. The plan is to install at least one detector looking upward inside the King's burial chamber at the bottom of the pyramid. The other detector could be positioned outside of the pyramid, looking at the upper part of the pyramid and a portion of the free sky at an angle. The exact location and configuration for this outside detector still has to be decided based on investigations of the expected muon flux via simulation.

2.2.1 Detector 1: Polyvinyl toluene plates

The first muon telescope consists of $61 \times 61 \times 2$ cm^3 plastic scintillator plates, made out of EJ-200 (Eljen Technology) PVT, stacked vertically along the Z-axis [50, 51, 45]. To reconstruct the muon hit coordinate, BCF-92 (Luxium Solutions, previously Saint-Gobain Crystals) WLS fibers with a diameter of 2 mm are installed in grooves [52]. These grooves are 3 mm deep and 1 cm apart, placed orthogonally at each surface side of the PVT plates. Hence, each PVT plane contains two times 60 WLS fibers. A detector plane with the embedded WLS fibers is shown in Figure 2.4. The orthogonal placement allows for data acquisition of the X- and Y-

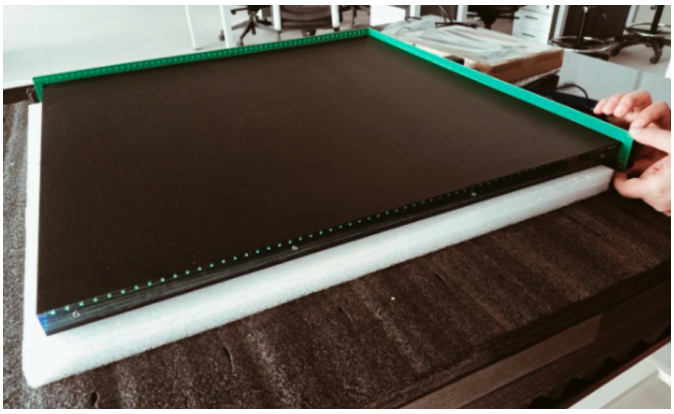


Figure 2.4: A detector PVT plate with the embedded WLS fibers and PCB boards. Figure taken from [49].

coordinate. These coordinates are later on used to reconstruct the trajectory of the muons. The scintillation light collected in these fibers is read out by S14160 (Hamamatsu) SiPMs, with a bias voltage of 41 V, which are being held by printed circuit boards (PCB) [49, 53, 54]. The original version of the muon detector consists of two PVT plates. Their surfaces are coated in a black paint, to limit the internal reflection of the scintillation light. The effective position resolution of the current detector configuration is still being confirmed via lab measurements,

but preliminary simulation studies imply that a position resolution better than 1 cm should be feasible [45]. The distance between the two plates is variable between 20 and 200 cm, allowing for the configuration of different angular resolutions and field of view [51]. The readout system of the detector is currently based on CAEN PETIROC2A-ASIC programmable data-acquisition boards, supported by DT5550W modules [55]. However, a custom-made readout electronics is currently under development to replace the CAEN system, as it has been difficult to operate [45]. This will be based on Weeroc ASICs in combination with field-programmable gate arrays (FPGA). The data will be stored in an ASCII file for further analysis. At present, the detector is located at the Egypt-Japan-University of Science and Technology (E-JUST) in Alexandria, Egypt. Here, the characterization of the system is in process.

2.2.2 Detector 2: Polystyrene bars

The second detector is built out of polystyrene scintillator bars, each measuring $2.5 \times 1 \times 100$ cm^3, manufactered by Amcrys in Ukraine [56]. The detector is composed of four layers, where each layer contains 36 bars [45]. A setup of the detector in the laboratory is shown in Figure 2.5. Each bar contains a 1.5 mm diameter BCF-91A (Luxium Solutions) WLS optical fiber which

Figure 2.5: Setup of detector 2 in the laboratory. Figure taken from [45].

is connected to a MicroFC-30035 (ONSemi, previously SensL) SiPM for light readout, as seen in Figure 2.6. The scintillator bars are covered in polytetrafluoroethylene (PTFE) tape to ensure light containment in the bars. Consecutive layers are placed orthogonally to allow data acquisition of the X- and Y-coordinate, with the top two layers at a certain distance from the bottom two layers. For a separation distance of 90 cm, Monte Carlo simulations predict a 1.82 sr geometrical angular acceptance. When all four plates are stacked on top of each other, the simulations predict a 5.71 sr angular acceptance. The readout system is FPGA-based, employing a Xilinx custom-made solution [57]. When every layer has at least one bar that detected an event, the event is accepted and recorded in an output file [45]. Otherwise, the trajectory reconstruction cannot be executed. Experimental tests have been performed to investigate the imaging capabilities of the muon detector. Currently, the detector is being prepared for in-situ measurements at the Astroparticle Physics Laboratory of the Horia Hulubeei National Institute for Nuclear Physics and Engineering (IFINHH) in Romania.

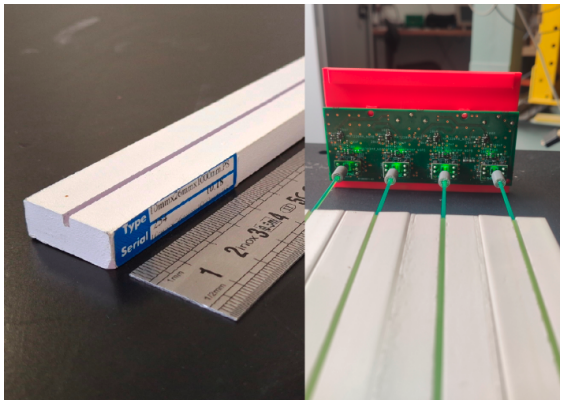


Figure 2.6: Left: scintillator bar with groove. Right: WLS fibers installed in the grooves and connected the SiPMs and PCB. Figure taken from [45].

2.3 Simulation framework

To identify structural anomalies and draw conclusions in muography, it is essential to compare measured data to accurate simulations. Simulations can be performed under various hypotheses—for example different densities or cavity dimensions—to determine the best agreement between observed data and simulation [16]. To ensure meaningful comparisons, the simulation should be as realistic as possible; the target, the interaction of muons with matter and the detector response should be modeled realistically. Generally, the simulation pipeline of any muography experiment can be characterized by the following phases:

1. Cosmic ray generation,
2. Passage of muons through matter,
3. Hit coordinate reconstruction,
4. Track reconstruction,
5. Image reconstruction.

In the following subsections, each simulation phase will be explained concisely in the context of the ScIDEP project.

2.3.1 Cosmic ray muon generation

The first step in the simulation process is the generation of cosmic-ray muons and background particles. Various cosmic-ray generators are available, and the choice depends on the requirements of the study and the subject that is being investigated, as some generators are adapted for underground targets for example. CR generators can broadly be divided into three categories: cosmic-ray air shower generators, parametric generators and special generators [25]. Cosmic-ray air shower (CRAS) generators simulate the complete cascade of secondary CRs produced by the interaction of primary CRs with the atmosphere, while parametric generators are based

on mathematical models of the CR muon flux. The parametrization of the latter category can either be derived from experimental data or from outcomes of simulations using CRAS generators. Lastly, special generators are developed for specific applications, such as underground or underwater muons.

Two widely used CRAS generators are CORSIKA and CRY [23, 24]. CORSIKA (COsmic Ray SImulation for KAscade) simulates the evolution of the cosmic-ray air showers in the atmosphere initiated by the primary CRs. All generated particles are tracked until they interact, decay or are absorbed. While this exensive tracking makes it a detailed simulation, it also makes it computationally intensive. In contrast, CRY is based on precomputed input tables, making it faster but more approximate: only protons are used as primary cosmic rays, the energy range is limited and the atmosphere is modeled in a simplified manner.

EcoMug (Efficient COsmic MUon Generator) is a parametric Monte Carlo generator tailored for muography applications, as it takes the muon distribution into account. With EcoMug it is possible to generate cosmic ray muons from a hemispherical, planar or cylindrical surface, in contrary to other generators, which usually only allow flat generation surfaces. This variety allows for a more efficient generation of nearly horizontal muons, particulary useful in case the detector views the target from an angle. The muon generation is implemented in such a way that the generated tracks follow the realistic angular and momentum distribution, regardless of the shape of the generation surface. Furthermore, EcoMug provides users the flexibility to generate cosmic ray muons according to user-defined parameterizations of their differential flux. The advantage of a parametric generator is generally a faster speed of execution and the possibility to pre-filter muons based on momentum or direction. A comparison of the performance of EcoMug, CORSIKA and CRY was done by [16].

Examples of special generators are muTeV and MUPAGE [25]. The former is developed to generate high-energy muons for underground or underwater laboraties, while the latter is mainly designed for underwater or underice neutrino experiments. This category of generators is not relevant for the objectives of the ScIDEP project and therefore will not be adressed in more detail.

2.3.2 Passage of muons through matter

After muons are generated, their interaction with both the pyramid and the detector has to be simulated accurately. Several simulations tools are available for this purpose, such as Geant4 or PUMAS [58, 59]. While Geant4 provides the possibility to accurately model the detector and pyramid, PUMAS does not offer this feature and is specifically meant for muon transport through the target. To realistically model the energy loss and potential scattering of the muons, the pyramid must be defined with precise dimensions and density. Similarly, the detector has to be modeled according to reality, incorporating the correct geometries and material properties. In addition, the detector response requires accurate simulation. When a muon traverses the scintillator and deposits energy, this energy is converted into photons, which spread over the scintillator. Photons crossing the WLS fibers are absorbed and re-emitted, after which they are transported via internal reflection in the fibers to the SiPMs. This process should be modeled with the specific optical properties. Furthermore, the relevant physical interactions themselves, that the muons or secondary particles undergo, have to be simulated accurately. The specific simulation model developed and used in this thesis is described in detail in Chapter 3.

2.3.3 Hit coordinate reconstruction

When a muon traverses a scintillator detector plate, the produced photons that cross a WLS fiber are transported into the fibers to the SiPMs, where the signal will be read out. A conceptual illustration of this is shown in Figure 2.7. As previously mentioned, each detector plate

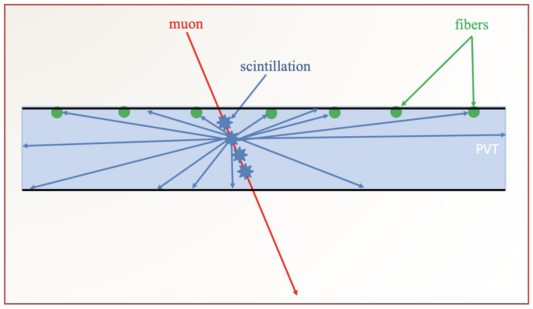


Figure 2.7: Muon traversing a PVT plate producing photons, some of which are captured and reflected through a WLS fiber. Figure taken from [51].

comprises 120 channels—60 for the X-dimension and 60 for the Y-dimension. Since the detector consists of one homogeneous plate, scintillation light will spread across the plate. Consequently, a single muon hit will typically result in multiple channels being activated. To determine the actual hit coordinate, the detector response must be digitized. Through a process of clustering, the hit coordinate is reconstructed from the signal distributed across the different channels. The reconstructed hit position is calculated through the center of gravity, as a weighted average of the fiber positions x_i and y_i:

$$x_{\text{hit}} = \frac{\sum_i^n w_{xi} x_i}{\sum_i^n w_{xi}} \quad \text{and} \quad y_{\text{hit}} = \frac{\sum_i^n w_{yi} y_i}{\sum_i^n w_{xi}} \,. \tag{2.1}$$

Here, x_{hit} and y_{hit} are the reconstructed x- and y-coordinate, with the weights w_{xi} and w_{yi} being either the number of detected photons per channel, or the total energy deposited per channel. A reconstructed hit coordinate is characterized by two parameters: the total energy deposit and the hit position [16]. If a muon traversing a scintillator plate happens to activate multiple channels, the ensemble of these channels is typically referred to as a cluster.

2.3.4 Track reconstruction

This list of reconstructed hit coordinates is passed to a tracking algorithm, which reconstructs the muon trajectories. A tracking algorithm uses mathematical methods to determine the most likely muon trajectory from a set of input coordinates. Its purpose is to identify the optimal combination of clusters while suppressing noise, such as secondary particles. The development and performance of a novel tracking algorithm for muography is explained in Chapter 4.

2.3.5 Image reconstruction

Once the directions of the CR muons are retrieved from the tracking algorithm, it is possible to reconstruct an image of the pyramid. The first step in this process involves constructing a

two-dimensional muon flux map as a function of the direction angles. After doing this for both the free-sky configuration and target configuration, a transmission map can be reproduced, using the definition in Equation 1.13. This map is then converted into a two-dimensional opacity map, where higher opacity corresponds to regions with greater density of thickness, and lower opacity may indicate cavities or less dense areas. By comparing the measured muon flux with the expected flux through a relative transmission map, anomalies can be revealed that may hint at unknown features or hidden cavities inside the pyramid. The process of image reconstruction is discussed in Chapter 5.

Chapter 3

Simulations

From generating muons and simulating their interactions with the pyramid to recording detector hits and storing the output: a muography simulation involves many sequential steps and demands a careful setup. To draw legitimate conclusions, the simulation has to be fine-tuned accurately according to the experiment and requires enough statistics. Both the detector and pyramid must be accurately modeled, and an appropriate simulation toolkit must be selected to handle muon interactions, as discussed in Section 3.1. Additionally, cosmic ray muons must be generated according to their realistic energy and angular distribution. Selecting a suitable CR muon generator depends on the specific objectives of the muographic study, as described in Section 3.2. One obstacle to muography is obtaining a sufficient number of detector hits: many CR muons generated in the atmosphere will never reach the detector, either because they will miss it or because they are absorbed in the target. Consequently, it will computationally be extensive—demanding significant power and time—to simulate and acquire enough muon hits. To counter this challenge, an optimized yet approximative simulation has been developed, which restricts the muon generation to an inverted cone aligned with the detector's field of view, as explained in detail in Subsections 3.2.1 and 3.2.2. Once the muons have been tracked through the pyramid, the relevant hit information must be recorded and stored, as discussed in Section 3.3.

3.1 Detector and pyramid characterization

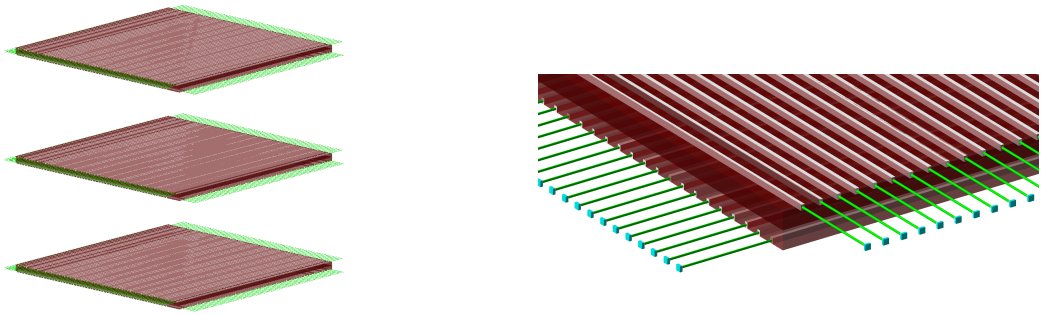
3.1.1 Geant4

To simulate the passage of CR muons through the pyramid and the detector, Geant4 was utilized, a state-of-the-art Monte Carlo toolkit for simulating the interaction of particles with matter [58]. Developed in C++ in an object-oriented framework, it allows users to model complex detector geometries and accurately track particle trajectories as they traverse various geometries and materials. Geant4 uses advanced algorithms to extensively model different physical interactions, including processes that result in multiple scattering and energy loss, leading to an accurate simulation of the muon's interaction within the object. The simulation parameters can be adjusted according to the experimental setup, including materials, detector characteristics, and relevant physics processes. This flexibility, in combination with its ability to precisely simulate complex particle interactions, makes Geant4 a powerful tool for studying muon interactions in detail.

A Geant4 simulation—commonly referred to as an application—is composed of several user classes, which allow customization of various aspects of the simulation [60]. User initialization classes are used to define and configure the experimental setup, detector geometry for example, and the physics processes before the start of a simulation run. User action classes, on the other hand, are responsible for defining behavior during the run itself. These may include run, event, tracking, and stepping processing, as well as primary particle generation. A run in Geant4 consists of a sequence of events, which all share the same detector configuration and physics list [61]. An event begins with the generation of one or more primary particles, after which each particle is tracked through the geometry, and secondary particles are produced through the earlier defined physical interactions. Each secondary particle is likewise tracked. This continues until the event's particle stack is empty, meaning that there are no more particles to be tracked. Particles may be removed from the stack if they exit the experimental boundaries or fall below a certain, user-defined, energy threshold for example. A track can be seen as a snapshot of a particle, it represents the current state of a particle as it moves through the geometries. It is updated every step, where a step captures the differential information of a particle over a segment of its trajectory. The simulation of a detector response is achieved using the sensitive detector class. Whenever a particle crosses a sensitive detector volume, a hit object will be generated. A hit is defined as a snapshot of the particle's physical interaction within the sensitive region of the detector. Hit objects can store step information such as position, time, and energy deposition during the step, as well as geometrical information such as volume identifiers. For each type of sensitive detector, a corresponding hit class is created to specify which information should be stored. Moreover, it is possible to set certain conditions, such that only hits meeting these predetermined requirements—particle type for instance—are stored in a hit collection, which can be accessed at the end of an event to extract the relevant data. The physics used in the Geant4 simulations can be configured either through a predefined model—known as a reference physics list—or by constructing a custom physics list. A widely known and currently Geant4's default reference physics list is `FTFP_BERT`. Each process in this list is constructed using cross sections and interaction models that define the detailed physical behavior. There is no implementation of optical photons in this physics list, so to simulate the full detector response this should be added explicitly.

3.1.2 Modelization

A detailed model of the first muon detector, as described in Subsection 2.2.1, was implemented for this thesis. Three tracking stations were modeled, as seen in Figure 3.1a. This resulted in a total of three tracking stations, one more compared to the configuration described in Subsection 2.2.1. The motivation for this extra tracking station is discussed in Chapter 4. The grooves were modeled into the tracking stations by subtracting volumes of $2 \times 3 \times 60$ m^3 from the plates. The tracking stations were spaced 30 cm apart and modeled as EJ-200 PVT, with the specific material properties, such as absorption length and scintillation yield, taken from the manufacturer's product sheet [50]. EJ-200 has a long optical attenuation length and fast timing, with an emission spectrum peaking at 425 n within a range of 400-500 nm [62]. These scintillator properties were included in detail in the GEANT4 material description. The WLS optical fibers were placed in the grooves and modeled as a BCF-92 formulation, again with the material properties provided by the manufacturer [52]. BCF-92 absorbs photons in the 359-458 nm range and reemits photons in the 465-502 nm range [62]. In Geant4, these fibers were modeled as plexiglass ($C_5H_8O_2$) with a density of 1.19 g/cm^3 and were then assigned the specified optical properties such as spectrum, refractive index, and decay time. A close-up of the WLS fibers and SiPMs is shown in Figure 3.1b. The SiPM windows were modeled as $3 \times 3 \times 0.3$ mm^3 glass plates and



(a) Detector model.

(b) Close-up of the WLS fibers and SiPMs.

Figure 3.1: Model in Geant4 of the scintillator detector described in Subsection 2.2.1, as well as a close-up of the placement of the WLS fibers and SiPMs. Colors are chosen arbitratrily for visualization purposes and do not represent the actual detector model.

the SiPMs themselves as $3 \times 3 \times 1.15$ mm^3 aluminum boxes. All definitions of elements and materials used in the simulation were extracted from the National Institute of Standards and Technology (NIST) database.

A three-dimensional model of the Pyramid of Khafre and its known internal structure was downloaded as an STL file [63], which included the burial chamber and both entrances that lead to it, as well as the subsidiary chamber. This model was modified in the 3D computer graphics software tools Blender and MeshLab, to ensure it was closed and suitable for import into Geant4 [64, 65]. During this process, the corridors were separated from the pyramid body. Both these models were saved in separate ASCII-format STL files and imported into Geant4 as a solid using the CADmesh tool [66]. Within Geant4, the corridors were subtracted from the pyramid body, resulting in the final geometrical pyramid structure as known today. The imported model in Geant4 is shown in Figure 3.2. This structure was modeled as limestone,

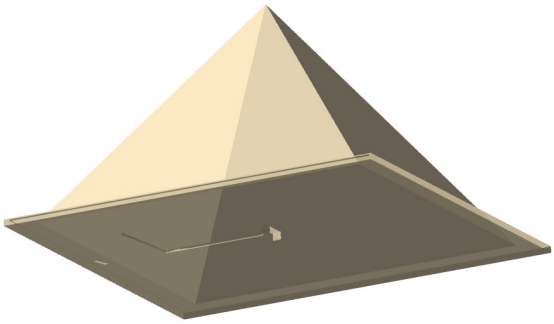


Figure 3.2: Model in Geant4 of the Pyramid of Khafre including its known internal features.

represented by calcium carbonate ($CaCO_3$) provided by the NIST material database. Since the lower corridor lies primarily below ground level, underneath the base of the pyramid, it is not

visible in the model shown in Figure 3.2. To include the underground parts, a ground volume should be implemented in Geant4. The ground can be modeled as standard rock or another material that more accurately represents the geological composition of the Giza plateau. The corridor structure can then be subtracted from this ground volume.

The detector and the pyramid were encapsulated in a shared World volume in the simulation. This World volume, which represents the experimental area of interest, was a cube of 500×500 m^3 and was composed of air. Consequently, the voids within the pyramid were filled with air. The global coordinate system is defined by the World volume, with the origin being the center of the cube. The pyramid was placed inside the World volume such that its base center aligned with the origin and the origin corresponded to ground level. As a result, because of the positioning of the King's chamber within the pyramid, the origin lied inside the burial chamber. The vertical axis of the pyramid aligned with the z-axis. The detector was placed around the origin of the world volume, with its first scintillator plate 9 cm above the origin and the two subsequent tracking stations always 30 cm below the preceding one.

For this thesis, the processes of interest were Bremsstrahlung, ionization, pair production, and multiple scattering for muons. Although additional processes—optical processes or interaction processes for secondary particles—were also included in a custom physics list, these were not relevant for this analysis, as only muons were tracked. The motivation for focusing solely on muons is given in Section 3.3.

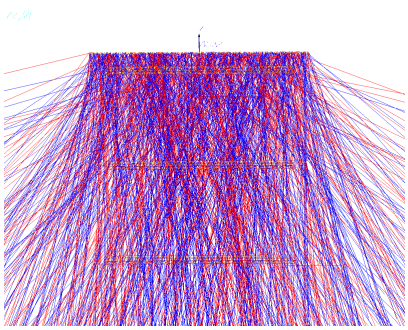
To record hit information when a particle traverses a scintillator plate, the detector plates were set as sensitive detectors in Geant4. A requirement was implemented such that only hits from primary muons with a non-zero energy deposition were stored, to prevent the recording of secondary particles. A complete simulation of the detector response—the optical photon propagation within the scintillator and the corresponding SiPM response—would require marking the SiPM windows as sensitive detectors, to check of many photons reach the SiPMs, and transform this into a realistic output. However, simulating the propagation of these optical photons significantly increases computational time, and was therefore omitted from the simulation. Nevertheless, it is important to note that in an actual experimental setup, the true positions are not directly observable, as only the signal from each channel is recorded. Thus, future studies will need to incorporate the hit reconstruction based on the realistic detector output, also investigating the influence of noise.

3.2 Cosmic ray muon generation

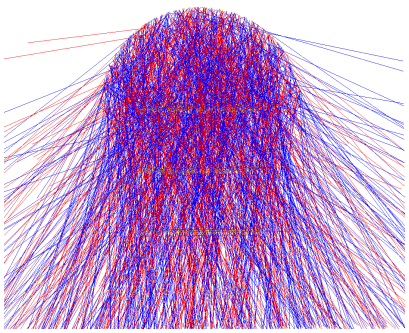
To generate the CR muons, various CR generators are available, as explained in Subsection 2.3.1. The most straightforward way to generate a particle in Geant4 is by employing a particle gun from the `G4ParticleGun` class [60]. It allows users to specify the particle type, momentum, direction, and position, among other variables. In the case of muography, the particle gun is a simplified setup, as it does not take the angular and energy distribution of the CR muons into account. Because of this, it should not be used to simulate muographic images intended to be compared to observed data. Nevertheless, the method still has its benefits: due to its simplicity, it is an ideal option for initial testing and debugging. In this thesis, the particle gun was employed to generate artificial muons for acquiring hit data to develop, test, and optimize the tracking algorithm (see Chapter 4). It was also used to scan the traversed length through the pyramid across all directions and to obtain a relationship between incoming and outgoing en-

ergies for a muon traversing the pyramid, two essential things to optimize the simulation (see Subsection 3.2.2).

For a more realistic CR muon generation, EcoMug was selected as the primary generator in the simulation. As a parametric generator, it offers both computational efficiency and flexibility in defining the generation surface. The planar and hemispherical generation surfaces from EcoMug were investigated, with two examples shown in Figure 3.3a and 3.3b respectively. Due



(a) Planar generation surface of size 70×70 cm^2 centrally positioned 5 cm above the uppermost detector plate.



(b) Hemispherical generation surface with a radius of 50 cm centered on the uppermost detector plate.

Figure 3.3: Planar and hemispherical generation surface of the EcoMug generator, shown for a generation of 1000 muons in Geant4. The red and blue lines correspond to the muon tracks.

to its computational efficiency, EcoMug is well-suited for producing preliminary simulations within the limited time frame of this research. Its seamless integration with Geant4 furthermore enhances its utility for this analysis. It is important to note that while EcoMug generates the muon's energy, position, and direction, the muon itself must still be fired through the Geant4 particle gun.

The first generation surface that was explored, was a hemisphere centered at the origin with a radius large enough to cover the entire pyramid. However, as mentioned in Chapter 1, most generated CR muons will never reach the detector, either because they will miss the detector or because they will be absorbed in the target. When generating 1 billion muons in Geant4 on a hemisphere covering the pyramid, only roughly 1500 events[1] were recorded. The simulation of these 1 billion muons traversing the pyramid took approximately 1 hour and 30 minutes. Hence, it would have taken an incredible amount of time to acquire enough data to reconstruct the image of the pyramid with a proper resolution and statistics. To solve this computational obstacle, the development of an optimized simulation is proposed, which limits the generation of muons inside an inverted cone comprising the field of view of the detector. EcoMug does offer constraints on the generated zenith and azimuthal angle, yet it does not offer a function to limit the generation of muons inside an inverted cone. The principle of the optimized simulation is to first generate the muons close to the detector, using a small generation surface, which decreases the probability of muons missing the detector planes. Muons that enter a region close enough to the detector will then be extrapolated to outside the pyramid, where they will be

[1] An event is defined as a muon hitting all three tracking stations, not to be confused with the event definition of Geant4 as explained in Subsection 3.1.1.

fired and tracked through the pyramid. Finally, muons that actually hit all three detector planes are accepted as events and their information is stored in a file for further analysis.

3.2.1 Small surface generation

The first step is to generate muons on a small generation surface. This generation surface should be sufficiently small, to ensure a significant fraction of the generated muons passes near the detector, yet large enough to enclose the entire detector field of view. A muon is defined as a nearby muon if its trajectory passes within the range of the detector plate surface plus an additional foreseen margin, as illustrated in Figure 3.4. The area of the detector plus margin will

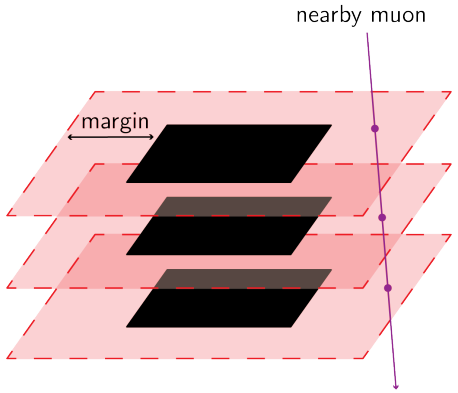


Figure 3.4: Simplified illustration of the three tracking stations (black) surrounded by an imaginary margin, depicted in red. The purple dots represent the intersection of the trajectory of the nearby muon with the corresponding margin areas.

be called the nearby area. It is important to remark that at this stage, the muon is not yet actually fired; however, the generation information such as position and momentum is available through the EcoMug generator. This information is used to approximate the trajectory of the muon as a straight line and calculate its intersection with the nearby area. If this intersection falls within the nearby area for all three detector planes, it can be reasonably argued that the muon is likely to hit all three detector planes. Consequently, a nearby muon will be extrapolated to outside the pyramid.

To find the optimal generation surface, resulting in the most number of hits in all three planes, a study was done using the free sky configuration. Hemispherical and planar generation surfaces were explored, differing in sizes and central positions. For each of these surfaces, 3 million muons were generated and fired. The hits on the uppermost detector plane are shown in Figure 3.5 for four configurations of generation surfaces. While a hemisphere with a radius of 40 cm positioned in the center of the uppermost detector plane results in more hits than a similarly placed hemisphere with a radius of 50 cm, it is obvious from the upper right plot that not the entire field of view of the detector is covered. This does not come as a surprise, since a detector plane with a surface size of 60×60 cm^2 needs a hemisphere with a radius of minimum 42.43 cm for full coverage. For a generation surface placed at a distance d above the uppermost detector plane, one needs to be careful that its size scales with d, such that it still covers the detector's field of view. To compare the hemisphere with a planar surface, a plane of size 70×70 cm^2 positioned 5 cm above the uppermost detector plane was modeled. By comparing the upper

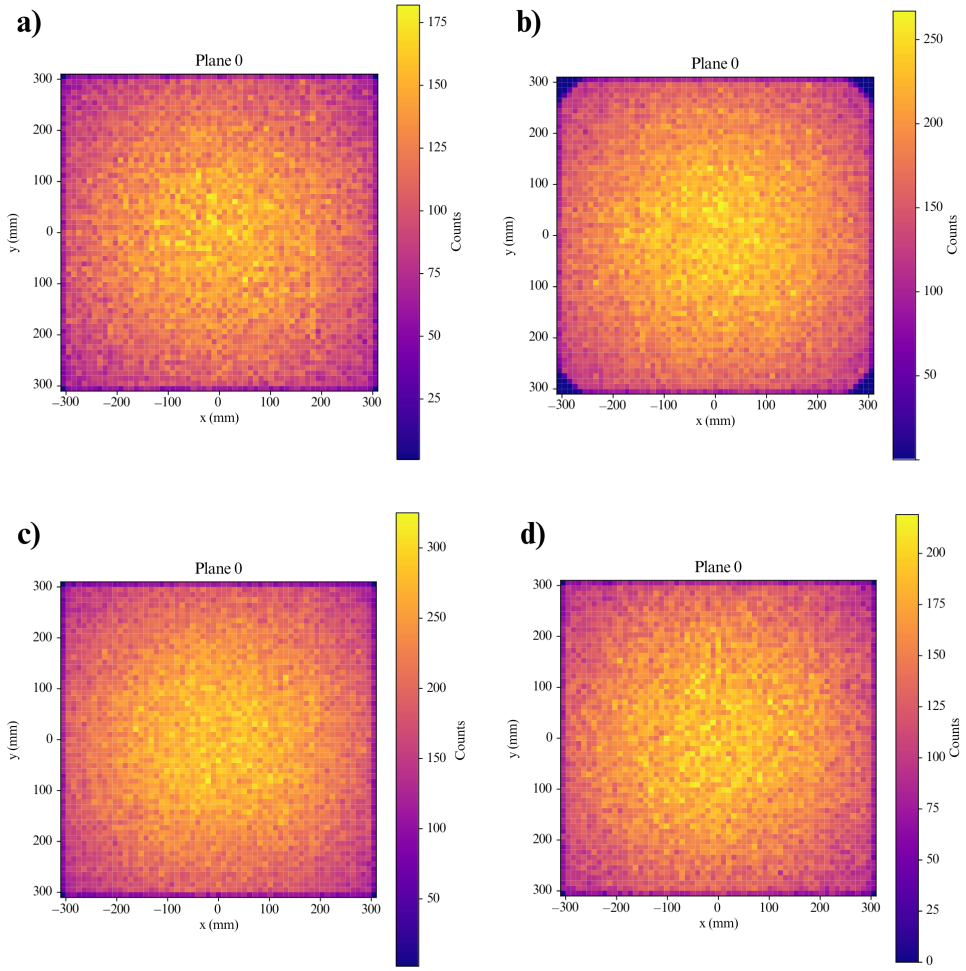


Figure 3.5: Number of hits on the uppermost detector plane for a) a hemisphere with a radius of 50 cm positioned at $(0, 0, 9)$ cm, b) a hemisphere with a radius of 40 cm positioned at $(0, 0, 9)$ cm, c) a plane of 70×70 cm^2 positioned at $(0, 0, 14)$ cm and d) a plane of 85×85 cm^2 positioned at $(0, 0, 19)$ cm.

left and lower left plots, it can be deduced that the planar surface results in more hits. A bigger planar surface of 85×85 cm^2 positioned 10 cm above the detector—as seen in the lower right plot—was also explored, however, this did not result in more hits. Consequently, the planar generation surface of 70×70 cm^2 positioned 5 cm above the uppermost detector plane was selected as the small generation surface. To find the ideal margin, another study was conducted in a later stage of this optimized simulation, as discussed in Section 3.3. Other notions of a muon passing near the detector could be defined, such as it passing in a volume (sphere or cube for instance) around the detector.

3.2.2 Muon extrapolation

If a generated muon is classified as a nearby muon, the next step is to extrapolate it to the outside of the pyramid. To determine a suitable position to generate the muon from, outside the pyramid, a large surface—like a plane or hemisphere—can be defined. Then, it is only a matter of calculating the intersection between the surface and the muon's trajectory—which is approximated as a straight line defined by the initial position and direction by Ecomug in the first step. While calculating the intersection with a plane is mathematically less complicated, a hemisphere offers the advantage of confining the generation point closer to the pyramid. In turn, this allows for a smaller World volume to be used in the simulated, which is preferred to reduce computational time.

When a muon passes through the pyramid, it will interact with the surrounding matter, resulting in a loss of energy. Evidently, the muon's energy when reaching the detector is lower than its energy when entering the pyramid. Thus, to extrapolate a muon from near the detector to outside the pyramid, this energy loss has to be accounted for, to fire the muon with the correct energy. The energy loss of a muon traveling through matter is a non-trivial calculation, as its interactions with this surrounding matter are stochastic. The average energy loss for a muon traveling through a medium is described by Equation 1.9. This expression highlights the complexity of the calculation: the energy loss per traversed unit of length is a function of the energy itself, along with the factors $a(E)$ and $b(E)$, which depend on the energy in a non-trivial way. It is due to these complex interactions that MC simulation tools such as Geant4 exist, to simulate these processes that are analytically unsolvable. The aim of this analysis, the extrapolation, is even more challenging: instead of calculating the energy loss, the energy gain has to be estimated. Doing this with the same level of accuracy as Geant4's forward simulation requires the calculation of the reverse physics processes. However, for muons, these reverse processes have not been implemented yet in Geant4. While specialized tools like PUMAS have been designed specifically for muography to work in backward mode, these are often hard to integrate into the Geant4 framework [59, 16]. Consequently, an alternative approach was pursued in this thesis.

This novel approach is an approximation of the reality, nevertheless, for producing preliminary images of the pyramid within the limited time frame of this research, it serves its purpose. From Equation 1.9, it can be deduced that the energy that the muon has before entering the pyramid—called the incoming energy E_{in}—can be related to the traversed length through the pyramid and the muon's energy when reaching the detector—called the outgoing energy E_{out}. To explore this relationship, a simulation was implemented in Geant4 to generate a two-dimensional data histogram. The principle is that, for a given traversed length and outgoing energy, the corresponding bin can be identified from which the incoming energy can be acquired. In the simulation, a limestone block with a length of 150 m and a cross section of 5×5 m^2 was modeled. A muon particle gun was positioned to shoot particles through the cross section into the block, as seen in the setup in Figure 3.6. Since the average energy of CR muons at sea level is around 4 GeV, and the relative energy loss is expected to be larger for low-energy muons, it was motivated to split the data into different energy histograms. The intervals for the incoming energy for the generated muons were selected as follows: $[0, 300]$ GeV, $[200, 2300]$ GeV, $[2000, 13000]$ GeV and $[10000, 55000]$ GeV. This way, the lower energy histograms can be reconstructed with a higher resolution capturing more variability, providing a more detailed relationship between traversed length, incoming, and outgoing energy. For each of these ranges, a run of 1 million events was executed. Per event, one muon was generated with a randomly assigned energy from the designated range. The overlap between the consecutive energy ranges

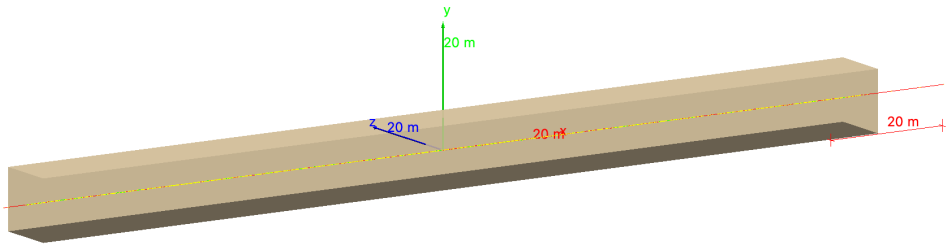
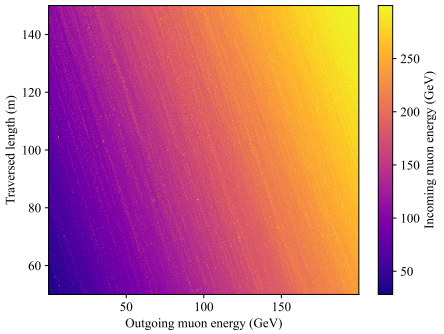


Figure 3.6: Model in Geant4 of the limestone block used to acquire the energy histograms. The red/yellow line corresponds to the trajectory of a muon.

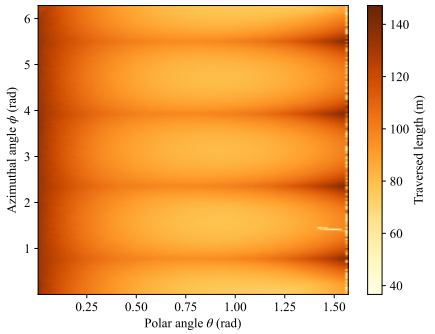
is intentional, as it accounts for the fact that the outgoing energy of a muon will be lower than its incoming energy. This will result in a data gap in the histogram for bins corresponding to large traversed lengths and outgoing energies. Consequently, when reconstructing the energy histograms, to remove this data gap, a cut was applied on the outgoing energy intervals, such that these correspond to $[0, 200]$ GeV, $[200, 2000]$ GeV, $[2000, 10000]$ GeV and $[10000, 50000]$ GeV. For outgoing energies exceeding 50000 GeV, it was observed that the relationship between the incoming and outgoing energy was relatively stable, exhibiting less variation in terms of the traversed length, compared to lower energy ranges. Hence, a constant value of 2000 GeV was added to estimate the incoming energy in this case. It should be noted that this approximation is quite rough and should be refined in the future[2]. For every generated muon, the energy E_{out} it had inside the limestone block after traversing a certain distance, was recorded in steps of one meter from 50 to 150 meters. For each combination of the traversed length at a step and corresponding energy E_{out}, the histogram was populated with the generated energy E_{in}, capturing this complex three-dimensional relation. As a result, given a specific value for the traversed length and outgoing energy E_{out}, the average incoming energy E_{in} can be extracted from the corresponding bin. An example of an energy histogram is shown for the range of $E_{out} = [0, 200]$ GeV in Figure 3.7a.

It is important to note that this approach only offers an approximation of reality, as the energy loss depends on the specific stochastic interaction processes involved. Nonetheless, as an initial method for reconstructing two-dimensional images of the pyramid in a relatively short time frame, this approach can be used to obtain a preliminary understanding of what to expect for this image reconstruction. Another method to determine the relationship between the incoming energy, outgoing energy, and traversed length could involve fitting a model to this relationship or solving the differential equation in Equation 1.9—but without the negative sign, to reflect the energy gain—through a numerical procedure. The benefit of fitting a model is that it would describe a continuous relationship rather than a binned one, yet the question remains which appropriate model to choose. On the other hand, the numerical procedure has the advantage that there is no necessity to guess a model, though it demands more elaborate calculations. Additionally, for both methods, it remains challenging to verify that the obtained model reflects reality. Furthermore, neither approach captures the stochastic nature of muon energy loss in a medium.

[2]Nevertheless, the outgoing energy in the final optimized simulations never exceeded 2000 GeV, thus this approximation did not affect the results.



(a) Incoming energy histogram in function of the outgoing energy and traversed length of a muon crossing the pyramid. This example is for an energy range of $E_{out} = [0, 200]$ GeV.



(b) Traversed length histogram in function of the azimuth ϕ and zenith (polar) angle θ.

Figure 3.7: Data histograms to acquire the incoming energy and traversed length to use for extrapolating muons outside of the pyramid.

To use the energy histogram, it is necessary to determine the distance traversed by the muon through the pyramid. In Geant4, the traversed length of a particle traveling through a certain volume can be calculated conveniently using the `SteppingAction` class, as done to generate the energy histogram. However, this method is unsuitable for the optimized simulation, as the traversed length is required before the muon actually traverses the pyramid volume. Instead, the data was also binned in a two-dimensional histogram, containing the traversed length as a function of the azimuth and zenith angle. In a simulation in Geant4 including the pyramid model, a particle gun was placed at the origin inside the King's chamber, practically at the place where the detector normally is positioned. A run was simulated with 100, 000 events, where per event a geantino was fired with a randomly assigned direction in the upper hemisphere, to scan the whole pyramid. A geantino is an auxiliary, imaginary particle in Geant4 that is transported through the geometry without interacting, making it beneficial for relatively fast simulations aiming to trace a structure or for debugging purposes [60]. The zenith and azimuth angles were divided into bins of 1°, where each bin was populated with the average traversed length. Due to the relatively low number of events, the histogram still contained some empty bins. These were filled by performing an extrapolation and the resulting histogram is shown in Figure 3.7b. This histogram effectively illustrates the influence of the pyramid's features—the corners and the top—on the traversed length, which are also expected to be present in the reconstructed muographic images. For $\theta \approx 0$, the traversed length is higher as this corresponds to the top of the pyramid. The same is visible for the corners of the pyramid ($\phi \approx \pi/4$, $\phi \approx 3\pi/4$, $\phi \approx 5\pi/4$ and $\phi \approx 7\pi/4$), where the maxima are reached for $\theta \approx \pi/2$ at the bottom corners of the pyramid. The small decrease in traversed length for $\theta \approx 1.40$-1.55 corresponds to the upper corridor that leads to the King's chamber. Given that CR muons which cross the detector plates will not always pass exactly through the origin, the acquired histogram is still an approximation. Nevertheless, considering that the detector is relatively small compared to the pyramid, the detector can be approximated as a point for this specific case.

Both the energy and length histograms were stored as a two-dimensional histogram in ROOT format, as this allows for easy importation and data extraction in Geant4.

3.3 Muon tracking through pyramid

All necessary information is now available to extrapolate nearby muons outside the pyramid, where they will be fired by a particle gun with the extracted incoming energy E_{in}. The muons were fired from a hemisphere of 200 m, centered around the origin, by calculating the intersection between this hemisphere and the muon's trajectory defined through its generated direction and position by EcoMug. After a muon is fired, it is tracked by Geant4 through the pyramid, calculating the interactions and energy losses at each step. An image illustrating the generation cone of the optimized simulation in Geant4 is shown in Figure 3.8. Only muons were tracked,

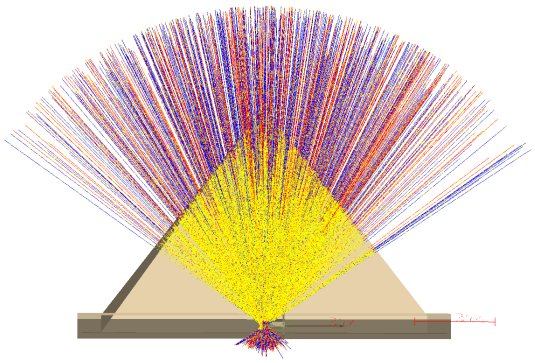


Figure 3.8: Simulation in Geant4 of the generation cone, where muons are tracked through the pyramid, shown for 1000 generated muons at the small generation surface.

killing all tracks from secondary particles, to speed up the simulation process. This means that the resulting hits will be absent of noise, hence the detector response is not fully simulated. However, as mentioned before, the optimized simulation is intended as an initial approach for reconstructing the two-dimensional images of the pyramid and requires refinement in the future. When a muon passes through all three planes, it is considered an event, and its generated information and hit information are stored in a ROOT file.

As discussed in Subsection 3.2.1, a margin was implemented to decide which muons should be extrapolated. This margin is motivated by the fact that muons will undergo interactions when traversing the pyramid, possibly resulting in a minor change of direction. As a result, a muon generated on the small generation surface, that would slightly miss the detector, might, after being extrapolated and tracked through the pyramid, still end up hitting all three detector planes. Including this margin essentially means that the inverted generation cone will be bigger than the field of view of the detector. The cone must be wide enough to include muons that slightly change course—to maximize the number of hits—but narrow enough to avoid an excessive generation of muons that will never reach the detector. To recover this optimal margin, the optimized simulation was run for a fixed number of 100000 muons, for different margins. Finally, a margin of 50 cm was selected as it provided the optimal balance between maximizing hits and minimizing unnecessary muon tracking.

Chapter 4

Muon Track Reconstruction

After simulating the muons passing through the pyramid, the result is a collection of hit coordinates recorded by the detector. To determine the directions from which they traversed the pyramid, their trajectories must be reconstructed from these hit coordinates. These directions are essential for the image reconstruction process discussed in the next chapter. To identify straight-line trajectories from a sparse set of hit coordinates, a reconstruction algorithm based on the Hough transform is employed. The Hough transform is a computational method for detecting geometric shapes such as lines, as clarified in Section 4.1. For this purpose, straight-line trajectories in 3D are represented using a non-redundant four-parameter model, introduced in Section 4.2. From this theoretical foundation, a track reconstruction algorithm was developed tailored to the ScIDEP project's data, as detailed in Section 4.3. To evaluate the performance of this algorithm, efficiency studies were conducted and are discussed in Section 4.4. Since the algorithm is designed to detect straight-line trajectories, its performance was first assessed under these ideal conditions, to explore the relationships and dependencies among various tracking parameters, as described in Subsection 4.4.2. Lastly, recognizing that in realistic scenarios the hit coordinates may be affected by smearing, an additional analysis was performed to investigate the effects of noise on the algorithm's performance, as presented in Subsection 4.4.3.

4.1 Hough transform

Given a set of hit points, the aim is to find the true muon trajectories through these points. At first sight, this might seem straightforward: fitting a straight line through a set of points is a widely studied and established procedure today. However, due to limited detector resolutions (time and space), the fact that primary muon hits cannot be distinguished from the other particles, and the possibility of multiple muons hitting the detector at the same time (i.e. in the same triggered event), it is necessary to employ a tracking algorithm that is capable of detecting multiple tracks at a time, while also excluding noise hits. A well-known computational method to extract features in an image is the Hough transform. This technique is named after Paul Hough, who patented the idea in 1962 to detect lines in two-dimensional binary images [67]. Later, it was extended to include the detection of other parametric shapes in images [68]. For the scope of this thesis, only the application regarding straight lines will be discussed.

The concept of the Hough transform is to transform an image into a parameter space (interchangeably called the Hough space) and, using a voting scheme, find the optimal line parameters, which describe the lines present in the image. Consider a line in 2D in an image space,

with the cartesian representation

$$y = mx + b \tag{4.1}$$

where m is the slope of the line and b the intersection of the line with the y-axis. The aim is to recover these line parameters m and b. When ranging over values of x and calculating the corresponding y-values, the collection of coordinates (x, y) will form the line. Equation 4.1 can be transformed into

$$b = -xm + y \tag{4.2}$$

This can be seen as the representation of a 2D line with slope $-x$ and vertical intersection y in parameter space. Similarly, when ranging over m and calculating the corresponding b-values, the collection of coordinates (m, b) will form the line. Hence, every fixed point (x, y) in the image space will result in a line $b = -xm + y$ in the parameter space. Every point on this line represents a line in the image space going through (x, y). Suppose there are two points in the image space, namely (x_1, y_1) and (x_2, y_2), and assume the line going through them is described by the equation $y = m_H x + b_H$, as depicted in Figure 4.1 on the left. When these points are transformed into the parameter space, each point maps to a line given by equations $b = -x_1 m + y_1$ and $b = -x_2 m + y_2$, respectively, as illustrated in Figure 4.1 on the right. These two lines intersect at the point (m_H, b_H), corresponding to the slope and intercept of the original line going through (x_1, y_1) and (x_2, y_2) in the image space. This concept can be extended: any set of collinear points in the image space will result in lines in the parameter space that intersect at a common point representing the parameters of the line on which the collinear points lie. Thus, by finding these values where many lines intersect, the line parameters (m, b) can be

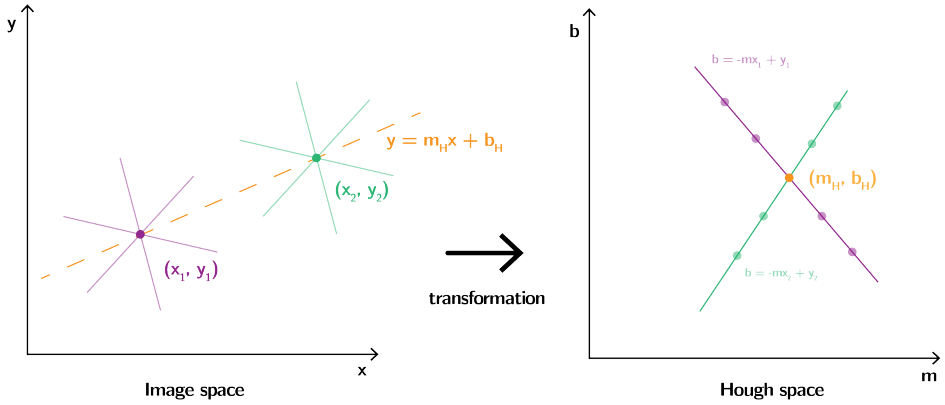


Figure 4.1: Illustrative picture of the principle of the Hough transform. The figure on the left depicts the image space with points (x_1, y_1) and (x_2, y_2), while the right figure depicts their transformation into the parameter space.

retrieved. Due to vertical lines giving rise to an unbounded slope ($m = \infty$) in the cartesian representation, the Hough transform is usually performed in a polar representation, known as the Hesse normal form:

$$x \cos\theta + y \sin\theta = \rho \tag{4.3}$$

Here, ρ is the shortest distance between the origin and the line, and θ is the angle between the horizontal axis and the perpendicular ρ, as illustrated in Figure 4.2. Analogous to before, now

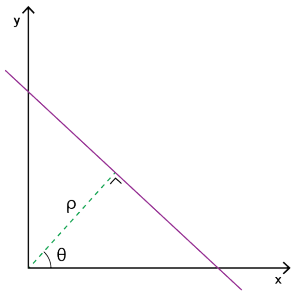


Figure 4.2: Definition of the parameters ρ and θ of the line representation determined by the Hesse normal form, as in Equation 4.3.

each point in the image space will represent a sinusoid in the parameter space. If n sinusoids in the parameter space intersect in a point (ρ_H, θ_H), then these values ρ_H and θ_H represent a line via Equation 4.3 going through n points in the image space. Summarized, the core principle of the Hough transform is to transform an image to a parameter space and find the points of concurrency.

To implement the Hough transform algorithm, both the image space and parameter space are discretized. The image space is usually naturally discretized into pixels, while the parameter space—for this example defined by the parameters ρ and θ—is discretized into a finite number of bins, denoted by n_ρ and n_θ. This discretized parameter space is known as the accumulator, consisting of $n_\rho \times n_\theta$ cells, where each cell stores a vote count, which is initially zero. Then, the algorithm iterates over each pixel (x, y) in the input image and, for a range of θ-values, calculates the corresponding ρ value using Equation 4.3. For each obtained (ρ, θ) pair, the vote count in the corresponding accumulator cell is increased by one. This process will result in a discretized sinusoid in the accumulator for each image point. After all points have been processed, cells with a high vote count—local maxima—correspond to a pair (ρ_H, θ_H) which represents the parameters of a potential line in the image.

The Hough transform offers several advantages, making it a widely used and recognized technique. First, its discretized voting scheme allows for line detection in noisy images [69]. Additionally, it can detect multiple lines at once and detect other parametric shapes, such as circles. These convenient features have led to its application in various fields, including traffic (lane detection), biometrics, object recognition and tracking, industrial (detection of pipe joints), medical image analysis, etc. However, a drawback is that the voting accumulator requires a significant amount of computational storage and time.

In two dimensions, a line is described by two parameters, resulting in a two-dimensional parameter space and corresponding Hough accumulator. Conversely, in three dimensions, a line requires a minimum of four parameters for its full representation, leading to a four-dimensional parameter space, which is more complex and requires more computational power. Consequently, other techniques and variations of the Hough transform were investigated for the three-dimensional reconstruction of the muon's trajectory. While a significant amount of research exists on 3D Hough-based or other line detection methods, most of these algorithms are based on edge images. Generally, these edge images consist of a large number of data points, as

each pixel in the image space has an assigned value. Hence, in that case, the local maxima in the accumulator will clearly stand out from the surrounding vote areas. In contrast, this research focuses on employing line detection using only a few hit coordinates per line, resulting in only a few votes per local maximum. In addition, any two points determine a straight line. Consequently, two hit coordinates is insufficient to reconstruct a muon's trajectory, as the votes for true tracks will not stand out to the votes of false combinations. For this reason, a third tracking station was modeled in the optimized simulation, as described in the previous chapter. Furthermore, due to this sparse hit data, it is not feasible to calculate a gradient image—a technique used in some of the existing line detection algorithms. This challenge led to the development of a new tracking algorithm, tailored specifically to the needs of this research.

4.2 Non-redundant line representation

To describe a line in three-dimensional space, at least four parameters are required. The most common representation of a line in 3D is the cartesian representation:

$$\vec{x} = \vec{p} + t\vec{b} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} + t \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} , \tag{4.4}$$

where $\vec{p}$ is an anchor point of the line and $\vec{b}$ is the direction of the line. This representation utilizes six parameters, making two redundant. In 1988, K. S. Roberts proposed a new line representation using four parameters, from now on called Robert's line representation [70]. It characterizes the direction by two parameters: the azimuth angle ϕ and the zenith angle θ, where θ defines the angle between the direction vector and the positive z-axis and ϕ the angle between the positive x-axis and the projection of the direction vector on the xy-plane, as shown in Figure 4.3 on the left.

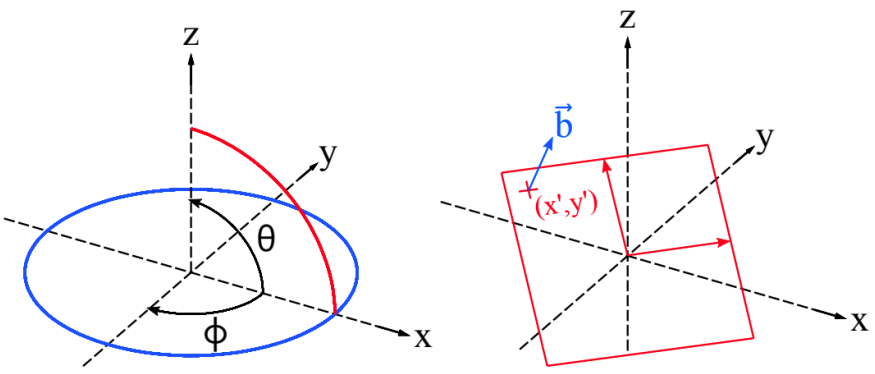


Figure 4.3: Illustration of the line parameters θ, ϕ, x' and y' as defined in Robert's line representation. The left figure demonstrates the azimuth ϕ and zenith θ angle, while the figure on the right shows the intersection point between the line and a perpendicular plane going through the origin. Figure taken from [71].

Then, the direction vector $\vec{b}$ can be written as

$$\vec{b} = \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = \begin{pmatrix} \cos\phi \sin\theta \\ \sin\phi \sin\theta \\ \cos\theta \end{pmatrix} , \tag{4.5}$$

already eliminating one of the redundant parameters. An ambiguity of this representation of the direction vector is that two anti-parallel direction vectors can describe the same line. To remove this, the parameters θ and ϕ are restricted such that the direction vectors are confined to the upper hemisphere, by requiring $0 \leq \theta \leq \frac{\pi}{2}$ and $-\pi \leq \phi \leq \pi$. After meeting this requirement, there remains an ambiguity in the xy-plane. This is resolved by implementing the restrictions $b_y \geq 0$ if $b_z = 0$ and $b_x = 1$ if $b_y = b_z = 0$. The two remaining degrees of freedom are used to describe the position of the line. A naive approach would be to pick an arbitrary anchor point, however, as this still involves three parameters, one of these redundancies can be eliminated. Robert's line representation defines two parameters x' and y' as the coordinates of the intersection of the line and a plane going through the origin with as its normal the line direction $\vec{b}$, as illustrated in Figure 4.3. The coordinate (x', y') is defined in a 2D coordinate frame in the plane itself. The coordinate frame is defined in such a way that it is unique and without special cases or singularities, with its definition described in Appendix A. For a line with an arbitrary reference point $\vec{p}$, the intersection coordinate (x', y') becomes:

$$x' = \left(1 - \frac{b_x^2}{1 + b_z}\right) p_x | \left(\frac{b_x b_y}{1 + b_z}\right) p_y | b_x p_z \tag{4.6}$$

and

$$y' = | \left(\frac{b_x b_y}{1 + b_z}\right) p_x + \left(1 - \frac{b_y^2}{1 + b_z}\right) p_y | b_y p_z \, . \tag{4.7}$$

4.3 Tracking algorithm

4.3.1 Procedure

The tracking algorithm for the detection of muon trajectories using the hit coordinates as input is based on the Iterative Hough Transform developed by Dalitz *et al.* [71], but tailored to the sparse data of this research. The proposed algorithm implements two kinds of accumulators to find the four parameters θ, ϕ, x' and y' describing a muon's track.

The tracking algorithm begins by applying a Hough transform to the ingested hit coordinates to obtain candidate trajectory directions. This is done using a $\theta\phi$-accumulator of size $n_\theta \times n_\phi$, where n_θ and n_ϕ stand for the number of bins along the θ- and ϕ-dimension, respectively. The amount of bins determines the inititial[1] angular resolution of the detected direction and can be tuned to optimize the performance of the tracking algorithm, as discussed in Section 4.4. For each unique pair of points from the input data, the direction vector defined by those two points is computed in terms of θ and ϕ, using the conventions described in Section 4.2. For every obtained (θ, ϕ), the vote of the corresponding cell in the $\theta\phi$-accumulator is incremented, if the computed value of θ does not exceed 80°. If $\theta > 80°$, the vote is discarded to avoid forming trajectories from hits within the same tracking station. This threshold could be tightened

[1]The choice for n_θ and n_ϕ does not reflect the definite angular resolution of the tracking algorithm, as a fit will be applied afterwards to refine the resolution, as explained later in the paragraph.

further; for example, also excluding pairs from consecutive tracking stations that exceed the detector's angular acceptance. Pairs of collinear points will give rise to identical θ and ϕ, resulting in regions of higher votes in the accumulator, which indicate the potential trajectory directions. Due to the discretized parameter space and reconstruction uncertainties, hit coordinates from the same trajectory may yield slightly different (θ, ϕ) values, possibly resulting in a spread of votes to neighboring cells. To account for this, a clustering algorithm is applied to the accumulator, to identify regions of high votes. This clustering algorithm is described in Section 4.3.2. After the clusters are detected, only clusters whose total vote count exceeds a certain threshold are retained. For each of these clusters, the weighted average of the (θ, ϕ) values is calculated, with as weights the vote count per cell within the cluster. This weighted average is denoted as (θ_H, ϕ_H) and represents a potential direction for a muon trajectory.

The second Hough transform is applied similarly. An xy-accumulator of size $n_{xy} \times n_{xy}$ is used, with n_{xy} being the number of bins along the x'- and y'-dimension. For every direction candidate (θ_H, ϕ_H), the intersection coordinate (x', y')—as defined through Robert's line representation in Section 4.2—can be calculated for every hit coordinate $\vec{p}$ in the input data, using Equation A.2 and A.3. For every computed value (x', y'), the vote of the corresponding cell in the xy-accumulator is incremented. Again, the votes are clustered, with the clustering algorithm applied to this accumulator playing a crucial role in the track reconstruction. While in theory, for a perfect line, an arbitrary point on the line will consistently yield the same value for (x', y'), this will not be the case in reality. Since (x', y') is obtained using (θ_H, ϕ_H), even a slight discrepancy from the actual trajectory's direction will result in differences in (x', y') for hit points belonging to the same trajectory. This discrepancy will be even more significant for the xy-accumulator, as it inherits the (θ_H, ϕ_H)-discrepancy caused by the discrete nature of the accumulator and weighted average of the detected cluster. In principle, the spreading of votes to neighboring cells can be reduced by reducing the number of bins n_{xy}. Nevertheless, this cannot be reduced arbitrarily, as a certain level of resolution needs to be maintained, and even for small xy-accumulators, vote spreading cannot be eliminated completely. Hence, clustering in this accumulator is an essential part of this tracking algorithm, with the discrepancy being one of the main reasons why the Iterative Hough Transform from Dalitz *et al.* had to be converted to meet the needs of this research [71]. For every detected cluster, a weighted average (x'_H, y'_H) is calculated in the same way as described before, which represents the potential anchor point of a muon trajectory.

This leads to a collection of candidate line parameters $(\theta_H, \phi_H, x'_H, y'_H)$. For each candidate line, the distance is calculated between this line and the hit points from the data input. If this distance is smaller than a certain threshold, the point is added to a collection X_{nearby}. For all the points in this collection, all the possible combinations are made of three hit points on different tracking stations, where combination i is denoted as a collection $X^i_{combination}$. This way, it is avoided that multiple hits from the same tracking station belong to the same line, in the case two hit coordinates from the same tracking station lie particulary close to each other for example. For each combination $X^i_{combination}$, a linear fit is computed through these points. To fit a line through three 3D points $\vec{x}_1, \vec{x}_2, \vec{x}_3$, the orthogonal least squares fitting method is used [71]. The anchor point $\vec{a}$ of the fitted line is calculated as the center of mass of the points

$$\vec{a} = \frac{1}{3} \sum_{i=1}^{3} \vec{x}_i \tag{4.8}$$

and the direction vector $\vec{b}$ can be calculated from principal component analysis (PCA): suppose

Q is the matrix derived from the centered hit points $\vec{q}_i = \vec{x}_i - \vec{a}$:

$$Q^T = (\vec{q}_1, ..., \vec{q}_n) \, . \tag{4.9}$$

Then $Q^T Q$ is defined as the scatter matrix of size 3×3. Now, the direction $\vec{b}$ of the fitted line is the eigenvector corresponding to the largest eigenvalue of $Q^T Q$. If the distance between the fitted line and the points in the collection $X^i_{combination}$ is smaller than a strict threshold, the points are immediately accepted as part of the same trajectory, and the corresponding parameters from the fitted line are saved in X_{track}. If the collection $X^i_{combination}$ satisfies a broader threshold[2], the line is regarded as a potential trajectory, and $X^i_{combination}$ is added to a collection $X_{candidate}$. In the end—when all line parameters have been iterated over—for the lines in $X_{candidate}$, it is verified if one of these hit points is already part of a detected line in X_{track}. If so, the line is dismissed, otherwise, the line is accepted as a trajectory and added to X_{track}. Finally, the output of the algorithm is the collection X_{track}, containing the detected line parameters and hit coordinates

4.3.2 Clustering algorithm

The clustering algorithm employed in the accumulators of the tracking algorithm is DBSCAN, Density-Based Spatial Clustering of Applications with Noise, by Ester *et al.* [72]. The principle idea behind the algorithm is to detect dense regions of data and to distinguish these from noise. Before delving into the algorithm itself, it is necessary to introduce some key concepts. The mathematical definitions of these concepts can be found in [72]. Here, the explanation will be constrained to 2D, but DBSCAN also applies to high-dimensional space. A core point is a point that lies in a dense region, where the density around a point is defined as the number of points within a specified radius ϵ. The point $\vec{p}$ lies in a dense region, if in its neighborhood—a circle with radius ϵ centered on $\vec{p}$—there are more points than a specified value $minPts$. On the other hand, a border point is defined as a point that does not meet the criteria to be a core point but does lie within a distance ϵ of at least one core point. A noise point is neither a core point nor a border point—it does not fall within a distance ϵ of any core point. An illustration of core, border and noise points in the case of $minPts = 5$ is shown in Figure 4.4. Two additional

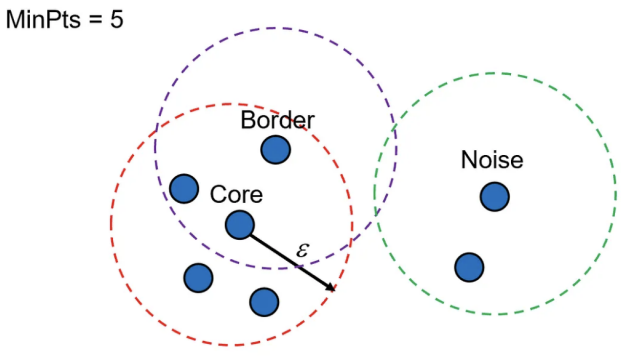


Figure 4.4: An illustration of core, border and noise points as defined by the DBSCAN algorithm; example for $minPts = 5$. Figure taken from [73].

[2]The strict and broader threshold are determined by assessing the tracking algorithm's performance for different values of the thresholds.

notions are essential for understanding the clustering process: density reachability and density connectivity.

- A point $\vec{p}$ is directly density-reachable from a point $\vec{q}$ if $\vec{p}$ lies in the ϵ-neighborhood of $\vec{q}$ and $\vec{q}$ is a core point.
- A point $\vec{p}$ is density-reachable from a point $\vec{q}$ if there exists a sequence of points $\vec{p}_1, \vec{p}_2, ..., \vec{p}_n$, where $\vec{p}_1 = \vec{q}$ and $\vec{p}_n = \vec{p}$, such that $\vec{p}_{i+1}$ is directly density reachable from $\vec{p}_i$.
- A point $\vec{p}$ is density connected to a point $\vec{q}$ if there exists a third point $\vec{o}$ such that both points $\vec{p}$ and $\vec{q}$ are density reachable from $\vec{o}$.

The clustering algorithm proceeds as follows: it selects an arbitrary point $\vec{p}$ in the dataset and finds all points that are density-reachable from it, using the specified parameters ϵ and $minPts$. If $\vec{p}$ is a core point, this results in the formation of a new cluster containing all density-connected points. If $\vec{p}$ is a border point or noise point, it is either added to an existing cluster—if reachable from a core point—or neglected, and the algorithm moves to the next point. In this context, a cluster is defined as a maximal collection of density-connected points. Additionally, if a point $\vec{p}$ belongs to a cluster and another point $\vec{q}$ is density-reachable from $\vec{p}$, then $\vec{q}$ is also a part of that same cluster.

The DBSCAN algorithm offers several advantages. One of its strengths is its robustness to outliers, as it can effectively identify and isolate noise points. Unlike many other clustering algorithms, the number of clusters does not need to be specified in advance and it is capable of detecting arbitrary shapes of clusters. DBSCAN is also relatively efficient, being designed for handling large databases. While it relies on only two parameters (ϵ and $minPts$), it can be quite sensitive to their selected value. Another drawback is that DBSCAN may perform poorly on datasets containing clusters of varying densities. This can be solved by either selecting $minPts$ in function of the sparsest cluster, or by exploring other algorithms, such as the algorithm HDBSCAN, which extends DBSCAN by allowing variation in the parameter ϵ.

4.4 Tracking performance

4.4.1 Algorithm parameters

The tracking algorithm relies on a broad set of input parameters, where each one will have its influence on the performance of the algorithm. Because of this diverse range of parameters, it can be challenging to keep track of which values to use for optimal performance. A clear overview of the parameters and their definition is given in Table 4.1. The parameters can be divided into three categories, depending on which component of the algorithm they influence: the accumulator, clustering, and distance parameters.

As the name suggests, accumulator parameters control the bin size of the accumulators, which in turn determines the resolution of the candidate line parameters $(\theta_H, \phi_H, x'_H, y'_H)$. The $\theta\phi$-accumulator must be refined enough to ensure that the direction angles (θ_H, ϕ_H) obtained through the Hough transform are sufficiently accurate. This accuracy is crucial because these angles are used to calculate (x', y'). As mentioned in the previous section, since the trajectories are not perfectly straight, (x', y') will still slightly vary when calculated for points on the same line. A finer resolution in the direction angles limits these variations and prevents the

Table 4.1: Definition of the tracking algorithm parameters.

Accumulator parameters	
n_θ	number of accumulator bins for polar angle θ
n_ϕ	number of accumulator bins for azimuthal angle ϕ
n_{xy}	number of bins for accumulator $x'y'$
Clustering parameters	
V_C	minimum votes a cell must have in order to be part of a cluster
V_R	minimum total votes a candidate region must have in order to be a cluster
ϵ	distance between two cells to be considered in the neighborhood of each other
$minPts$	number of points that must be in the neighborhood of a point to be considered core point
Distance parameters	
d_1	points which are considered close to the Hough-track
$d_{accepted}$	points are considered part of the fitted track $\rightarrow$ track is accepted immediately
$d_{candidate}$	points are considered part of the fitted track $\rightarrow$ track is possible candidate

spread of votes to neighboring cells. The same discussion holds for the xy-accumulator: since $(\theta_H, \phi_H, x'_H, y'_H)$ represent the initially reconstructed line, its accuracy is important, as higher precision allows for tighter distance thresholds, which in turn reduces the number of false tracks. On the other hand, an overly refined accumulator can be counterproductive: when bins are too small, votes that should accumulate in the same cell might spread across neighboring cells. This will specifically be the case in the presence of noise, as will be discussed in Subsection 4.4.3. Therefore, a balance must be found: the accumulator should be refined enough to offer an adequate resolution, but not so fine that related votes are fragmented. In this context, the adjective *related* will be used to denote hit points stemming from the same muon trajectory.

Clustering parameters influence both the spatial density—defined in Subsection 4.3.2—and the vote density of the clusters. These parameters may differ between the $\theta\phi$-accumulator and the xy-accumulator. For the radius ϵ, again balance is necessary. It should be large enough—especially in the xy-accumulator, where more vote spreading occurs—to ensure that related cells are clustered together. However, if ϵ is too large, votes from unrelated points might be merged into the same cluster. This will affect the weighted average of the cluster, which defines the line parameters, either (θ_H, ϕ_H) or (x'_H, y'_H). This will lead to a discrepancy between the actual trajectory and the candidate track, which if too large, causes points on the true trajectory not to be associated with the candidate track anymore. The parameter $minPts$ is set to 1, as this parameter includes the point $\vec{p}$ itself. This allows a cell to form a cluster on itself, without requiring neighboring cells. A filtering mechanism could be applied to only perform clustering on cells with a minimum required number of votes V_C, however, efficiency studies showed this is best avoided, as explained in the next subsection. Hence, all cells with at least one vote were considered a core point. While this increases the number of initial candidates for the clustering algorithm, the parameter V_R ensures that only meaningful clusters are selected. Since V_R represents the minimum number of votes a cluster must have, it is intuitive to set $V_R = 3$, as each vote represents a hit point and there are three related hit points per muon trajectory.

Lastly, the distance parameters define the thresholds to decide whether a candidate line is accepted as a valid muon trajectory. As for the other parameters, these thresholds also require careful tuning. The first threshold d_1 must be large enough to ensure that all related points lying

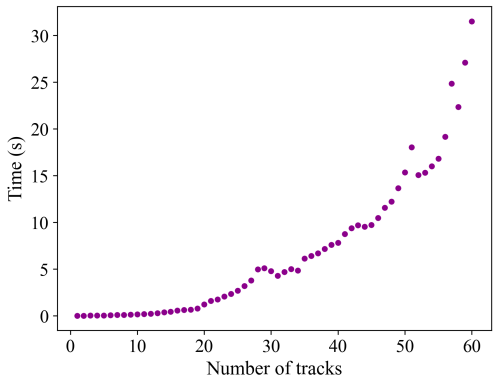
on a candidate line fall within this distance and are added to the collection X_{nearby}. Given the limited resolution of the reconstructed $(\theta_H, \phi_H, x'_H, y'_H)$, this threshold cannot be too small. On the other hand, if it is too big, many irrelevant points might be added to X_{nearby}. While this is not critical, as the other distance thresholds $d_{candidate}$ and $d_{accepted}$ will filter out false combinations, it does increase the number of combinations $X^i_{combination}$ the algorithm must consider, increasing the computational time. In principle, the strict threshold $d_{accepted}$ can be set as small as required, as it was implemented as an extra measure to selectively accept well-fitting reconstructed tracks, which will prevent interference in other xy-accumulators or interchangeability with other trajectories. If $d_{accepted} = 0$, this strict threshold is not applied, and all selections are made based solely on $d_{candidate}$. For the value of threshold $d_{candidate}$ a trade-off must be made between being sufficiently small to prevent the detection of false tracks, yet large enough to ensure all related hit points are considered a part of the fitted line. This is a crucial parameter to fine-tune, balancing between the rejection of false tracks while maximizing the algorithm's efficiency.

Achieving the right balance between different parameters is crucial. In some cases, even small adjustments to a parameter value can have a subtle yet significant impact on the performance, and the relationship between parameters is not always immediately obvious. To investigate these complex behaviors, a performance study of the algorithm was conducted. In this context, performance refers to both the efficiency in correctly identifying true tracks and the ability to avoid detecting false ones. The performance was studied for *ideal tracks*, representing perfectly straight lines. Finally, the influence of noise and smearing on the tracking algorithm is investigated.

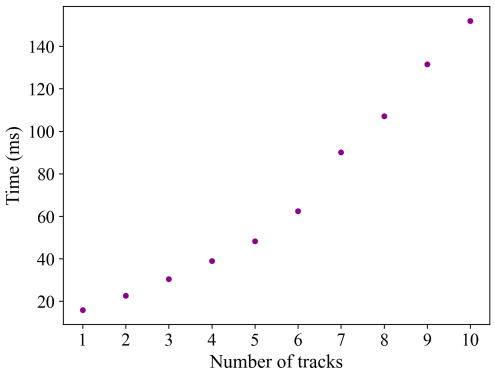
4.4.2 Ideal tracks

Ideal tracks refer to muon trajectories defined by their exact hit coordinates as they pass through the detector, without taking the detector response and secondary particles into account. This means the hit coordinates represent the precise location where the muon traversed the detector plates, rather than the reconstructed hit coordinates obtained through the procedure described in Subsection 2.3.3. Consequently, the muon trajectories can be regarded as quasi-perfectly straight lines, showing minimal signs of smearing or measurement uncertainties. Nonetheless, slight deviations may still occur due to potential scattering in the detector material.

The first thing that can be investigated is the computational time required by the tracking algorithm as a function of the input size. Here, the input size is represented in number of tracks, with each track consisting of three hit points. Figure 4.5a shows the computational time as a function of the number of input tracks being processed by the algorithm. This plot was generated using a sample of 60 events and, while it displays the absolute time rather than the average computational time, it provides a useful initial indication of how computational time scales with input size. The computational time remains relatively steady when fewer than 10 to 20 tracks are being fed into the algorithm. Beyond this point, a noticeable increase occurs. This jump is expected, given the algorithm's nature of combinatory mathematics and iteration loops. It involves calculating all possible combinations of hit point pairs to determine (θ, ϕ), constructing the sets $X^i_{combination}$ and iterating over candidate directions (θ_H, ϕ_H) to retrieve the corresponding (x'_H, y'_H). Additional looping is performed at the end to determine which candidate tracks will be accepted as a trajectory. The total runtime of the algorithm is also sensitive to parameter choices, as more clustering or larger distance thresholds are expected to prolong the calculation. Nevertheless, the general trend will remain consistent and is sufficient for the



(a) Computational time for a sample of 60 events. The time is displayed in seconds.



(b) Zoom-in of the left plot for the first 10 events. The time is displayed in milliseconds.

Figure 4.5: Computational time of the tracking algorithm in function of the number of tracks it processes at once.

analysis conducted here. The parameters used to generate Figure 4.5 are listed in Table B.1 in Appendix B. The leap in runtime beyond 10 to 20 tracks suggests a trade-off: while processing more tracks at once is more productive, it may slow down the algorithm significantly. While this leap is not ideal, it is also not an obstacle: each track comprises only three hit points, and splitting the input is manageable. It has to be noted that this algorithm is not suitable for large-scale applications such as images or large point clouds. Feeding too many tracks at once will cause the accumulator to become oversaturated, leading to excessive clustering and decreased performance, resulting in missed or falsely identified tracks. Based on this turning point and the computational time, the algorithm is ideally applied to batches of tracks, limiting the batch size to 10 tracks. A zoomed-in view of the computational time for the first 10 tracks is displayed in Figure 4.5b, with the exact values presented in Table 4.2. From six tracks onward, the computational time begins to rise faster with respect to the number of input tracks, suggesting that it is optimal to limit the number of tracks being fed into the algorithm to a maximum of 5 at a time. To better understand where the performance starts to degrade, the tracking efficiency

Table 4.2: Computational time in function of tracks, for the values plotted in Figure 4.5b.

Tracks	time (ms)
1	15.9
2	22.6
3	30.4
4	38.9
5	48.3
6	62.5
7	90.1
8	107
9	131
10	152

was analyzed for the same sample of 60 events. The results, based on the same parameters used in the timing study, are shown in Figure 4.6. For this sample of ideal tracks, the tracking

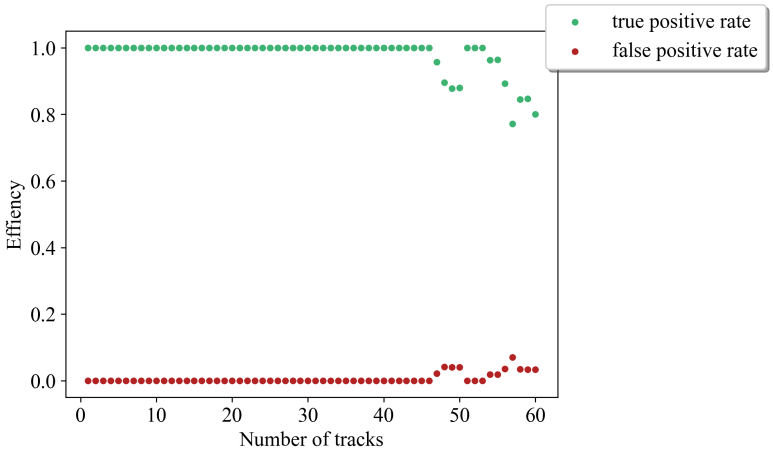


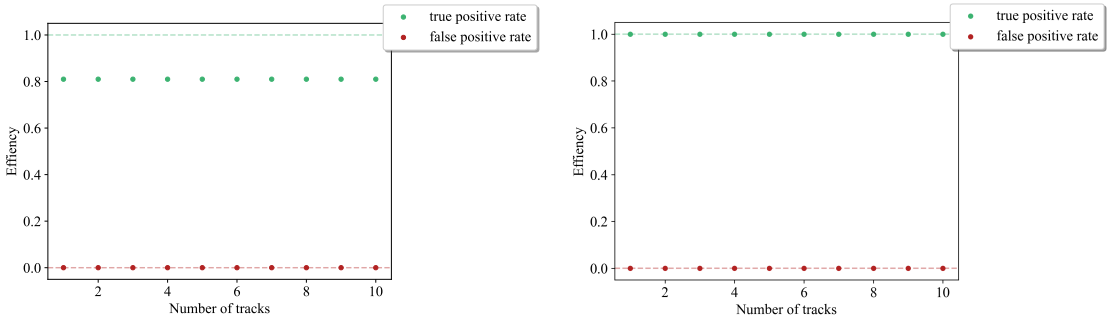
Figure 4.6: Tracking algorithm efficiency in function of number of tracks it processes at once, for a sample of 60 events.

algorithm is able to detect all tracks when up to 46 tracks are being fed into the algorithm at once. Beyond this point, the algorithm starts to decrease in performance, either failing to detect all tracks, detecting false tracks, or both. This arises from the increased density of hit points, which makes it more likely that unrelated points lie close together. Consequently, the distance threshold becomes less effective, causing incorrect combinations to form false tracks. If these false tracks are immediately accepted as valid tracks, their associated hit points are removed from the dataset, preventing correct tracks from being identified later on. Additionally, the accumulators get oversaturated. To study the efficiency in more detail, the input was limited to 10 tracks at a time. A sample of 100 events was processed by the tracking algorithm in ten different configurations: first, one track was being fed into the algorithm at a time; second, two tracks at a time; continuing up to ten tracks at a time. As a result, the average tracking performance could be inspected as a function of the number of tracks simultaneously processed. This procedure was repeated for multiple combinations of the tracking parameters to evaluate whether subtle parameter changes significantly affect the outcome and to explore the relationship between parameters. From these studies, the optimal parameter values were identified for reconstructing the complete dataset acquired through simulations. As there are countless possible parameter combinations, only the most relevant results are presented and discussed in this section.

In the case of ideal tracks, minimal spreading is expected in the $\theta\phi$-accumulator for the votes corresponding to hit points lying on the same track. Hence, it seems reasonable to set the clustering threshold V_C equal to 2 or 3, such that only cells with a relatively high number of votes are considered. The performance was evaluated using $V_C^{\theta\phi} = 2$, with the other parameter values provided in Table B.2[3]. As seen on the performance plot in Figure 4.7a the efficiency of detecting true tracks is smaller than 100%, having a constant value of 81% regardless of the number of tracks. This constant value hints that the same tracks are systematically missed, likely due to

[3]All parameters for the efficiency plots are presented in Appendix B. To maintain the readability of the text, this appendix will not be mentioned repeatedly.

their votes being distributed across two or three different accumulator cells. When decreasing V_C to 1, the efficiency reaches 100%, as seen in Figure 4.7b, indicating that all tracks are then correctly identified. These performance plots were conducted using high-resolution accumulators,

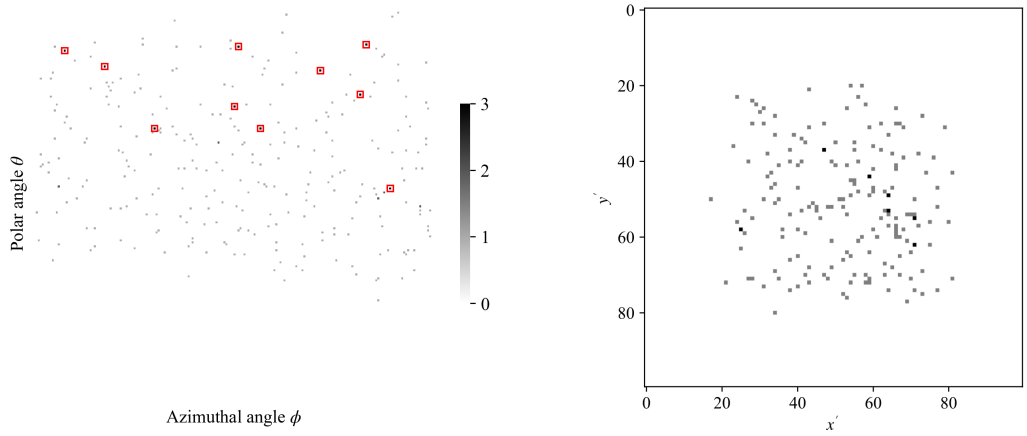


(a) Efficiency for $V_C = 2$.

(b) Efficiency for $V_C = 1$.

Figure 4.7: Tracking efficiency compared for parameter $V_C = 2$ and $V_C = 1$ while keeping the other parameters constants. The other parameters used in the tracking algorithm to create these plots are given in Table B.2 and B.3.

where the $\theta\phi$-accumulator had a size of 360×720 and the xy-accumulator size 200×200. To reduce the spreading of votes across multiple cells, lower-resolution accumulators are proposed. While this will result in a lower resolution of the obtained direction and position candidates, this can be countered by loosening the distance thresholds. Moreover, since the number of tracks being processed simultaneously is relatively small, it is assumed that the accumulators will still be detailed enough to make the distinction between collinear hits and bad votes. Figures 4.8a and 4.8b show examples of a $\theta\phi$-accumulator of size 200×200 and an xy-accumulator of size 100×100, respectively. As shown in Figure 4.8a, the true track candidates are clearly separa-



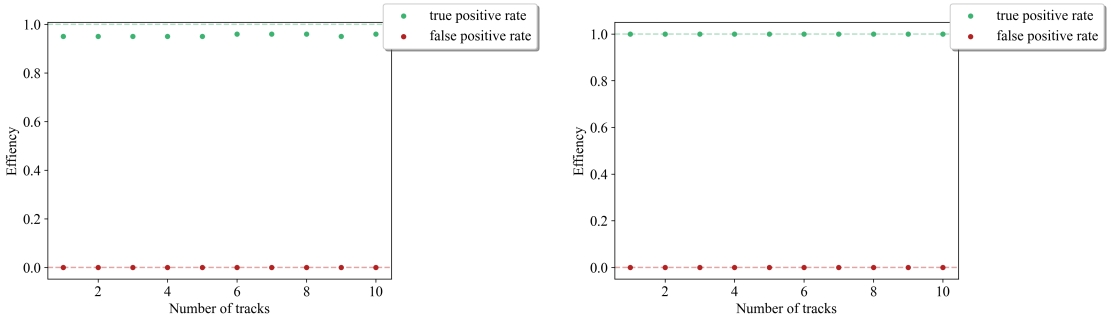
(a) $\theta\phi$-accumulator of size 200×200

(b) xy-accumulator of size 100×100

Figure 4.8: Example of a $\theta\phi$-accumulator and xy-accumulator for a vote distribution of 10 tracks..

ble from bad votes in the $\theta\phi$-accumulator. However, the xy-accumulator (Figure 4.8b) already

begins to show signs of oversaturation, containing multiple, ambiguous high-vote regions. Using these smaller dimensions for the accumulator, the efficiency was recalculated for $V_C = 2$, with other parameters provided in Table B.4. While the efficiency has improved compared to the high-resolution case, it is still smaller than 100%, reaching approximately 95-96%, as seen in Figure 4.9a. When reducing V_R from 3 to 2, the 100% efficiency is recovered, as observed in Figure 4.9b. However, since V_R reflects the minimum number of votes a cluster must have, it is



(a) Tracking efficiency for $V_C = 2$ and $V_R = 3$.

(b) Tracking efficiency for $V_C = 2$ and $V_R = 2$.

Figure 4.9: Tracking efficiency compared for parameter V_C = 2 and V_R = 3 and 2, while keeping the other parameters constants. The other parameters used in the tracking algorithm to create these plots are given in Table B.4 and B.5.

more intuitive to set $V_R = 3$, considering that each trajectory consists of exactly three hit points. Given that the efficiency remains more or less constant for the configuration in Figure B.4, this suggests that the same tracks are still consistently missed. This is likely due to their votes being spread over three separate cells, resulting in clusters where each cell receives only a single vote. To detect these sparse clusters, the threshold V_C should be decreased to 1.

A similar discussion applies to the xy-accumulator. However, contrary to the $\theta\phi$-accumulator, setting $V_C = 2$ and $V_R = 2$, does not yield 100%, as visible in Figure 4.10. A plausible explana-

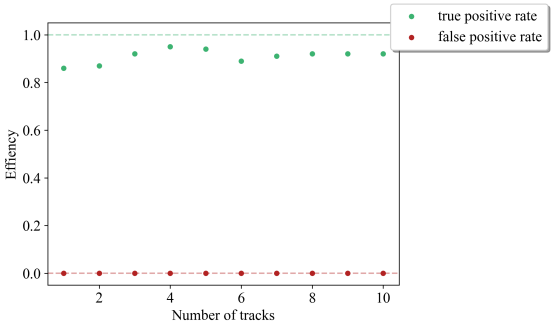
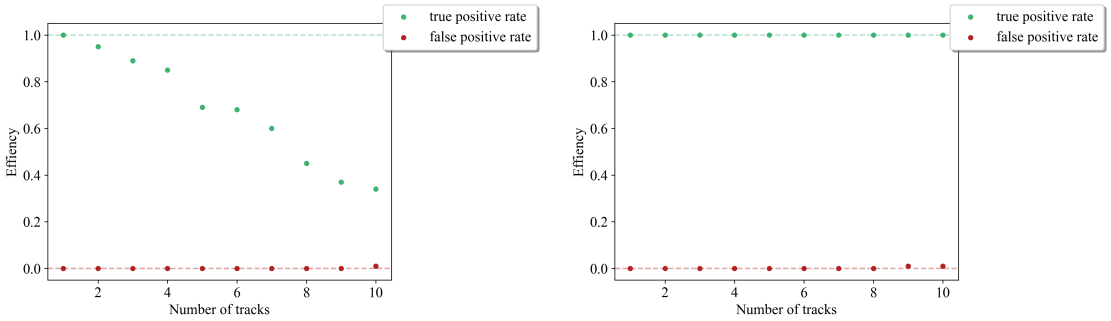


Figure 4.10: Tracking efficiency for $V_C = 2$ and $V_R = 2$ for xy-accumulator clustering. The other parameters used in the tracking algorithm to create this plot are given in Table B.6.

tion for the lower efficiency is that the xy-accumulator inherently exhibits more vote spreading than the $\theta\phi$-accumulator, due to its dependence on the direction candidates obtained from the $\theta\phi$-accumulator. This motivates the choice of $V_C = 1$ and $V_R = 3$ in the xy-accumulator. To

account for this vote spreading, the clustering parameter ϵ should be sufficiently large. Nevertheless, it is also evident that an excessively large epsilon will have a reverse effect, as demonstrated in Figure 4.11a for $\epsilon_{\theta\phi} = 10$. In this configuration, the efficiency drops significantly as



(a) Tracking efficiency for ϵ = 10 and d_1 = 50

(b) Tracking efficiency for ϵ = 10 and d_1 = 700.

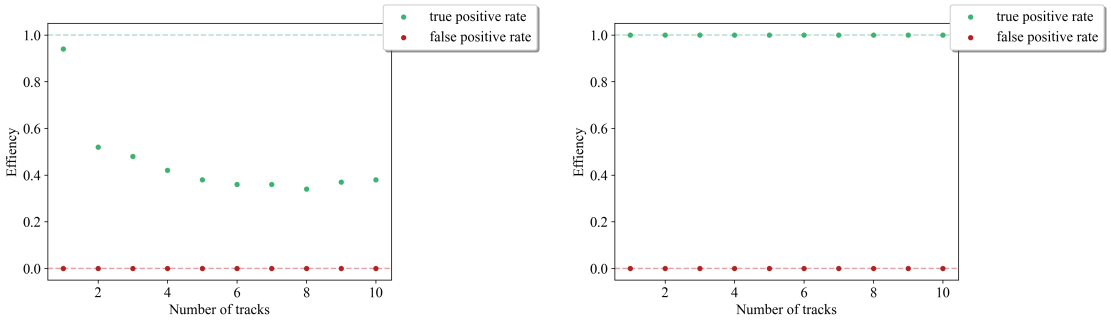
Figure 4.11: Tracking efficiency compared for parameter ϵ = 10 and d_1 = 10 and 700, while keeping the other parameters constants. The other parameters used in the tracking algorithm to create these plots are given in Table B.7 and B.8.

the number of tracks increases. This behavior is intuitive: a larger radius means more points farther apart will be grouped into the same cluster, eventually leading to sparsely populated clusters that can span large portions of the accumulator. When calculating the weighted average of such a cluster, the resulting line parameters will not correspond to a valid track anymore. As seen in Figure 4.11b, increasing the distance threshold d_1 to 700 mm[4] restores the track detection efficiency to 100% Interestingly, the number of false tracks remains close to 0%, presumable because the values of $d_{accepted}$ and $d_{candidate}$ used in this configuration remain sufficiently strict to prevent incorrect tracks from being formed. Nevertheless, setting $d_1 = 700$ is highly undesirable, as such a large threshold will result in nearly all hits points being included in the collection X_{nearby}, involving a lot of combinatorics to construct the corresponding candidate collections $X^i_{combination}$. This undermines the core objective of the algorithm, to minimize the combinatory operations and reduce computational cost.

The final category of parameters to explore is the distance parameters. In Figure 4.12a, the tracking efficiency is shown for a distance threshold of $d_1 = 10$ mm. This threshold proves to be too low, as not all tracks are detected, which is likely due to the finite resolution of the Hough accumulators. Increasing the threshold to $d_1 = 30$ mm restores full efficiency, as illustrated in Figure 4.12b. Evidently, relaxing the thresholds $d_{accepted}$ and $d_{candidate}$ will eventually lead to an increase in the detection of false tracks. The final decision for the algorithm parameters used in the track reconstruction of all simulated data, is presented in Table B.12.

This discussion on parameter tuning could go on indefinitely since numerous factors influence the efficiency, yet these observations highlight the main parameters that must be balanced carefully. Selecting optimal values involves a trade-off: stricter thresholds may miss some tracks but may improve overall track quality by filtering out less reliable trajectories, whereas looser thresholds may increase efficiency but risk the detection of false trajectories.

[4]The units of this threshold are inherited from the units of the hit coordinates.



(a) Tracking efficiency for $d_1 = 10$.

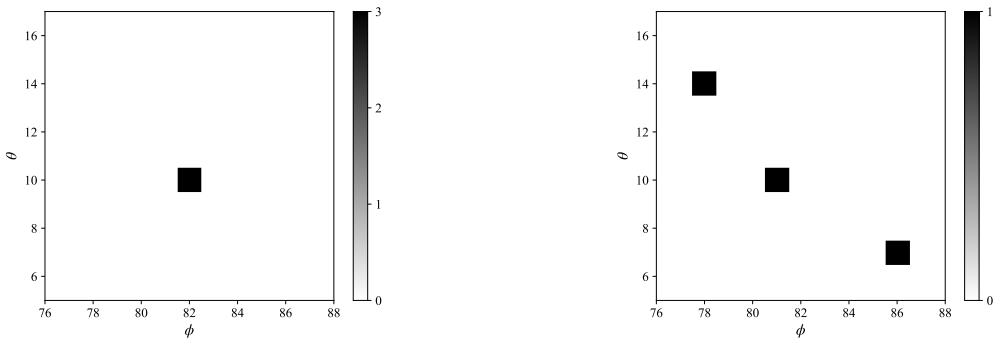
(b) Tracking efficiency for $d_1 = 30$.

Figure 4.12: Tracking efficiency compared for parameter $d_1 = 10$ and $d_1 = 30$, while keeping the other parameters constants. The other parameters used in the tracking algorithm to create these plots are given in Table B.9 and B.10.

4.4.3 Noise and smearing

As mentioned in Subsection 3.1, the scintillator detector plates were defined as sensitive detectors in the simulation. This setup provided the exact hit coordinates corresponding to the points where the muon intersected the scintillator plates. Moreover, secondary particles were not tracked, resulting in a clean, noise-free data acquisition. In the actual experimental setup, however, these exact coordinates will not be available. Instead, hit coordinates will have to be reconstructed from the signals registered by the SiPMs (Subsection 2.3.3). This reconstruction process may lead to a slight smearing of the muon's trajectory, introducing an uncertainty based on the resolution of the detector. Additionally, noise hits may arise from secondary particles or dark noise, a characteristic of the SiPMs. In principle, the tracking algorithm should be able to make a distinction between noise hits and muon hits, provided the level of noise remains within acceptable limits. However, dark noise specifically will influence the reconstructed hit coordinate. Track smearing can result in more vote spreading in the accumulator, making it more challenging to identify clusters corresponding to real trajectories. Therefore, the final tracking algorithm applied to experimental data must be carefully tuned to take these factors into account. This subsection presents preliminary studies investigating the impact of noise and smearing on the track reconstruction performance.

To evaluate the tracking resolution under more realistic conditions, a single muon trajectory was artificially smeared 1000 times. The smearing was implemented using a Gaussian distribution with a standard deviation of 11 mm, based on preliminary detector resolution studies [54]. An example of the impact of smearing on the vote distribution in the $\theta\phi$-accumulator is shown in Figures 4.13a and 4.13b, which display the ideal and smeared cases, respectively. The smeared hit coordinates were then processed by the tracking algorithm and the reconstructed average direction $(\theta_{reco}, \phi_{reco})$ was compared to the true track direction $(\theta_{actual}, \phi_{actual})$. It has to be noted that this procedure essentially assesses how well a straight line fits the smeared hit points, as the final stage of the tracking algorithm always involves a linear fit through the identified candidate points in $X^i_{combination}$. Nevertheless, it is still interesting to compare this resolution to the ideal tracking resolution discussed in Chapter 5. Using the same algorithm parameters as for the ideal case, the reconstruction failed to detect 844 out of 1000 smeared tracks. This was expected, as the original parameters were optimized to utilize straight-line trajectories.



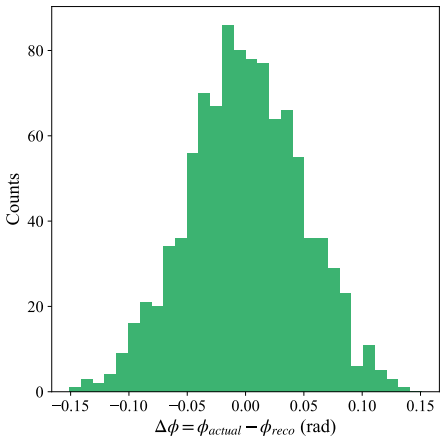
(a) $\theta\phi$-accumulator for an ideal track.

(b) $\theta\phi$-accumulator for a smeared track.

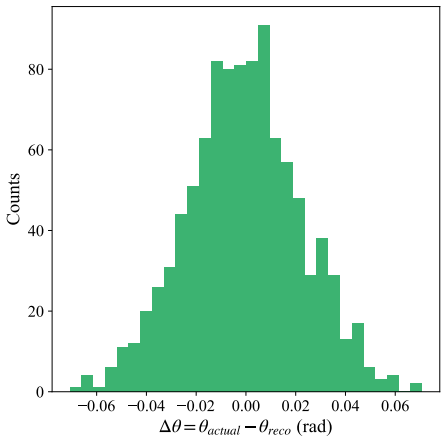
Figure 4.13: Comparison between the vote distribution of an ideal and smeared track in the $\theta\phi$-accumulator. The visualization is zoomed-in on the vote distribution of the track.

First of all, an improvement to the tracking algorithm was introduced: periodic clustering in the $\theta\phi$-accumulator. For tracks with the azimuth angle ϕ close to π or $-\pi$, calculating the directions for pairs of related hit points may result in votes at opposites side of the accumulator, due to the constrained range of $\phi \in [-\pi, \pi]$, especially in the case of smearing. To takes this into account, the (θ, ϕ) votes were converted to direction vectors on the unit sphere and the clustering algorithm was performed in 3D on these vectors. This adjustment had not yet been implemented in the ideal case, since votes for these related hit points generally populate the same cell, and preliminary results of the algorithm were satisfactory. However, this periodic clustering should be applied to future versions for consistency and accuracy. A drawback of this approach is that the clustering parameter ϵ becomes less intuitive, as it now represents a distance in 3D space rather than a distance between accumulator cells. The parameters used for the reconstruction of the smeared tracks are listed in Table B.11 in Appendix B. These were adapted from the ideal case, to account for the smearing effects. With these adjustments, only 5 out of 1000 tracks remained undetected. These 5 missed tracks likely result from excessive spreading of the votes in the accumulator or a poor match between the smeared hits and the proposed candidate line. Figures 4.14a and 4.14b show the distribution of the differences between the reconstructed and actual trajectory direction. The standard deviations of these distributions are 48.64 mrad for $\Delta\phi$ and 22.69 mrad for $\Delta\theta$. As will be discusseed in the next chapter, this is considerably larger than those obtained in the ideal case. However, this comparison should be interpreted with caution, as the results presented here are based on a single smeared trajectory, hence not reflecting the variability of the dataset. Additional studies are required to draw more general conclusions.

When a single smeared track is processed by the algorithm, the effects of smearing remain manageable by the tracking algorithm. The challenge arises when multiple smeared tracks, or a combination of smeared tracks and noise, are fed into the algorithm simultaneously. Figure 4.15 demonstrates the $\theta\phi$-accumulator for three ideal tracks (Figure 4.15a) and for three smeared tracks (Figure 4.15b). For visualization purposes, the accumulator size used in this case is 100×100. As observed in the figures, smearing causes the votes to become scattered over the accumulator: no cell contains multiple votes and no clear local maxima are visible. On the contrary, the ideal $\theta\phi$-accumulator contains concentrated votes from related hit points, forming distinct peaks. This lack of pattern in the smeared case makes it difficult for the algo-

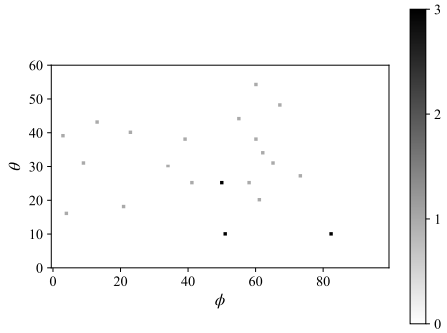


(a) Tracking resolution for azimuthal angle ϕ.

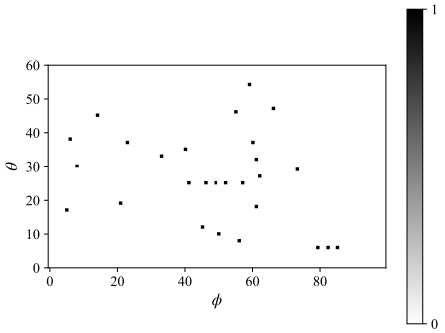


(b) Tracking resolution for zenith angle θ.

Figure 4.14: Obtained tracking resolution for a track that was smeared a 1000 times, for Gaussian smearing with a standard deviation of 11.



(a) $\theta\phi$-accumulator for three ideal tracks.

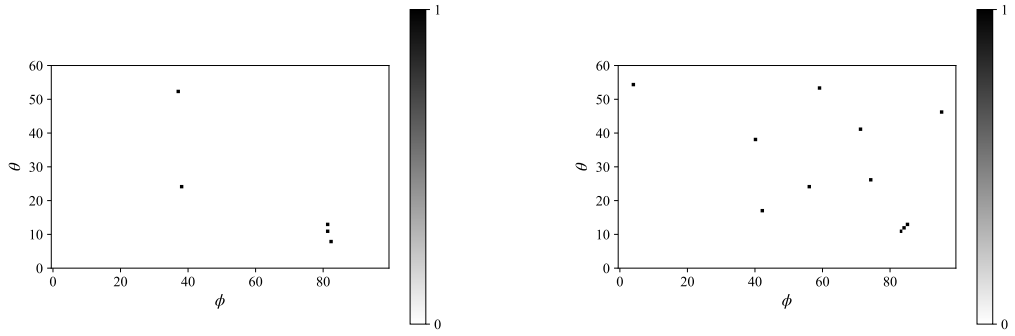


(b) $\theta\phi$-accumulator for three smeared tracks.

Figure 4.15: Comparison between the vote distribution of three ideal and smeared track in the $\theta\phi$-accumulator. For visualization purposes, the accumulator is 100×100, where the range of θ is limited to 60 as there were no votes outside this range.

rithm to reliably identify clusters, limiting its effectiveness. This highlights the need for further improvements.

For noise, the perspective remains slightly more optimistic. Artificial noise was introduced by randomly selecting a position coordinate within a detector plate using a uniform distribution. Figure 4.16a displays the $\theta\phi$-accumulator of a smeared track combined with one artificial noise hit, while Figure 4.16b shows the accumulator for the same track—yet smeared differently—with three artificial noise hits. Despite each cell containing only a single vote, the votes



(a) $\theta\phi$-accumulator for one smeared track and one noise hit.

(b) $\theta\phi$-accumulator for one smeared track and three noise hits.

Figure 4.16: Comparison between the vote distribution in the $\theta\phi$-accumulator for one artificial noise hit and three artificial noise hits, and in both cases one smeared track.

associated with the smeared track remain distinguishable from those originating from the noise hits. This highlights the power of the accumulator approach in the tracking algorithm, separating trajectories from noise. Finally, the case of three smeared tracks combined with three noise hits was studied. Figure 4.17 displays the corresponding $\theta\phi$-accumulator where the votes originating from noise are shown in black, while the votes from related muon hits are shown in white. Again, a lack of pattern is observed, which comes as no surprise, given that this effect

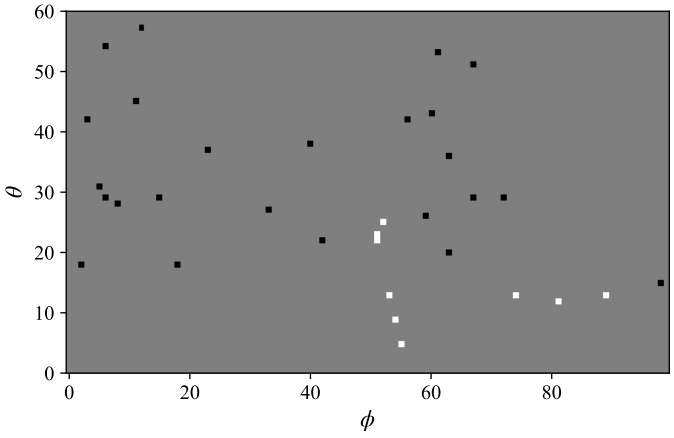


Figure 4.17: $\theta\phi$-accumulator for three smeared tracks and three additional noise hits. The votes originating from noise are shown in black, while the votes from related muon hits are shown in white. The grey area indicates cells without votes.

was already present in the case of three smeared tracks alone.

Over time, a more realistic simulation—one that takes into account the full detector response, reconstructed hit coordinates, and tracking of secondary particles—will need to be implemented, as this is essential to better characterize noise behavior and the extent of track smearing. Additional parameters, such as the energy deposited by the hits, could also be incorporated, to help discriminate between muon hits and noise. Moreover, strategic cuts could be applied to filter out poorly reconstructed or heavily smeared tracks, improving the overall tracking quality. Finally, insights from alternative tracking algorithms could be explored to further enhance the performance.

Chapter 5

Image Reconstruction

With the optimized simulation and tracking algorithm in place, the full simulation workflow can now be executed to obtain a muographic image of the pyramid. For this analysis, three simulation scenarios were carried out: a free sky configuration, a configuration including the pyramid as it is currently known, and a configuration of the pyramid with an additional room. The resulting hit coordinates from each configuration were processed using the tracking algorithm to reconstruct the muon trajectories. As a first step, a quality assessment of the reconstructed tracks was performed to ensure that they accurately reflected the originally generated muon directions. This evaluation is presented in Section 5.1. Once the directions were successfully reconstructed, the image reconstruction of the pyramid could proceed. Firstly, a transmission map of the pyramid was generated by comparing the free sky and current pyramid configurations, as discussed in Section 5.2. This transmission highlights the structural features of the pyramid. Then, a relative transmission map was generated by comparing the current pyramid model with the modified model that includes an additional void. This analysis aims to explore the sensitivity limits of the detector in identifying an excess in muon flux that could indicate the presence of hidden cavities, as detailed in Section 5.3.

5.1 Simulation and evaluation of muon tracks

To generate detailed images of the muon transmission through the Pyramid of Khafre, it is essential to acquire a sufficient amount of data so that the features of the pyramid and possibly unknown features are distinct. Typically, high-resolution muographic imaging requires datasets containing millions of muon hits, collected over several weeks or months of measurement [37, 42]. Fortunately, data generated through optimized simulations is usually faster, making it possible to gain a first insight into the expected image of the pyramid. Three simulation configurations were performed in this analysis: free sky, pyramid, and pyramid with an additional room.

The free sky data acquisition did not require the optimized version of the simulation, as the small generation surface was already adequate to yield a sufficient number of hits. To maintain consistency with the optimized version, the same small generation surface—a planar surface of 70×70 cm^2 positioned 5 cm above the uppermost detector plate—was used. From this surface, 40 million muons were generated, resulting in approximately 10 million recorded events. The exact details of the data acquisition are provided in Table 5.1. For the transmission through the

Table 5.1: Details of the simulation time (rounded to the minute), total amount of recorded events and track reconstruction for the three configurations: free sky, pyramid, and pyramid with an additional room. The simulation time is the time it took to run a batch of 10 million muons.

	Simulation time	Total recorded events	Undetected tracks	False tracks
Free sky	29 min	10,504,923	2.2%	0.033%
Pyramid	4 h 13 min	3,079,505	0.20%	0.022%
Room	3 h 34 min	3,239,904	0.14%	0.022%

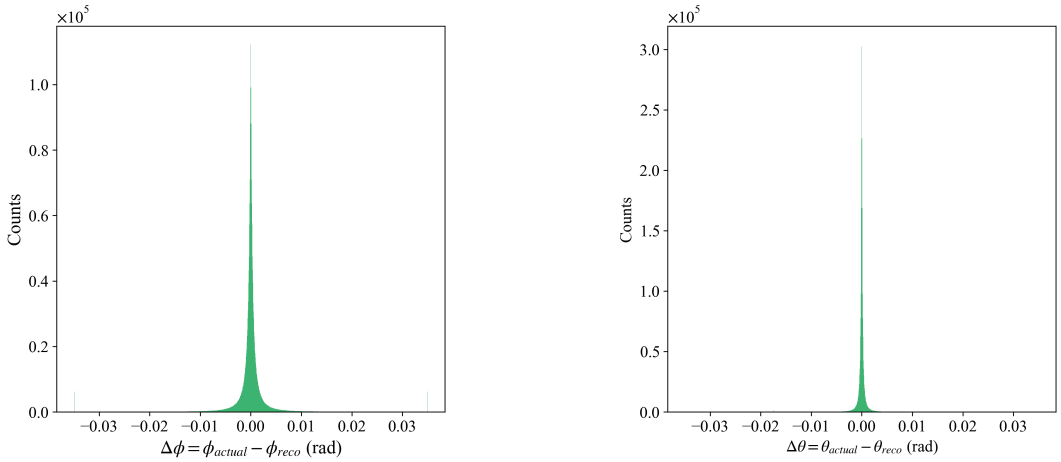
pyramid—both the simulation of the pyramid as known today, as the pyramid with an extra room, both referred to as the target configuration—the optimized simulation was used. In these cases, 100 million and 105 million muons were simulated, respectively, on the small generation surface. This resulted in the acquisition of approximately 3 million muons for each configuration, with the details in Table 5.1. The simulation was run in batches of 10 million muons, as this took around 3 to 4 hours for the target configuration, to avoid possible crashes resulting in data loss. Initially, the average simulation time of these batches was considered as a runtime measure. However, it was observed that the simulation times varied significantly per run, despite processing the same number of muons. These variations were linked to fluctuations in the computational load on the laptop used for the simulations, often running other processes such as track reconstruction simultaneously. To ensure consistency, a representative runtime was measured by running one simulation under certain conditions—while no other programs were active on the laptop—and is provided in Table 5.1, for each configuration.

To calculate the muon transmission, as defined in Equation 1.13, it is necessary to determine the muon flux, which requires knowledge of the data acquisition time. Technically, the detector efficiency is also required. However, it is assumed that the efficiency remained constant across all simulation runs, reducing the factor ϵ_{FS}/ϵ to unity. While EcoMug can estimate the total acquisition time to generate a desired number of muons on a given generation surface, even accounting for constraints on azimuth and zenith angles, this estimation could not be applied in this particular analysis. The optimized simulation used the small generation surface from EcoMug but afterwards generated the muons from a manually defined hemispherical surface. The effective generation area on this hemisphere is limited by the intersection of the inverted generation cone. As a result, EcoMug's built-in time estimation was not applicable to the target configurations. To approximate the acquisition time, the computational runtime of the simulations was used instead, as given in Table 5.1. Notably, the free sky simulation executed significantly faster than the target simulations—a pattern consistent with real experimental conditions—making the runtime a reasonable basis for estimating relative rates. However, this approach provides only an order-of-magnitude estimation, rather than absolute values of the flux. Therefore, the term *muon rate* is adopted throughout this chapter, instead of muon flux.

Although the generated information on the muons is available through the simulation, the purpose was to reconstruct images under conditions that closely resemble those of the real experimental setup. Therefore, the track reconstruction was applied to all hit data. The reconstruction was performed in batches of three events at a time, as this was found to provide the best performance in terms of speed. The finalized parameters for the full track reconstruction used across all simulated datasets are listed in Table B.12 in Appendix B. The algorithm's per-

formance was evaluated for all three datasets to ensure reconstruction quality before the image reconstruction. The detected tracks were classified as either true or false tracks, by utilizing the EventID information, available through the simulation. The efficiency in identifying true tracks and the rate of falsely detected tracks are presented in Table 5.1. The false positive rate remained consistently low across all datasets, ranging from 0.022% to 0.033%. The number of undetected tracks also remained small for the target configurations, however, for the free sky dataset, the missed detection rate exceeded 2%, corresponding to approximately 200,000 tracks not being reconstructed. This may partly be attributed to limitations in the accumulator used during the track reconstruction, particularly its lack of periodicity in ϕ. While this might account for some of the missed detections, the tenfold increase compared to the target configurations is notable and cannot be fully explained at this stage. Nonetheless, this result highlights the need for periodic clustering in the $\theta\phi$-accumulator and further refinement of the tracking algorithm.

To further assess the quality of the track reconstruction, the tracking resolution was evaluated by comparing the true track directions to those detected by the algorithm. Figure 5.1 shows the distribution of angular difference for the pyramid dataset, with false tracks excluded. The left-hand plot (5.1a) illustrates the difference in azimuth angle ϕ, while the right-hand plot (5.1b) shows the difference in zenith angle θ. To focus on the central distribution, the range of the plots



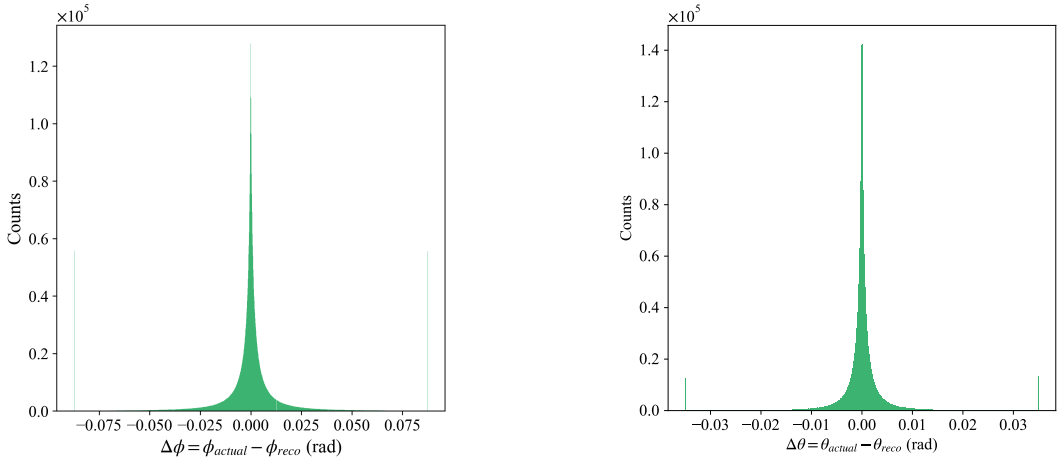
(a) Resolution for the azimuth angle ϕ.

(b) Resolution for the zenith angle θ.

Figure 5.1: Tracking resolution of the tracking algorithm for the pyramid configuration as known today.

was constrained to $[-2°, 2°]$. As seen in Figure 5.1a, a degree of overflow is present, indicating large deviations between the reconstructed and actual track angle. It is also notable that the distribution for the azimuth angle is broader than for the zenith angle. When calculating the standard deviation within this plotted range—excluding overflow—values of 3.85 mrad for ϕ and 1.27 mrad for θ were obtained. The broader distribution for ϕ could be explained by the lower resolution of ϕ in the $\theta\phi$-accumulator and possibly the lack of periodic clustering. A similar pattern is observed in the configuration with the additional room. The distributions of this case can be found in Figure C.1 in Appendix C. The standard deviations here are 3.85 mrad for ϕ and 1.29 mrad for θ, indicating a consistent angular resolution across these target

configurations. In contrast, the free sky configuration again shows a notably different behavior. Figures 5.2a and 5.2b reveal broader distributions in both ϕ and θ, suggesting a decrease in an-



(a) Resolution for the azimuth angle ϕ.

(b) Resolution for the zenith angle θ.

Figure 5.2: Tracking resolution of the tracking algorithm for the free sky configuration.

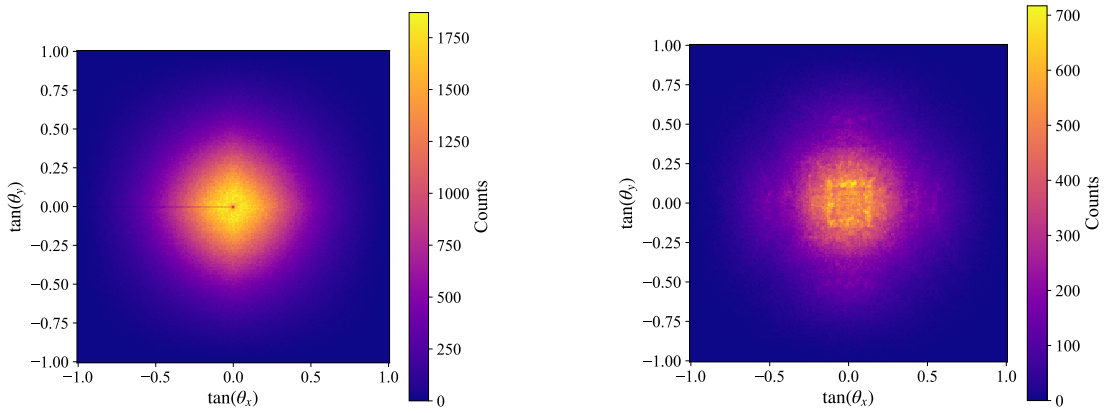
gular precision. Furthermore, there is a significant overflow in the azimuth angle distribution, though it should be noted that this dataset contains a larger number of events, which could contribute to the increased spread. The standard deviation for the free sky configuration—again excluding overflow and false tracks—is 11.4 mrad for ϕ and 4.38 mrad for θ. Although these values remain relatively low, within acceptable resolution limits, the discrepancy with the target configurations is still particular. The reason remains unclear. One hypothesis is that the larger dataset size may introduce more variability since it contains about three times as many hits and is not influenced by muon absorption. All datasets were reconstructed using identical algorithm parameters, so reconstruction differences cannot explain the observed broadening. To investigate this further, the tracking algorithm could be rerun with additional flags to trace potential anomalies. However, due to the large computational time of reconstructing these free sky tracks, such an analysis is left for future studies.

5.2 Muon transmission through the pyramid

After acquiring the hit data and performing the track reconstruction, the final phase is to reconstruct the image of the pyramid. This is typically done in terms of direction angles θ and ϕ and more specifically, using the components $\tan(\theta_x)$ and $\tan(\theta_y)$, where θ_x and θ_y correspond to the azimuthal projections of the direction vector on the xz-plane and yz-plane, respectively. Given a direction vector defined as in Equation 4.5, these components can be calculated as:

$$\begin{cases} \tan(\theta_x) = b_x / b_z = \tan(\theta)\cos(\phi) \\ \tan(\theta_y) = b_y / b_z = \tan(\theta)\sin(\phi) \end{cases} . \tag{5.1}$$

The image of the muon count distribution was reconstructed as a two-dimensional histogram with 200 bins for both dimensions, and the results for the free sky and pyramid are shown in Figures 5.3a and 5.3b, respectively. In the free sky configuration, the muon rate distribution



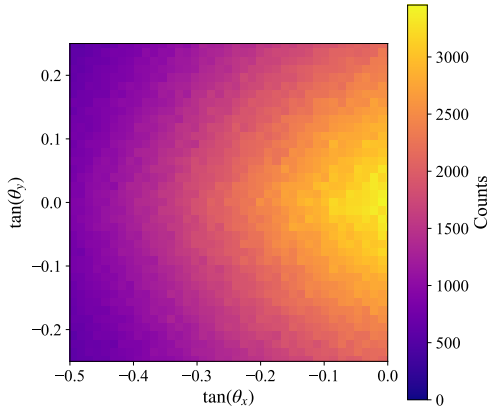
(a) Histogram from the free sky acquisition.

(b) Histogram from the pyramid acquisition.

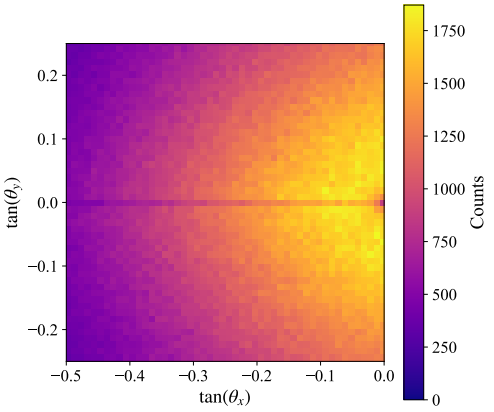
Figure 5.3: Two-dimensional histograms of the number of detected muons at the detector position for the free sky and pyramid configuration. The dimension of the histograms is 200×200 bins.

appears fairly uniform, as expected in the absence of any structures. However, a notable absence of counts is observed in the region around $\tan(\theta_y) \approx 0$ and $\tan(\theta_x) \leq 0$, particularly near the origin. This is more clearly visible when zooming in, as shown in Figure 5.4b. This region corresponds to muons arriving from the fourth quadrant of the xz-plane and to vertical muons. This drop in counts is not physical, as confirmed by comparison with the distribution of the generated muon directions in Figure 5.4a. The discrepancy becomes even more visible when analyzing the difference between the reconstructed and generated images, which clearly shows the deficit of reconstructed tracks in this region, as visible in Figure 5.4c. A likely explanation is that the clustering algorithm used in the track reconstruction was not periodic, which may have introduced a bias or inefficiency in that angular sector. Regarding the deficit at the origin, this can presumably be attributed to the redundancy of ϕ for vertical muons, i.e., $\theta = 0$. To address this in the accumulator, calculated directions with θ falling below a small threshold $\epsilon \sim 0$ should be concentrated into a single bin, independent of the corresponding ϕ value. These deficits also explain the tenfold increase of undetected tracks for the free sky dataset as discussed in the previous section, since the muon flux is maximal for nearly vertical muons. While such discrepancies are unfortunate, they highlight the importance of conducting these muographic simulations to identify and correct deficiencies in the tracking algorithm.

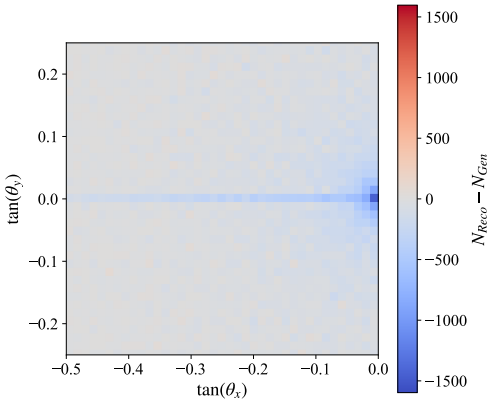
On the contrary, the reconstructed image of the pyramid displays a more structured pattern, consistent with the expected absorption of muons by the pyramid's internal geometry. The region with lower counts (pink-purple) is curved inwards toward the origin, while the region around the origin (orange-yellow) forms a square-like shape with distinct corners. Notably, the ϕ-assymetry is absent here. This may be due to the smaller amount of data, but it is more likely due to better tracking performance, as supported by the values in Table 5.1 discussed in the previous section. Since nearly vertical muons enter near the apex, the distance traveled through the pyramid before reaching the detector will be larger, resulting in a lower transmission for



(a) Muon distribution for the generated directions.



(b) Muon distribution for the reconstructed directions.



(c) Difference of the muon distribution between the generated and reconstructed directions.

Figure 5.4: Comparison of the muon distribution between the generated directions and the reconstructed directions for the free sky configuration. Zoom-in on region $-0.25 < \tan(\theta_y) < 0.25$ and $\tan(\theta_x) \leq 0$.

this direction.

To better interpret this reconstructed structure, the pyramid muon rate is compared to the free sky rate by calculating the transmission according to Equation 1.13, using the simulation times as acquisition times, as mentioned in the previous section. First, to have a reference, a two-dimensional map of the traversed length through the pyramid was reconstructed as a function of the muon direction. This map was created using the same simulation employed to generate the length histogram described in Subsection 3.2.2. It was constrained to the region in the field of view of the detector and is demonstrated in Figure 5.5. The cross-like structure in this map can be explained by the features of the pyramid. Muons entering the pyramid from the corners, corresponding to the diagonals in the image, will travel a longer distance than muons entering from the sides of the pyramid. Furthermore, for the region in the detector's field of view, the largest traversed distance will be for muons entering near the top, corresponding to

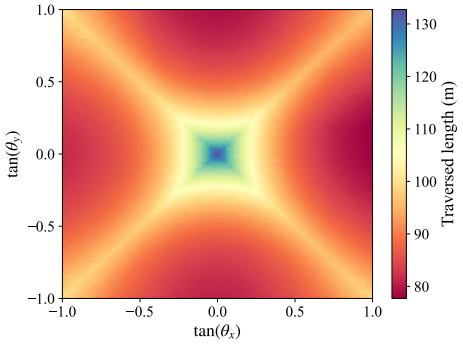
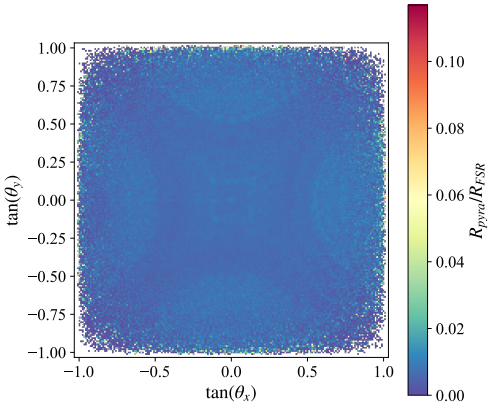
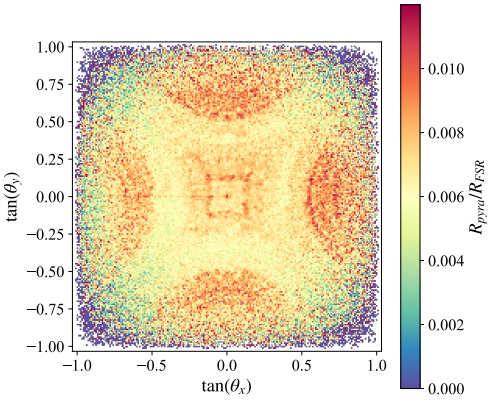


Figure 5.5: Two-dimensional map of the traversed length through the pyramid as a function of the muon direction, for the pyramid as known today. The map is constrained to the detector's field of view.

the blue-green region of the map. While this can serve as a benchmark for the muon transmission image, it has to be remarked that a transmission map also takes the muon spectrum into account, which will result in a more complicated relationship. The resulting transmission map for the pyramid is shown in Figure 5.6a. As expected, the transmission values are less than one, reflecting the absorption of muons by the pyramid. While the absolute values may not exactly match reality, due to the simulation-specific assumptions and acquisition time approximation, the order of magnitude is consistent with expectations. In this initial transmission map, some structural features begin to emerge, but the overall image remains relatively uniform compared to the traversed length map in Figure 5.5. This is largely due to the edges: as visible in Figure 5.6a, the outer regions exhibit some bins with relatively higher transmission, primarily due to



(a) Transmission map.



(b) Masked transmission map: $T < 0.012$.

Figure 5.6: Two-dimensional maps of the muon transmission through the pyramid as known today. The dimension of the histograms is 200×200 bins.

low statistics. Although a substantial amount of data was acquired, the muon flux decreases with increasing zenith angle θ, resulting in fewer events for large values of θ. To enhance the appearance of the structural features, a mask was applied to filter out higher transmission val-

ues, selecting $T < 0.012$. With this threshold applied, the features of the pyramid become much clearer, as demonstrated in Figure 5.6b. In particular, the cross-shaped pattern is visible: the four regions at the sides of the map containing relatively higher transmission (red-colored) compared to the central region and diagonals. These diagonals correspond to the path through the corners of the pyramid, where muons will traverse more limestone, resulting in stronger attenuation. The central region exhibits a more complex behavior, showing alternating regions of higher and lower transmission. This pattern is due to the interplay between the geometry of the pyramid and the angular distribution of the incoming muon flux. In the origin of the map, an increase in transmission is observed. However, this effect is an artifact resulting from the less reliable track reconstruction in the free sky dataset and consequently biased muon distribution. This was also confirmed by reconstructing the transmission map for the generated muon directions, as shown in Figure C.2 in Appendix C. Ignoring this artifact, the transmission near the origin is lower, which corresponds with the physical expectations, since this region correlates to nearly vertical muons, entering near the apex of the pyramid. Because these muons travel through a greater thickness of limestone, they experience more absorption, resulting in lower transmission. Moving slightly outward from the center, a localized increase is observed, which can be attributed to the decrease in the traversed length. Beyond this point, the transmission decreases again until it begins to rise once more at large inclination angles, apart from the diagonals. This trend reflects the competition between the path length and the muon flux distribution: while the flux decreases with increasing zenith angle, the distance traveled through the pyramid also shortens. Additionally, for higher inclination angles, the average energy of the muons increases, which may further contribute to the increased transmission observed. The very low transmission in the corners of the map is attributed to limited data in those regions.

5.3 Muographic search for a hidden void

The purpose of the ScIDEP project is to scan the Pyramid of Khafre for potential hidden voids, to gain a deeper understanding of its internal structure, and in turn, the construction of the pyramids as a whole. Hence, an interesting aspect to determine in this analysis is the smallest void that could realistically be detected with the current detector model, as well as estimating the time required for data acquisition to observe a muon excess associated with such a void.

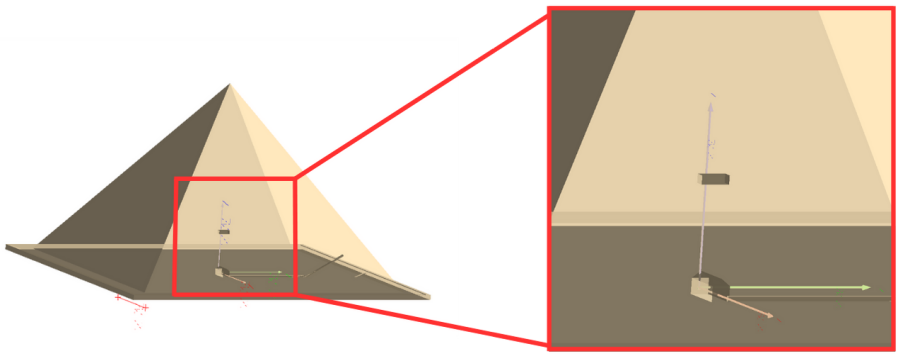


Figure 5.7: Geant4 model of the Pyramid of Khafre with an additional room placed above the King's chamber, at position $(5\ \text{m},\ 0\ \text{m},\ 30\ \text{m})$.

As shown in the previous section, muon transmission is generally lower in the region directly above the detector compared to the lateral regions at the edges of the detector's field of view. Therefore, one expects that a void placed at a lower inclination angle would be easier to detect, as more muons are observed in this region, leading to more statistics. To explore the detection limits, a hypothetical hidden chamber with dimensions of $3 \times 8 \times 2.5$ m^3 was placed into the pyramid at a height of 30 m, with an offset of 5 m along the x-axis. These dimensions are based on the sizes of the existing rooms within the pyramid. This corresponds to the chamber being positioned above the King's chamber, with its longest side oriented orthogonally to that of the King's chamber, as illustrated in the Geant4 model in Figure 5.7. The muon distribution of the pyramid with this additional room is shown in Figure 5.8. At first glance, the distribution ap-

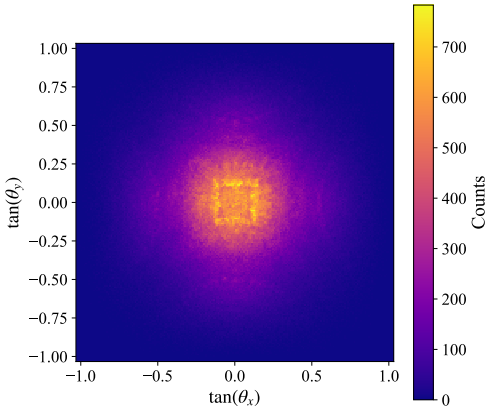


Figure 5.8: Two-dimensional histogram of the number of detected muons at the detector position for the pyramid with additional room configuration. The dimension of the histogram is 200×200 bins.

pears very similar to the case without the hidden room, with no notable difference to the naked eye. This is expected, given the relatively small size of the void in comparison to the overall mass of the pyramid, and the limited data acquisition.

To evaluate whether the presence of such a hidden chamber can be detected, the relative transmission is calculated, as defined in Equation 1.18. This way, the measured transmission can be compared to the expected one. Again, a two-dimensional map was generated of the traversed length through the pyramid as a function of the muon direction, as shown in Figure 5.9. Comparing this map with the map of the traversed length through the pyramid without hidden structures (Figure 5.5), the room is clearly identifiable as the red rectangular spot, located to the right of the center, in the region of reduced traversed length. This can be used as a reference to locate the room in the relative transmission map. In this analysis, only simulated data is available, as measured data acquisition is not ongoing yet. Accordingly, $T_{expected}$ represents the simulated transmission through the known structure of the pyramid, while $T_{measured}$ includes the effect of the simulated hidden chamber. The resulting relative transmission is shown in Figure 5.10a. In this image, 100 bins were used in each dimension, to compensate for the limited statistics. Nevertheless, especially at the outer edges, several bins exhibit high relative transmission values. In principle, the relative transmission should be close to 1, with values exceeding 1 in angular regions corresponding to a void. A trained eye might spot a faint, lighter-blue region slightly shifted to the right of the center in Figure 5.10a. However, due to the low statistics at

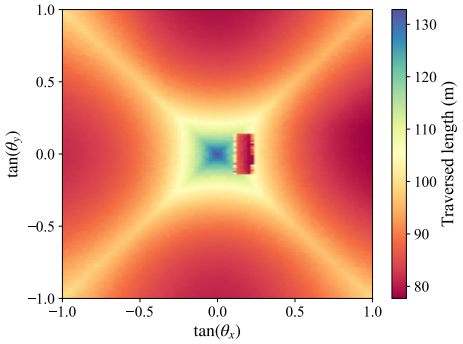
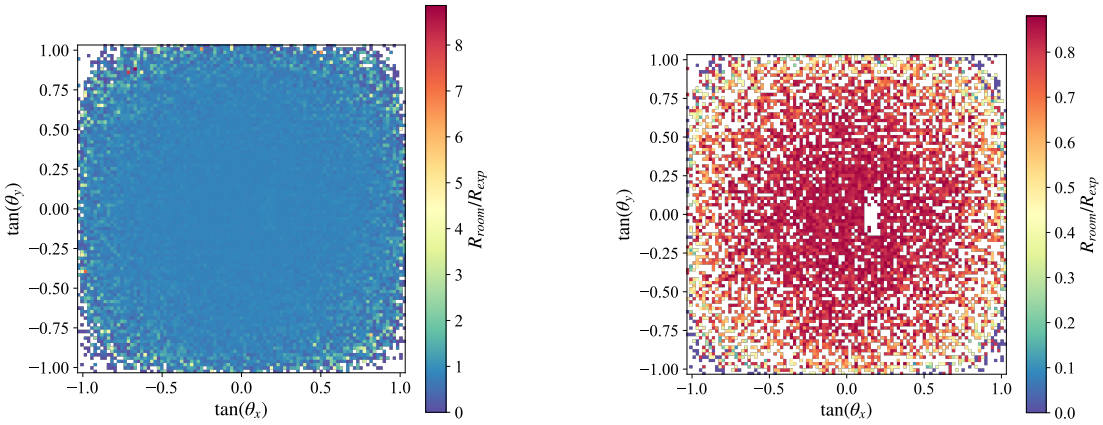


Figure 5.9: Two-dimensional map of the traversed length through the pyramid as a function of the muon direction, for the configuration of the pyramid with an additional void.

the image boundaries, this signal remains vague. To enhance visibility, a mask can be applied to selectively display regions where the relative transmission falls below a certain threshold, potentially revealing the shape of the void. Initially, a mask of $R < 1$ was tested. In theory, this should highlight the region corresponding to the room, by excluding the excess. However, the resulting image remained fairly uniform. When testing other thresholds, a value of $R < 0.88$ produced a visible contour where the data vanishes, as observed in Figure 5.10b. This contour



(a) Relative transmission map.

(b) Masked relative transmission: $R < 0.88$.

Figure 5.10: Two-dimensional relative transmission maps of the pyramid with additional room configuration vs. pyramid as known today. The dimension of the histograms is 100×100 bins.

corresponds with the simulated room's location. Although the order of magnitude of the relative transmission is reasonable, the value of 0.88 for the mask proves that the absolute value appears to be underestimated. This highlights the limitation of using the simulation runtimes as an approximation for the realistic acquisition times. While this approach provides a useful initial estimate of the order of the muographic quantities, it is not a fully accurate approximation. A more precise estimation could be achieved by incorporating EcoMug's estimated acquisition time and scaling it according to the actual surface area from which muons are generated outside

the pyramid.

Due to the low statistics, a cut was applied to the data, restricting the angular range to $-0.4 < \tan(\theta_x), \tan(\theta_y) < 0.4$, and the number of bins was reduced to 50. Figure 5.11 shows the resulting relative transmission map for this region, where a drop of magnitude is observed in the color bar, compared to the one from 5.10a, due to the low statistic edge bins being excluded. In this zoomed-in map, the room becomes visible even without applying a mask. The purple

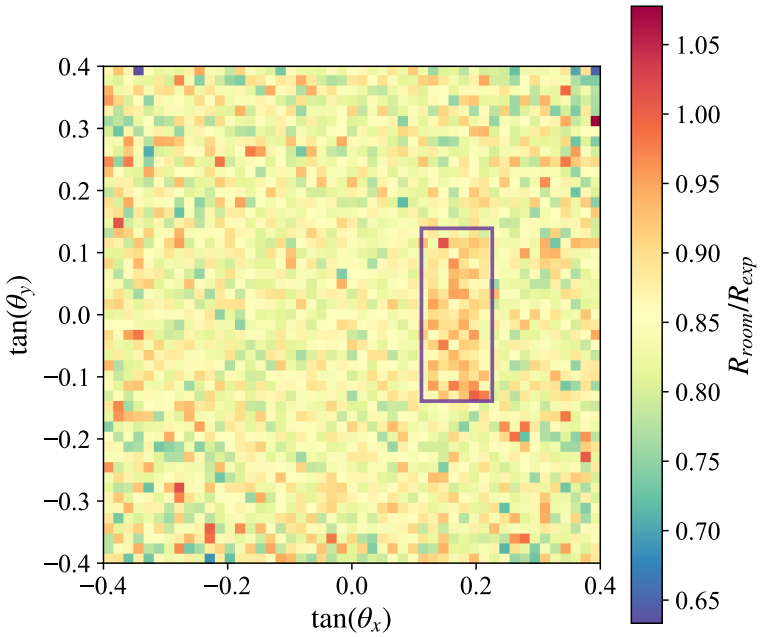


Figure 5.11: Two-dimensional relative transmission map of the pyramid with an additional room configuration vs. the pyramid as known today, after a cut was applied: $-0.4 < \tan(\theta_x), \tan(\theta_y) < 0.4$. The dimension of the histogram is 50×50 bins. The angular range of the room is indicated by the purple rectangle

rectangle in the figure denotes the angular region corresponding to the actual room dimensions: $0.112 < \tan(\theta_x) < 0.226$ and $-0.139 < \tan(\theta_y) < 0.139$. To determine whether this region truly shows a muon excess, the flux difference between the expected and measured configuration was calculated. The initial approach was to fix $\tan(\theta_y)$ (e.g., at 0, within the range of the visible excess) and plot the flux difference profile along $\tan(\theta_x)$, and vice versa. The flux difference here is more specifically the difference in the flux density between the simulation with the hidden room and the one with the pyramid as known today. This density-based comparison was necessary for a proper normalization because simulation runtimes only approximate the actual acquisition time. Instead of analyzing a single bin, however, a range of bins was selected, and the total flux within this range was used to calculate the difference. The results are presented in Figure 5.12, where the top plot shows the difference profile along $\tan(\theta_x)$, and the right plot the profile along $\tan(\theta_y)$. The dashed purple lines indicate the angular ranges used for $\tan(\theta_x)$ and $\tan(\theta_y)$, derived from the actual span of the simulated room. Ideally, one would determine these ranges based on the observed excess, but this was not possible here due to the limited statistical significance. The light purple shaded regions in the top and right plots correspond to

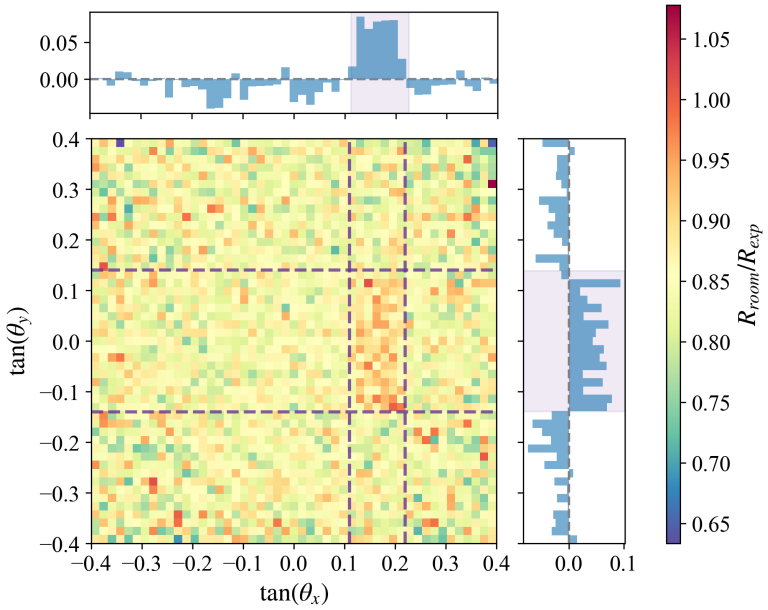
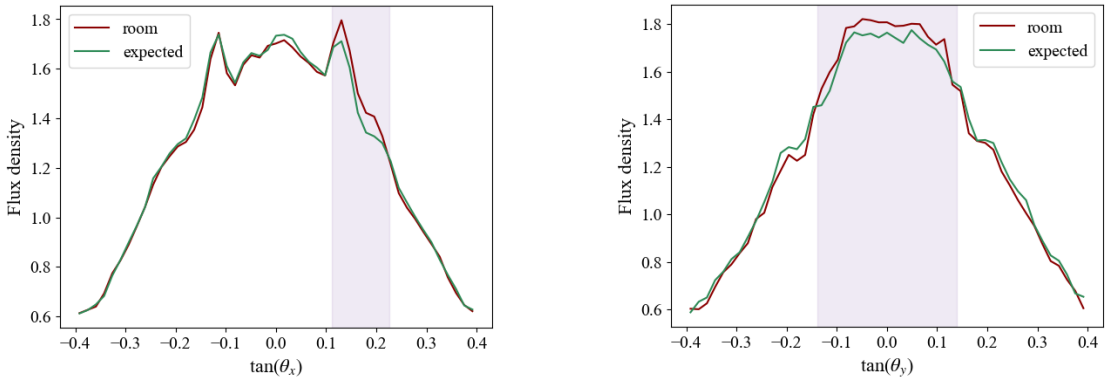


Figure 5.12: The same relative transmission map as in Figure 5.11. The top plot shows the profile of the total flux density difference between the pyramid with an extra room and the pyramid as known today, for $-0.14 < \tan(\theta_y) < 0.14$, along the range of $\tan(\theta_x)$. The right plot shows the profile of the total flux density difference for $0.11 < \tan(\theta_x) < 0.22$, along the range of $\tan(\theta_y)$. The dashed purple lines indicate the angular ranges used for $\tan(\theta_x)$ and $\tan(\theta_y)$ to calculate the flux differences. The light purple shaded regions in the top and right plot correspond to the angular extent of the simulated void

the angular extent of the simulated void. In theory, the flux difference should be around zero if the measured configuration matches the expected model. While the observed excess within the room's angular region is too small to draw a strong statistical conclusion, a clear trend emerges: within the room's range, a muon excess is observed, whereas outside this range, there is a deficit. To further explore this observation, the actual muon (density) flux profiles were plotted for the same angular ranges of $\tan(\theta_x)$ and $\tan(\theta_y)$, as shown in Figure 5.13. Within the purple-shaded region—the region of the room—an increased flux is observed for the simulation with the hidden room, underlining the presence of a localized muon excess. From the current analysis, it can be concluded—aside from statistics—that a clear muon excess is observable in the region corresponding to the simulated void. To determine whether this excess is statistically significant, however, more simulated data must be acquired.

The dimensions of the simulated void were based on the existing chambers in the Pyramid of Khafre, taken slightly smaller than the known subsidiary chamber. This choice was motivated by the expectation that any yet-undiscovered rooms are likely to be smaller and to test the limits of the detector. However, larger voids remain of interest, as the ScanPyramids collaboration discovered a corridor over 30 m in the Pyramid of Khufu, as mentioned in Subsection 2.1 [42]. In general, regardless of the size of the simulated void, additional simulation data is required to obtain more detailed results and to perform a meaningful statistical analysis. To accelerate the exploration process, multiple voids could be inserted simultaneously in future sim-



(a) The total flux density profile along $\tan(\theta_x)$, for $-0.14 < \tan(\theta_y) < 0.14$.

(b) The flux density profiles along $\tan(\theta_y)$, for $0.11 < \tan(\theta_x) < 0.22$.

Figure 5.13: The flux density profiles, for the same ranges as illustrated by the dashed purple lines in Figure 5.12.The light purple shaded regions correspond to the angular extent of the simulated void.

ulations, provided their angular regions do not overlap. Overlapping regions would introduce ambiguity in the excess profiles unless a three-dimensional reconstruction is feasible. In that case, overlapping angular regions could be disentangled through tomographic imaging techniques [74]. Hence, an important future step is to include the second detector in the simulation framework. Furthermore, using both detectors simultaneously—each placed at different viewpoints—would allow for data acquisition from multiple angles, enabling three-dimensional reconstruction of the pyramid's interior. For even more detailed simulations, a future model could include the limestone cap of the pyramid. Additionally, incorporating a realistic detector response is essential for producing transmission images that accurately reflect the experimental conditions. To ensure meaningful comparisons between simulated and measured data, it is crucial that the analysis pipelines for both remain as consistent as possible. Alongside this, the muon tracking algorithm should be reviewed and refined, especially in light of the discrepancies observed in the free sky data. Moreover, acquisition times should be determined more accurately to estimate how long real measurements would take in order to achieve sufficient data for statistical precision. This could be explored by incorporating EcoMug's predicted acquisition time and adjusting it based on the actual surface from which muons are generated outside the pyramid.

Chapter 6

Conclusion and Outlook

Over a century after Victor Hess's discovery of cosmic rays, cosmic ray muons have proven to be a powerful tool for imaging structures on Earth through muon radiography. Due to their relativistic properties, cosmic ray muons can penetrate large, dense materials, rendering them ideal candidates for non-invasive imaging of the internal structure of these objects. In particular, transmission muography estimates the opacity of an object by analyzing the attenuation of the muon flux along specific paths. By comparing the number of muons passing through a structure with a free sky reference, a density map of the structure can be produced. This technique has already proven its value across a wide range of disciplines, from volcanology to archaeology—a famous example being the ScanPyramids collaboration's discovery of a large void in the Great Pyramid of Giza. The Egyptian pyramids continue to captivate both researchers and the public, as there is no definite consensus on the precise methods of their construction. The ScIDEP project aims to shed new light on this mystery by employing transmission muography to investigate the internal structure of the Pyramid of Khafre, the second-largest of the Pyramids of Giza. Despite being only slightly smaller, the interior of the Pyramid of Khafre appears significantly simpler. Using scintillator-based tracking detectors placed both inside and outside the pyramid, the objective is to acquire two- and three-dimensional maps of the Pyramid of Khafre, in search of hidden structures.

A key challenge in muography is the low probability that muons will traverse the detector plates after passing through the target: many CR muons will either miss the detector or be absorbed in the target. Consequently, it is computationally demanding to simulate and acquire a sufficient number of muon events. To address this, an optimized simulation was developed to constrain the generation of muons inside an inverted cone comprising the field of view of the detector. First, the pyramid and the detector components were modeled using Geant4, state-of-the-art for simulating particle interactions with matter. A 3D model of the pyramid as known today was imported and modeled as limestone. The scintillator plates were defined as sensitive detectors, enabling to record muon hit data when traversing a detector plate. After characterizing the experimental setup, a suitable cosmic ray muon generator was required. The choice of EcoMug was driven by its computational speed, adaptability, and seamless integration with Geant4. The optimized simulation started by generating muons on a small surface positioned close to the detector, maximizing the probability that a muon would pass through all three detector planes. When a muon passed sufficiently near all three planes, it was extrapolated to outside the pyramid, which required the muon's energy prior to entering the pyramid. Since the energy loss depends on the muon's energy and its stochastic interactions

along its path—making it a complex quantity to calculate—an approximative method was implemented: the relationship was captured in a two-dimensional histogram, correlating outgoing energy and path length to the corresponding incoming energy. An additional histogram was created to find the traversed length as a function of the muon's direction. Finally, using these histograms, muons could be generated outside the pyramid. These extrapolated muons were tracked through the pyramid and, if a muon passed through all three detector planes, its hit data and generated information were recorded for further analysis for the next stage.

This next stage focused on the reconstruction of the muon's trajectory. For this purpose, a new tracking algorithm was developed, based on the Hough transform, a computational method for detecting geometric shapes. The muon's trajectory was described by Robert's line representation and two accumulators were implemented: one to determine the line's direction (θ, ϕ), and another to identify the coordinate (x', y'). The clustering algorithm DBSCAN was applied to the accumulators to identify clusters of maximal votes—corresponding to likely track candidates—followed by a fit to refine the resolution. The algorithm was evaluated for *ideal tracks*, referring to muon trajectories derived from the true hit coordinates (i.e., not reconstructed from the fiber signals). While computational time increases significantly with the number of input tracks, this could be managed by splitting the input into batches. A sample of 100 events was processed in 10 different configurations to assess the algorithm's performance in function of the number of input tracks. This was done for various values of the tracking parameters, where results showed that subtle changes in these values could significantly affect the performance. Finally, the influence of noise and smearing on the performance of the algorithm was tested. While the presence of noise appears manageable, smearing was found to be a harder challenge, indicating that further refinement of the algorithm will be necessary for its application to realistic data.

Finally, the optimized simulation and track reconstruction were integrated to reconstruct two-dimensional muographic images of the Pyramid of Khafre. Simulations were run for three configurations: free sky, the current model of the pyramid, and the pyramid with a hypothetical additional void. Approximately 10 million events were simulated for the free sky scenario and about 3 million for each of the two pyramid configurations. While the false track rate remained below 0.33%, the rate of undetected tracks in the free sky data was—with a value of 2.2%—notably higher compared to the pyramid configurations. This discrepancy was traced back to limitations in the track reconstruction, specifically the lack of periodic clustering in the $\theta\phi$-accumulator and a lack of proper treatment for vertical tracks. These findings highlight the critical role of simulations in validating and refining the reconstruction algorithm. Using the reconstructed tracks, a transmission map was generated by comparing the *muon rate* distributions between the free sky and pyramid configuration. Here, the muon rate was defined as the number of detected muons divided by the simulation runtime. The resulting transmission map showed realistic magnitudes (less than 1), where the structures present in the map corresponded to the known features of the pyramid. Lastly, to investigate the detector's limits in terms of sensitivity, a room was simulated above the King's Chamber. In this configuration, a localized muon excess was observed in the relative transmission map, which corresponded to the angular span of the hypothetical room. While definitive conclusions cannot be made because of limited statistics, these results offer a first insight into the detector's potential for resolving hidden structures and highlight promising directions for future refinement and studies.

The ScIDEP project emphasizes the value of muography in non-invasive archaeological exploration to uncover new insight into ancient structures, and illustrates the importance of in-

terdisciplinary and international scientific collaboration. Looking forward, additional data is required to draw statistically significant conclusions. This can be achieved either by extending the simulation runtime or by employing alternative simulation techniques. To accelerate the analysis, multiple void hypotheses could be evaluated simultaneously. Another step forward involves simulating the second detector and determining its optimal placement outside the pyramid. By modeling the expected muon flux on this detector, an ideal positioning can be selected to maximize signal strength. Consequently, more accurate estimations of the data acquisition times are essential to determine how long actual data collection would need to last to reach sufficient statistical precision. Additionally, with two detectors in operation, three-dimensional reconstruction becomes feasible, and the development of a tomographic reconstruction algorithm will be necessary. Incorporating a realistic detector response—including hit reconstruction from the SiPM signals, and detector efficiency—is also vital for generating transmission images that accurately reflect the experimental conditions. Furthermore, for more precise modeling, the model of the pyramid could include the limestone cap at the top of the pyramid. Lastly, another crucial development is the tracking algorithm. In particular, the $\theta\phi$-accumulator should be refined, in light of the discrepancies observed in the free sky data, and the influence noise should be accounted for by including secondary particles in the simulation. Just as pyramids are built stone by stone, each refinement is a building block in the muographic research of the ScIDEP project.

Appendix A

Robert's line representation: coordinate frame

As discussed in Section 4.2, Robert's line representation of the three-dimensional line is defined by four parameters, θ, ϕ, x', and y'; where θ and ϕ determine the direction of the line. The parameters x' and y' are defined as the coordinates of intersection of the line and a plane going through the origin with as its normal the line. The coordinate (x', y') is defined in a 2D coordinate frame in the plane itself. This coordinate frame is determined in such a way that it is unique and without special cases and singularities. The way this frame is defined is by rotating in the plane defined by $\vec{z}$ and $\vec{b}$ (thus rotating around an axis $\vec{z} \times \vec{b}$), by an angle $\arccos(\vec{z} \cdot \vec{b})$. The rotation matrix A is:

$$A = \begin{pmatrix} 1 - \frac{b_x^2}{1+b_z} & -\frac{b_x b_y}{1+b_z} & b_x \\ -\frac{b_x b_y}{1+b_z} & 1 - \frac{b_y^2}{1+b_z} & b_y \\ -b_x & -b_y & b_z \end{pmatrix} \tag{A.1}$$

Then the unit vectors of the coordinate frame are given by the columns of A. For a 3D line with an arbitrary reference point $\vec{p}$, the intersection coordinate (x', y') can be calculated as follows:

$$x' = \vec{p} \cdot \hat{x}' = \vec{p} \cdot A\hat{x} = \left(1 - \frac{b_x^2}{1+b_z}\right) p_x - \left(\frac{b_x b_y}{1+b_z}\right) p_y - b_x p_z \tag{A.2}$$

and

$$y' = \vec{p} \cdot \hat{y}' = \vec{p} \cdot A\hat{y} = -\left(\frac{b_x b_y}{1+b_z}\right) p_x + \left(1 - \frac{b_y^2}{1+b_z}\right) p_y - b_y p_z \,. \tag{A.3}$$

Reversely, when given (x', y'), a line anchor point can be calculated as

$$\vec{p} = x' \cdot \begin{pmatrix} 1 - b_x^2/(1+b_z) \\ -b_x b_y/(1+b_z) \\ -b_x \end{pmatrix} + y' \cdot \begin{pmatrix} -b_x b_y/(1+b_z) \\ 1 - b_y^2/(1+b_z) \\ -b_y \end{pmatrix} , \tag{A.4}$$

and the direction vector as

$$\vec{b} = \begin{pmatrix} b_x \\ b_y \\ +\sqrt{1 - b_x^2 - b_y^2} \end{pmatrix} . \tag{A.5}$$

Appendix B

Tracking algorithm parameters

In Subsection 4.4.2, the influence of several tracking parameters on the performance of the tracking algorithm was investigated. This appendix contains all values used to generate the corresponding efficiency figures.

Table B.1: Parameters used for Figure 4.5.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
180	360	100	1	3	2	1	1	2	3	1	50	5	1

Table B.2: Parameters used for Figure 4.7a.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
360	720	200	2	3	2	1	1	3	3	1	50	5	1

Table B.3: Parameters used for Figure 4.7b.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
360	720	200	1	3	2	1	1	3	3	1	50	5	1

Table B.4: Parameters used for Figure 4.9a.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
200	200	100	2	2	2	1	1	3	3	1	50	5	1

Table B.5: Parameters used for Figure 4.9b.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
200	200	100	2	2	2	1	1	3	3	1	50	5	1

Table B.6: Parameters used for Figure 4.10.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
200	200	100	1	3	2	1	2	2	3	1	50	5	1

Table B.7: Parameters used for Figure 4.11a.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
200	200	100	1	3	10	1	1	3	3	1	50	5	1

Table B.8: Parameters used for Figure 4.11b.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
200	200	100	1	3	10	1	1	3	3	1	700	5	1

Table B.9: Parameters used for Figure 4.12a.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
200	200	100	1	3	2	1	1	3	3	1	10	5	1

Table B.10: Parameters used for Figure 4.12b.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
200	200	100	1	3	2	1	1	3	3	1	30	5	1

In Subsection 4.4.3, the tracking resolution was evaluated by smearing a single track 1000 times and reconstructing the resulting smeared tracks. The values of the trackings parameters used for this reconstruction are provided in Table B.11.

Table B.11: Parameters used to determine the tracking resolution in case of smeared tracks, in Subsection 4.4.3.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
200	200	100	1	3	0.8	1	1	2	5	1	60	30	10

In Section 5.1, the track reconstruction was performed on the three simulation dataset. The algorithms used for this reconstruction are provided in Table B.12.

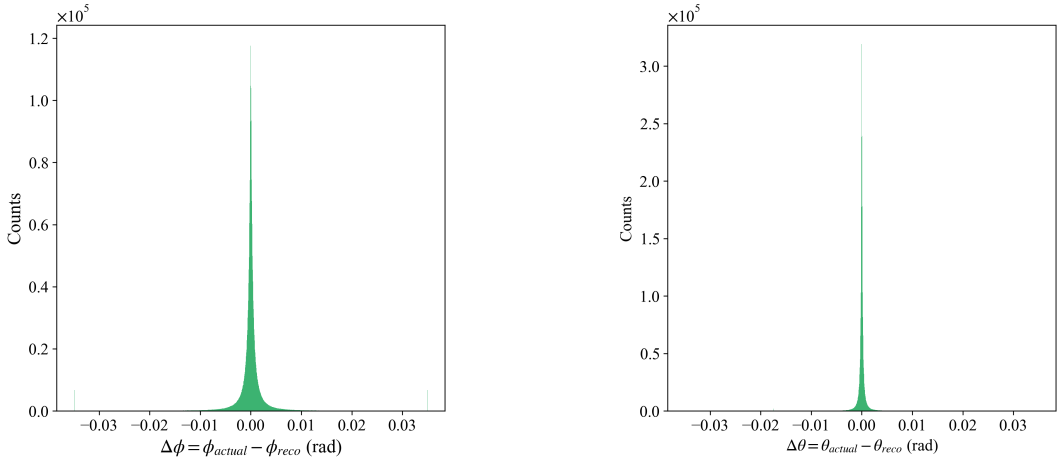
Table B.12: Parameters used to reconstruct the muon trajectories.

Accumulator			Clustering $\theta\phi$				Clustering xy				Distances		
n_θ	n_ϕ	n_{xy}	V_C	V_R	ϵ	$minPts$	V_C	V_R	ϵ	$minPts$	d_1	$d_{accepted}$	$d_{candidate}$
200	200	100	1	3	2	1	1	2	3	1	50	5	1

Appendix C

Tracking resolution

In Section 5.1, the tracking resolution was acquired by comparing the reconstructed directions to the actual muon directions, for all three configurations. The distribution of the angular differences for the configuration of the pyramid containing an additional room are presented in Figure C.1. The left-hand plot (C.1a) illustrates the difference in azimuth angle, while the right-hand plot (C.1b) shows the difference in zenith angle.



(a) Resolution for the azimuth angle ϕ.

(b) Resolution for the zenith angle θ.

Figure C.1: Tracking resolution of the tracking algorithm for the pyramid configuration with an additional room.

To verify that the transmission map of the pyramid in Section 5.2 should not exhibit a deficit near the origin of the map, the transmission map was reconstructed using the generated muon directions, as shown in Figure C.2.

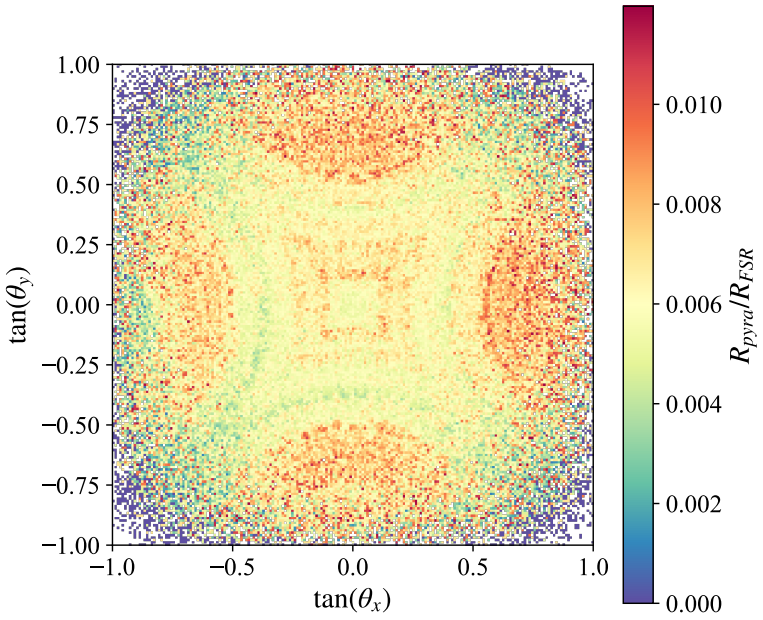


Figure C.2: Two-dimensional map of muon transmission through the pyramid as known today, based on the generated track directions. The dimension of the histogram is 200 $\times$ 200 bins. A mask to the transmission of $T < 0.012$ was applied.

Bibliography

[1] OpenAI. *ChatGPT (5 28 version)*. https://chat.openai.com. [Large language model]. 2025.

[2] V. Hess. *On the Observations of the Penetrating Radiation during Seven Balloon Flights*. 2018. arXiv: 1808.02927 [physics.hist-ph]. URL: https://arxiv.org/abs/1808.02927.

[3] *Muon Imaging*. TECDOC Series 2012. Vienna: International Atomic Energy Agency, 2022. ISBN: 978-92-0-142722-9. URL: https://www.iaea.org/publications/15182/muon-imaging.

[4] S. Hayakawa, K. Ito, and Y. Terashima. "Origin of Cosmic Rays". In: *Progress of Theoretical Physics Supplement* 6 (Feb. 1958), pp. 1–92. ISSN: 0375-9687. DOI: 10.1143/PTPS.6.1. eprint: https://academic.oup.com/ptps/article-pdf/doi/10.1143/PTPS.6.1/5269796/6-1.pdf. URL: https://doi.org/10.1143/PTPS.6.1.

[5] S. Navas et al. "Review of particle physics". In: *Phys. Rev. D* 110.3 (2024), p. 030001. DOI: 10.1103/PhysRevD.110.030001.

[6] V. S. Ptuskin. "The origin of cosmic rays". In: *Physics-Uspekhi* 53.9 (Dec. 2010), p. 958. DOI: 10.3367/UFNe.0180.201009k.1000. URL: https://dx.doi.org/10.3367/UFNe.0180.201009k.1000.

[7] Y. Ohsawa. "Ultrarelativistic particle acceleration in collisionless shock waves". In: *Physics Reports* 536.4 (2014). Ultrarelativistic particle acceleration in collisionless shock waves, pp. 147–254. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2013.11.004. URL: https://www.sciencedirect.com/science/article/pii/S0370157313004316.

[8] T. K. Gaisser. *Cosmic rays and particle physics*. 1990. ISBN: 978-0-521-33931-5.

[9] P. K. F. Grieder. "Chapter 1 - Cosmic Ray Properties, Relations and Definitions". In: *Cosmic Rays at Earth*. Ed. by P. K.F. Grieder. Amsterdam: Elsevier, 2001, pp. 1–53. ISBN: 978-0-444-50710-5. DOI: https://doi.org/10.1016/B978-044450710-5/50003-8. URL: https://www.sciencedirect.com/science/article/pii/B9780444507105500038.

[10] L. Cimmino. "Principles and Perspectives of Radiographic Imaging with Muons". In: *Journal of Imaging* 7.12 (Nov. 2021). DOI: https://doi.org/10.3390/jimaging7120253. URL: https://jais.andromedapublisher.org/index.php/JAIS/article/view/470.

[11] N. Lesparre et al. "Geophysical muon imaging: Feasibility and limits". In: *Geophysical Journal International* 183 (Dec. 2010), pp. 1348–1361. DOI: 10.1111/j.1365-246X.2010.04790.x.

[12] S. Schnetzer. *Cosmic Rays*. Accessed: 04-05-2025. 2023. URL: https://www.physics.rutgers.edu/~steves/Rutgers_summer_program/Cosmic_Rays.html.

[13] D. J. Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. New York, NY: Wiley, 2008.

[14] M. M. Duarte. *Muon tomography: application of image reconstruction algorithms within the LouMu project*. Dec. 2023.

[15] D. C. Giancoli. *Physics: Principles with Applications*. Physics: Principles with Applications. Pearson, 2014. ISBN: 9781292057552.

[16] M. Al Moussawi. "Muography: Using Cosmic Rays as an Imaging Tool for Volcanology and Cultural Heritage Applications". Prom. : Giammanco, Andrea ; Cortina Gil, Eduardo. PhD thesis. Louvain-la-Neuve, Belgium: Université catholique de Louvain, Jan. 2024. URL: http://hdl.handle.net/2078.1/286002.

[17] D. E. Groom, N. V. Mokhov, and S. I. Striganov. "Muon stopping power and range tables 10-MeV to 100-TeV". In: *Atom. Data Nucl. Data Tabl.* 78 (2001), pp. 183–356. DOI: 10.1006/adnd.2001.0861.

[18] N. Su et al. "A Comparison of Muon Flux Models at Sea Level for Muon Imaging and Low Background Experiments". In: *Frontiers in Energy Research* Volume 9 - 2021 (2021). ISSN: 2296-598X. DOI: 10.3389/fenrg.2021.750159. URL: https://www.frontiersin.org/journals/energy-research/articles/10.3389/fenrg.2021.750159.

[19] A. Tang et al. "Muon simulations for Super-Kamiokande, KamLAND, and CHOOZ". In: *Phys. Rev. D* 74 (5 Sept. 2006), p. 053007. DOI: 10.1103/PhysRevD.74.053007. URL: https://link.aps.org/doi/10.1103/PhysRevD.74.053007.

[20] M. Guan et al. *A parametrization of the cosmic-ray muon flux at sea-level*. 2015. arXiv: 1509.06176 [hep-ex]. URL: https://arxiv.org/abs/1509.06176.

[21] E. V. Bugaev et al. "Atmospheric muon flux at sea level, underground, and underwater". In: *Phys. Rev. D* 58 (5 July 1998), p. 054001. DOI: 10.1103/PhysRevD.58.054001. URL: https://link.aps.org/doi/10.1103/PhysRevD.58.054001.

[22] T. Hebbeker and C. Timmermans. "A compilation of high energy atmospheric muon data at sea level". In: *Astroparticle Physics* 18.1 (2002), pp. 107–127. ISSN: 0927-6505. DOI: https://doi.org/10.1016/S0927-6505(01)00180-3. URL: https://www.sciencedirect.com/science/article/pii/S0927650501001803.

[23] D. Heck et al. *CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers*. Tech. rep. 1998, pp. 1–90. URL: https://in2p3.hal.science/in2p3-00005094.

[24] C. Hagmann, D. Lange, and D. Wright. "Cosmic-ray shower generator (CRY) for Monte Carlo transport codes". In: *2007 IEEE Nuclear Science Symposium Conference Record*. Vol. 2. 2007, pp. 1143–1146. DOI: 10.1109/NSSMIC.2007.4437209.

[25] D. Pagano et al. "EcoMug: An Efficient COsmic MUon Generator for cosmic-ray muon applications". In: *Nucl. Instrum. Meth. A* 1014 (2021), p. 165732. DOI: 10.1016/j.nima.2021.165732.

[26] L. Bonechi, R. D'Alessandro, and A. Giammanco. "Atmospheric muons as an imaging tool". In: *Reviews in Physics* 5 (Nov. 2020), p. 100038. ISSN: 2405-4283. DOI: 10.1016/j.revip.2020.100038. URL: http://dx.doi.org/10.1016/j.revip.2020.100038.

[27] Z. Y. He et al. "Feasibility and optimization study of a two-dimensional density reconstruction method for large-object muography". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1061 (2024), p. 169138. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2024.169138. URL: https://www.sciencedirect.com/science/article/pii/S0168900224000640.

[28] R. Kaiser. "Muography: overview and future directions". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 377.2137 (2019), p. 20180049. DOI: 10.1098/rsta.2018.0049.

[29] E. P. George. "Cosmic Rays Measure Overburden of Tunnel". In: *Commonwealth Engineer* 42.12 (1955), pp. 455–457.

[30] R. N. I. Altameemi et al. "Determination of muon absorption coefficients in heavy metal elements". In: *Journal of Radiation Research and Applied Sciences* 12.1 (2019), pp. 281–288. ISSN: 1687-8507. DOI: https://doi.org/10.1080/16878507.2019.1652965. URL: https://www.sciencedirect.com/science/article/pii/S1687850721000340.

[31] G. Bonomi et al. "Applications of cosmic-ray muons". In: *Prog. Part. Nucl. Phys.* 112 (2020), p. 103768. DOI: 10.1016/j.ppnp.2020.103768.

[32] A. Ariga et al. "Nuclear Emulsions". In: *Particle Physics Reference Library: Volume 2: Detectors for Particles and Radiation*. Ed. by Christian W. Fabjan and Herwig Schopper. Cham: Springer International Publishing, 2020, pp. 383–438. ISBN: 978-3-030-35318-6. DOI: 10.1007/978-3-030-35318-6_9. URL: https://doi.org/10.1007/978-3-030-35318-6_9.

[33] T. Avgitas et al. "Prototype Cherenkov detector characterization for muon tomography applications". In: *Nuclear Physics B* 1006 (2024), p. 116645. ISSN: 0550-3213. DOI: https://doi.org/10.1016/j.nuclphysb.2024.116645. URL: https://www.sciencedirect.com/science/article/pii/S0550321324002116.

[34] G. Saracino et al. "Imaging of underground cavities with cosmic-ray muons from observations at Mt. Echia (Naples)". In: *Scientific Reports* 7.1 (Apr. 2017), p. 1181. DOI: 10.1038/s41598-017-01277-3.

[35] A. Anastasio et al. "The MU-RAY experiment. An application of SiPM technology to the understanding of volcanic phenomena". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 718 (2013). Proceedings of the 12th Pisa Meeting on Advanced Detectors, pp. 134–137. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2012.08.065. URL: https://www.sciencedirect.com/science/article/pii/S0168900212009680.

[36] L. Cimmino, G. Baccani, P. Noli, et al. "3D Muography for the Search of Hidden Cavities". In: *Scientific Reports* 9 (2019), p. 2974. DOI: 10.1038/s41598-019-39682-5.

[37] V. Tioukov, K. Morishima, C. Leggieri, et al. "Hidden chamber discovery in the underground Hellenistic necropolis of Neapolis by muography". In: *Scientific Reports* 13 (2023), p. 5438. DOI: 10.1038/s41598-023-32626-0. URL: https://doi.org/10.1038/s41598-023-32626-0.

[38] G. Liu et al. "High-precision muography in archaeogeophysics: A case study on Xi'an defensive walls". In: *Journal of Applied Physics* 133 (Jan. 2023), p. 014901. DOI: 10.1063/5.0123337.

[39] G. Saracino et al. "Applications of muon absorption radiography to the fields of archaeology and civil engineering". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 377.2137 (2019), p. 20180057. DOI: 10.1098/rsta.2018.0057.

[40] A. Giammanco et al. "Toward using cosmic rays to image cultural heritage objects". In: *iScience* 28.3 (Mar. 21, 2025). DOI: 10.1016/j.isci.2025.112094. URL: https://doi.org/10.1016/j.isci.2025.112094.

[41] V. Seglins and A. Kukela. "Unknown Deformations on the Facades of the Pyramid of Khafre at Giza Pyramid Complex in Egypt". In: *Procedia Engineering* 161 (2016). World Multidisciplinary Civil Engineering-Architecture-Urban Planning Symposium 2016, WMCAUS 2016, pp. 1780–1784. ISSN: 1877-7058. DOI: https://doi.org/10.1016/j.proeng.2016.08.777. URL: https://www.sciencedirect.com/science/article/pii/S1877705816330065.

[42] K. Morishima et al. "Discovery of a big void in Khufu's Pyramid by observation of cosmic-ray muons". In: *Nature* 552.7685 (Nov. 2017), pp. 386–390. ISSN: 1476-4687. DOI: 10.1038/nature24647. URL: http://dx.doi.org/10.1038/nature24647.

[43] Guardians Egypt. *The Pyramid of Khafre*. https://guardians.net/egypt/pyramids/Khafre/KhafrePyramid.htm. Accessed: 05-05-2025.

[44] M. Lehner. *The Complete Pyramids: Solving the Ancient Mysteries*. London: Thames & Hudson, 1997. ISBN: 9780500050842.

[45] M. Gamal et al. *The ScIDEP Muon Radiography Project at the Egyptian Pyramid of Khafre*. 2025. arXiv: 2503.21959 [physics.ins-det]. URL: https://arxiv.org/abs/2503.21959.

[46] L. W. Alvarez et al. "Search for Hidden Chambers in the Pyramids". In: *Science* 167.3919 (1970), pp. 832–839. DOI: 10.1126/science.167.3919.832. eprint: https://www.science.org/doi/pdf/10.1126/science.167.3919.832. URL: https://www.science.org/doi/abs/10.1126/science.167.3919.832.

[47] S. Procureur et al. "Precise characterization of a corridor-shaped structure in Khufu's Pyramid by observation of cosmic-ray muons". In: *Nature Communications* 14 (Mar. 2023). DOI: 10.1038/s41467-023-36351-0.

[48] T. Schumacher, P. Pugacheva, H. Allam, et al. "Confirmation of the ScanPyramids North Face Corridor in the Great Pyramid of Giza using multi-modal image fusion from three non-destructive testing techniques". In: *Scientific Reports* 15 (2025), p. 9275. DOI: 10.1038/s41598-025-91115-8.

[49] M. Tytgat et al. "The ScIDEP Project at the Egyptian Pyramid of Khafre". In: *Journal of Advanced Instrumentation in Science* 2024.1 (Aug. 2024). DOI: https://doi.org/10.31526/jais.2024.470. URL: https://jais.andromedapublisher.org/index.php/JAIS/article/view/470.

[50] Eljen Technology. *EJ-200 General Purpose Plastic Scintillator*. Technical Datasheet. Accessed: 20-03-2025. Eljen Technology, n.d. URL: https://wiki.jlab.org/ciswiki/images/4/49/EJ-200.pdf.

[51] R. Kouzes et al. *Novel Muon Tomography Detector for the Pyramids*. 2022. arXiv: 2202.08166 [physics.ins-det]. URL: https://arxiv.org/abs/2202.08166.

[52] Luxium Solutions. *Fiber Product Sheet*. Technical Datasheet. Accessed: 20-03-2025. Luxium Solutions, n.d. URL: https://luxiumsolutions.com/sites/default/files/2021-11/Fiber-Product-Sheet.pdf.

[53] Hamamatsu Photonics. *S14160/S14161 Series MPPCs*. Technical Datasheet. Accessed: 2025-03-20. Hamamatsu Photonics, n.d. URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99_SALES_LIBRARY/ssd/s14160_s14161_series_kapd1064e.pdf.

[54] Z. Wang et al. "Simulations for the ScIDEP Muon Radiography Project at the Egyptian Pyramid of Khafre". To be submitted to *IEEE Transactions on Nuclear Science*. 2025.

[55] CAEN S.p.A. *DT5550W - Desktop Waveform Digitizer*. https://www.caen.it/products/dt5550w/. Accessed: 05-05-2025. n.d.

[56] Z. Wang. "The ScIDEP Project at the Egyptian Pyramid of Khafre". Conference presentation, 2024 International Workshop on Cosmic-Ray Muography (Muography2024), Santa Fe, New Mexico, USA, 2024.

[57] A. Gindac et al. "Open-source SoC-based Off-the-shelf Industrial-Grade Low-Cost Logic Analyzer". In: *Romanian Journal of Information Science and Technology* 1.1 (2021), pp. 117–126. URL: https://www.romjist.ro/full-texts/paper683.pdf.

[58] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[59] V. Niess. "The PUMAS library". In: *Computer Physics Communications* 279 (Oct. 2022), p. 108438. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2022.108438. URL: http://dx.doi.org/10.1016/j.cpc.2022.108438.

[60] Geant4 Collaboration. *Geant4 User's Guide for Application Developers*. Version 11.2, Accessed: 15-05-2025. CERN. 2024. URL: https://geant4-userdoc.web.cern.ch/UsersGuides/ForApplicationDeveloper/fo/BookForApplicationDevelopers.pdf.

[61] M. Asai. *Introduction to Geant4*. https://cds.cern.ch/record/491492/files/p107.pdf. Accessed: 20-05-2025. Hiroshima, Japan, 1997.

[62] S. Aly et al. *Simulation Studies of a Novel Muography Detector for the Great Pyramids*. 2022. arXiv: 2202.07434 [physics.ins-det]. URL: https://arxiv.org/abs/2202.07434.

[63] Wikimedia Commons contributors. *Pyramid of Khafre.stl*. https://commons.wikimedia.org/wiki/File:Pyramid_of_Khafre.stl. Accessed: 05-05-2025. 2024.

[64] Blender Online Community. *Blender - a 3D modelling and rendering package*. Blender Foundation. Stichting Blender Foundation, Amsterdam, 2018. URL: http://www.blender.org.

[65] P. Cignoni et al. "MeshLab: an Open-Source Mesh Processing Tool." In: vol. 1. Jan. 2008, pp. 129–136. DOI: 10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136.

[66] C. M. Poole et al. "A CAD interface for GEANT4". In: *Australasian Physical & Engineering Sciences in Medicine* 35.3 (Sept. 2012). Epub 2012 Sep 6, pp. 329–334. DOI: 10.1007/s13246-012-0159-8.

[67] P. V. C. Hough. *Method and means for recognizing complex patterns*. Dec. 1962.

[68] O. Duda R and P. E. Hart. "Use of the Hough transformation to detect lines and curves in pictures". In: *Commun. ACM* 15.1 (Jan. 1972), pp. 11–15. ISSN: 0001-0782. DOI: 10.1145/361237.361242. URL: https://doi.org/10.1145/361237.361242.

[69] A. S. Hassanein et al. "A Survey on Hough Transform, Theory, Techniques and Applications". In: *CoRR* abs/1502.02160 (2015). arXiv: 1502.02160. URL: http://arxiv.org/abs/1502.02160.

[70] K. S. Roberts. "A new representation for a line". In: *Proceedings CVPR '88: The Computer Society Conference on Computer Vision and Pattern Recognition* (1988), pp. 635–640. URL: https://api.semanticscholar.org/CorpusID:21080758.

[71] C. Dalitz, T. Schramke, and M. Jeltsch. "Iterative Hough Transform for Line Detection in 3D Point Clouds". In: *Image Processing On Line* 7 (2017). https://doi.org/10.5201/ipol.2017.208, pp. 184–196.

[72] M. Ester et al. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD)*. 1996, pp. 226–231. URL: https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf.

[73] R. Yehoshua. *DBSCAN: Density-Based Clustering*. Accessed: 2025-05-29. 2023. URL: https://ai.plainenglish.io/dbscan-density-based-clustering-aaebd76e2c8c.

[74] A. D. Bross et al. *Tomographic Muon Imaging of the Great Pyramid of Giza*. 2022. arXiv: 2202.08184 [physics.ins-det]. URL: https://arxiv.org/abs/2202.08184.