Modeling high energy atmospheric physics and lightning inception

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnificus, prof.dr.ir. F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op donderdag 15 maart 2018 om 13.30 uur

door

Casper Rutjes

geboren te Tilburg

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

voorzitter:	prof.dr.ir. G.M.W. Kroesen
1^e promotor:	prof.dr. U.M. Ebert
2^e promotor:	prof.dr. O. Scholten (Rijksuniversiteit Groningen)
co-promotor:	dr. E. Camporeale (CWI)
onafhankelijke leden:	prof.dr.ir. W. van Saarloos (Universiteit Leiden)
	prof.dr. Y. Yair (IDC Herzliya)
	prof.dr.ir. O.J. Luiten
adviseur(s):	dr.ir. S. Nijdam

Contents

Part I

Introduction and background

Chapter 1

Research topic, contributions and structure

First the environment - the thundercloud - is introduced in which the natural phenomena studied in this thesis occur. Secondly, the research topics are addressed, which are twofold. Thirdly, the contributions and structure of the thesis is given.

1.1 Environment of studied phenomena - the thundercloud

Thunderclouds are meteorological systems where significant electric charge is separated by strong convection. The strong convection brings warm air, which includes water vapor, to great heights. During the ascent into cooler regions, the water vapor condenses and later freezes, releasing all its latent heat. This heat powers the convection further. The ice particles formed at great heights eventually fall down again through the regions of ascending droplets. In these regions, charge is separated between the two opposite moving populations of particles. Much like a balloon rubbed by woolen cloth, charge is transferred on collision from one to the other. Net positive electrical charge is put on the ascending particles and negative charge on the descending particles. Simply stated, the thundercloud acts as a giant power supply where the latent heat of water vapor is converted into electrical energy, reaching over 100 MV of potential difference, see figure 1.1 for illustration. In a first order approximation, thunderclouds are dipoles with a main positive layer above a main negative layer. A second order includes the screening layers, upper negative and lower positive. In reality, and in particular outside updraft regions, the system can be much more complex, see figure 1.1. This large electric potential leads directly and indirectly to powerful natural phenomena which are studied in this thesis.
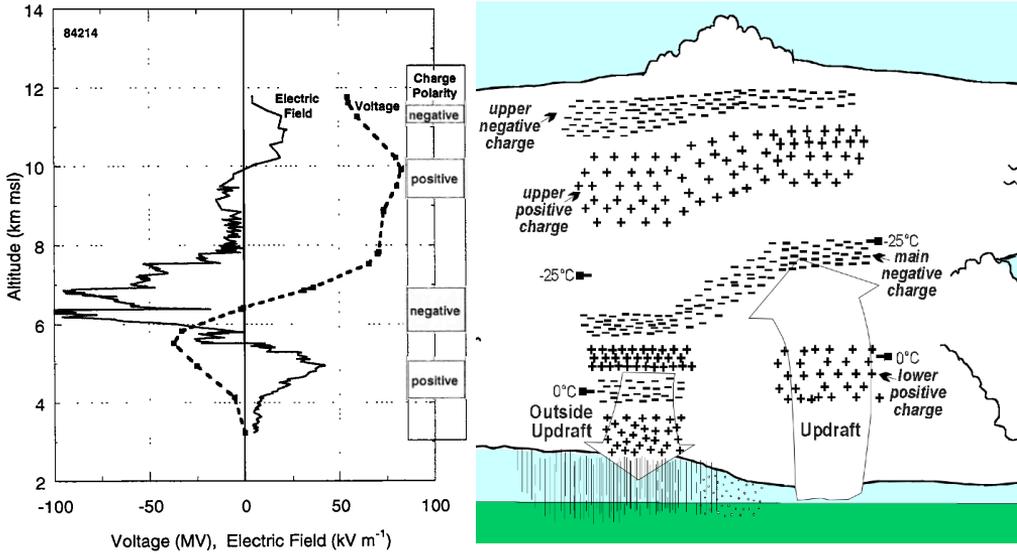
Figure 1.1: Balloon sounding of electric field inside updraft region (on the left) and illustration of charge structure inside thundercloud (based on multiple balloon soundings), from [1].

1.2 Two research topics

1.2.1 Lightning initiation

Thunderclouds produce one of the most familiar types of discharges, known as *lightning*. Lightning is an electric discharge that develops in different phases. The last phase, called the return stroke, is observed as a flash and heard as thunder. In the first phase preliminary discharges try to start, which is far from trivial and sometimes referred to as *the lightning initiation problem*. Roughly half of this thesis is devoted to contributions to the lightning initiation problem.

1.2.2 Understanding of high energy radiation from thunderclouds and lightning

After the first phase, lightning further develops and can produce flashes of high energy radiation, including the emission of neutrons. But also, much longer lasting glows of high energy radiation have been observed from thunderclouds, including again the emission of neutrons. The presence of neutrons is stressed, as that indicates that the photon energy must have exceeded 10 MeV. Roughly half of this thesis is devoted to contributions to the understanding of high energy radiation from thunderstorms and lightning, and, in particular, of neutron emission and propagation.

1.3 Contributions and structure

The thesis is divided in five parts:

   I Introduction and historical background (of part II, III and IV)

  II Model development, reduction and evaluation

 III Lightning initiation

 IV Neutrons from thunderclouds and lightning

  V Conclusions (of part II, III and IV), references and Dutch summary

Each research topic can be divided into two, but both can be again divided into two, of which half is of predictive nature (specific to the subject) and the other half of model development, reduction and evaluation nature (applicable to more general problems). Therefore the main body consists of three collections (part II, III and IV) of stand alone (in preparation or published) scientific works. In that respect, part (II) collects work that can be used for problems in part (III) or part (IV) or both.

In total the main body of this thesis consists of nine chapters, of which by current date four are published in international journals. Next to the chapters presented in this thesis, I collaborated and contributed in the joint 'Cosmic Lightning' research project, initiated by prof. dr. U. Ebert (Centrum Wiskunde & Informatica) and prof. dr. O. Scholten (Kernfysisch Versneller Instituut, University of Groningen). In this collaboration, we also worked on how to use the cosmic air shower footprints, altered in thunderstorm conditions, to diagnose thundercloud electric fields. And how to use LOFAR as a new type of lightning mapping array. This collaboration resulted in an additional three (+1 submitted publications) in international journals and five proceedings. The full list of publications is given at the end of my Curriculum Vitae.

Chapter 2

History of the lightning initiation problem

This chapter introduces the history of the lightning initiation problem, including scientific jargon as electric breakdown, electron avalanche and streamers.

2.1 The 'problem'

The first occurrence, to my knowledge, of stating a 'problem' of lightning initiation, is by Crabb and Latham [2] in 1974, when they start their introduction by saying:

> *"The problem of lightning initiation within thunderclouds has provoked considerable discussion in recent years. The majority of investigators have favoured the suggestion that lightning is triggered by the emission of positive corona from the surface of a raindrop highly deformed by strong electric fields; so that the surface field at the regions of high curvature will be appreciably magnified over the ambient value."* [2]

Crabb and Latham [2] refer here to several works that propose solutions on increasing the field to values above the ambient thundercloud field. This is needed, because already from the early B25 airplane measurements through active thunderclouds in 1944/1945 by Gunn [3] and later with instrumented rockets by Winn and Moore [4], no ambient fields were measured above 0.4 MV/m, which is well below electric breakdown. Electric breakdown occurs when the electric field surpasses the so-called classical breakdown threshold $E_k \approx 3$ MV/m (at standard temperature and pressure). The breakdown threshold E_k is defined as the field where the number of electron impact ionizations of air molecules equals

the number of electron attachments and, to be precise, the breakdown threshold scales with air density. Also in more sophisticated balloon measurements, from around the year 2000 [1, 5, 6], ambient fields were always lower than 20% of classical breakdown. The lightning initiation problem is thus the problem to start an electric discharge, when the ambient field is far below the breakdown threshold.

The problem is constrained to only the start of the discharge, in the electron avalanche phase where electrons and ions grow exponentially up to the point where they produce a space charge electric field above breakdown and the field enhancement of the dielectric is not necessary anymore. From the point on that the discharge propels itself, by its own space charge, it is called a streamer. Streamers have a characteristic thin charge layer and a strong electric field ahead of the streamer tip. Streamers come in two polarities, *positive* streamers propagate in the direction of the electric field and *negative* streamers propagate against the direction of the electric field. Negative streamers are intuitively easier to understand, as they are a natural transition from the electron avalanche, where the thin charge layer has a net negative charge and *pushes* electrons out. In the positive streamer, the thin charge layer has a net positive charge and *pulls* electrons in, and therefore electrons ahead of the streamer must by supplied by another source, either by a large amount of initial background ionization, or by photoionization. In the case of photoionization, photons are created in the electron avalanches towards the streamer and liberate non-locally new free electrons ahead of the streamer, to repeat the process. Modeling electron avalanches, avalanche-to-streamer transition and streamers are introduced in more depth in section 4.

2.2   Hydrometeors enhance fields

Hydrometeors, which are any water containing meteors in the cloud, liquid or frozen, are dielectric. When dielectrics are embedded inside an ambient field, they will electrically polarize. The created field screens the field inside the dielectric, where it adds destructively to the ambient field, and near some edges it enhances the field, where it adds constructively to the ambient field. Loeb [7] was the first who hypothesized about discharge inception near water droplets, to solve the problem of too low ambient fields. However, in the debate referred to by Crabb and Latham [2], investigators knew that spherical drops only enhance the electric field by a factor less than 3, too little to overcome the gap between measured fields and the breakdown threshold. The enhancement can be much stronger, for example near regions of high curvature, thus hydrometeors must be deformed in more elongated or sharper shapes, was the working hypothesis. According to the investigation in [2], deformation by the ambient field was favored in 1974, but the authors argue that deforming droplets by collisions with

other droplets works better. They conducted a series of experiments where they start preliminary discharges (streamers) from remnants of collided droplet. The water remnants are much more elongated and enhanced the field by significant factors. They found that streamers can start for ambient electric fields ranging from 0.5 down to 0.25 MV/m, for remnant water structures ranging in length from 8 to 20 mm, respectively.

Along a different but equivalent approach, frozen hydrometeors appear naturally in a variety of shapes, including needles, prisms, plates and graupel aggregates. Published in the same journal and volume as Crabb and Latham [2], Griffiths and Latham [8] proposed initiation from ice hydrometeors, claiming that discharges can start in fields as low as 0.4 MV/m (rescaled to STP). However, Griffiths [9] puts a disclaimer on the measurements, as currents of the order of 0.1 μA or larger could not be obtained from the samples if temperature was below -18°C. This restriction was considered to be caused by the electrical properties of ice, the surface conductivity of which is highly temperature dependent. This was a serious issue, as lightning was observed to start at altitudes well above the -18°C temperature level i.e., at lower temperatures. And at those great heights also the colliding water droplet hypothesis would not work, as all hydrometeors would be frozen. After these publications, the problem remained untouched for decades. Yes, there were multiple exotic hypotheses how lightning could start, but all were dis-proved. It took more than 30 years after the measurements of Griffiths and Latham [8], until investigators looked back into streamer initiation from ice hydrometeors, as that hypothesis was actually the only one which made sense. Petersen et al. [10] conducted a large series of streamer initiation measurements at different temperatures from ice hydrometeors, see their results in figure 2.1.

The new measurements of Petersen et al. [10] were in full agreement with the ones of Griffiths and Latham [8], but now they found that lower temperatures can work:

> "Observations of lightning have demonstrated that it can begin at altitudes with environmental temperatures well below −18°C. Previous laboratory studies have cast doubt on the possibility that ice hydrometeors in such cold environments can contribute to lightning initiation, but this doubt appears unnecessary. The results of the present study indicate that simulated ice hydrometeors subject to temperatures as low as −38°C can undergo individual positive streamer discharges, with single crystals undergoing several successive individual discharges if the electric field is continually increased. " [10]

From the measurements of Petersen et al. [10] it was most likely that lightning starts from ice hydrometeors, which can enhance the ambient field to values above breakdown and start streamers even in very cold environments (down to
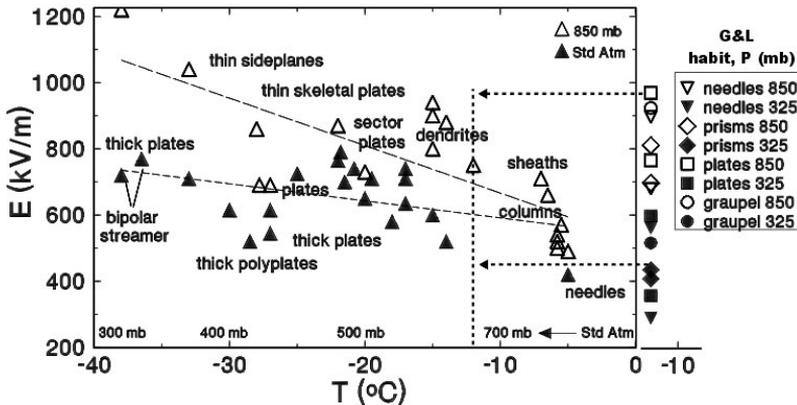
Figure 2.1: Threshold electric fields for single non-repeating positive streamer discharges from ice crystals of various types as function of temperature and pressure. Open triangles are threshold values obtained at a laboratory pressure of 850 hPa and solid triangles are values at standard pressure for temperatures shown. At the right side of zero °C the data of Griffiths and Latham [8] is shown for comparison, which are all obtained at temperature of -12°C. Figure from [10].

$-38°C$). So problem solved? No, actually all this time a critical requirement for starting a discharge was overlooked: seed electrons.

2.3 Electrons are needed

In the experiments of [2, 8, 10] the setups were likely accompanied by a radioactive source to supply initial electrons. Although not written in their manuscripts, by personal communication with Dan Petersen at the American Geophysical Union (AGU) in 2014, I got the confirmation that at least his measurements in [10] where conducted this way. Without an electron (either initially present or created by other means) in the region above breakdown, nothing happens. In the laboratory this is called jitter, before a discharge really starts, there is a random waiting time and sometimes a few unsuccessful attempts to start. To avoid jitter, experimentalists sometimes include artificial sources of electrons, for an easier start of their discharges or they perform their experiments in a repetitive way. In thunderclouds however, free electrons are hard to find. Even though they are naturally created by cosmic rays, the average electron density is very low (less than 1 m^{-3}). This is a very small number, in particular compared to the region above breakdown of a sharp ice hydrometeor, which is of the order of 1 cubic millimeter (10^{-9} m^3).

I show in my thesis that very rare and energetic cosmic rays, from far outside

our solar system, can provide enough free electrons in so-called extensive air showers.

Chapter 3

History of neutron research from thunderclouds and lightning

This chapter introduces the historical origin of neutron research from thunderclouds and lightning, including scientific jargon as leader stepping, runaway electrons and terrestrial gamma-ray flashes. The chapter is an adapted version of 'A natural neutron source' written by C. Rutjes and U. Ebert, published in Physics World Focus on Neutron Science in October 2017.

3.1 Discovery of correlation

The link between thunderstorm and neutron science is not an obvious one. Indeed, it took a Nobel laureate to spot it. By the time he made the connection, Willard Libby was already a highly regarded scientist thanks to his profound work on radiocarbon dating. This method which became a standard tool for archaeologists, and earned Libby the Nobel Prize for Chemistry in 1960 stemmed from the observation that when cosmic rays impinge on the Earth's atmosphere, they produce a shower of particles, including neutrons. These neutrons can react with atmospheric nitrogen to create carbon-14, or radiocarbon ($^1_0$n + $^{14}_7$N $\rightarrow$ $^{14}_6$C + $^1_1$p), which enters the food chain when plants absorb the resulting radioactive carbon dioxide. Libbys insight was to realize that in living organisms, carbon-14 is constantly refreshed together with other carbon isotopes. However, when an animal or plant dies, the ratio of radiocarbon to the stable isotopes carbon-12 and carbon-13 decays with a half-life of 5730 years making it possible to estimate the age of objects made from formerly living matter with a high degree of accuracy.

By 1973, decades had passed since Libby published this Nobel-winning work, and he was entering the autumn of his career. He was still thinking about neutrons, though, and as he was examining tree rings with a colleague, H R Lukens, the pair noticed interesting fluctuations in the amount of radiocarbon in each ring. These fluctuations could not be explained by variations in the cosmic ray flux. Instead, Libby and Lukens found a surprising correlation with thunderstorm activity. The effect was not negligible: they estimated that thunderstorms could account for up to one percent of the neutrons produced in the atmosphere. A few years after Libby's death in 1980, G N Shah and colleagues [11] presented convincing measurements to back up the suggestion that lightning produces neutrons. Shah estimated that between 10 and 100 million neutrons are produced per stroke. But how?

3.2 False starts and potential sources

From laboratory studies, researchers knew that intense electrical discharges through polymer fibers could produce neutrons at 2.45 MeV, probably by deuterium fusion ($_1^2\text{H}+_1^2\text{H} \rightarrow _2^3\text{He}+_0^1\text{n}$). Hence, the first attempts to explain neutron production by lightning focused on fusion. Visible lightning strikes (as shown in the third panel of the figure 3.1) can reach temperatures of up to 30 000 K, and it was assumed that this would, in combination with natural deuterium in water vapor, do the trick. This theory dominated the literature for a long time, even though a number of independent measurements proved otherwise. The problem was that if one excludes deuterium fusion as a neutron source, one has to explain what is providing the energy to release neutrons from the nuclei of typical air molecules. For nitrogen, the binding energy is 10.5 MeV, while for oxygen it is 15.6 MeV.

Hints for a completely different (and correct) explanation came from far outside the laboratory. Since the 1960s, satellites equipped with gamma-ray detectors have monitored compliance with nuclear-test ban treaties here on Earth; later, similar spacecraft were launched that use gamma-ray flashes to study the cosmic realm. In 1993, however, more sensitive detectors recorded gamma-ray flashes coming from Earth that had nothing to do with weapons. These so-called terrestrial gamma-ray flashes (TGFs) are microsecond-to-millisecond-long pulses of photons with up to 40 MeV of energy (see figure 3.2), and in 1996 it was found that TGFs can be related to individual lightning strokes.

Since then, research on high-energy atmospheric physics has gained momentum. Ground-, balloon- and plane-based observations measured fluxes of gamma rays with energies as high as tens of MeV, more than enough to free neutrons from nitrogen. These non-satellite observations also revealed a new phenomenon, much dimmer and of longer duration, that came to be known as gamma-ray glow and that is also accompanied by neutron release. But regardless of whether they
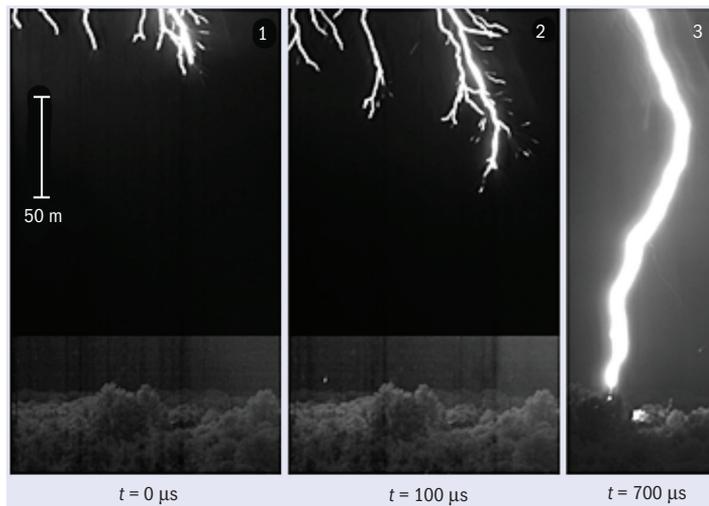
Figure 3.1: These frames show the growth of a lightning leader at 0 and 100 μs (left and middle) and the visible return stroke at 700 μs (right). Each frame had an exposure time of 99 μs. In the middle panel, we can see space stems forming that have no visible connection to the glowing lightning leaders; they will become part of the leader somewhat later. The formation and integration of a space stem into a leader constitutes a step in the lightning-propagation process. Figure from [12].
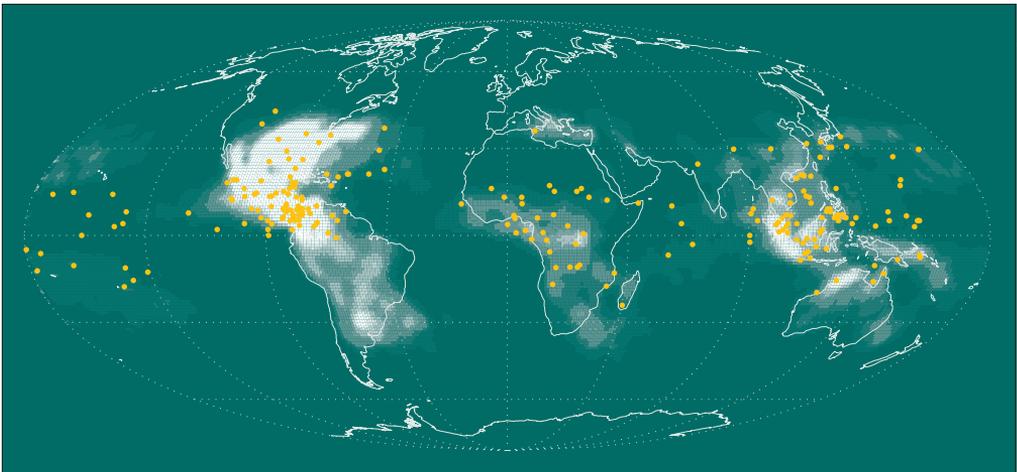
Figure 3.2: Location of terrestrial gamma-ray flashes (TGFs) across the globe shown as yellow dots, based on coincident measurements by the ground-based World Wide Lightning Location Network and the space-based Reuven Ramaty High Energy Solar Spectroscopic Imager. White shading indicates lightning frequency. Figure from [13].

are long and dim or short and intense, these gamma-ray phenomena pushed nuclear fusion off the table. As L P Babich argued very clearly in [14], the only way to create neutrons in a thunderstorm in detectable numbers is by photonuclear reaction of gamma rays with nitrogen, and, to a lesser degree, with oxygen.

3.3   Shifting the problem

So far, so good: we understand that neutrons can be generated in a thunderstorm by gamma rays with energies well above the 10.5 MeV photonuclear threshold of nitrogen. If you look more closely, however, you will see that this explanation merely shifts the problem. Gamma rays in our atmosphere are mostly the result of bremsstrahlung radiation, which occurs when energetic electrons and positrons (collectively known as leptons) collide with air molecules. So how does a thunderstorm generate substantial numbers of leptons with > 10 MeV energies? What role do the electric fields inside the storm play, and which fields at which stage of storm evolution are responsible?

To answer these questions, we need to dive into the current theory of lightning physics. First, we need to understand how free electrons move in a thunderstorm environment. In vacuum, electrons can easily be accelerated by external electric fields; this is how particle accelerators in the laboratory typically work, from Browns tubes up to synchrotron facilities such as DESY. However, in air, electrons also lose energy in collisions with air molecules. Hence, as long as the field does not exceed a threshold of about 0.2 MV/m at standard pressure and temperature, friction from inelastic or ionizing collisions with air molecules balances the acceleration provided by the field, and electrons drift with a field-dependent velocity rather than accelerating continuously (see figure 3.3). On the other hand, if the electric field exceeds a value of about 25 MV/m, friction is always smaller than the acceleration provided by the field, and all electrons enter a run-away mode.

In fact, electric fields in clouds cannot remain at such a high value for long, as classical electric breakdown sets in at a field of 3 MV/m. At that point, free electrons can gain enough kinetic energy from the field to liberate more free electrons when they collide with air molecules, and hence to set off ionization avalanches and eventually to create a plasma that cancels out the external electric field.

So can such a low field nevertheless support relativistic electrons, which have kinetic energy of 500 keV or more? To answer this question, we need to look again at figure 3.3, which shows the friction an electron experiences in air at standard temperature and pressure as a function of the electron energy. This friction reaches a maximum at an electron energy of about 200 eV and then decreases, before increasing again when the electrons attain MeV energies and start radiating substantial amounts of gamma rays by bremsstrahlung. Figure 3.3 also
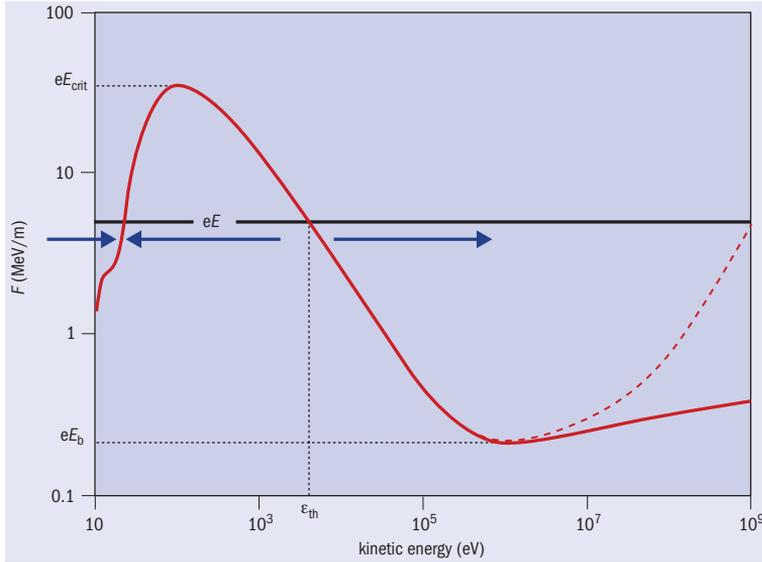
Figure 3.3: Schematic of the friction force on electrons in air at standard temperature and pressure as a function of electron energy, in a double logarithmic plot. The solid red curve is due to inelastic scattering of the electrons with air molecules, while the dashed red curve includes the effects of bremsstrahlung (i.e. X- and gamma-ray) emission. In the presence of an electric field E, electrons also experience an accelerating force eE (black line). If $E > E_{\mathrm{crit}} = 25$ MV/m, all electrons will experience runaway acceleration; if $E < E_{\mathrm{b}} = 0.2$ MV/m, none will. For field strengths between these extremes, outcomes depend on the electron's initial energy: electrons with energies below a threshold ϵ_{th} reach a steady state in the eV range, while those above the threshold keep accelerating, run away and reach MeV energies. Blue arrows indicate how electrons are accelerated or decelerated in such a field.

indicates an acceleration force eE in an electric field E. If this force is larger than the friction, electrons gain energy, and if it is smaller, they lose energy; this is indicated by the blue arrows. For the field indicated in the figure, the electrons clearly fall into two populations: electrons with a lower initial energy approach a steady mean energy in the eV range, while electrons with higher initial energies are accelerated into the runaway regime, reach MeV energies and start to radiate gamma rays. But how do electrons get into this runaway regime within the moderate fields measured inside thunderstorms?

Recall that in the previous section, we mentioned that there are two different thunderstorm phenomena: terrestrial gamma-ray flashes and gamma-ray glows. These phenomena produce neutrons on different time scales, and according to our present understanding, they are associated with different physical mechanisms.

Careful measurements have found that gamma-ray glows actually occur before lightning activity starts, when large volumes of air inside a developing thunderstorm have built up an electric field exceeding the lower runaway threshold of 0.2 MV/m. Cosmic rays shooting into the atmosphere create particle showers, including a continuous flux of relativistic electrons that gain additional energy in this electric field, and form relativistic run-away electron avalanches. This process can continue as long as the field is present and electrons are refreshed, producing a dim gamma-ray glow of long duration. Recent work by Chilingarian and colleagues [15, 16], shows clear evidence of thunderstorm-correlated boosts in the flux of high-energy electrons, gamma rays and neutrons on a timescale of minutes. Remarkably, these fluxes are not correlated with lightning strokes, but are actually competing with lightning as a discharge mechanism. As a result, lightning tends to snuff out gamma-ray glows [17, 18].

Terrestrial gamma-ray flashes, on the other hand, are correlated with lightning leaders. While our eyes just see one lightning flash, high-speed cameras resolve how the lightning channel grows from the cloud to the ground, as shown in Figure 3.1. The growing plasma channel shown in the left and middle panels is called a lightning leader. (It should be noted that leaders propagate for much longer distances inside clouds, in a more horizontal direction, before the visible cloud-to-ground lightning starts. However, during that stage they can only be detected with radio antennas, not by eye or by cameras.) Negatively charged lightning leaders propagate in steps, creating the characteristic zig-zag pattern familiar to anyone who has ever witnessed a big electrical storm. This stepping sometimes occurs through the formation of so-called space stems that initially have no visible connection to the glowing lightning leaders (middle panel of figure 3.1), although they will become part of the leader somewhat later. The formation of a space stem and its integration into a leader constitutes one leader step.

It is now thought that during such a step, the very transient electric field near the leader tip could be high enough to accelerate even thermal electrons into the run-away regime in an intensive explosion − a so-called "cold runaway" − and thereby produce a TGF. But compared to gamma-ray glows, TGFs are much harder to study. In gamma-ray glows, the charged particle densities are so low that they do not change the local electric field, which makes the problem linear: one can simply model the particles developing in a given external (thundercloud) field. For TGFs, the situation is very different. The process of leader stepping is increasingly well measured, but the physical mechanism behind it is not well understood. One thing is clear: it is deeply nonlinear. To propagate, the leader both enhances the ambient electric field at its tip and interacts with its streamer corona, as well as with the mysterious space stems.

3.4 Neutrons as a diagnostic tool

This brings us back to neutrons. As it turns out, the very phenomenon that guided Libby to a connection between tree rings and thunderstorms could actually become a handy probe for building a deeper understanding of lightning leaders and the production of TGFs. The reasons are manifold. One is that neutrons are neutral and therefore do not affect the electric fields of the lightning leaders or the overall thunderstorm. Their presence is also a signature of gamma-rays significantly exceeding the 10.5 MeV threshold of neutron production in air. Neutrons have a much longer lifetime and spread out more isotropically than gamma rays, which makes them much easier to detect. An example is the 2013 study of how neutron bursts at ground level correlate with cloud-to-ground lightning of different polarities. In this work, Toropov and colleagues [19] saw that neutrons are only emitted from negative cloud-to-ground lightning, and not from the positive variety. This is consistent with satellite observations showing that TGFs are produced by negative lightning leaders when they step.

The connection between neutrons and thunderstorms has a long history, but the use of neutrons as a tool to research them is still in its infancy. Even so, it is clear that neutrons give us an interesting window to study the highly intense bursts of energetic radiation from thunderstorms known as TGFs. In the future, they might also help us understand how lightning propagates in steps. Such information could be key to protecting our increasingly vulnerable infrastructure from lightning strikes and similar (potentially destructive) processes that occur in plasma and high-voltage technology.

Chapter 4

Modeling and simulation in multiscale dynamics

This chapter introduces concepts and strategies for modeling and simulating electric discharges and high energy radiation. Scientific concepts and phenomena, such as breakdown and streamers, electron runaway and terrestrial gamma-ray flashes, have already been introduced in chapter 2 and chapter 3, respectively.

Phenomena studied in this thesis evolve on multiple scales, which is important to note as usually theories of physical phenomena are built and valid only in a limited range. In addition, a research field or an individual researcher sometimes favors a certain convention, terminology and/or unit to express the physics. This may be because his or her expertise lies in a confined scale range, but can be confusing and sometimes even inconsistent with others. To illustrate the multiscale nature of the research discussed in this thesis an example per physical quantity is given, with the focus on what this multiscale nature implies for the use of models.

4.1   Energy

Both in lightning and high energy radiation from lightning and thunderclouds, electrons are the important particles, but electron energies vary a lot. At room temperature electrons are around 0.03 eV, while in the radiation from lightning or the thundercloud electrons have energies well above 10^7 eV. More than enough to create gamma-rays by bremsstrahlung, observed as short bursts in so-called terrestrial gamma-ray flashes (TGFs) and of longer duration in so-called gamma-ray glows depending on their correlation with lightning, see also chapter 3. Many

orders higher on the energy spectrum, (inter)galactic cosmic rays can reach energies up to 10^{20} eV, resulting in extensive air showers (EAS) up to billions of secondary particles, discussed in chapter 11.

For bridging 9 to 22 orders of magnitude, physics of multiple disciplines are needed. Low-energy electrons in the discharge are described *classically*, while electrons that accelerate and produce the high-energy radiation must be described *relativistically*, because their energy is above the electron rest mass of 511 keV. Using the terms 'low'- and 'high'-energy might be ill-defined, as for the discharge physicist MeV energy is considered high-energy, but for the cosmic ray physicist MeV energy is considered to be low-energy. Here the definition of a "low-energy electron" would be an electron below 200 eV based on the maximum in the friction curve, see figure 3.3. This low-energy electron classically drifts in an electric field and interacts with shell electrons of molecules. At the far right side, a "high-energy electron" is an electron above 1 MeV, where the friction curve increases again (see figure 3.3). A high-energy electron moves relativistically (close to the speed of light) and interacts with the nuclei of atoms (e.g. bremsstrahlung) and only with the shell electrons as if they where not bound (no molecular structure). Electrons with energy between 200 eV and 1 MeV, they are somewhat intermediate and complex to describe, as they can not be described fully-classical nor with high energy approximations.

4.1.1 Low-energy electron interactions

Air is an insulating gas, meaning that it has a significant electrical resistance, as there are almost no charge carriers available. If however a strong enough voltage is applied, the few electrons which may be around can ionize molecules and free more electrons, which in their turn free other electrons. To understand this process one must consider all electron interactions with molecules. Electron densities in the electron avalanche and streamer phase at standard temperature and pressure (STP) or lower result in air than is *weakly ionized*, meaning that the electron (and ion density) $n_e < 10^{20}$ m^{-3} is small compared to the neutral density $(2.5 \times 10^{25}$ m$^{-3})$. In weakly ionized gases it is fair to assume that electrons only interact with neutrals and thus electron-ion and electron-electron collisions may be neglected. An electron can interact with a neutral molecule M in several ways, for which the important ones are:

- Elastic collision, $e + M \rightarrow e + M$, where the total kinetic energy does not change. Actually, as the molecule is much more heavy than the electron, more than 50,000 times, the electron energy approximately does not change. The momentum however, is largely transferred to the molecule, that is to say the electron scatters almost isotropically away from the molecule.

- Excitation collision, $e + M \rightarrow e + M^*$, where some of the energy of the

electron is transferred to internal energy of the molecule. The excitation energy can be removed by collisional or radiative quenching, or by de-excitation transferred back to the electron.

- Ionization collision, $e + M \rightarrow e + M^+ + e$, where the electron kinetic energy is large enough to overcome the ionization threshold and the remaining energy is distributed over two electrons.

- Attachment collision, $e + M \rightarrow M^-$, where the electron is captured by the molecule. This can only happen for electronegative molecules. Oxygen is an important electronegative molecule in air and can preform attachments in two-ways, two-body or so-called dissociate attachment, $e + O_2 \rightarrow O^- + O$, and three-body attachment, $e + O_2 + M \rightarrow O_2^- + M$, which typically dominates in air at standard temperature and pressure.

All these electron interactions are directly caused by the electron and are thus local interactions. Another important interaction in air is photoionization. Photoionization occurs when electrons excite molecules, which in turn emit photons with enough energy to ionize other molecules. The electron that excites a molecule thus creates a non-local ionization at another location. Usually this is modeled without time delay, which means that photons are approximated as traveling infinitely fast.

4.1.2   High-energy electron interactions and behavior

The energy scale for high-energy electrons can be separated in two regimes, separated roughly at 100 MeV, see figure 4.1. Below roughly 100 MeV (and above 200 eV), electron interaction with bound electrons by impact ionization is most important. For electron energies above 100 MeV nuclear collisions, mainly bremsstrahlung creating photons, is most important, see figure 4.1. These photons in turn can create again electrons and other particles like positrons and neutrons (if they have sufficient energy). In other words, beams of high-energy electrons are typically accompanied by photons, positrons and neutrons which feedback from one to the other. In beams with max 20 MeV particle energy, electrons are most abundant (when counted only above 200 eV), important for TGFs and gamma-ray glows. In beams with higher particle energies, above 20 MeV, photons are must abundant and stabilize to around 80% of the beam population (when counted only above 200 eV), important for energetic TGFs and EASs. The emerging research field of high energy atmospheric physics (HEAP), which includes terrestrial gamma-ray flashes and gamma-ray glows from thunderclouds, operates in the intermediate energy scale (below 100 MeV), which is driven by the electron physics, but usually observed as photons (as photons reach largest distances and then always end up to be the most abundant particle). Cosmic particle physics on the other hand, which includes extensive air showers, operates mainly in the highest energy regime (above 100 MeV).
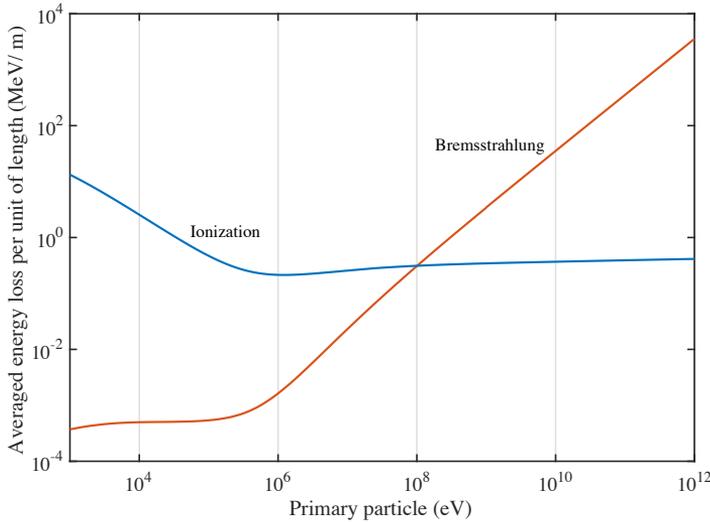
Figure 4.1: Energy loss per unit of length of primary electron in air at STP.

Unequal splitting

In high energy atmospheric physics (HEAP), ionization is the interaction which is most significant to the electrons. Interestingly, the higher the energy of the incoming electron the more unequal is the *expected* splitting of the energy over the two exiting electrons. It remains an intrinsically random interaction, so any energy splitting is possible, but the expectation value is increasing towards a more beamed behavior. Here, the convention is to call the exit electron with the highest energy still the primary, even though the particles are fundamentally indistinguishable. That is to say the primary electron, by convention, will thus not lose more than half of its energy and the secondary gets always less than half of the incoming energy. If one would drop this convention, the factor one-half ($\frac{1}{2}$) appears at a different location in the equation, to reflect the symmetry factors in the interaction.

For illustration of how the unequal splitting creates different behavior, consider ionizing N_2. If a 16 eV electron or 20 eV electron starts the ionization, first the $I = 15.6$ eV must be subtracted, when the remaining energy is divided equally among the two exiting electrons, then the primary has lost a factor 98% or 89% of its energy, respectively. So low-energy electrons lose almost all of their energy. But, if a 1 MeV or 10 MeV electron starts the ionization, the expected secondary energy turns out to be only around 8 eV (actually, for all energies above 100 keV, see figure **??**) and the total energy loss is just 23.6 eV, so the primary just lost a fraction of 2.4×10^{-3} or 2.4×10^{-4} of its energy. In other words, high-energy electrons are likely to undergo interactions that hardly change their momentum, therefore this interaction channel can be approximated
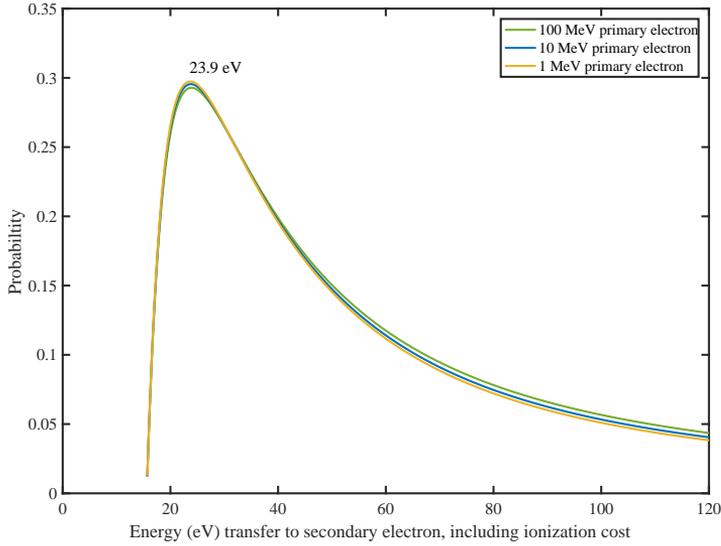
Figure 4.2: Energy transfer of three different primary energies, based on relativistic binary-encounter Bethe RBEB [20] for nitrogen. Most probable energy of the secondary is around 8 eV, after subtracting the ionization cost of 15.6 eV.

as an average energy loss per unit of length, with a so-called *friction* or *stopping-power*. I do not use the term stopping-power, as the term "power" is misleading, as one would expect the unit [energy per unit of time], and not [energy per unit of length] (what it is used for).

Low-energy cutoff

Not mistaken, friction cannot be used for interaction channels which do significantly alter the primary particle, because of so-called straggling. So the averaged energy-loss presented for example in figure 4.1 may give the wrong impression. For example for ionizing N_2, there is still a non-negligible probability that the primary energy is more equally distributed among the two exiting electrons, which would lead to a large energy loss (up to 50%). The separation between the significant and the not-significant channel is usually denoted by the so-called *low-energy cutoff* ϵ_c. For the channel creating a secondary between 0 eV and ϵ_c a continuous approximation is used (by construction of the definition of ϵ_c), and the channel of creating secondaries between ϵ_c to the maximum energy is treated by explicit interactions. Usually ϵ_c must be small (e.g. less than 1%) compared to primary energy minus ionization cost.

Møller scattering

In the case where the low-energy cutoff is much larger than the ionization threshold, an extra approximation becomes available. For example, a 10 MeV electron and low-energy cutoff $\epsilon_c = 10$ keV, which is fair as it is only 0.1% of its energy. In this example, the interaction of producing secondary electrons below 10 keV is averaged out (and treated as friction) and producing electrons above 10 keV must be calculated explicitly as collisions. These collisions are practically free-free elastic interactions. Free-free in the sense that the electron collides with another electron (at rest) and without any influence of the atom or molecule. The free-free elastic interaction can be solved analytically from the Dirac equation, first derived by (C. Møller, 1932) and therefore called Møller scattering.

4.2 Space and density

In space, the interaction length changes with particle energy. Thermal electrons interact and stop in air on micrometer scale (10^{-6} m), while a beam of high-energy electrons (40 MeV) travels more than 120 meters (10^2 m) before they are stopped. In addition to the spatial extent, also the relative density can be very different. Before lightning starts, electron densities are as low as 1 m^{-3}. Inside a streamer, only some 10 ns later, the electron density can be 10^{20} m^{-3}, so as much as 20 orders of magnitude larger. In cases with low density, particles should be described individually, in the so-called *particle-perspective*. Individual electrons propagate and interact in an intrinsically random manner. In other cases, when densities are very high and the fluctuations average out, the particle perspective becomes computationally unfeasible, and it is better to describe the group of particles in the so-called *fluid-perspective*. Of course, the latter is only valid if the fluctuations really average out, that is to say the group of particles is well described by a single average energy, which is not always the case.

4.2.1 Particle perspective

Particles are usually parametrized by their energy ε. Particles move and interact with background medium (here air), with density n_b. The probability of interaction is expressed as a *cross section* as function of energy. Cross sections [in units of area] are defined such that for given density n_b the product σn_b gives the average number of collisions [per unit of length]. In other words, cross sections are the rate of interactions per unit unit of length per unit of density. Then, the expected length between two collisions is given by the *mean free path* λ,

$$\lambda(\varepsilon) = \frac{1}{\sigma(\varepsilon)n_b}. \tag{4.1}$$

Equivalently, given the velocity v of the particle, the *collision frequency* is,

$$\nu(\varepsilon) = v(\varepsilon)\sigma(\varepsilon)n_{\mathrm{b}}, \tag{4.2}$$

with the expected time between collisions $\tau = \nu^{-1}$. The steps (between collisions) are *exponentially distributed*, both in space and in time, with average λ or τ. The combined effect of multiple different collisions is again exponentially distributed, with a mean of $\lambda_{\mathrm{tot}}^{-1} = \sum_i \lambda_i^{-1}$ and $\delta t_{\mathrm{tot}}^{-1} = \sum_i \delta t_i^{-1}$. That is to say, the total cross section is just the sum of the cross sections of all possible interactions $\sigma_{\mathrm{tot}} = \sum_i \sigma_i n_{\mathrm{rel}}(i)$, weighted by the density fraction $n_{\mathrm{rel}}(i)$ for the corresponding interaction.

Space- or time-oriented simulations

In the Monte Carlo method, particles are stepped, from one collisions to the other, using random numbers. When ξ is an uniform random variable between $(0,1]$ the step-length ℓ can be sampled as,

$$\ell = -\log(\xi)\lambda, \tag{4.3}$$

and the time step δt as,

$$\delta t = -\log(\xi)\tau, \tag{4.4}$$

A particle simulation is now either *space-oriented* or *time-oriented*, equation 8.14 or 8.15 respectively. Sampling the step-length (or time-step) between collisions, using using equation 8.14 (or 8.15), assumes that the mean does not change during that step. The mean can change due to a change in particles energy, material composition or material density. To limit the error made in this approximation, null-collisions have to be used.

Null-collisions

The exponential distribution is *memoryless*, meaning that at any moment in space (or time) the probability for interaction remains the same. Or practically, at any moment before a collision takes place, equation 8.14 or 8.15 could be drawn again to get the new step. At the intermediate location the particle does not undergo a real interaction, but a so-called *null-collision*. Both in space-oriented and time-oriented simulations null-collisions can be used, to circumvent errors due to change in particles energy, material composition or material density.

   The exact way to implement null-collisions is to investigate the minimum mean free path λ_{min} (or minimum mean time interval $\tau_{\mathrm{min}} = \nu_{\mathrm{max}}^{-1}$) for any expected energy ϵ. At each sampled location a second random number determines which type of interaction; real (e.g. ionization, elastic, ect.) or null-collision (no collision), based on the probabilities calculated with the current particle energy.

In space-oriented simulations null-collisions are also used to cope with internal discrete boundaries, usually of different materials and/or different densities, and scoring purposes. After each step the distance to the closest (internal) boundary d_{wall} is calculated and steps are only accepted if they are smaller than this distance. If a sampled step-length is larger than d_{wall}, the particle is just moved to the boundary (with there a null-collision) and stored, enabling in the next sampling to use the cross section and density of the other material.

Synchronizability

By construction, space-oriented simulations are not synchronous in time. Usually, a single particle is simulated over its entire life-time before going to the next particle. This procedure makes it impossible to incorporate self-consistent particle interactions, such as a space charge electric field, because than information must by available synchronously. A big advantage of asynchronous simulations is that, besides the ability to include boundaries, particles step much less, minimizing the overhead due to null collisions. Particle codes used in this thesis that are space-oriented are EGS5 [21], FLUKA [22], Geant4 [23] and MC-PEPTITA [24].

In time-oriented simulations null-collisions enable *synchronizability*, where every particle is at the same moment in time. This is enforced by letting a particle step and collide up to the point that the remaining time t_{left} plus the sampled time δt to the next collision is larger than Δt, after which the particle is moved a time $\Delta t - t_{\text{left}}$ (without a collision) and is stored. Before continuing the same particle, first all other particles are moved over the time step Δt.

After the time step Δt all particle data is available and can be used to calculate for example the space charge electric field. For electrical discharge simulations space charge is usually important. The particle code used in this thesis that is time-oriented is DIPIC3D in chapter 6, which uses particle_core from [25].

4.2.2 Fluid perspective

Drift-diffusion-reaction equation

In cases where the particle densities are large enough, it is computationally more efficient to interpret particles as densities. From cross section data one can derive average properties of an ensemble of particles, expressed in the so called transport and reaction data. In first order models, drift $\mu(E)$, diffusion $D(E)$, attachment $\eta(E)$ and ionization $\alpha(E)$ coefficients are used. The coefficients are all a function of the electric field E, within the *local field approximation*. The drift-diffusion-reaction equation for the electron density n_e is given by,

$$\frac{\mathrm{d}n_e}{\mathrm{d}t} = \nabla \cdot (\mu \mathbf{E} n_e + D \nabla n_e) + (\alpha - \eta) \mu E n_e + S_{\text{ph}}, \qquad (4.5)$$

where S_{ph} is the photoionization source. Ions are usually assumed to be immobile on the nanosecond streamer timescale and therefore the equation for the ion density n_i has only the reaction term,

$$\frac{dn_i}{dt} = (\alpha - \eta)\mu E n_e + S_{ph}. \tag{4.6}$$

A discharge simulation is only self-consistent if the electric field created by the charges is also incorporated. In streamer simulations in air at around STP, it is a valid approximation that any magnetic field contribution can be neglected and that the field may act instantaneously (no retarded fields - electrostatic approximation). In that case, the field equation reduces to the Poisson equation for the electric potential,

$$\nabla \cdot (\epsilon \nabla \phi) = q(n_i - n_e), \tag{4.7}$$

where ϵ is the dielectric permittivity, ϕ the electric potential and q the elementary charge. The electric field is,

$$E = -\nabla \phi. \tag{4.8}$$

Streamer code which uses the fluid description in this thesis is DieLightning [26], see chapter 10.

4.3   Time

The initiation of lightning, i.e. streamer inception, lasts only tens of nanoseconds 10^{-8} s. However, in self-consistent discharge models the actual time to update the electric field is relatively frequent (every 10^{-12} s or so), because of the very small dielectric screening time (due to physics) or numerical instabilities (due to the discretization). The computation time is therefore very important to keep in mind when designing calculations. For example in streamer simulations, the bottleneck is usually the calculation of the electric field, when solving the Poisson equation on a certain grid. State-of-the-art fast field solvers, e.g. Afivo [27, 28], take around one second to solve the field in 3D, which result in around 16 minutes per nanosecond and 10 hours to simulate 40 ns of streamer evolution. More common slower field solvers, e.g. in [26, 29, 30], can easily take 15 seconds to one minute to calculate the field in 2D cylindrical symmetry (not even 3D!), which yields for a simpler simulation two weeks to complete. This example also highlights that when simulating beyond streamers, model reduction must be applied. An example of this is streamer tree modeling, as was introduced in [31].

Part II

Model development, reduction and evaluation

Chapter 5

Discharge inception from positive tip

We present a fluid simulation of discharge inception near an ice particle in a thundercloud, that includes the emerging streamer with its space charge effects, and we include the frequency dependent dielectric permittivity $\epsilon(\omega)$ of ice and discuss its effects. The discharge inception occurs for $\alpha d \approx 10$, much lower than the traditional Raether-Meek-criterion. We explain qualitatively and quantitatively why the old Reather-Meek criterion does not apply to positive streamer inception near a dielectric or conducting surface.

5.1   Introduction

To understand the inception of streamer discharges near dielectric and conductive bodies is a fundamental problem in gas discharge physics. It is relevant for atmospheric pressure plasmas during the start-up phase or when they are driven in pulsed mode, for high voltage technology as well as for lightning inception. We concentrate on inception from small protrusions or particles. For example, rain droplets hanging on high voltage electric energy transmission lines can start noisy corona discharges [32], and ice coverage of high voltage insulators can trigger surface flashover [33]. On the other hand, the inception of streamer discharges from droplets and ice particles is now much studied in the context of lightning inception [26, 29, 30, 34, 35].

The general problem can be stated as follows: There is a fixed boundary between solid and gas. An external electric field is applied, and this field is enhanced near a protrusion or corner of the conducting or dielectric solid. Can we predict whether a streamer discharge will start growing in the enhanced field near the solid boundary? In the present chapter, we limit the analysis to positive discharges that typically start easier than negative ones; 'positive' means that negative charges move through the gas towards the solid boundary and positive ones away from it.

The oldest and simplest approach to this problem assumes that a streamer discharge incepts after a critical number of avalanche multiplication lengths α^{-1}, where $\alpha(E)$ is the Townsend impact ionization coefficient in the electric field E; in air and other attaching gases, the parameter α should actually be read as $\alpha_{\mathrm{eff}} = \alpha - \eta$, where $\eta(E)$ is the attachment coefficient. Theoretical [36] and experimental work [37–39] (and later [40, 41]) suggest that a discharge would incept after an avalanche development length d with $\alpha d = 18$ to 20. This is known as the Raether-Meek criterion for discharge inception. But when is it actually applicable?

This chapter contains two main results. First we present a fluid simulation of discharge inception near an ice particle in a thundercloud, that includes the emerging streamer with its space charge effects, and we include the frequency dependent dielectric permittivity $\epsilon(\omega)$ of ice and discuss its effects. The discharge inception occurs for $\alpha d \approx 10$. Second, we explain qualitatively and quantitatively why the old Reather-Meek criterion does not apply and we revisit an earlier analysis [42] to indicate the effect of diffusion.

5.2   Streamer inception from a dielectric object

5.2.1   Implementation of the simulation with permittivity $\epsilon(\omega)$

For our simulations we developed a diffusion-drift-reaction model of streamer discharges in local field approximation with cylindrical symmetry and with pho-
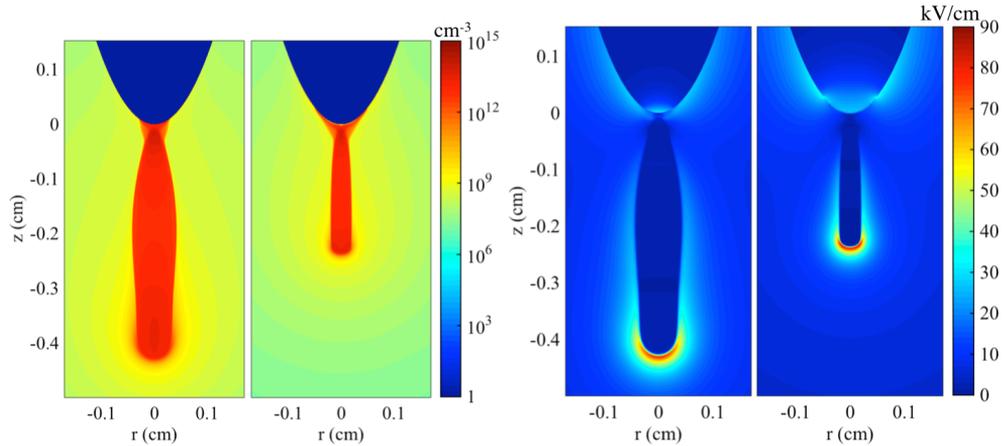
Figure 5.1: Electron density (left two panels) and electric field (right two panels) of a positive streamer after 46 ns simulated with constant dielectric permittivity 93 (first and third panel), and with the dielectric function of ice $\epsilon(\omega)$ [47] (second and fourth panel).

toionization included. The model is described in [43, 44] and it was also used for simulation of lightning inception in its streamer phase from dielectric ice particles [26]. To accurately capture the boundary conditions on the electrode and on curved dielectric surfaces the Ghost Fluid Method was implemented, based on [45, 46]. Neumann boundary conditions are applied in the radial direction and Dirichlet boundary conditions in the axial direction for the electric potential. The Poisson equation is coupled to the transport and reaction equations for electrons and ions, in which Neumann boundary conditions are applied on all the boundaries. The length of the simulation domain is 8.5 cm and its diameter is 4 cm.

Dielectric bodies are polarized by an external electric field according to their dielectric permittivity. When the electric field changes fast enough, the dipoles inside the dielectrics cannot follow these changes and their dielectric permittivity drops, eventually to 1 for electric fields that change infinitely fast. The effect of the frequency dependent dielectric permittivity can be important, for example, for the inception of streamer discharges from ice particles in thunderclouds [26]. In contrast to ice, the dielectric permittivity of water changes on sub-nanosecond time scale which can be important for discharges developing in water in electric fields with sub-nanosecond rise-time [47].

Once a streamer starts developing, the field produced by the charge separation at its head becomes comparable and opposite to the local field, hence it is eventually screened in the interior of the streamer. This dynamics evolves on the time scale of nanoseconds, and the icy dielectric cannot adjust to the changes. Essentially, the dielectric reacts to the field produced by the plasma of
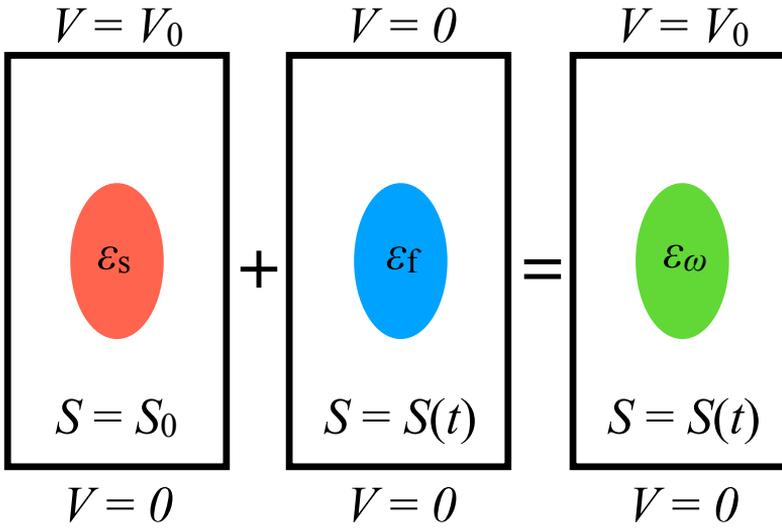
Figure 5.2: The total electric field experienced by the plasma in a discharge is the sum of the fast changing electric field (due the fast changing source $S = S(t)$) and the static background field (due the static source $S = S_0$ and boundary conditions $V = V_0$). When the latter is calculated the dielectric constant is set to ϵ_{s} with s for slow-timescale, and when the former it is set to ϵ_{f} with f for fast-timescale, resulting in a total calculation for varying ϵ_ω with frequency ω.

the streamer with the dielectric permittivity of 3 [48]. On the other hand, the dielectric responds to the static external field with the dielectric permittivity of 93. To implement such a dependence on frequency we consider the total electric field as a superposition of the rapidly changing electric field and the static electric field, which is a valid approximation due to the linearity of the Poisson equation. Each of the fields is to be calculated separately with different dielectric permittivities, as illustrated in figure 5.2, where V is the electric potential, and V_0 is the applied potential that generates the static external field.

In figure 5.1, we compare simulations with the dielectric permittivity of ice set to 93 with simulations with the correct dielectric function $\epsilon(\omega)$ as described above and illustrated in figure 5.2. In both cases we are able to initiate a discharge. In the case of the correct frequency dependent permittivity, the streamer discharge develops with only half of the velocity and stays much thinner. Figure 5.1 also shows that the electric field penetrates deeper into the dielectric, when $\epsilon(\omega)$ of ice is taken into account.

The imaginary part of the dielectric permittivity has to be also, generally speaking, taken into account, as it determines the losses due to the non-zero conductivity of the dielectric. However, for water and ice the imaginary part of the dielectric permittivity is small and even minimal (for ice) for fields changing
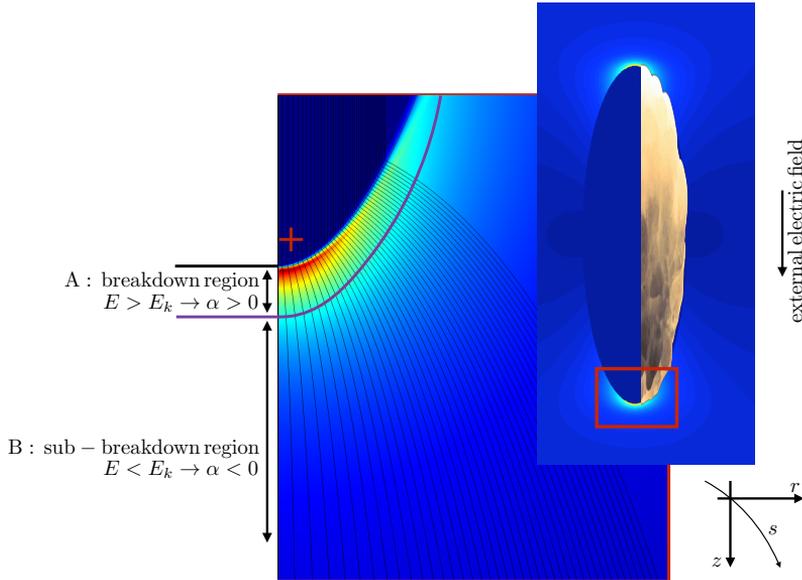
Figure 5.3: Illustration of the electric field enhancement by the example dielectric ellipsoid, for conditions see text. (Figure 5.4 shows the physical quantities, like electric field values, on the symmetry axis of the present plot.) Visible is the non-uniformity of the electric field and the electric field lines converging towards the positive tip. Electrons drift along these field lines towards the positive tip and grow in number where $\alpha > 0$, denoted as region A (the breakdown region). The coordinate system is cylindrical, with z and r pointing away from the dielectric and $z = 0$ at the dielectric tip, and the coordinate s denoting a curved electric field line that could act as a trajectory for an electron.

on the nanosecond timescale [48].

5.2.2   A closer look on how the streamer starts

At the tip of the dielectric ellipsoid the external electric field is strongly enhanced, as illustrated in figure 5.3. Figure 5.4 displays several physical quantities as a function of coordinate s along the electric field line with $r = 0$ (the symmetry axis) on a log-log scale, as the variations are huge. The electric field $E(s)$ is maximal at the positive tip (at $s = 0$), in this example it exceeds more than 7 times the breakdown field. But the electric field $E(s)$ decreases rapidly away from the tip, and thus the region of field enhancement is small, and so is the region above breakdown where $E(s) > E_k$ and $\alpha > 0$. The region above breakdown is denoted as A in figure 5.3 and 5.4. In the sub-breakdown region B, the electric field is such that $\alpha < 0$, therefore we plotted the absolute value of the effective ionization coefficient. As the breakdown region A is small, the

number of electron multiplications of a potential electron avalanche is limited.

We will revisit the derivation of the Reather-Meek criterion in the next section, but for the sake of further explaining figure 5.4 we here already define a few terms. The term αd that defines the electron multiplication in a homogeneous field, is generalized to the ionization integral

$$K(s, s') = \int_{s'}^{s} \alpha(E(\sigma)) \, d\sigma \tag{5.1}$$

in an in-homogenous field $E(s)$. We now analyze the avalanche formation from a single electron and we neglect diffusion; therefore the electron avalanche stays a point object, which is an over-simplified and incorrect representation, but serves for building a rationale in the formal derivations that will follow in the next section.

In this line of reasoning, the ionization integral $K(0, s)$ (equation 5.1) could be interpreted as the expected number of multiplications of an electron avalanche starting at s and drifting to $s' = 0$ (the positive tip). $K(s)$ is maximal at the interface between region A (breakdown) and B (sub-breakdown), where $\alpha = 0$. In this example the maximal ionization integral $K \approx 10$, well below the Raether-Meek criterion of $K > 18 - 21$, as was introduced in the introduction. In other words, in the current approximation there is no individual electron avalanche capable of developing into a negative streamer before it hits the positive tip. The number of electrons in the avalanche if started by a single electron is given by $N_e(s) = \exp K(0, s)$, see figure 5.4, and it never exceeds 10^5 for any initial location s. The maximum is achieved for the starting precisley on the axis and on the edge of the breakdown region A.

An important observation is that in the sub-breakdown region B of figure 5.4, where $\alpha < 0$, $K(0, s)$ can still be significant. In region B even for relative large s up to $s = 5 \times 10^{-3}$ m, $K(0, s) \approx 10$. The reason is that α is far less negative in region B than positive in region A. In sub-breakdown conditions K eventually always becomes negative far away from the dielectric, but in this example K becomes negative outside the region plotted in figure 5.4. Where $K(0, s)$ is negative, the number of electrons in the avalanche $Ne(s)$ is lower on the dielectric surface than at the start point, as attachment is stronger than impact ionization. Another important observation is that the further away the electron starts, the longer it takes before it arrives. The time that an electron (avalanche) drifts from s to s' is approximated by the mobility coefficient $\mu(|E|)$ times the electric field E integrated over the path,

$$T(s, s') = \int_{s'}^{s} \frac{d\sigma}{v(\sigma)}, \quad v(s) = -\mu(E(s)) \, E(s). \tag{5.2}$$

The time $T(0, s)$ from point s to the dielectric surface 0 is also plotted in figure 5.4.
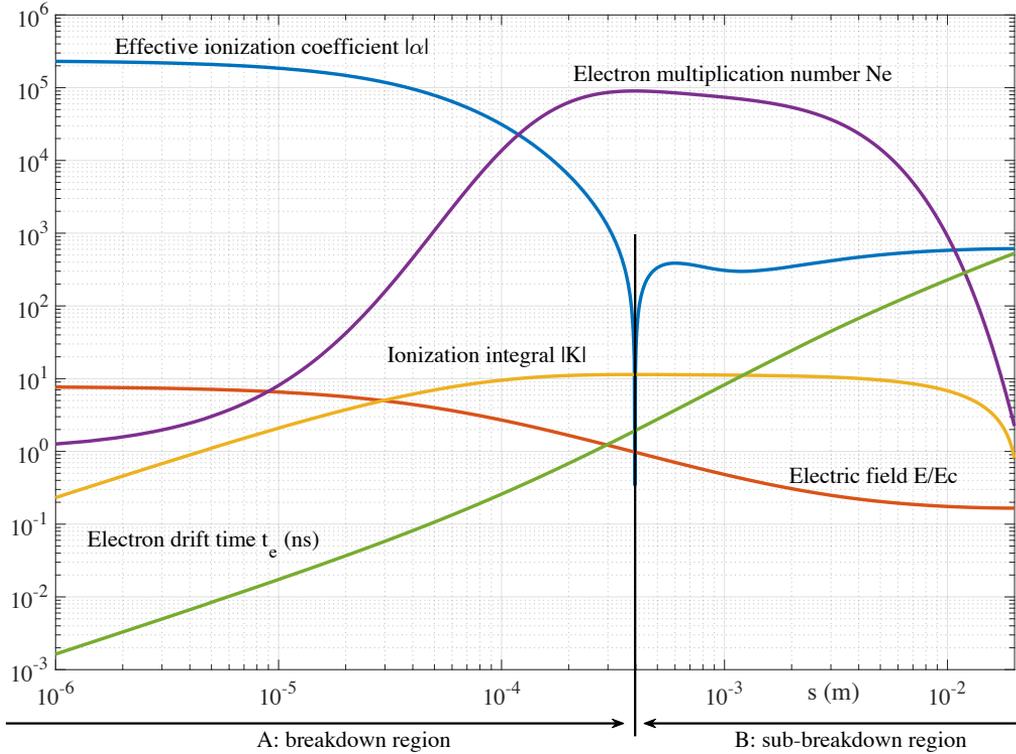
Figure 5.4: Several physical quantities are illustrated on the symmetry axis of the example dielectric ellipsoid (see text for properties) as a function of the position z with $z = 0$ the dielectric surface, on log-log scale. In red the electric field $E(z)$ in units of breakdown field E_c, which is > 1 in region A (the region above breakdown) and < 1 in region B (the sub-breakdown region). In blue the absolute value of the effective ionization coefficient $|\alpha| = |\alpha(E(z))|$, which is positive in region A and negative in region B. In yellow the absolute value of the ionization integral $|K(0, z)| = |\int_z^0 \alpha(E(z'))\mathrm{d}z'|$, K is positive in A and B in the part visible (and will become negative for larger z) and $|K|$ is maximal at the interface between A en B. In purple the exponential of K, $Ne = \exp K$ is drawn, which also maximal at the interface between A en B. Lastly in green the electron drift time $T(0, z)$ (ns), in the figure denoted as t_e, is the time that electrons take to drift from position z to $z = 0$.

5.3   The Reather-Meek criterion revisited, and the role of the dielectric boundary

5.3.1   The fluid model and some approximations

To formalize the ad-hoc analysis above, we start with the classical fluid model in local field approximation that was also used in the simulations. The continuity equation for the electron density n_e is

$$\partial_t n_e + \nabla \cdot (\mathbf{v} n_e - D \nabla n_e) = \alpha |\mathbf{v}| n_e, \quad \mathbf{v} = -\mu \mathbf{E}, \qquad (5.3)$$

where α is the effective ionization coefficient, including attachment. Here we will analyze a model with impact ionization only, without photo-ionization. Below we will discuss different approximations and solutions of the electron equation.

When the electron distribution is known, the ion generation is determined by

$$\partial_t n_+ = \alpha |\mathbf{v}| n_e, \qquad (5.4)$$

where positive and negative ions can be summarized into one number density n_+, as long as they are approximated as immobile. The equation can be integrated in time:

$$n_+(\mathbf{r}, t) - n_+(\mathbf{r}, 0) = \alpha(\mathbf{r}) \, |\mathbf{v}(\mathbf{r})| \int_0^t n_e(\mathbf{r}, \tau) \, d\tau \; . \qquad (5.5)$$

This means that the local ion density is determined by the time integrated local electron flux $|\mathbf{v}| n_e$ times α, i.e. by the total electron density that has passed a particular point within the time interval t, multiplied by α.

During the avalanche phase, space charge is negligible by definition, therefore we analyze the density equations in a time independent background field $\mathbf{E}$. The drift velocity field can be approximated as incompressible

$$\nabla \cdot \mathbf{v} = 0, \quad \mathbf{v} = -\mu \mathbf{E}, \qquad (5.6)$$

when space charges in the avalanche phase are neglected (hence $\nabla \cdot \mathbf{E} = 0$), and when the mobility coefficient μ is approximately constant ($\partial_E \mu \approx 0$) in the relevant field range.

5.3.2   Electron and ion dynamics when diffusion is neglected

In a homogeneous field

When the initial condition consists of a single electron at location $\mathbf{r}_0$

$$n_e(\mathbf{r}, 0) = \delta^3(\mathbf{r} - \mathbf{r}_0), \qquad (5.7)$$

and when diffusion is neglected, the solution in a constant electric field $\mathbf{E}$ directed in the negative x direction is

$$
\begin{aligned}
n_e(\mathbf{r}, t) &= e^{\alpha(x-x_0)}\, \delta(x - x_0 - vt)\, \delta(y - y_0)\, \delta(z - z_0), & (5.8) \\
n_+(\mathbf{r}, t) &= \alpha\, e^{\alpha(x-x_0)}\, \delta(y - y_0)\, \delta(z - z_0) \quad \text{for } x_0 < x < x_0 + vt. & (5.9)
\end{aligned}
$$

Note that the electron and the ion density are clearly different: all electrons are localized at $x = x_0 + vt$, while the ions are distributed over the propagation line $x_0 < x < x_0 + vt$ of the electrons. Of course, the spatial integral is the same,

$$
\int_{-\infty}^{\infty} n_e(x, t)\, dx = e^{\alpha vt} = \int_{-\infty}^{\infty} n_+(x, t)\, dx, \tag{5.10}
$$

as charge is conserved.

    If on the other hand the distribution of electrons would be stationary ($\partial_t n_e = 0$), the solution would be $n_e(x, t) = c\, e^{\alpha x}$ with a constant c, and the ion density would have the same shape, but linearly increasing in time $n_+(x, t) = \alpha vt\, c\, e^{\alpha x}$, as long as their mobility can be neglected.

In an inhomogeneous field

Let us now analyze the electron and ion density in a three dimensional configuration where the electric field changes in space. To focus on the production and transport of charges in the field, we use two approximations: We still neglect the electron diffusion, so that the continuity equation for the electrons is

$$
\partial_t n_e + \nabla \cdot (\mathbf{v} n_e) = \alpha v n_e, \tag{5.11}
$$

and we assume that the electron mobility depends only in a negligible manner on the local electric field; therefore the flow field $\mathbf{v}$ is incompressible (5.6).

    The electrons are convecting in this incompressible flow and they are multiplying due to impact ionization α. If there is no ionization reaction, and if the electron density is homogeneous initially, it will stay homogeneous. This is a direct consequence of the incompressibility: where the field lines converge, the velocity will increase such as to keep the electron density unchanged. However, the ion production will increase where field lines converge. This can be seen from the following calculation.

    We introduce curvilinear coordinates with coordinate s in flow direction and transverse area $A(s)$ of the flux tubes; the incompressibility of the flow (5.6) is then expressed by

$$
A(s)\, v(s) = \text{const.} \tag{5.12}
$$

Integrating over small volumes of flux tubes with transverse area A and using (5.12), one finds

$$
\begin{aligned}
& \partial_t(A n_e) + \partial_s(A v n_e) = \alpha A v n_e & (5.13) \\
\Rightarrow\ & \partial_t n_e + v \partial_s n_e = \alpha v n_e & (5.14) \\
\Rightarrow\ & n_e(s, t) = g\, e^{\int_{s_0}^{s} \alpha\, ds'}, \quad \partial_t g + v(s) \partial_s g = 0. & (5.15)
\end{aligned}
$$

The convective or material derivative $D_t = \partial_t + v(s)\partial_s$ is well known from hydrodynamics, and the equation for g can be integrated. The final result for an initial electron density $n(s_0, 0)$ can be expressed by

$$n_e(s, t) = n_e(s_0, 0)\, e^{K(s, s_0)} \quad \text{for } t = T(s, s_0),\tag{5.16}$$

where we used the previously introduced definition of the ionization integral (5.1) and of the drift time (5.2) that an electron needs to drift from s_0 to s.

The result can be interpreted as follows: an initial electron density $n(s_0, 0)$ on a flux line parameterized by s_0 is transported along the flux line towards s within the time $T(s, s_0)$, and the electron number increases by the ionization factor $\exp K(s, s_0)$, independently of the convergence of flux lines.

For the ion density we get

$$n_+(s, t) - n_+(s, 0) = \alpha(s)\, v(s) \int_0^t n(s, \tau)\, d\tau\ .\tag{5.17}$$

In contrast to the electrons, the generation rate of the ions does depend on the velocity field: where field lines converge, the electron drift velocity v is larger, and hence the ion generation rate αv increases, because both factors α and v increase.

When the electron density in the gas ahead of the surface is stationary, the ion density in the gas increases linearly in time.

The integral over time in (5.17) can be substituted by an integral over the flux line in space. Then for an arbitrary electron density $n(s_0, 0)$ at time 0, the accumulated ion density at coordinate s is

$$n_+(s, t) = \alpha(s) \int_{s_0}^s ds'\, \frac{v(s)}{v(s')}\, e^{K(s, s')}\, n_e(s', 0), \quad \text{where } t = T(s, s_0).\tag{5.18}$$

The spatial integration range from s_0 to s is limited by time t through $t = T(s, s_0)$. And $v(s)$ and $A(s)$ are related through $v(s)A(s) = \text{const.}$ (5.12), as long as the electron flow $\mathbf{v}$ is incompressible.

5.3.3   Transition into a streamer must include diffusion

The classical Raether-Meek criterion answers the question: How far does an electron avalanche need to travel through an ionizable medium until it develops some significant space charge and transits into a negative streamer? The main contribution to the self-consistent field generation in this case is due to the electrons as they are more concentrated in space than the ions. However, when diffusion is neglected, no reasonable estimate of the field enhancement can be made, as the electrons are too localized. E.g., if the avalanche starts at $t = 0$ with a single electron at location s_0, then after time $t = T(s, s_0)$ there are $e^K(s, s_0)$ electrons located in the point s.

For a realistic approximation of the space charge effects during an avalanche to streamer transition, electron diffusion should be taken into account. Here we recall the derivation of the Raether-Meek criterion in a homogeneous field, including electron diffusion, as discussed in [42].

When the field $\mathbf{E}$ is constant, drift velocity $\mathbf{v}$, diffusion D and ionization rate α are constant. Equation (5.3) with initial condition (5.7) can then be solved explicitly as,

$$n_e(\mathbf{r}, t) = e^{\alpha|\mathbf{v}|t} \; \frac{e^{-(\mathbf{r}-\mathbf{r}_0-\mathbf{v}t)^2/(4Dt)}}{(4\pi Dt)^{3/2}}, \tag{5.19}$$

see discussion of [42]. The physical interpretation is as follows: The single initial electron creates an avalanche where the total particle number grows as $e^{\alpha|\mathbf{v}|t}$, where the center of the avalanche moves with velocity $\mathbf{v}$, and where the avalanche width grows due to diffusion with characteristic length $\sqrt{4Dt}$. This means that a fixed density level n in the leading edge $(+)$ or in the tail $(-)$ of the distribution moves with velocity,

$$\mathbf{v}_\pm = \mathbf{v}\left(1 \pm \sqrt{4D\alpha/|\mathbf{v}|}\right), \quad \text{or for } v = |\mathbf{v}|: \quad v_\pm = v \pm 2\sqrt{D\alpha v}. \tag{5.20}$$

These different velocities are derived in [49]. The ion density generated within time Δt is explicitly, using (5.5) and (5.19):

$$n_+(\mathbf{r}, t) = \int_0^t \alpha|\mathbf{v}| \, d\tau \, e^{\alpha|\mathbf{v}|\tau} \; \frac{e^{-(\mathbf{r}-\mathbf{r}_0-\mathbf{v}\tau)^2/(4D\tau)}}{(4\pi D\tau)^{3/2}}. \tag{5.21}$$

This equation cannot be integrated analytically, but arbitrary moments of z^n in propagation direction or of r^n in the transversal direction can be calculated; they are given and discussed in [42].

The avalanche transits into a streamer when the field created by the avalanche becomes comparable to the applied electric field. Here the diffusion is vital, because otherwise the electric field would stay unbounded. Simulations in [42] showed that the self generated field becomes important when it is about 3% of the background field, and the transition from avalanche to streamer is illustrated in great detail in Fig. 1 of [42].

The field generated by the electron avalanche is maximal at approximately the radius of $\sqrt{4D\Delta t}$ where it has the value

$$E_{max}^{electrons}(t) = 0.4276 \; \frac{e}{\epsilon_0} \frac{e^{\alpha|\mathbf{v}|t}}{16\pi Dt}. \tag{5.22}$$

As recalled above, the field generated by the ions can be approximated as well. In summary, analysis of [42] yields the following results:

- The Meek number αd depends on the electric field and on the diffusion constant. For STP air (with the transport and reaction coefficients available in 2006 [42]), the Meek number just above the breakdown field is

about 21 and drops to 16 for increasing field, see Fig. 6 in [42]. For STP nitrogen, i.e., without attachment, the value of αd for the avalanche to streamer transition diverges for very small fields, and has values between 12 and 21 for fields above 40 kV/cm, depending on the field and on the value of $D\alpha/v$, see Fig. 3 in the article.

- To use only the electric field created by the electrons (5.22) is a pretty good approximation, see Fig. 3 (symbols for simulation, thin solid line for full field and dashed line for field of the electrons only, all for dimensionless $D = 0.1$).

- The Meek number depends logarithmically on the initial electron density, on the gas density and on the electron diffusion constant.

5.4   Conclusion

For the formation of a positive streamer near a conducting or dielectric surface, it is not the electrons that count, but the ions. The electrons pass through the high field region and flow towards the surface. If the surface is conducting or has a high dielectric permittivity, the electron charge is completely or largely screened. It is the accumulation of ions in the region ahead of the surface that will provide the space charge effects needed to launch a positive streamer in the direction away from the surface. We also saw that when the correct frequency dependent permittivity is used, the streamer discharge develops with much lower velocity and stays much thinner.

   The formation of a negative streamer in the gas far from any surfaces depends on the space charge effect of the electrons (5.16) and is known as the Raether-Meek criterion, while the formation of a positive streamer near a surface depends on the space charge effect of the ions (5.17). There is no reason why these two criteria should be identical, or in other words: There is no reason why the Raether-Meek-criterion should be applicable to positive streamer inception near a dielectric or conducting surface.

Chapter 6

Discharge Particle in Cell 3D

Prior to streamer formation, when the numbers of charged particles in the relevant region are still relatively low, a fluid discharge model is not appropriate. We have developed a particle model to study the initial phase, specifically the stochastic nature of positive streamer inception near dielectrics and other surfaces. The model is motivated by lightning initiation near ice particles in thunderclouds, but can be applied more generally to jitter in discharge inception. The model is designed such as to easily continue with 3D streamer simulations in the Afivo framework. This would enables the first 3D streamer simulations that start with realistic initial distributions of electrons and ions.

6.1 Problem setting

In thunderstorms, positive streamers are the precursors for positive lightning leaders. Positive streamers can be initiated from hydrometeors such as water droplets, graupel or ice needles, which locally enhance the thundercloud electric field above the critical field. For a discharge to start, some initial electrons are required, which in a humid cloud could be provided by extensive air showers [26]. Typically, streamers are modeled with a deterministic fluid model (i.e. drift-diffusion-reaction coupled with Poisson), often in 2.5D (axisymmetric) [26, 29, 30, 34, 35], although 3D simulation have also been demonstrated [27, 28].

The Raether-Meek criterion [36–39] is typically used to determine whether a discharge can start. The criterion states that the integral

$$K = \int \alpha_{\text{eff}}(E(s)) \mathrm{d}s, \tag{6.1}$$

of the effective ionization coefficient $\alpha_{\text{eff}} = \alpha(E) - \eta(E)$ should be larger than some empirical value M_0, typically between 18 and 21. An electron avalanche can then grow to about $\exp(M_0) > 10^8 - 10^9$ electrons, after which space charge effects become important and the avalanche transit into a streamer.

Avalanche to positive streamer transition, from dielectric surfaces (hydrometeors) works differently than avalanche to negative streamer transition, see chapter 5.

As discussed in chapter 5, the Raether-Meek criterion does not directly apply to the avalanche to positive streamer transition near dielectric surfaces, as individual electron avalanches that move towards the surface may not be sufficient to transform into a streamer (not satisfying the Raether-Meek criterion) but a discharge can still start. The discharge can start because in this configuration multiple avalanches can accumulate positive space charge, either by enough preionization or photoionization. Space charge accumulates in the form of immobile positive ions in front of the surface and negative surface charge of electrons that hit the surface. These positive ions, because of their relatively large mass, are assumed to be immobile for some tens of nanoseconds. However, the unavoidable ion drift imposes that the growth (accumulation) of space charge must be fast enough such that ions do not drift away.

Chapter 5 discusses a time-dependent revised criterion for the case of preionization, without photoionization, According to this criterion, streamer can start in much lower electric fields or from smaller hydrometeors, without requiring that somewhere $K > M_0$. But a smaller hydrometer means a smaller region of local field enhancement with values above breakdown. For typical hydrometeors this region is at most of the order of a cubic millimeter. Hence only individual electron avalanches with their intrinsically random nature are entering the breakdown area sequentially. On these scales, a deterministic fluid description is actually not valid. Therefore, we have developed a stochastic particle model to

study the behavior of the system described above, to calculate the probability of streamer inception for a given hydrometeor, electric field and initial electron density.

That hydrometeors can start with ionization integral values much less than the Raether-Meek criterion, $K \ll M_0$, was already stated in [26], were it was demonstrated that for ellipsoidal shapes $K \approx 10$ can suffice. In addition, with $K \approx 10$ and an ellipsoidal hydrometeor, there is a unique and optimal relation between the size and sharpness of a hydrometeor as function of thundercloud electric field [26]. For a subbreakdown field of 15% or less a hydrometeor of length around 6 cm or larger is needed [26]. Babich et al. [30, 35] state however that the 6 cm example simulated in [26] is extremely large and not expected. They [30, 35] demonstrate that the influence of charge on the hydrometeor is significant and can soften the relation between hydrometer size and sharpness and thundercloud electric field. In other words Babich et al. [30, 35] state that with the charge amounts expected on the hydrometeor, streamers can start from much smaller hydrometeors for the same sub breakdown thundercloud electric fields. However, they derive the amount of charge needed to start a streamer from the requirement that $K \geq M_0$ and take for $M_0 = 20$ which is, as already explained above, a much higher bound than needed for a positive streamer inception from a dielectric.

6.2 Methodology

A particle in cell simulation uses discrete particles to simulate the microphysical interactions and uses cells (a grid) to calculate the non-local interactions, here due to the electric field. The Discharge Particle in Cell 3D (DIPIC3D) code uses `particle_core` library [25, 50] and `Afivo` library [27, 28] and is built to incorporate dielectrics. In this technical section, we explain the connection between the two libraries and the incorporation of dielectrics on a programming level.

6.2.1 DIPIC3D

DIPIC3D is built to run in parallel on shared memory with OpenMP. For parallel implementation on the particle side, multiple object copies of type `PC_t`, denoted here as `pc`, from the `particle_core` library exist and are updated in parallel. `pc` contains a particle database with their kinetic information and routines for microphysical interactions with corresponding cross sections (particle behavior). On the mesh side there is only one object, denoted here as `td`, of type `a3_t` from the `Afivo` library which contains all information about the grid, densities and fields. Routines acting on `td` defined in `Afivo` library are executed in parallel. For the sake of illustration, we omit this parallel behavior in the future explanation.

In figure 6.1 an illustration of the information flow is given. In the center in blue, `particle_core` defines the `PC_t` object type, which includes all member variables and implementations of the particle behavior (functions and routines)

to simulate the movement of discrete electrons and their interactions with the ambient gas. The main particle routine to invoke the particle evolution is,

```
call pc%advance(dt) ! advances all particles in pc a time step dt
```

where all particles are advanced (move and collide) over a time step `dt`. Besides the regular set-up configuration, the `PC_t` object `pc` acquires information about the domain (`pc%outside_check`), the dielectric shapes (`pc%inside_check`) and particle acceleration function (`pc%accel_function`) due to the electric field calculated in DIPIC3D; DIPIC3D is illustrated in red in figure 6.1. `Afivo`, illustrated in green in figure 6.1 defines the `a3_t` object type (denoted here as `td`) which is the mesh data structure and the `mg_t` object type (denoted here as `mg`) which includes the multi-grid solver settings. `Afivo` further provides utility routines for operations acting on the objects type `a3_t` and `mg_t`. Notably, after every ionization or attachement in `PC_t` object `pc`, a callback is initiated in DIPIC3D, which uses the Afivo `a3_interp0_to_grid` routine to add ions to the mesh,

```
!> record the number of positive ions created
!> executed after each ionization
subroutine ionization_add_pos_ion(pc, my_part, c_ix, c_type)
  class(PC_t), intent(inout)  :: pc
  type(PC_part_t), intent(in) :: my_part
  integer, intent(in)         :: c_ix, c_type
  if (check_in_domain(my_part%x)) then
  call a3_interp0_to_grid(td, my_part%x, i_ion, my_part%w, .true.)
  end if
end subroutine ionization_add_pos_ion

!> record the number of negative ions created
!> executed after each attachment
subroutine attachment_add_neg_ion(pc, my_part, c_ix, c_type)
  class(PC_t), intent(inout)  :: pc
  type(PC_part_t), intent(in) :: my_part
  integer, intent(in)         :: c_ix, c_type
  if (check_in_domain(my_part%x)) then
  call a3_interp0_to_grid(td, my_part%x, i_ion, -1.0d0*my_part%w, .true.)
  end if
end subroutine attachment_add_neg_ion
```

which is a zeroth order projection. The extra check `check_in_domain(my_part%x))`, enables the possible costum user setting that the particle domain is larger than the gridded domain. In addition, after each `pc%advance(dt)` all electrons are reprojected to the mesh. Reprojected means that the electron density of previous timestep `dt` is overwritten. The external and internal boundaries, with logical functions (outside and inside checks), are controlled by DIPIC3D based on the user settings and chosen dielectric objects, illustrated in red, see figure 6.1.
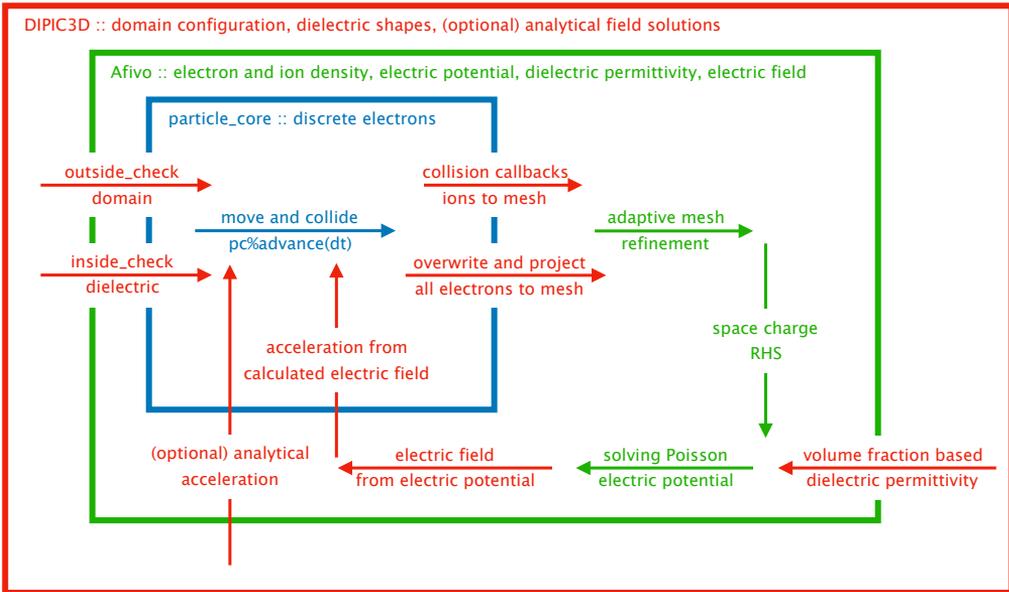
Figure 6.1: DIPIC3D information flow chart illustrating of the combination of particle_core code and afivo, extended to incorporate dielectrics

Currently, there is the option to include a dielectric sphere, rectangular cuboid, ellipsoid, cylinder and cylinder with single or double tip (pencil shape). In addition the user can provide a custom dielectric shape. A dielectric shape definition is just an analytical logical function which provides positional true or false for inside or outside the dielectric shape, plus a dielectric permittivity. Currently all dielectric shapes in the domain must have the same dielectric permittivity.

6.3   Preliminary results

This chapter will be extended for publication, in this section we present the current preliminary results.

6.3.1   Probability of streamer inception

Ice needle

For the background electric field of 16% of classical breakdown (0.5 MV/m at STP) we have simulated inception from an ice needle, here calculated without solving for space charge and using only the static dielectric permittivity of ice. The ellipsoidally shaped ice needle is 16 by 1 mm and has no charge. A realistic initial electron seed of only 100 cm^{-3} is used to highlight the fluctuations in

streamer inception. The electric field is directed in the negative z-direction, such that electrons drift upwards: at the bottom tip towards the dielectric and at the top away from the dielectric surface. The field is enhanced only very locally and thus electrons have to drift a significant path length to reach the breakdown volume. Along their path they have the probability to undergo an attachment reaction, which practically makes the charge immobile as the heavy ions do not move in the simulation. Only those electrons that successfully approach the breakdown region can start electron avalanches. The highest multiplication is precisely on the symmetry axis. Integrating the effective ionization coefficient in the region above breakdown on this axis gives

$$\int \alpha(E(z)) - \eta(E(z))\,\mathrm{d}z = 9, \tag{6.2}$$

which implies that the highest multiplication is around $\exp 9 \approx 10^4$. This single avalanche number is much too small to start a streamer.

Figures 6.2, 6.3 and 6.4 show the evolution of a random example. In figure 6.2 the first 0, 4 and 8 ns evolution is given, where we observe the electron drift. In figure 6.3 we see that in this example between 13.1 ns and 13.15 ns the closest electron got attached and no longer drifts towards the dielectric. For this example it takes up to 19.25 ns until the second electron reaches the breakdown region and starts an electron avalanche, see figure 6.4. In this electron avalanche new electrons are created at a distance by photoionization, visible by the spherical cloud of dots around the avalanche. These potentially start again electron avalanches and together accumulate positive ions and negative surface charge, i.e. space charge to initiate a streamer.

In figure 6.5 we present 100 repetitions of the previously discussed ice needle example. We stopped the simulation either when 10^6 electrons are reached or after 50 ns. We observe that 58% of them exceed 10^6 electrons within the 50 ns simulation time. For those that do not start, we see an exponential decrease in the number of electrons due to attachment. The reason is that most of the domain has an approximately uniform electric field, as the dielectric enhancement is only very local and thus the attachment rate is almost constant. For those examples that start multiplying we see large fluctuations as most electrons reach the tip on a non-optimal path, i.e., off symmetry-axis. From equation 6.2 we know that the theoretical maximum multiplication is around 10^4 and thus reaching 10^6 implies multi avalanche accumulation.

Charged droplet

In this example we calculated the inception probability for charged spherical droplets without an external electric field, for three different radii of 0.5, 1.0 and 1.5 mm at STP. We started with an initial electron density of 1000 cm^{-3} and varied the charge on the droplet between 10 pC to 1300 pC. Inception was
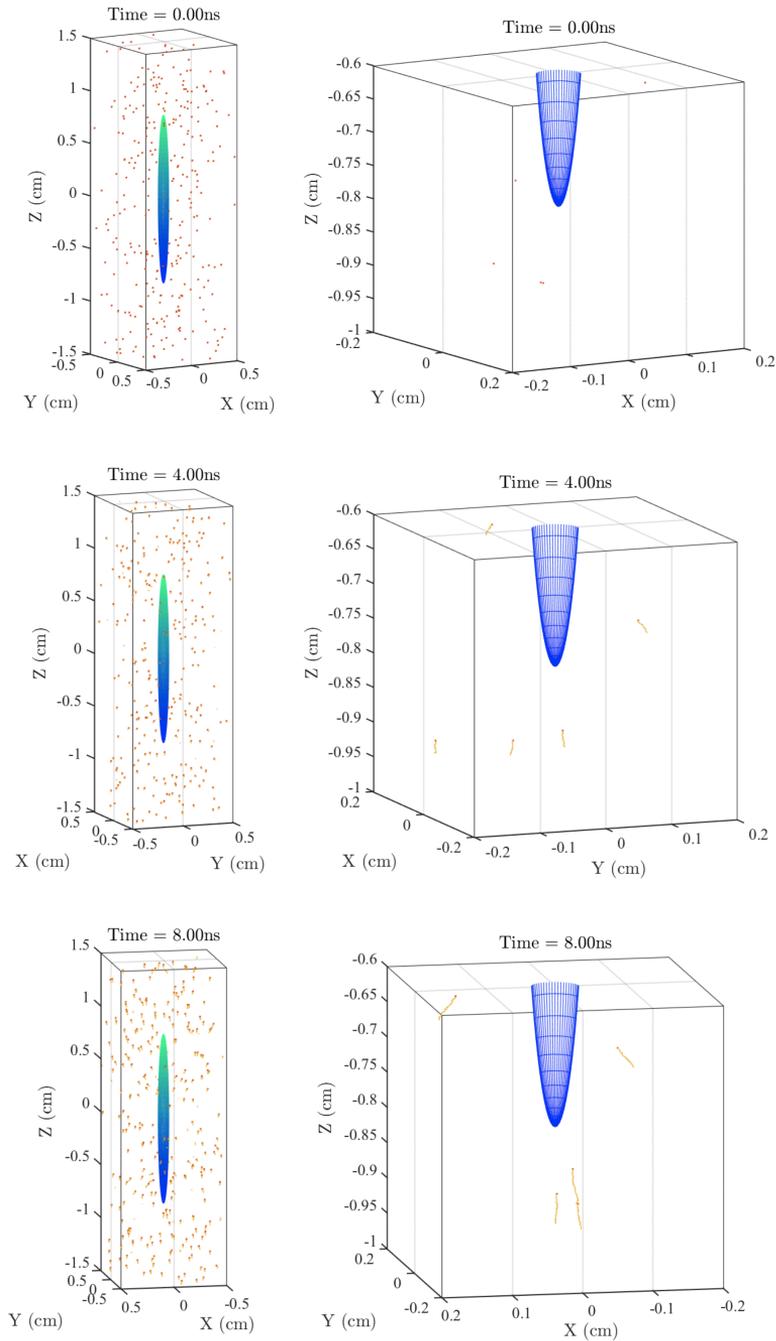
Figure 6.2: Evolution of electrons in the electric field enhanced by the dielectric ellipsoid for three time steps, 0 ns, 4 ns and 8 ns. Electrons are printed as red dots and their trajectory as yellow trails (just for visualization). On the left side the full domain is visible and on the right side a zoomed version of the same simulation.
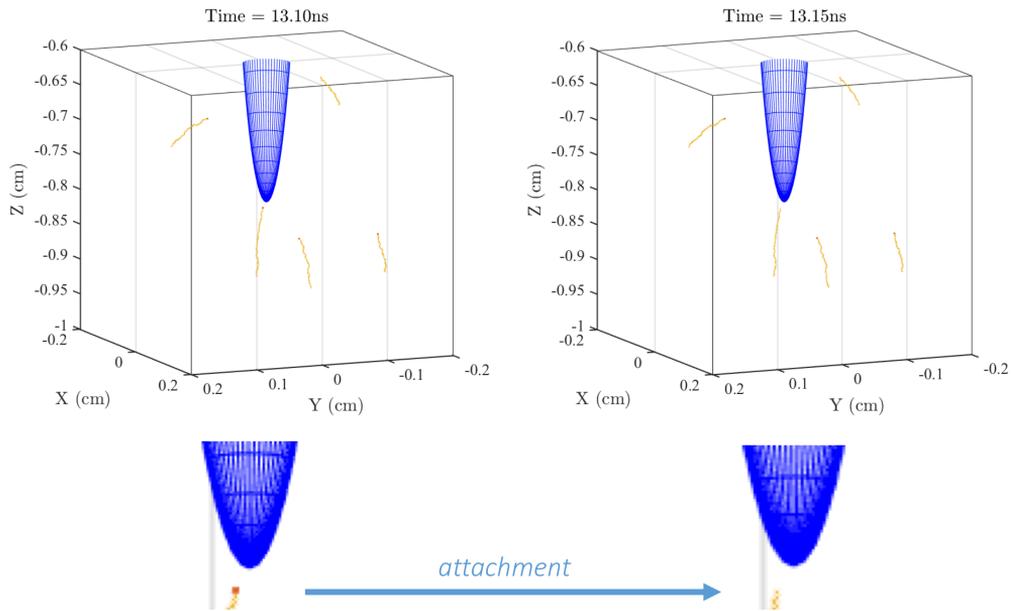
Figure 6.3: A detail of the example also shown in figure 6.2, showing that the first electron undergoes an attachment interaction between 13.1 ns and 13.15 ns.
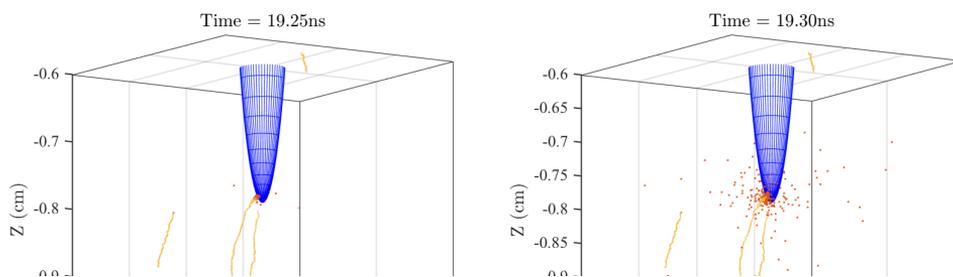
Figure 6.4: At 19.25 ns the second electron reaches the breakdown region and starts an electron avalanche. The electron avalanche creates new electrons away from the tip by photoionization, which can start new electron avalanches.
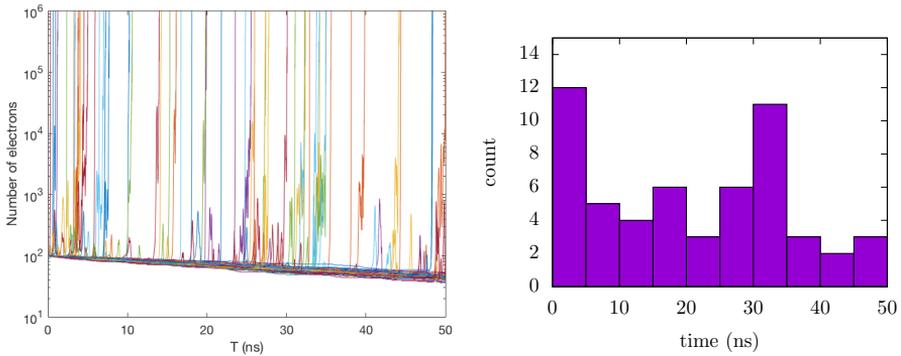
Figure 6.5: Left: electron number as function of time for one hundred repetitions of the example illustrated in figure 6.2 to 6.4. 58% of the calculations exceed 10^6 electrons within 50 ns. Right: histogram of the times at which avalanches reached a size of 10^6 electrons.

counted when the space charge electric field of the electrons and ions creates a total electric field which is twice of the initial value. This implies that the space charge electric field was approximately aligned and equal or larger than the initial electric field. If the system did not go into inception it was stopped after 50 ns. Figure 6.6 gives the percentage of simulations (out of 200) which counted as inception before reaching 50 ns.

We observe trivially in figure 6.6 that the larger the charge the higher is the probability to start, but we observe also that the larger the radii, the broader is the probability distribution. Our stochastic simulations show that discharges can start for considerably lower charges. The deterministic threshold for the minimum charge on droplets needed is calculated in [51]. They found 300 pC, 1000 pC and 2000 pC for the radii 0.5, 1.0 and 1.5 mm, respectively.

6.4 Conclusions and outlook

We identified that for realistic initial conditions as low as $\leq 10^3$ cm^{-3} a stochastic approach to streamer initiation near ice particles and similar dielectrics is needed. We combined and extended available codes to build a Discharge Inception Particle In Cell 3D (DIPIC3D) code that can include dielectrics. We see that discharges start with great jitter and potentially off symmetry axis. Further development of DIPIC3D is needed, in particular an update of automatic selection of streamer inception (to scan configurations) and a validation with experiments is desired.
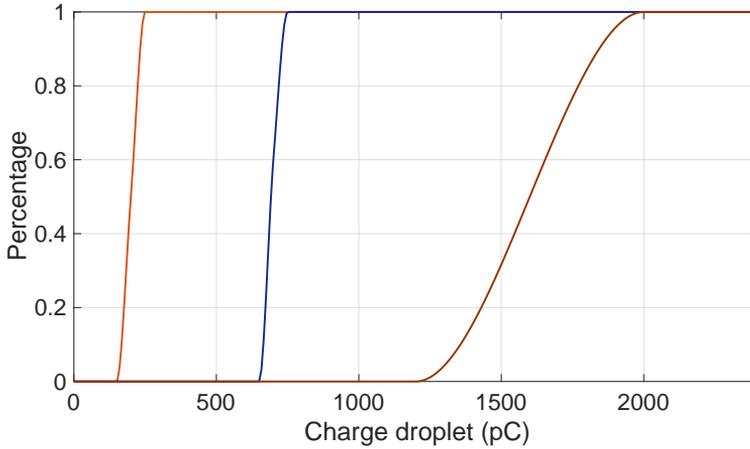
Figure 6.6: Percentage of simulations that created an electric field enhancing the initial electric field on the droplet surface by a factor of 2, within 50 ns. The initial field is created by the indicated charge on a droplet of radius, from left to right: 0.5, 1.0 or 1.5 mm.

6.A   Volume fraction based epsilon

In the case that a dielectric boundary goes through a cell, we use a volume fraction based epsilon, following [52]. The fractional epsilon ϵ_f is given by,

$$\epsilon_f = \frac{\epsilon_{\text{in}}\epsilon_{\text{out}}}{\epsilon_{\text{out}}f + \epsilon_{\text{in}}(f-1)}, \tag{6.3}$$

where f is the fraction of the volume inside the dielectric with epsilon ϵ_{in} embedded in the ambient medium with epsilon ϵ_{out}. The routine executed on grid cells is given by,

```fortran
!> routine to map epsilon to the grid
!> rarg = [tree label, volume fraction option, epsilon inside, epsilon outside]
subroutine set_eps_box(box,rarg)
  implicit none
  type(box3_t), intent(inout) :: box
  real(dp), intent(in)        :: rarg(:)
  real(dp) :: f
  integer :: i_eps
  integer :: i,j,k, nc

  nc        = box%n_cell
  i_eps = int(rarg(1)) ! tree label
```

```fortran
  do k = 0, nc+1
    do j = 0, nc+1
      do i = 0, nc+1
        if (rarg(2)<0) then ! no volume fraction chosen
          if (dielectric_check(a3_r_cc(box,[i,j,k]))) then
            box%cc(i, j, k, i_eps) = rarg(3) ! epsilon inside
          else
            box%cc(i, j, k, i_eps) = rarg(4) ! epsilon outside
          end if
        else ! volume fraction for epsilon
          f = vol_frac_inside(box,[i,j,k])
          box%cc(i, j, k, i_eps) = rarg(3)*rarg(4) / &
          (rarg(4)*f + (1-f)*rarg(3))
        end if
      end do
    end do
  end do
end subroutine set_eps_box
```

where `vol_frac_inside(box,cc_ix)` uses a Monte Carlo volume integration routine,

```fortran
real(dp) function vol_frac_inside(box,cc_ix)
  type(box3_t), intent(in) :: box
  integer, intent(in) :: cc_ix(3)
  real(dp)            :: h, rc(3),r(3), in_counts
  integer             :: i,j,k

  h = box%dr ! width of the cell
  rc = a3_r_cc(box,cc_ix) ! center of the cell

  in_counts = 0.0d0
  do i = 1,8
    do j = 1,8
      do k = 1,8
        r(1) = rc(1) + i*h/8.0d0 - h/2.0d0 - h/16.0d0
        r(2) = rc(2) + j*h/8.0d0 - h/2.0d0 - h/16.0d0
        r(3) = rc(3) + k*h/8.0d0 - h/2.0d0 - h/16.0d0
        if (dielectric_check(r)) in_counts=in_counts+1
      end do
    end do
  end do

  vol_frac_inside = in_counts / 512.0d0
end function vol_frac_inside
```

where `dielectric_check` is the logical function defining the dielectric shape.

Testing the dielectric field calculation

The electric field $\vec{E}$ of a dielectric sphere with dielectric permittivity ϵ is known analytically,

$$\vec{E}(\vec{r}) = \vec{E}_b + \frac{\epsilon - 1}{\epsilon + 2} \frac{R_{\text{sp}}^3}{r^3} \left[\frac{3(\vec{E}_b \cdot \vec{r})\vec{r}}{r^2} - \vec{E}_b \right], \tag{6.4}$$

where $\vec{E}_b$ is the background electric field and R_{sp} the sphere radius. We have numerically evaluated several spheres, with several dielectric permittivities, and compared them with the analytical solution. In figure 6.7 an example is given for a dielectric sphere of radius $R_{\text{sp}} = 1$ mm and dielectric permittivity of $\epsilon = 10^4$ on the z-axis parallel to $\vec{E}_b = 10^5 \hat{z}$ V/m. In the bottom panel of 6.7 the absolute error is given, which is maximal just at the edge and there less than 1%.
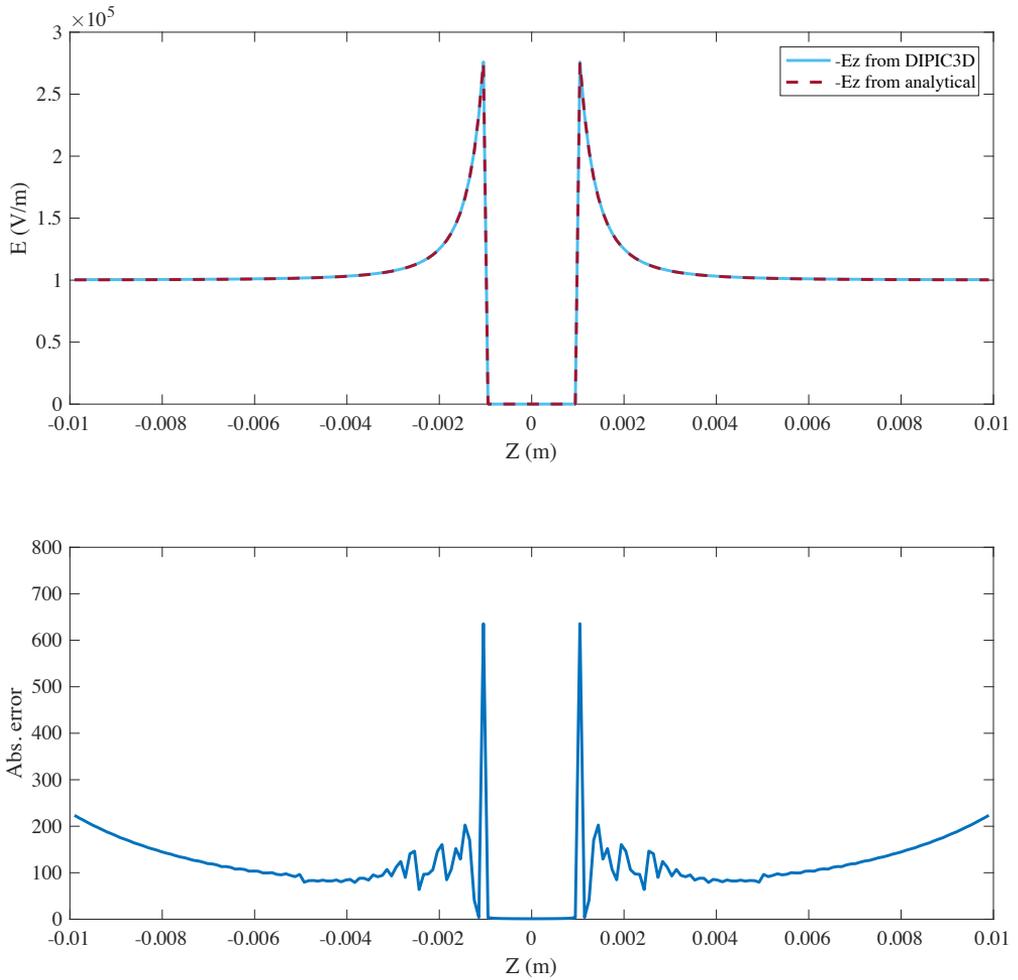
Figure 6.7: Comparison of the field calculation of DIPIC3D with the analytical solution for a sphere with dielectric permittivity of 10,000. Upper panel: direct comparison. Lower panel: absolute numerical error in V/m.

Chapter 7

Evaluation of Monte Carlo methods in high atmospheric physics I

The emerging field of high energy atmospheric physics (HEAP) includes Terrestrial Gamma-ray Flashes, electron-positron beams and gamma-ray glows from thunderstorms. Similar emissions of high energy particles occur in pulsed high voltage discharges. Understanding these phenomena requires appropriate models for the interaction of electrons, positrons and photons of up to 40 MeV energy with atmospheric air. In this chapter we benchmark the performance of the Monte Carlo codes Geant4, EGS5 and FLUKA developed in other fields of physics and of the custom made codes GRRR and MC-PEPTITA against each other within the parameter regime relevant for high energy atmospheric physics. We focus on basic tests, namely on the evolution of monoenergetic and directed beams of electrons, positrons and photons with kinetic energies between 100 keV and 40 MeV through homogeneous air in the absence of electric and magnetic fields, using a low energy cut-off of 50 keV. We discuss important differences between the results of the different codes and provide plausible explanations. We also test the computational performance of the codes. The supplementary material contains all results, providing a first benchmark for present and future custom made codes that are more flexible in including electrodynamic interactions.

7.1   Introduction

7.1.1   Phenomena in high energy atmospheric physics

Thunderstorms have been observed to produce Terrestrial Gamma-ray Flashes (TGFs) [53] and electron-positron beams [54, 55]. Signals lasting longer than TGFs such as x- and gamma-ray glows or thunderstorm ground enhancements (TGEs) have also been observed near thunderclouds, from balloons, planes, or high mountains [56–61]

Two possible theories are currently under discussion, as reviewed by Dwyer et al. [62], to create these phenomena by run-away electrons [63], which may further grow in the form of so called relativistic run-away electron avalanches (RREA), introduced by Gurevich et al. [64].

The first theory has been called the Cold Runaway theory [65] where thermal electrons are accelerated into the run-away regime within the strong electric fields of a transient discharge. Theoretical literature first focussed on the phase of the streamer discharge [66–68], and later on leader discharges [69–73]. Cold runaway is certainly at work in high energy emissions from nanosecond pulsed discharges [74–77] and during the formation of long sparks [78–84] in high voltage and pulsed plasma technology.

The second theory is the Relativistic Feedback Discharge model by Dwyer [85]. It is based on sustaining the RREA multiplication of the relativistic electrons in sufficiently high electric fields within a thunderstorm, by feedback of photons and positrons creating new avalanches [86–88]. The first electrons are typically supplied by cosmic particles from the sun or from other galactic or extragalactic sources. High energy seed electrons might also origin from lightning leaders, from radioactive decay or from some mixed form of electron sources.

An extreme case both of cold or of RREA would be a Relativistic Runaway Electron Front where the density of runaway electrons is high enough to provide electric screening behind the ionization front [89].

We remark as well that a sufficiently energetic cosmic particle can create an extensive air shower with very high electron density in the shower core even in the absence of any electric fields; such densities were used by Dubinova et al. [26] to explain lightning inception; and these air showers were also used to measure electric fields in thunderstorms [90, 91]. Radioactive decay is another source of high energy particles in the atmosphere.

All these phenomena require tracing the propagation of energetic electrons, photons, and also positrons through air, as well as modeling their interaction with air molecules and the subsequent scattering and energy loss or even total loss of the primary particles, together with the generation of secondary particles.

7.1.2 The multiple scales in energy and length

There are two basic problems for simulating these high energy phenomena in our atmosphere, related to the wide range of scales in energy and length.

First, the models have to bridge energy scales from tens of MeV down to thermal energies of tens of meV ($300\,K \rightarrow 0.03\,\mathrm{eV}$), i.e., over 9 orders of magnitude. At the upper edge of this energy range, models developed by the high energy physics community (e.g., for CERN) exist where it should be noted that they were originally developed for even higher particle energies, and for the interaction of energetic particles with metals rather than with air — though radiation medicine now also develops models for the penetration of energetic radiation into biological tissue [92–94], which consists mostly of similarly light molecules as air, but in the liquid rather than the gaseous state. In the low energy regime, models by the low temperature plasma physics community should be used, with cross-sections listed, e.g., on the community webpage [95].

Second, in particular, for cold run-away models, there are two widely separated spatial scales: the source region with high and time dependent self-consistent electrodynamic fields where electrons are accelerated, and the wide propagation distances from the source to detectors in space or on ground where electric fields can be neglected.

Here we focus on the second problem, namely the beam propagation towards detectors where the final products are characterized by energy spectra and arrival times, and the source properties must be reconstructed from this data, e.g. in the work by Østgaard et al. [96]. Accurately modeling the transport from the source to the very remote detector is, together with some knowledge of the source, thus very important to deduce production altitude, beaming angle or light curves of TGFs and associated electron beams from space data [54, 97–100].

7.1.3 Content and order of the present study

To model particle beams in air far from the source, some researchers use general purpose Monte Carlo (MC) codes developed by large collaborations like Geant4 (used by Carlson et al. [102] and by Skeltved et al. [103]) or FLUKA (used by Dubinova et al. [26]). On the other hand, to model, e.g., the radiation sources with their external or even self-consistent time dependent electric fields, other researchers develop custom made codes in small groups or as individuals, where the cross sections and numerical methods may come from already validated theory (e.g. Sarria et al. [24], Kohn et al. [72]).

While they are necessary for the understanding of the full physical phenomena, custom made codes are difficult to validate, especially if they are not made available by open access. Differences between one code and another may be explained by at least the following four factors:

- The choice of the included physics, as a compromise between correctness

Table 7.1: Codes used in this benchmark, their validity range (usable energy interval) and relative performance (normalized to the fastest code), possible inclusion of electric and magnetic fields (**E** & **B**) and self-consistent fields due to space charge. It should be noted that the synchronous particle tracking in GRRR, for the possible inclusion electric fields due to space charge, and the simulation without low energy cut-off approximation in MCPEP limits their performance. See for more descriptions Sect. 7.3.[a] 10 keV is the lowest energy advised in the manual, but in this study we found that this is too low, see Sect. 7.5.3.[b] 250 eV minimum for electrons and positrons and 10 eV minimum for photons.[c] Not out of the box, but there are add-ons or umbrella codes with provide this feature, for example CORSIKA [101]. [d] The magnitude of the electric field will be limited by the choice of the low energy cut-off.

Code	Validity range (eV)	Relative perform.	**E** & **B**	Space charge
EGS5	$[10^4, 10^{11}]^a$	4.02	N & N	N
FLUKA	$[10^4, 10^{11}]$	1.03	N^c & N^c	N
Geant4L	$[10^2, 10^{12}]^b$	1.17	Y^d & Y	N
Geant4D	$[10^2, 10^{12}]^b$	1.00	Y^d & Y	N
GRRR	$[10^4, 10^7]$	12.4	Y^d & Y	Y
MCPEP	$[10, 10^8]$	102	N & Y	N

and feasibility.

- Cross sections, that can come from theory, measurements or both. In most cases the cross section data has a certain uncertainty.

- Numerical and coded implementation, e.g. numerical integrations, interpolations, roundoff errors and bugs.

- The performance, as faster codes can run more particles in the same time, which results in more accurate statistics.

Even if it is possible in principle to determine the differences between the physical models and between the numerical methods, it may be very complicated (if not impossible)

- to estimate the uncertainties associated with a certain choice of physical models,

- to estimate the uncertainty propagation and accumulation of all input through the full multiscale models, and

- to review all source codes (if available) to find any mistakes and possible numerical problems.

In general it is found that software is underrepresented in high energy physics literature in spite of its significant contribution to the advancement of the field [104].

Therefore, we here strive to provide a comparison standard for the particle codes, as simple and as informative as possible, by only considering their physical outputs. We have chosen standard tests for the core parts of all codes: the evolution of monoenergetic and monodirectional beams of photons, electrons and positrons through homogeneous air and without electric or magnetic fields. We elaborate our standard tests in the methodology section 7.4.

The targeted energy interval for high energy atmospheric physics in this study is from tens of keV to tens of MeV, bounded above by the observed maximal energy in a TGF [105, 106]. Typically a low energy cut-off is chosen for two reasons:

1. The codes developed for accelerator or cosmic-ray applications use typical energies well above 1 MeV, larger than the rest mass of electrons and positrons. For these energies relativistic approximations are accurate, ionization potentials are negligible, and electron impact ionization is essentially a free-free elastic collision (i.e., similar to a collision of two free electrons). These approximations limit the validity of the codes at lower energies.

2. The mean free path of particles decreases and the number of particles increases with decreasing energy. Simulating with or without a low energy cut-off can make a difference of minutes to months of simulation time. Therefore a low energy cut-off is wanted for computational reasons.

The different implementations of the low energy cut-off, as reviewed in Sect. 7.3 cause significant differences in the results, see Sect. 7.5. These differences increase when electric fields are added, see Sect. 7.6 and puts an extra restriction on the value of low energy cut-off [103].

This chapter is organized as follows: Sects. 7.2 and 7.3 review the particle interactions and the codes included in this study. Sect. 7.4 describes the methodology we used to compare the codes. Sect. 7.5 contains a discussion of important differences between the results of the tested codes, and in Sect. 7.6 the implications of adding electric fields are discussed. Finally we conclude and give a list of recommendations for High Energy Atmospheric Physics simulations in Sect. 7.7.

7.2 Overview of interactions and approximations

In High Energy Atmospheric Physics (HEAP) it is usually assumed that the density of the considered high energy particles is too low to directly interact which each other, therefore they only interact with the background medium

which is here air molecules. In addition for some 'self-consistent' codes, like GRRR (see Sect. 7.3.4), charged particles can interact non-locally due to the electric fields they produce. But for the present study these interactions are turned off, resulting in a linear problem. This means that the number of particles at the end of the beam is proportional to the particle number in the initial beam, and that different beams simply add up according to the superposition principle. Below we summarize the interactions considered for electrons, positrons and photons in HEAP. In these interactions the target molecule M and its resulting state are explicitly given, but for the MC model of the high energy particles, these molecules (or ions) act as a random background.

7.2.1   Electrons and positrons

Electrons and positrons above 50 keV (which is the low energy cut-off in our study) behave almost identically; they scatter elastically on molecules M, they ionize them, and they create bremsstrahlung on collisions with molecules:

$$e^{\pm} + \text{M} \rightarrow \begin{cases} e^{\pm} + \text{M}, & \text{elastic (Rutherford)}, \\ e^{\pm} + e^{-} + \text{M}^{+}, & \text{ionization}, \\ e^{\pm} + \gamma + \text{M}, & \text{bremsstrahlung}, \end{cases} \qquad (7.1)$$

with cross sections that only slightly dependent on the incoming particle type.

In addition, when positrons come to rest, they usually annihilate with a bound electron,

$$e^{+} + \text{M} \rightarrow 2\gamma + \text{M}^{+}, \quad \text{annihilation}, \qquad (7.2)$$

and produce two photons of 511 keV. The standard implementation is that, when a positron drops below the low energy cut-off, it comes at rest immediately (in space and time). In reality the positron will come to rest over some distance and time, forming positronium (e.g. an $e^{+}e^{-}$ bound state) before annihilation. The positronium has a lifetime depending on the spins of the positron and electron [107], forming a singlet or triplet state with lifetimes of 124 ps or 139 ns (in vacuum), respectively. If the triplet state is formed in a medium like air, the lifetime permits pick-off annihilation where an opposite spin electron from the medium will annihilate in singlet orientation before the triplet-oriented electron can collapse and annihilate with the positron, thus resulting in again 2 photons (instead of 3). Thus, besides a small time delay, the magnitude of 511 keV line in the photon spectrum is not changed. None of the codes with the settings used in this benchmark include positronium.

In the eV regime, the interactions are getting more complex, as molecular excitations and dissociations need to be taken into account explicitly.

Friction (or stopping-power) for electrons and positrons

Usually, the energy transfer in an ionization collision of electrons and positrons with molecules is of the order of 10 eV, hence it causes only a small energy loss for a particle with energy above the keV range. By introducing a so-called low energy cut-off ε_{cut} , 'high' and 'low' energy particles and interactions can be decoupled. In this approximation, interactions producing secondary particles below the low energy cut-off are approximated as friction, while interactions with secondary particles above the cut-off are included explicitly.

Let ε_1 be the energy of the primary particle and ε_2 the energy of the secondary particle. The cross section $\sigma_k(\varepsilon_1)$ (in units of area) gives the probability of the primary particle to undergo an interaction labeled k. The differential cross section $d\sigma_k(\varepsilon_1, \varepsilon_2)/ d\varepsilon_2$ (in units of area per energy) gives the probability of a primary particle to produce a secondary particle within the infinitesimal energy interval $[\epsilon_2, \epsilon_2 + d\epsilon_2]$ for the interaction k.

The secondary energy ε_2 can take values between the minimum ε_{min} (of the order of eV and the primary is not sensitive for the precise value) and the maximum ε_{max} (of the order ε_1), depending on the interaction. For ionization $\varepsilon_{max} = \varepsilon_1/2$ as the primary by convention is defined to be the final particle with the highest energy. For bremsstrahlung we have $\varepsilon_{max} = \varepsilon_1$.

Now the energy range of the secondary particles is decomposed into two parts: the first part from ε_{min} to ε_{cut} is implemented as a friction, and the second part from ε_{cut} to ε_{max} is implemented by discrete collisions.

The friction F_k of interaction k is defined as

$$F_k(\varepsilon_{cut}, \varepsilon_1) = N \int_{\varepsilon_{min}}^{\varepsilon_{cut}} \left(\varepsilon_{loss}(\varepsilon_2) \frac{d\sigma_k(\varepsilon_1, \varepsilon_2)}{d\varepsilon_2} \right) d\varepsilon_2, \qquad (7.3)$$

where N is the number density of molecular collisions targets M, and ε_{loss} the energy loss of the primary which is of the order of ε_2 plus the ionization energy. The resulting friction on the primary is given by the sum of all considered interactions,

$$F(\varepsilon_{cut}, \varepsilon_1) = \sum_k F_k(\varepsilon_{cut}, \varepsilon_1). \qquad (7.4)$$

For electrons and positrons in the energy regime important for HEAP, the resulting friction is almost completely determined by the ionization part, as illustrated in Fig. 7.1. Especially if only the friction with $\varepsilon_{cut} = 50$ keV is considered (solid line), there the energy loss due to bremsstrahlung is more than two orders smaller than the energy loss due to ionization.

We remark that the friction is also frequently called the stopping-power for historical reasons, though it has the dimension of friction (energy/length) rather than of power (energy/time).
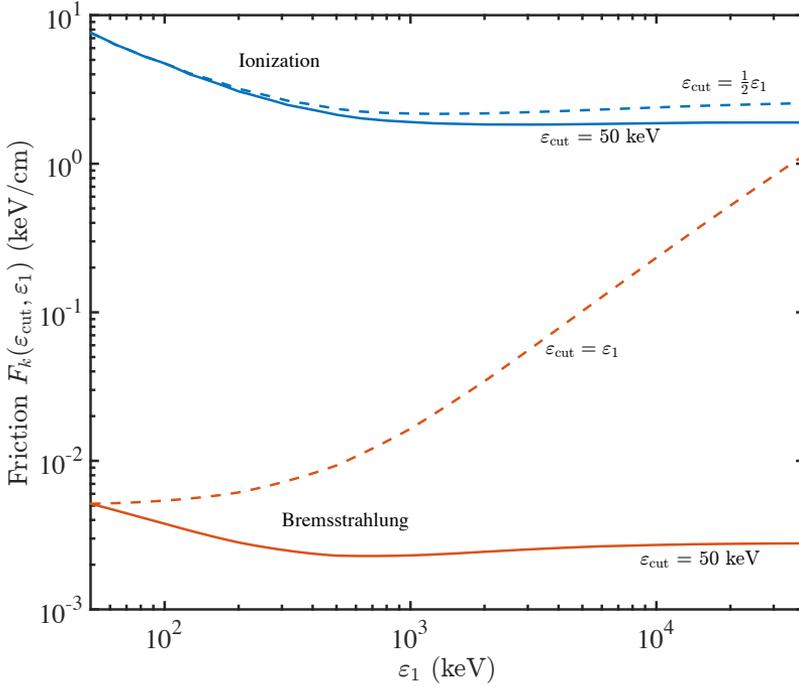
Figure 7.1: Friction $F_k(\varepsilon_{\mathrm{cut}}, \varepsilon_1)$ for electrons per interaction (Bremsstrahlung in red and ionization in blue), for two different low energy cut-offs, $\varepsilon_{\mathrm{cut}} = 50$ keV (solid line) and $\varepsilon_{\mathrm{cut}} = \varepsilon_{\mathrm{max}}$ (dashed line). The resulting friction is the sum of the two contributions, which in the energy regime of HEAP the resulting friction is dominated by the ionization (please, note the log-scale). The data are from [108, 109] for an air density of 1.293×10^{-3} g cm^{-3} corresponding to 1 bar and 273 K as used in this study.

Straggling

In a simple implementation of the low energy cut-off, the primary particle suffers a uniform (and deterministic) friction $F(\varepsilon_{\mathrm{cut}}, \varepsilon_1)$, as given in Eq. (7.4). This means that now only the energy of the primary particle is altered, but not its direction. A greater concern is that the accuracy of the assumed uniform energy loss is a matter of length scale. If the scale is much smaller than $\varepsilon_1/F(\varepsilon_{\mathrm{cut}}, \varepsilon_1)$, only a few interactions have taken place. On such a small length scale the real energy loss distribution (if one had considered all interactions explicitly) among the population would have a large spread. This effect is called straggling, and it was first studied by Bethe and Heitler [110].

One way to mimic the real energy distribution is by implementing a stochastic friction, as is done in FLUKA and Geant4L. Basically the energy loss of the primary particle is as if it would be modeled by real low energy collisions below the cut-off, but without creating the secondary particles and without altering

the direction of the momentum. The different implementation of the low energy cut-off (i.e., different implementations of the friction) is one of the significant differences we see in the studied programs, as discussed in the results section 7.5.

Continuous Slowing Down Approximation

Using the friction equation (7.3) over the whole range of secondary particle energies, hence with $\varepsilon_{\text{cut}} = \varepsilon_{\text{max}}$, the expectation value of the maximal penetration depth of a primary particle into a medium can be calculated in the so-called Continuous Slowing Down Approximation (CSDA). Integrating the friction over distance ℓ up to the point where the particle has lost all its primary energy ε_1,

$$\int_{\ell(\varepsilon_1)}^{\ell(0)} F(\varepsilon_{\text{max}}, \varepsilon(\ell)) \; \mathrm{d}\ell = \int_{\varepsilon_1}^{0} F_{\text{tot}}(\varepsilon_{\text{max}}, \varepsilon) \; \frac{\mathrm{d}\ell}{\mathrm{d}\varepsilon} \; \mathrm{d}\varepsilon = \varepsilon_1, \qquad (7.5)$$

defines one CSDA range through

$$\text{CSDA}(\varepsilon_1) = \ell(\varepsilon_1) - \ell(0). \qquad (7.6)$$

One CSDA range is thus the maximal length that primaries can penetrate into a material. Due to feedback from secondaries (e.g. electron -¿ photon -¿ electron) the complete avalanche can survive longer. As we describe in the methodology section 7.4, we choose half a CSDA range as the optimal detector distance to compare the differences in outputs of the codes as comprehensively as possible.

7.2.2 Photon interactions

The typical photon interactions are

$$\gamma + \text{M} \rightarrow \begin{cases} \gamma + \text{M}, & \text{elastic (Rayleigh)}, \\ e^- + \text{M}^+, & \text{ionization (by absorption)}, \\ \gamma + e^- + \text{M}^+, & \text{ionization (by Compton)}, \\ e^+ + e^- + \text{M}, & \text{pair production}. \end{cases} \qquad (7.7)$$

Photons have no charge, and therefore they lose energy much less gradually than electrons and positrons. In a typical inelastic interaction of a photon, the energy loss is significant.

Photon attenuation

The most important interaction for low energies (below 30 keV) is photo-absorption, and for the highest energies (above 40 MeV) it is pair production; in both cases the photon completely disappears. Inbetween, where Compton scattering is most important, the energy loss per interaction is still significant; the expectation value for the energy loss of the primary photon grows from 5% (at 30

keV) to above 90% (at 1 MeV). The Continuous Slowing Down Approximation is thus not appropriate for photons, as photons do not continously lose small amounts of energy, in contrast to electrons and positrons, but they lose a substantial fraction of their energy after some free path. Consecutively, for most energies (certainly above 1 MeV and below 30 keV) the photon intensity I can be approximated by an exponential decay or attenuation,

$$I(\ell) = I(0) \exp(-\ell/\mu),$$ (7.8)

where $\mu(\varepsilon)$ is the attenuation-coefficient depending on energy (and material).

In this work we need to estimate an appropriate detector distance (the exponential decay does not appear explicitly in any model), and we use two e-folding lengths (i.e., the inverse of half the attenuation-coefficient) as the optimal detector distance to compare the output differences, as described further in the methodology section 7.4.

7.3 Overview of codes

In Table 7.1 we have summarized the codes used in this benchmark. In this chapter we give more detailed descriptions.

7.3.1 EGS5

EGS5 (Electron-Gamma Shower version 5, developed by [21]) is a general purpose software package for the Monte Carlo simulation of the coupled transport of electrons, positrons and photons in an arbitrary geometry. It is the next version after EGS4 that was released by Nelson et al. [111] with a history that dates back to 1960's. The user controls an EGS5 simulation by means of an input text file for settings and a written FORTRAN user code, to which the rest of the FORTRAN source files are appended and compiled as one. In the user code several subroutine calls create, establish and initiate the cascade. Two important subroutines HOWFAR & AUSBGAB, which should be written inside the user-code are to specify the geometry and the output of the results. EGS5 can simulate particles from a few keV up to several hundred GeV, depending on the material. There is a limited option for including magnetic fields, and no option to include electric fields. All interactions of equations (7.1), (7.2), and (7.7) are implemented, in this work with a low energy cut-off of 50 keV. In the user manual of [21] a minimum low energy cut-off of 10 keV is advised, but we noticed that for the bremsstrahlung cross sections relativistic limits are applied, which results in a too low production of photons, see Sect. 7.5.3. Friction is implemented uniformly, without straggling effect (that is to say without fluctuations in the energy loss). The input file and user code, used in this work, can be found in the supplementary material. Please see the documentation of [21] for a detailed overview of the implemented physics.

7.3.2   FLUKA

FLUKA (developed by [22], copyright to INFN and CERN 1989-2011), is a general purpose tool for calculations of particle transport and interactions with matter. FLUKA is able to simulate the interaction and propagation in matter of roughly 60 different particles, including photons from 100 eV and electrons and positrons from 1 keV to thousands of TeV, neutrinos, muons of any energy, hadrons of energies up to 20 TeV (up to 10 PeV by linking FLUKA with the DPMJET code) and all the corresponding antiparticles, and neutrons down to thermal energies. FLUKA includes recent datasets, published by [112]. The program can handle magnetic and electric fields, although not self-consistently (i.e., the charged particles do not produce magnetic or electric fields). The program, written in FORTRAN, reads in so called user-cards, in which the user defines the geometry, materials and detectors. The user card, used in this work, can be found in the supplementary material. All interactions of equations (7.1), (7.2), and (7.7) are implemented, f in this work with a low energy cut-off of 50 keV. Friction in FLUKA is modeled with universal fluctuations, mimicking the straggling effect, meaning that the primary particle loses its energy as if it would undergo random collisions. But the direction of its momentum is not changed and no secondary particles are produced. Please see the documentation 'FLUKA Manual' at `www.fluka.org` for a detailed overview of the implemented physics.

7.3.3   agostinelli2003geant4

Geant4 is an open source toolkit to simulate the passage of particles through matter, developed by a wide international collaboration lead by the CERN. It is coded in C++, following an object oriented philosophy. It can simulate the transport of almost all known particles, and can include electric and magnetic fields [23]. We use the version 10.2 released in December 2015. In Geant4, the user can choose between six main models for the treatment of electrons, positrons and photons, with different performances and accuracies. One can also specify the implementation of the friction, to take into account energy losses below the low energy cut-off. For this study we are using two Geant4 configurations, that are detailed below. All Geant4 codes are available in the supplementary material. References and details for these models are presented in the 'Geant4 Physics reference manual' available at `http://geant4.web.cern.ch`.

Geant4D

Geant4D uses the default model, but in addition we deactivated the fluctuations of the continuous energy loss, i.e. the energy losses are applied uniformly, without straggling effect. This choice is for benchmark purposes, to identify the effect of straggling.

Geant4L

Geant4L uses the Livermore model, which uses cross sections from the EPDL and EEDL databases, provided by the Lawrence Livermore National Laboratory. The detailed implementation is provided in [108, 109]. The 'Universal fluctuation model' is activated to include the straggling effect in the implementation of friction.

7.3.4   The GRanada Relativistic Runaway (GRRR) code

Developed by A. Luque at the Astrophysics Institute of Andalusia (IAA-CSIC), the GRanada Relativistic Runaway (GRRR) code was designed to investigate the self-consistent interaction between electrons in the limit of very intense Relativistic Runaway Electron Avalanches (RREA). This investigation, presented in [89], concluded that due to the interaction between electrons in the avalanche RREAs saturate into a steady-state propagating Relativistic Runaway Ionization Front (RRIF). As the GRRR code was implemented with that specific goal in mind, its scope is narrower than the general purpose codes (EGS5, FLUKA, Geant4) analyzed in this chapter. It only follows the evolution of high-energy electrons, and includes a limited set of interactions between these electrons and the embedding medium. Electron ionization and Rutherford scattering are modeled discretely, and in this work down to a low energy cut-off of 50 keV. The friction for these interactions is uniform, without straggling effect. Bremsstrahlung collisions with nuclei are modeled deterministically by friction, in other words: as continuous radiative losses. The supplemental material of [89] contains further details about the physical model underlying the GRRR code. In the supplement material of this work the input files are given for the presented benchmark tests. The full source code for GRRR is available at `https://github.com/aluque/grrr`. However, presently the code is mostly undocumented so we advise potential users to contact the author.

7.3.5   MC-PEPTITA

The Monte Carlo model for Photon, Electron and Positron Tracking In Terrestrial Atmosphere (MC-PEPTITA) by Sarria et al. [24] is a Fortran 90 code that simulates the propagations of TGF and associated electron/positron beams within the Earth environment, from the production altitude at 10 to 20 km to satellite altitude. To simulate the quasi-exponential atmospheric density profile and the Earth's magnetic field, it uses the NRLMSISE-00 and IGRF-11 models [108, 109]. It is optimized to run in this environment, whereas some other codes (e.g., Geant4) can only handle layers of constant density. Concerning the interactions between particles and matter, it mainly uses the EPDL and EEDL cross section sets [108, 109], except for inelastic scattering of electrons

and positrons where the GOS model is used. The interactions are simulated similarly to PENELOPE [113], with equivalent numerical methods. MC-PEPTITA does not include any continuous energy losses: the particles are followed discretely down to the lowest possible energies allowed by the models used, with exception of bremsstrahlung where the minimal energy is set to 100 eV.

7.4   Methodology

We focus on the evolution of monoenergetic and directed beams of electrons, positrons and photons with kinetic energies between 100 keV and 40 MeV through homogeneous air in the absence of electric and magnetic fields, using a low energy cut-off of 50 keV. Providing a first benchmark, in the case when the fields are turned off. Assuming sufficiently low densities of high energy particles, arbitrary particle beams can be decomposed into such monoenergetic and directed beams.

The electron, positron and photon beams propagate through air, consisting of 78.085% nitrogen, 20.95% oxygen and 0.965% argon. We use a constant and homogenous air density of 1.293×10^{-3} g cm^{-3} corresponding to 1 bar and 0 degree Celsius. For all programs we choose a low energy cut-off of 50 keV, below which all particles are removed. For most programs, this low energy cut-off is also the threshold to treat collisions discretely or continuously, with two exceptions: MC-PEPTITA handles all collisions explicitly, and GRRR uses continuous radiative loss (bremsstrahlung). During the simulation electrons, positrons or photons above the low energy cut-off can be created (except for GRRR, which only models electrons), and are then followed as well until they also drop below the low energy cut-off. If considered in the program, positrons dropping below the low energy cut-off can produce photons by annihilation above the low energy cut-off.

We use ideal flat surface detectors, perpendicular to the primary particle beam. On a detector, the type, kinetic energy, position and arrival time of the arriving particles are recorded. After detection, the particles are removed from the program, thus we do not measure backscattered particles that have already been detected. Depending on the program, other secondary particles are created with a very low probability (e.g. neutrons by photo-nuclear interactions), but we do not record them in the output. First, we study the particle number of all particles as function of propagation distance (attenuation). Second, for one specific distance, (depending on particle type and initial energy) we proceed to a detailed analysis of energetic, spatial and temporal distribution. Complementarily we also benchmark the performance (i.e., the simulation completion time) of the programs used in this study.

7.4.1   The number of particles versus distance (attenuation)

We study the particle number of all particles as a function of beam propagation distance, up to of one CSDA range for electrons and positrons and of four times the inverse of the attenuation coefficient (four e-folding lengths) for photons. This range is divided in several distances (roughly 20) or data points. For each distance (or data point), we perform a new simulation. Each simulation with ten thousand particles in the initial beam, for beams of electrons, positrons and photons with energies of 0.1, 0.4, 1, 10 and 40 MeV. The particle numbers are therefore derived under the assumption that the detectors are impenetrable. This means that back scattering is excluded, and that the particle number therefore is lower than in a passing avalanche in air only.

We added a $\pm 1/\sqrt{n_i}$ relative error expected from the Monte Carlo methods (n_i being the number of counts in the ith bin). In this way we performed roughly 1800 simulations, namely circa 300 simulations per program: for 3 particle types, 5 initial energies and on average 20 distances per beam. GRRR only considers electrons while the energy loss due to production of photons is implemented as a continuous energy loss. The relevant results are given and discussed in Sect. 7.5. In addition all data of this part are visualized and available in the supplementary material.

7.4.2   Spectral analysis

We performed detailed simulations with 1 million particles per beam for one specific distance per beam. For electrons and positrons, the detection distance was chosen as half of the CSDA range. This gives most information in one plot, since the primary particles are still alive, while there is a significant number of secondary particles produced. For photons, the inverse of half the attenuation coefficient (two e-folding lengths) is chosen as the distance for the detailed study. At the detector we analyze the kinetic energy, the radial distance from the symmetry axis and the time of arrival. The spectra are binned using the FreedmanDiaconis rule in the log-domain and rescaled to numbers per primary. As also for the attenuation study, we added a $\pm 1/\sqrt{n_i}$ relative error expected from the Monte Carlo methods (n_i being the number of counts in the ith bin). We performed roughly 90 different simulations (circa 15 simulations per program: 3 particles and 5 initial energies). The relevant results are given and discussed in Sect. 7.5. In addition all data of this part are visualized and available in the supplementary material.

7.4.3   Performance benchmark

As a complement, we also tested how much time the codes needed to complete the simulations. We did not try to do an in-depth performance benchmark of the codes, but we think this is an interesting piece of information for someone

who is seeking for a code to be used in the HEAP context. Since the programs are written in different languages (Fortran, C++ and Python) and may be run on different machines with different architectures, we normalized the completion time with respect to a reference computer configuration.

The simulation starting with one million 1 MeV electrons is used as the test case because it is feasible for all the evaluated codes, and it takes a completion time that is neither too short, nor too long. More details are given in the supplementary material. The normalized results are discussed in Sect. 7.5.5.

7.5 Results

Most tests show similar outputs for the different codes, to within deviations of $\pm 10\%$, see supplementary material. Here we focus on important differences between the results of the codes, and we provide several plausible explanations.

7.5.1 Straggling

For electrons and positrons below 1 MeV, the data clearly show the effect of straggling, as discussed in Sect. 7.2.1. For example in the 400 keV electrons beam shown in Fig. 7.2, EGS5, Geant4D and GRRR do not include straggling, therefore the maximal electron energy is too small and the drop of the energy spectrum towards this maximal energy too steep. Geant4L, MCPEP and FLUKA show the correct spectrum, but for different reasons. MCPEP simulates without a low energy cut-off (and thus without friction). Geant4L and FLUKA use a stochastic implementation of the friction called universal fluctuations. Basically the friction is not applied uniformly to all particles of the same energy equally, but a distribution of energy losses in time mimics the random nature of the collisions. Only the direction change is considered negligible.

The same effect is also seen for electron and positron beams with energy above 10 MeV, in the scenario where bremsstrahlung is treated as continuous. GRRR shows an unphysical drop in the electron spectrum at high energies, as illustrated in Fig. 7.3. The reason is that the energy loss by bremsstrahlung is mostly above the low energy cut-off, see Fig. 7.1, meaning that the energy loss of the electrons and positrons is mostly due to discrete 'hard' collisions and thus ill-approximated by uniform averaged friction. Nevertheless we found that the total integrated energy is similar. This approximation is also used by others in the community like [70, 71].

7.5.2 Opening angle

High energy photons penetrate the medium much deeper than electrons and positrons, and therefore small differences in opening angles after Compton collisions are more important. In inelastic collisions photons always lose a significant
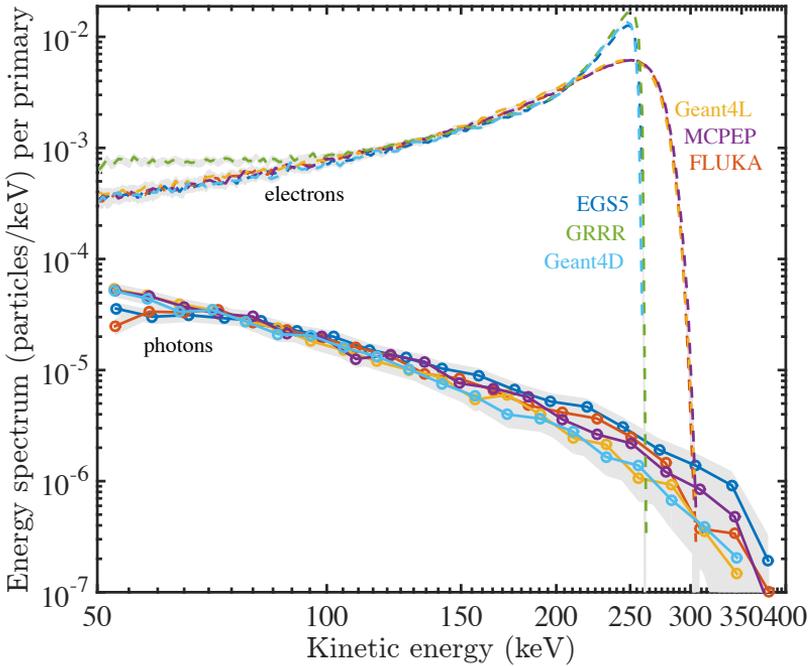
Figure 7.2: Products of a beam of 400 keV electrons after a propagation distance of 0.5 times their CSDA range which is 1.9 m in air at 1 bar and 273 K. The electrons have now a maximal energy of 250 to 300 keV depending on the code, but the total integrated energy is equivalent. The difference in electron distribution is due to straggling by ionization, see Sect. 7.5.1

amount of energy, as discussed in Sect. 7.2.2, and therefore they get a significant opening angle.

MCPEP simulates all collisions explicitly (others use a friction - which does not change the primary direction). The energy spectra agree between these codes, but Fig. 7.4 illustrates, that the radial and temporal spectra vary: MCPEP shows a wider photon beam and substantially later photon arrival times.

7.5.3 Bremsstrahlung

We saw that EGS5 uses an ultra-relativistic approximation in the treatment of bremsstrahlung and thereby we question the validity at lower energies, as discussed in Sect. 7.3.1). For the primary electron, in the energy regime important for HEAP, bremsstrahlung is negligible compared to ionization (see Fig. 7.1) and we thus do not see a difference there, but in the production of photons there is a significant difference, as can be seen in Fig. 7.5.
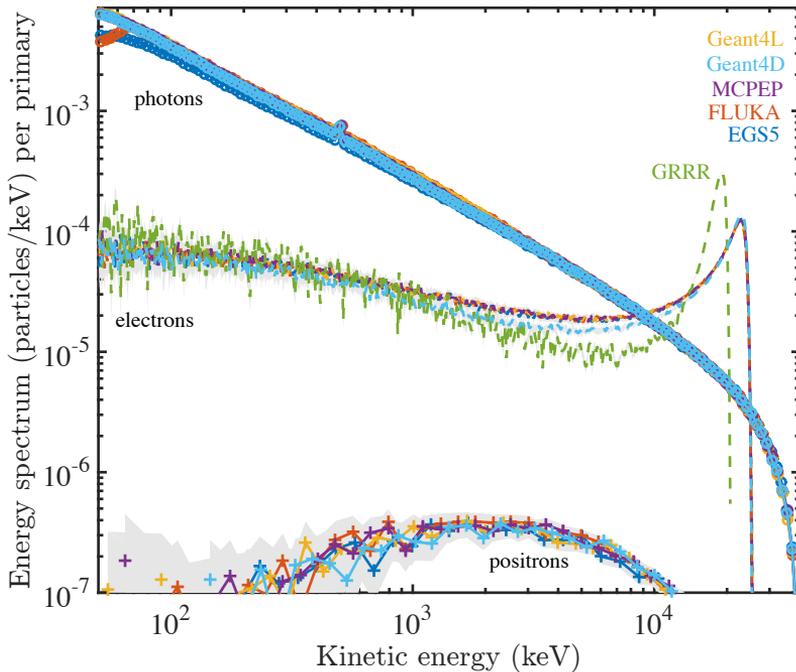
Figure 7.3: The same as in Fig. 7.2, but now for 40 MeV electrons. The propagation distance of 0.5 times their CSDA range is now 63.8 m (1 bar and 273 K). Now not only electrons and photons, but also positrons have been produced. The difference in electron distribution is due to straggling by bremsstrahlung, see Sect. 7.5.1

7.5.4 Other differences

Other differences we have found are listed below.

- For the electron and positron beams we see in the energy spectrum of FLUKA below 70 keV a dip in the number of photons. Fig. 7.3 shows an example.

- For the electron beams $\leq$ 1 MeV (but not in the positrons or photon beams) we see a difference in the longest arrival times ($>$ 100 ns) for photons between the programs FLUKA and EGS5 compared to Geant4D and Geant4L. GRRR does not model photons, and MCPEP is completely different because of the opening angle, see Sect. 7.5.2.

- GRRR shows a slight higher count (less than 15% higher) than the other codes for the number of electrons in the avalanche as function of distance. Fig. 7.5 shows an example. In the energy spectrum we see that these electrons are in the low energy tail of the spectrum, see for example Fig. 7.2.
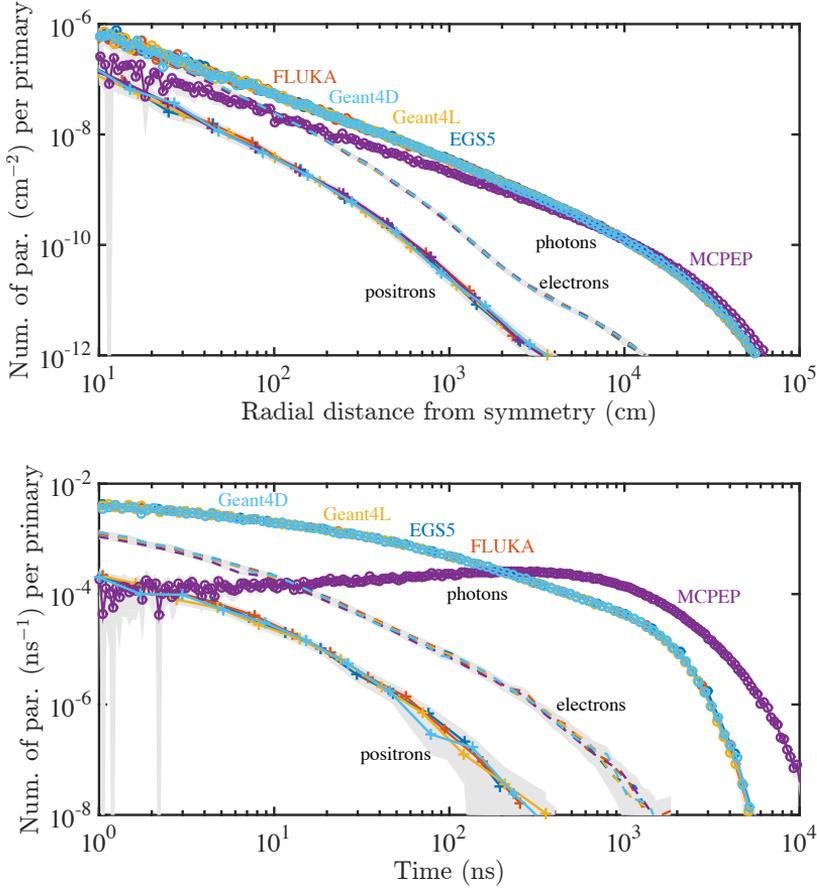
Figure 7.4: Products of a beam of 10 MeV photons at a distance of $1/(0.5\ \mu)$ which corresponds to 756 m (1 bar and 273 K). Particle number per primary as a function of the radial distance from the symmetry axis (above), and of arrival time (below).

- For the electron and positron beams we see a difference in the shortest arrival times (< 1 ns) for electrons and positrons between the programs FLUKA, EGS5 and MCPEP compared to Geant4D, Geant4L and GRRR.

7.5.5   Performance

The performances in terms of completion time of the codes are presented in Tab 7.1. On one hand, we see a clear difference of performance between MC-PEPTITA (simulations with a low energy cut-off as low as possible) and the rest. As said in the introduction, the low energy cut-off is generally introduced to speed up the simulation. Moreover, MC-PEPTITA was not optimized to run with a constant density and without magnetic field, and is then making a signif-
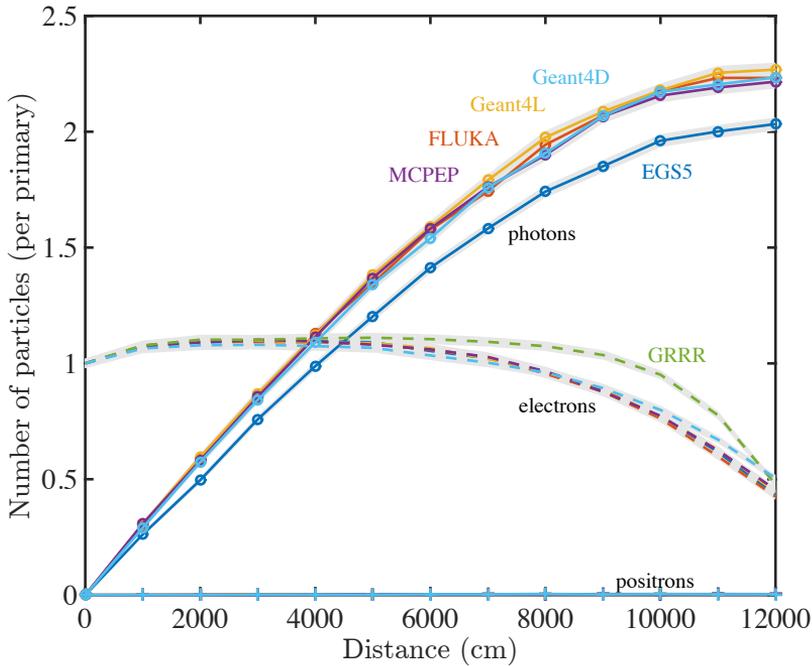
Figure 7.5: Products of a beam of 40 MeV electrons, as detected by 12 detectors at 10 to 120 m distance in 1 bar and 273 K. The detectors are impenetrable to hinder backscattering; therefore a new simulation is run for every detector distance.

icant amount of useless (but time consuming) calculations for this benchmark case. On the other hand, the choice to simulate all particles synchronously (to include self-consistent electric fields) slows the simulations significantly down, as seen for GRRR.

Concerning codes developed by wider collaborations, Geant4 and FLUKA show similar and best performances, but EGS5 is about 4 times slower. We can also note that in Geant4, the use of the energy straggling costs about 20% more computation time than turning it off.

7.6 The effect of electric fields

In this study we have provided benchmarks in the absence of electric fields, applicable to custom codes when the fields are turned off. The programs reviewed in this study are at least able to simulate the simplest case of particle beam evolution in air, in the wide distance from the particle source to detectors in space and on ground. However, as discussed in the introduction, the particles are initially accelerated by electric fields in the thunderstorm, either by weaker fields in larger regions in the Relativistic Feedback regime, or by strong and very

localized self-consistent electric fields in the Cold Runaway regime. We here give a short outlook on the range of validity of the presented models in these cases. In general, it can be expected that electric fields will magnify all differences in choice and implementation of cross-sections to a certain extent, because particles not just lose energy and drop eventually below the energy cut-off, but charged particles can also be reaccelerated and reappear in the ensemble.

To be specific, we recall the definition of the three characteristic electric fields and electron energy regimes of the problem (giving field values for air at standard temperature and pressure (STP)). For electrons with energies in the eV regime, the classical breakdown field is $E_k \approx 32$ kV/cm. For higher fields electron avalanches are formed, but their energies typically do not exceed the range of several eV, as their friction increases with energy. The electron friction increases up to an electron energy of approximately 200 eV where the critical electric field $E_c \approx 260$ kV/cm is required to balance friction — as long as the approximation of the electron ensemble by classical friction is valid. For electron energies above 200 eV the friction decreases to a minimum that is balanced by an electric field of $E_b \approx 2.8$ kV/cm, called the break-even field, at an electron energy of about 1 MeV.

Clearly two limitations to using a particle model with a low energy cut-off are immediately visible. First, if the electric field is above the critical electric field of 260 kV/cm ($E > E_c$) in a sufficiently large volume, the two populations of electrons with energies below and above 200 eV are strongly coupled and essentially all electrons can be accelerated into the runaway regime, to 1 MeV and beyond. Second, if the electric field is below the critical field, but above the classical breakdown field ($E_k < E < E_c$), the population of electrons in the eV regime (the so-called thermal electrons) can grow strongly, and eventually 'tunnel' into the run-away regime; we will come back to this effect below.

On the other hand, for electric field strengths below the break-even field ($E < E_b$), all electrons, regardless of initial energy, will eventually stop as the friction force of air is stronger than the accelerating force of the electric field.

Finally, when the electric field is above the break-even and below the classical breakdown field ($E_b < E < E_c$), the use of the energy cut-off of 50 keV (or even lower) can have strong implications: For an electron energy of 50 keV, friction and electric acceleration force balance each other when the field is 7.8 kV/cm. So in classical approximation one would estimate that at lower fields the inclusion of the cut-off is justified. However, this classical approximation neglects the stochastics of the actual process. Due to the randomness of free paths and scattering events, electrons actually can 'tunnel' into energy regimes that they could not reach in the classical approximation, an effect similar to the straggling effect discussed earlier.

Skeltved et al. [103] have observed this effect: For all fields between 4 and 25 kV/cm, they found that energy spectrum and mean energy of runaway electrons depended on the low energy cut-off, even when it was chosen between

250 eV and 1 keV. They also found – not surprisingly – that the differences become most apparent when the electric field force approaches the friction force corresponding to the low energy cut-off.

A related observation was made by Li et al. [114] when they found electron runaway from a negative streamer even though the maximal electric field at the leader tip was well below the critical field E_c.

Future studies on how to choose the low energy cut-off for given fields are desirable to optimize computations between efficiency and accuracy.

7.7   Conclusions

The goal of this work is to provide standard tests for comparing the core part of Monte Carlo simulations tools available for HEAP. We focused on the propagation of electrons, positrons and photons through air, in the absence of electric and magnetic fields. We compare the output at half the CSDA range for electrons and positrons, and at two e-folding lengths (the inverse of half the attenuation coefficient) for photons. We have run these tests for 0.1, 0.4, 1.0, 10 and 40 MeV initial energy for the several codes (Geant4, EGS5, FLUKA, GRRR, and MC-PEPTITA) used by the co-authors. The outputs show equivalent results, but there are important differences one can identify. Especially the different implementations of the friction are causing observable effects. First we see that straggling is important in the energy regime of HEAP and should be included in the simulations. Secondly the opening-angle of photon beams are very sensitive to the low energy cut-off. Thirdly we noticed that EGS5 has an ultra-relativistic approximation for bremsstrahlung which is not appropriate in the energy regime of HEAP. Last but not least there is a big difference in completion time between programs, mainly depending on the low energy cut-off and the synchronous implementation of the code. Adding electric fields will only increase these differences further and limits the value of the low energy cut-off. All results are published as supplementary material, and they can then be used by anyone to benchmark their custom made codes, with the fields switched off. The next step is provide benchmarks including fields and finding the optimal low energy cut-off for simulations in HEAP.

7.8   Recommendations

- Check where possible custom made codes to well established general purpose codes, we provide benchmarks in the energy regime of HEAP, in the case of zero field.

- Make your custom made code available to other researchers.

- For electrons and positrons below 1 MeV straggling should be included.

- For electrons and positrons above 10 MeV radiative loss should not be implemented with uniform friction.

- Photon production (due to bremsstrahlung) by electrons and positrons in energy regime of HEAP is under-estimated by EGS5.

7.9 Code and/or data availability

Figures of all output are available in the supplementary material. All raw data, circa 2 gb in compressed form, can be downloaded on request. In addition, the input files for reproducing the tests done in this benchmark are given for EGS5, FLUKA, Geant and GRRR, including links to the main source files. MC-PEPTITA simulations can be requested, contact David Sarria (david.sarria.89@gmail.com). MC-PEPTITA program was developed under a contract of Centre National D'Edtudes Spatiales (CNES) and Direction Gnrale de l'Armement (DGA), whose permissions are required in order to get access to the source code. Details of the performance tests are also available in the supplementary material, including the reference code 'pidec.cpp', used for normalizing the different computer architectures.

7.A Performance benchmark

As complementary, we also want to test how much time the different codes need to complete an equivalent simulation. We do not pretend to do an in-depth performance benchmark of the codes, but we think this is an interesting piece of information for someone who is seeking for a code to be used in the HEAP context. Since the programs are written in different languages (Fortran, C++ and Python) and may be run on different machines with different architectures, we normalized all the completion time with respect to a reference computer configuration. We normalized the results with the fastest code, intermediate results are given in Tab. 7.2.

7.A.1 Procedure

First, one need to calculate the normalization factor N_{user}, using the c++ code 'pidec.cpp', written by Xavier Gourdon, and provided in the this supplementary material. It computes 8 digits of pi after a given digit position called n. It should be compiled using the GNU g++ compiler with no options, in particular no optimization options (eg '-O3'). The time taken to complete it with $n = 1000000$ (usually about 10-20 minutes) is called t_{user}. The code itself outputs it in the terminal, and it is equivalent to the 'user time' given by the 'time' bash command. The reference time t_0 is set to 1162 seconds, and the normalization factor is then given by $N_{user} = t_{user}/t_0$.

Table 7.2: Summary of the performance (completion time).

Code	GEANT4D	GEANT4L	MC-PEPTITA
CPU	Q9650 3.0Ghz		
'pidec' execution time (s)	1 162 s		
Normalization factor	1		
Simulation time (s)	206	241	21 040
Normalized time (s)	206	241	21 040

Code	EGS5	FLUKA	GRRR $dt = 25$ ps	GRRR $dt = 2.5$ ps
CPU	Xeon E-3 1271 3.6Ghz		Xeon X7350 2.9Ghz	
'pidec' time (s)	596 s		1 362 s	
Norm. fact.	1.95		0.85	
Sim. time (s)	425	109	3 017	34 451
Norm. time. (s)	829	213	2 564	29 283

The one million 1 MeV electron beam simulation is used as the comparison case. If the considered code is parallelized, it should run on one single thread, but any compilation options can be used to make it as fast as possible. In any case, one should make several runs and get an average time to minimize the estimation error. This will give a simulation completion time that must be multiplied by N_{user} to get the normalized completion time.

7.B Supplement comparison

We study electron, positron and photon avalanches through air, consisting of 78.085% nitrogen, 20.95% oxygen and 0.965% argon, with kinetic energies starting with 0.1 MeV, 0.4 MeV, 1.0 MeV, 10 MeV and 40 MeV. We use a constant and homogenous density of 1.293×10^{-3} g cm^{-3}. The distances considered up to one CSDA range for electrons and positrons and four times the inverse of the attenuation coefficient (four e-folding lengths) for the photons. For each distance we preform at least 10k initial particles. The spectra are simulated at one half CSDA range for electrons and positrons and two times the inverse of the attenuation coefficient (two e-folding lengths). The spectra are simulated with 1 million initial particles.

——	EGS5 program
——	FLUKA program
——	GEANT4L program
——	MCPEP program
——	GRRR program
——	GEANT4D program
- -	marker for electrons
o	marker for photons
+	marker for positrons
▨	95% noise interv.
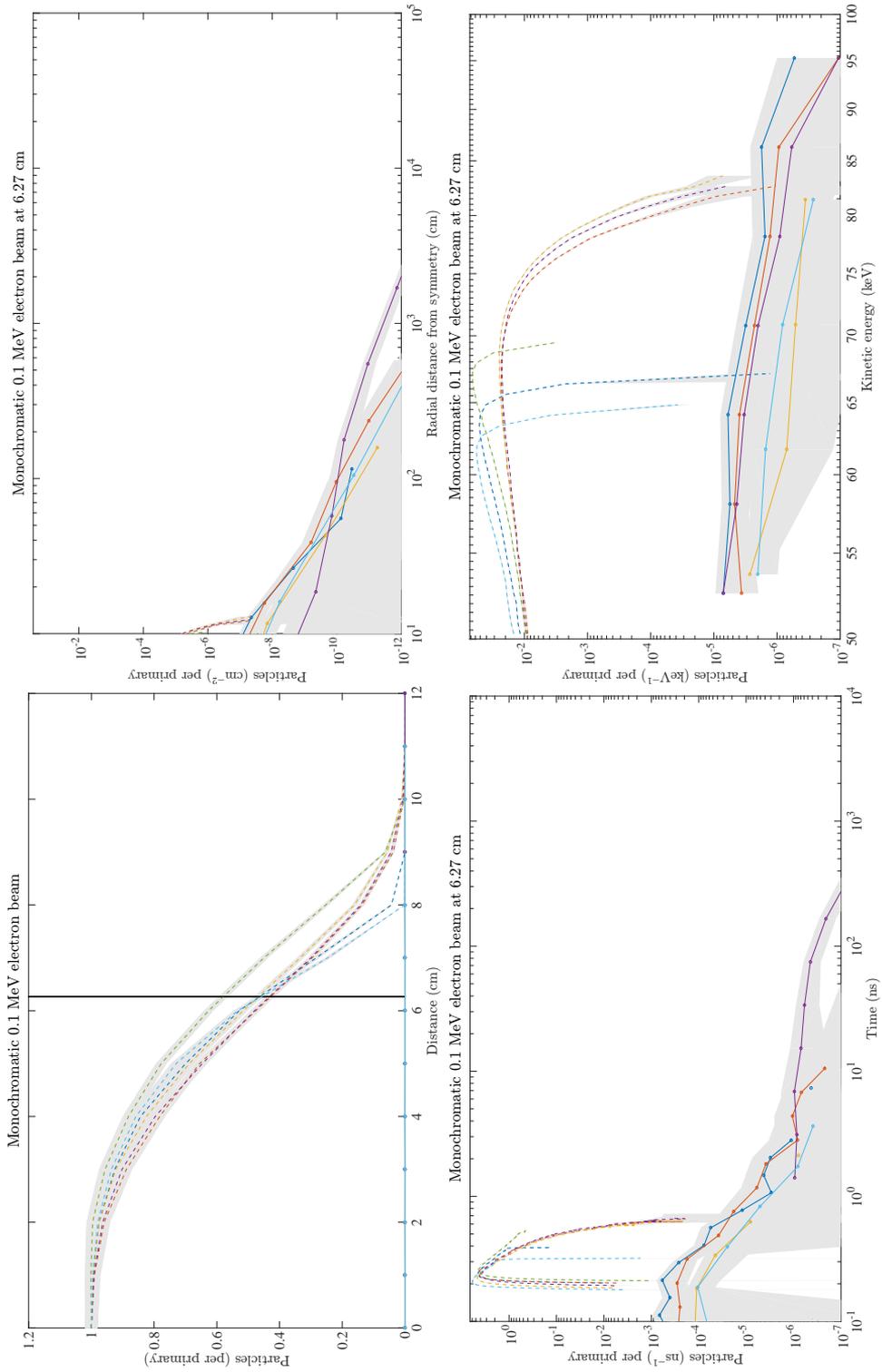
Figure 7.6: 100 keV electron beam.
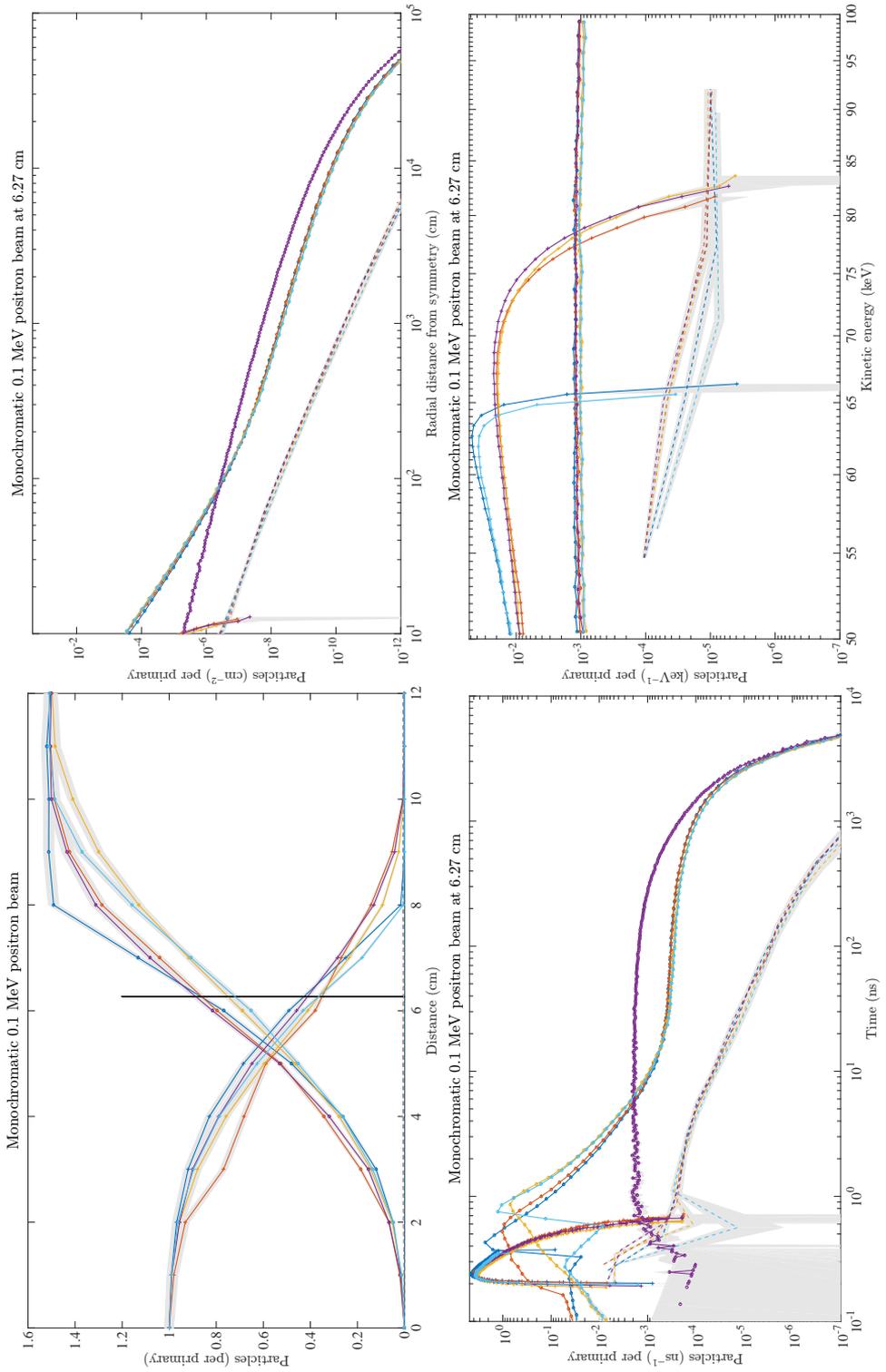
Figure 7.7: 100 keV positron beam.
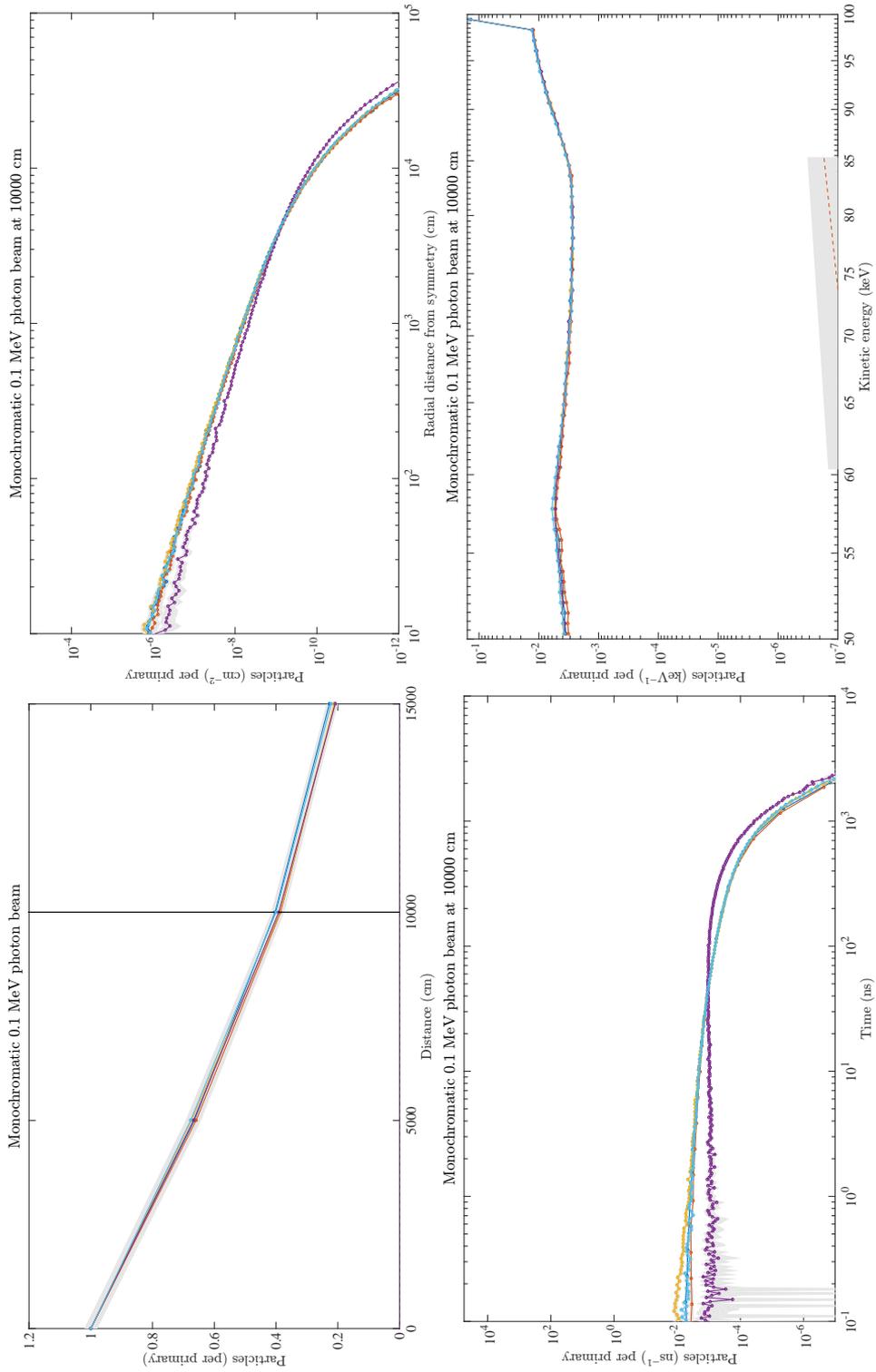
Figure 7.8: 100 keV photon beam.
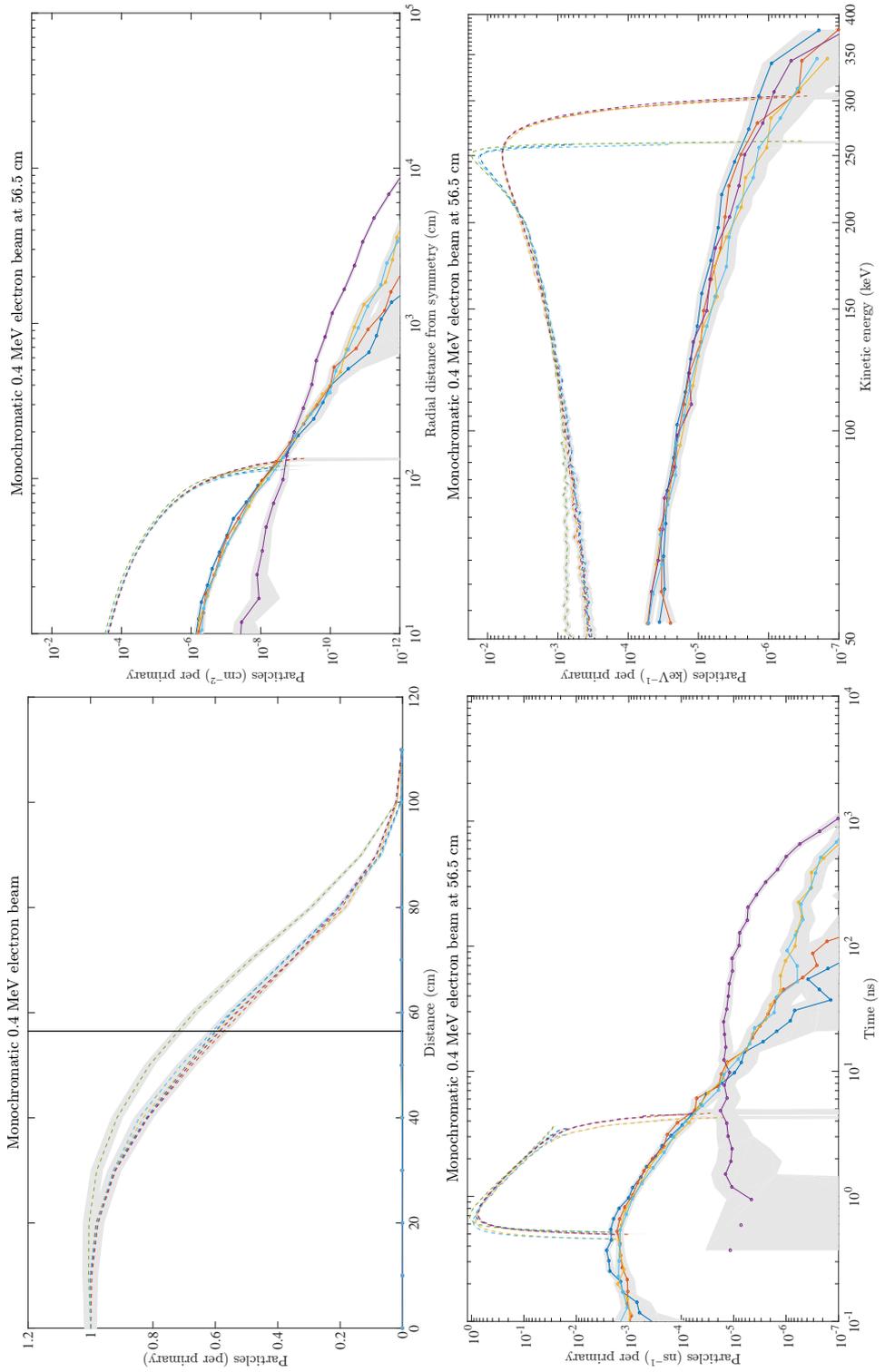
Figure 7.9: 400 keV electron beam.
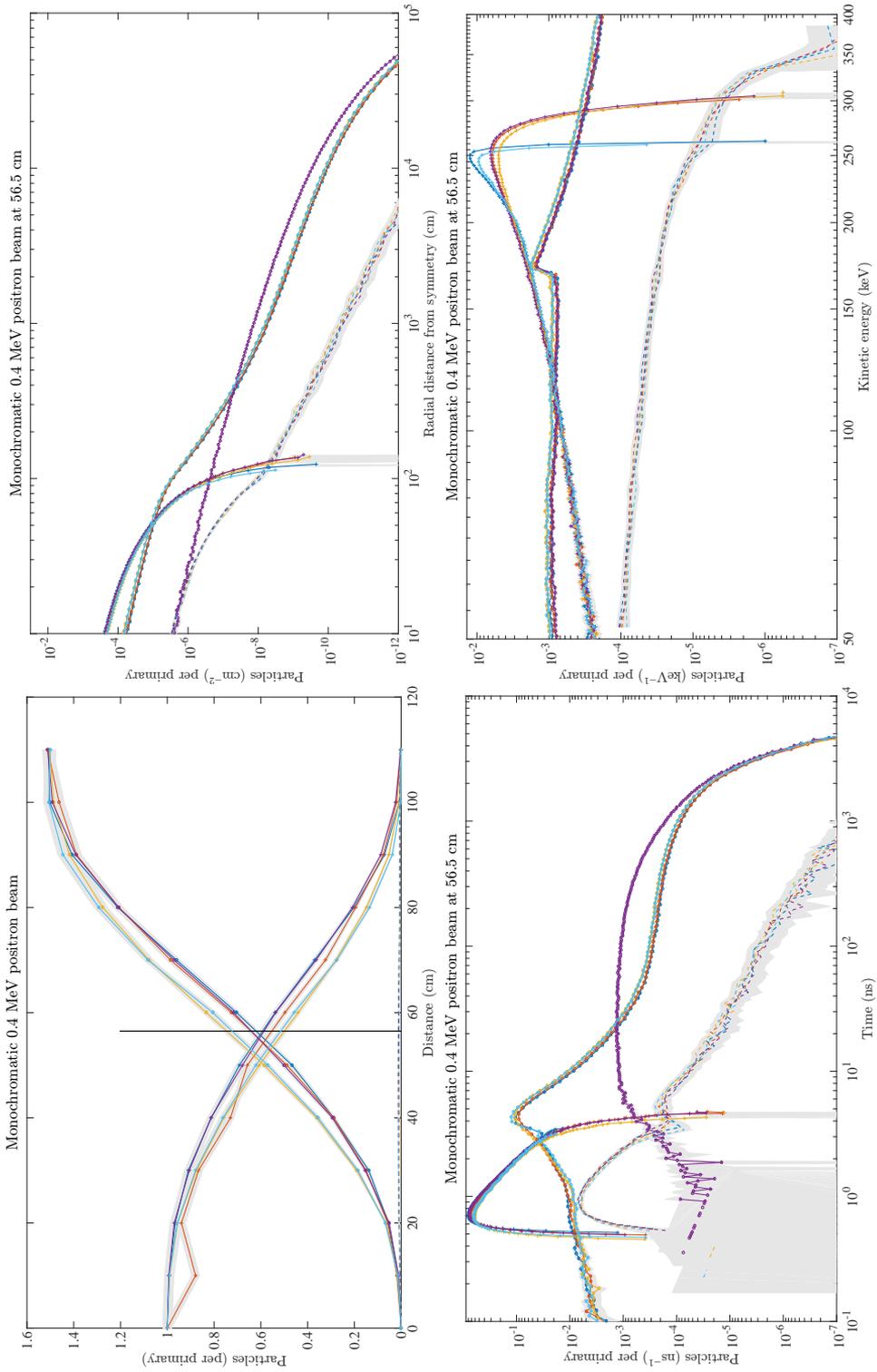
Figure 7.10: 400 keV positron beam.
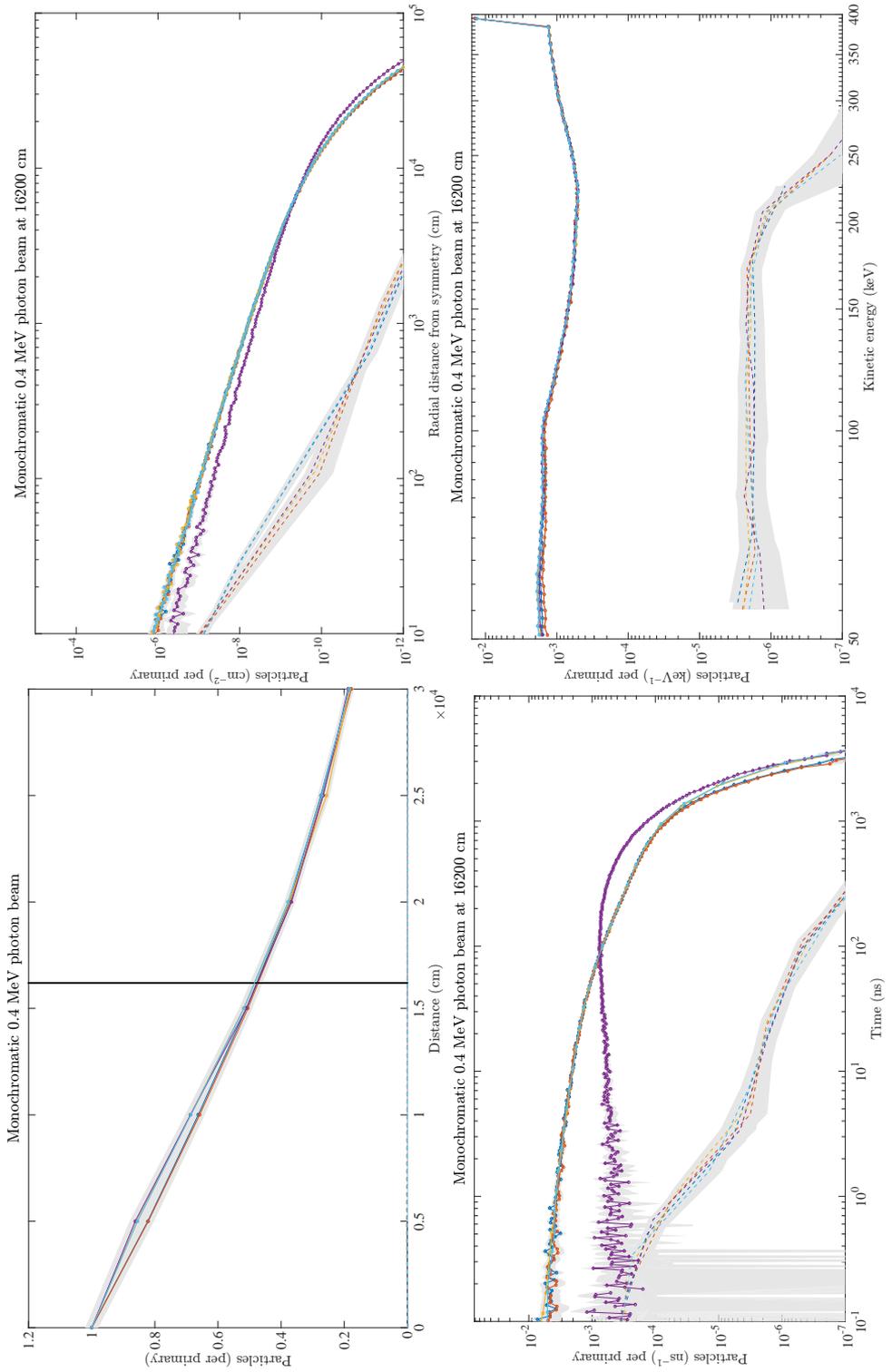
Figure 7.11: 400 keV photon beam.
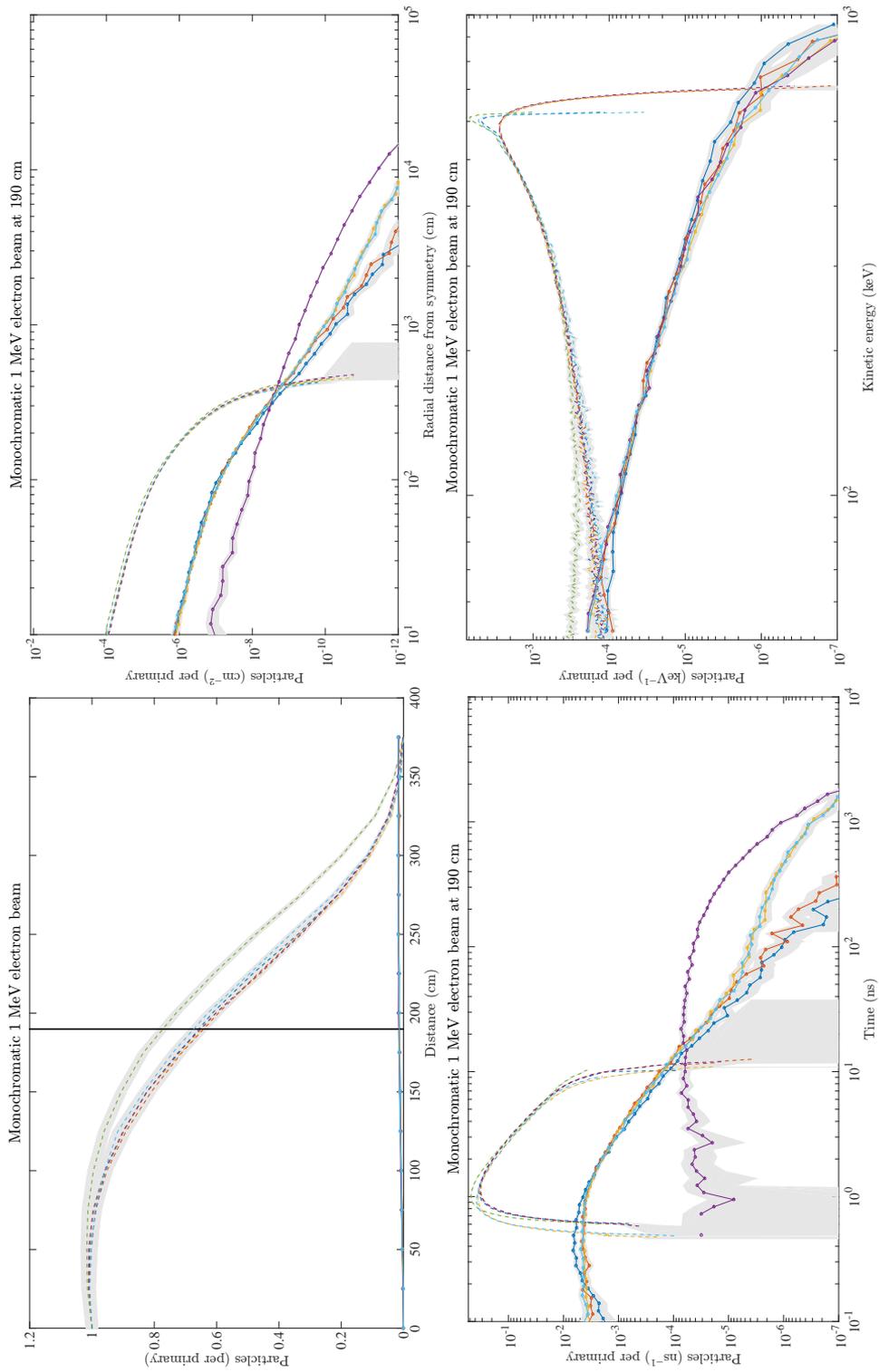
Figure 7.12: 1 MeV electron beam.
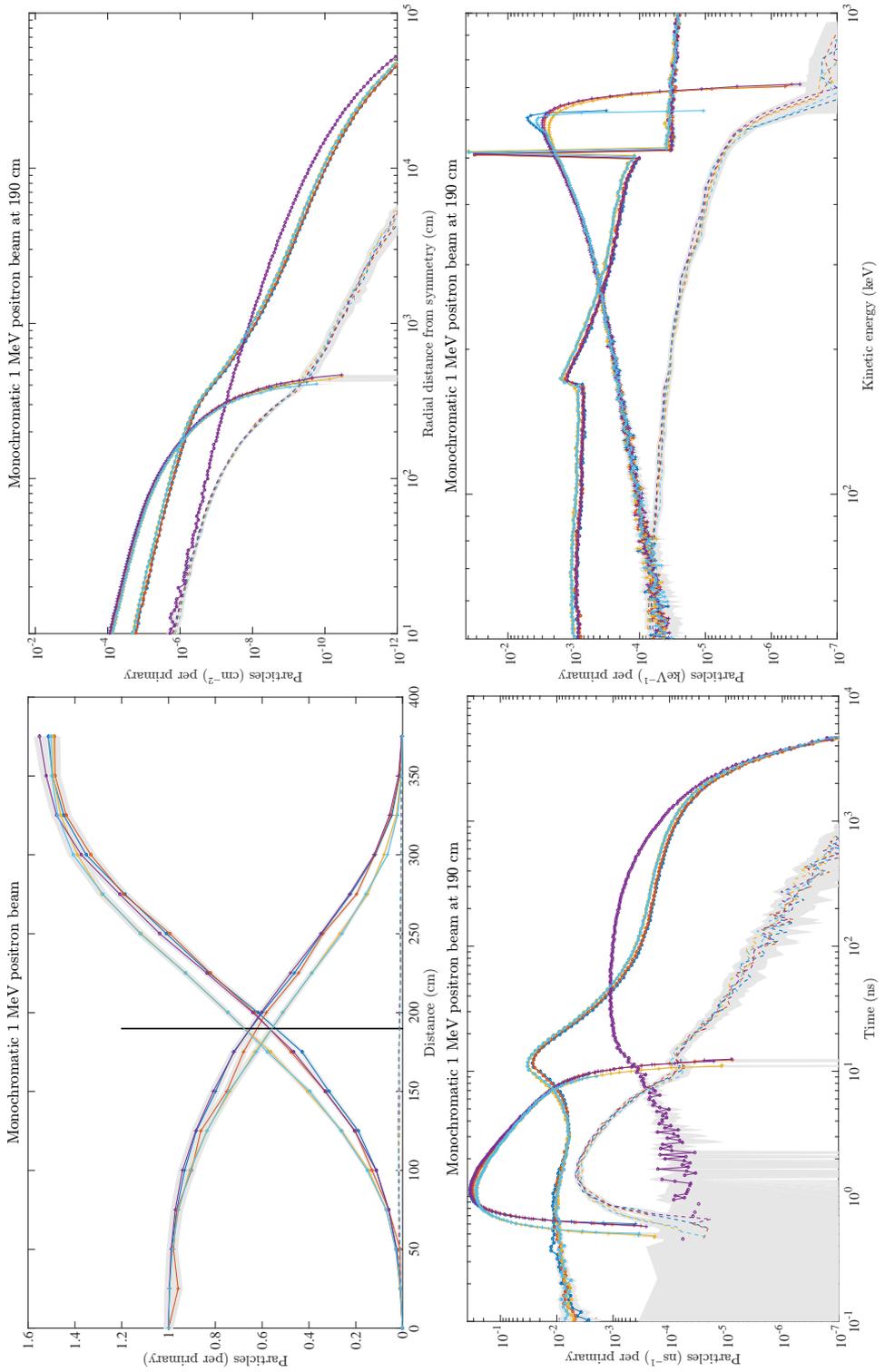
Figure 7.13: 1 MeV positron beam.
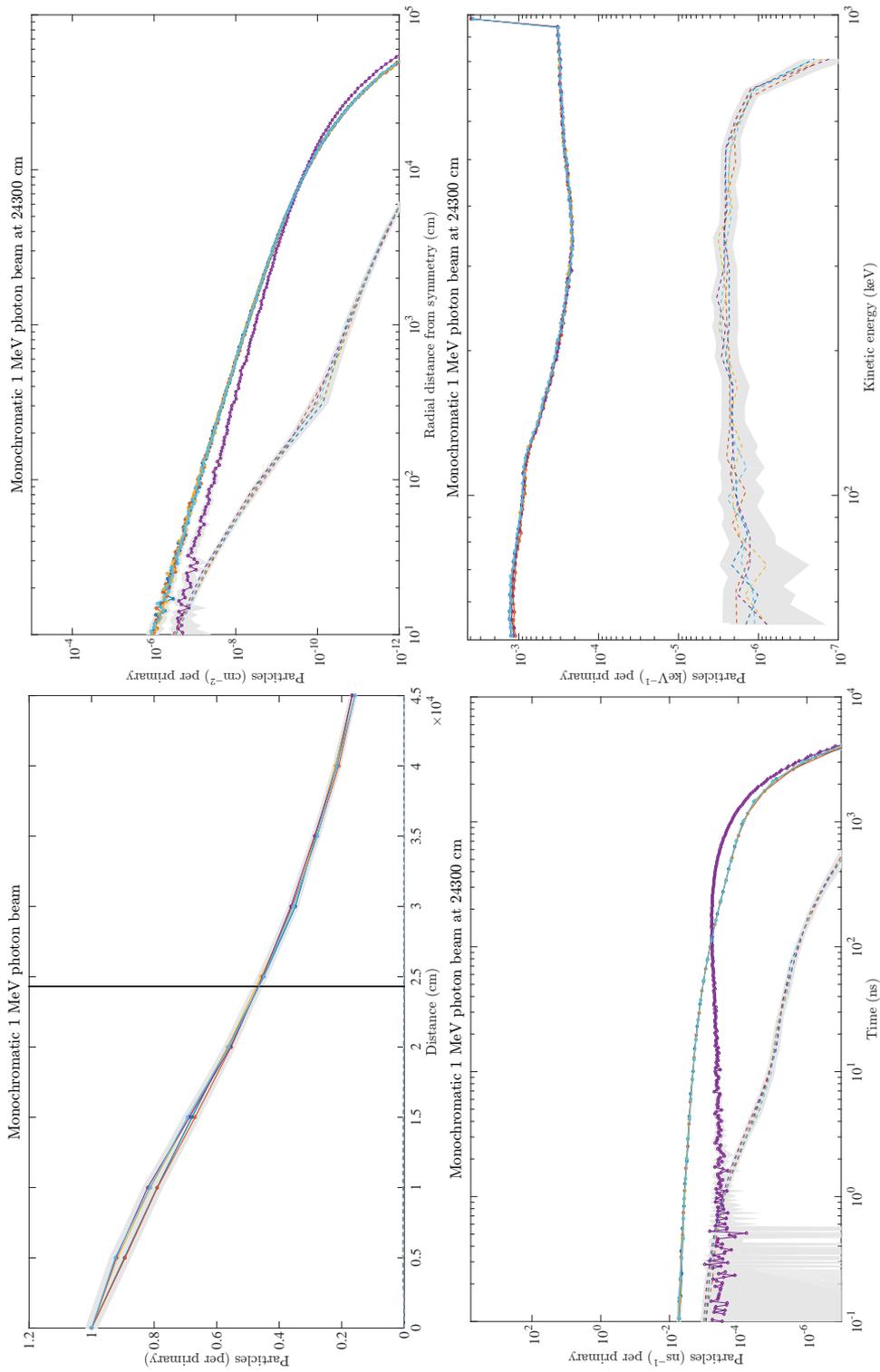
Figure 7.14: 1 MeV photon beam.
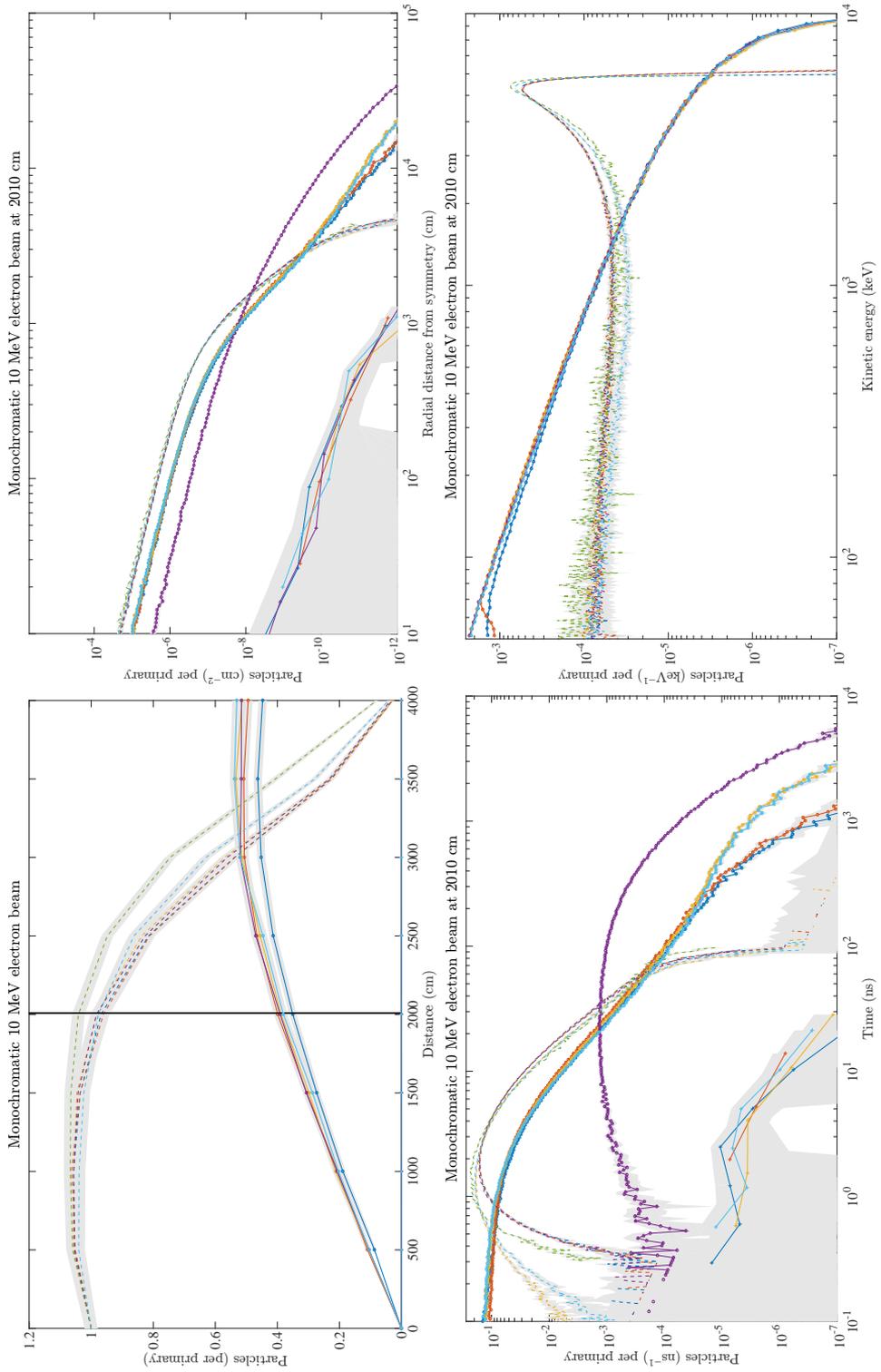
Figure 7.15: 10 MeV electron beam.
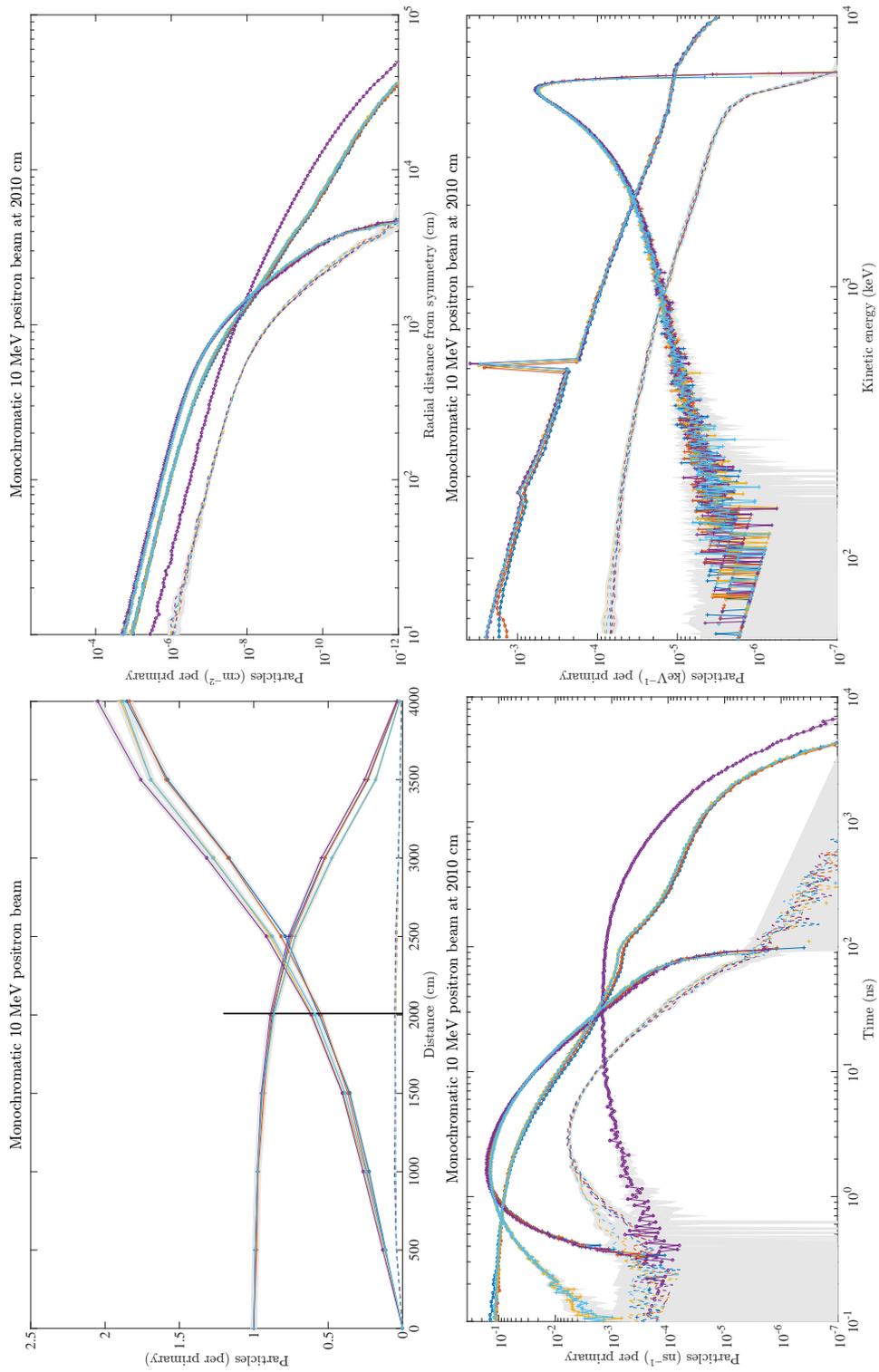
Figure 7.16: 10 MeV positron beam.
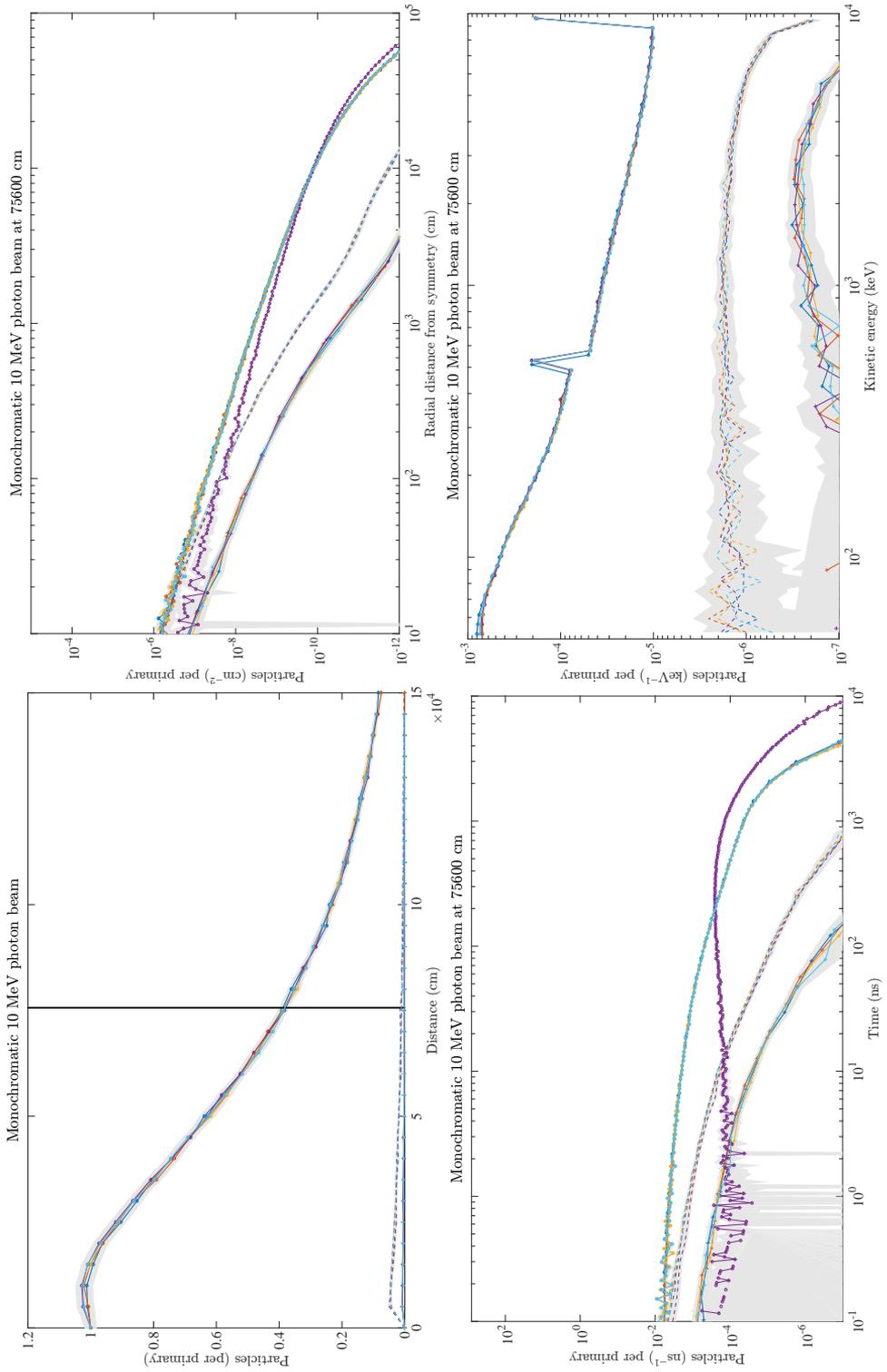
Figure 7.17: 10 MeV photon beam.
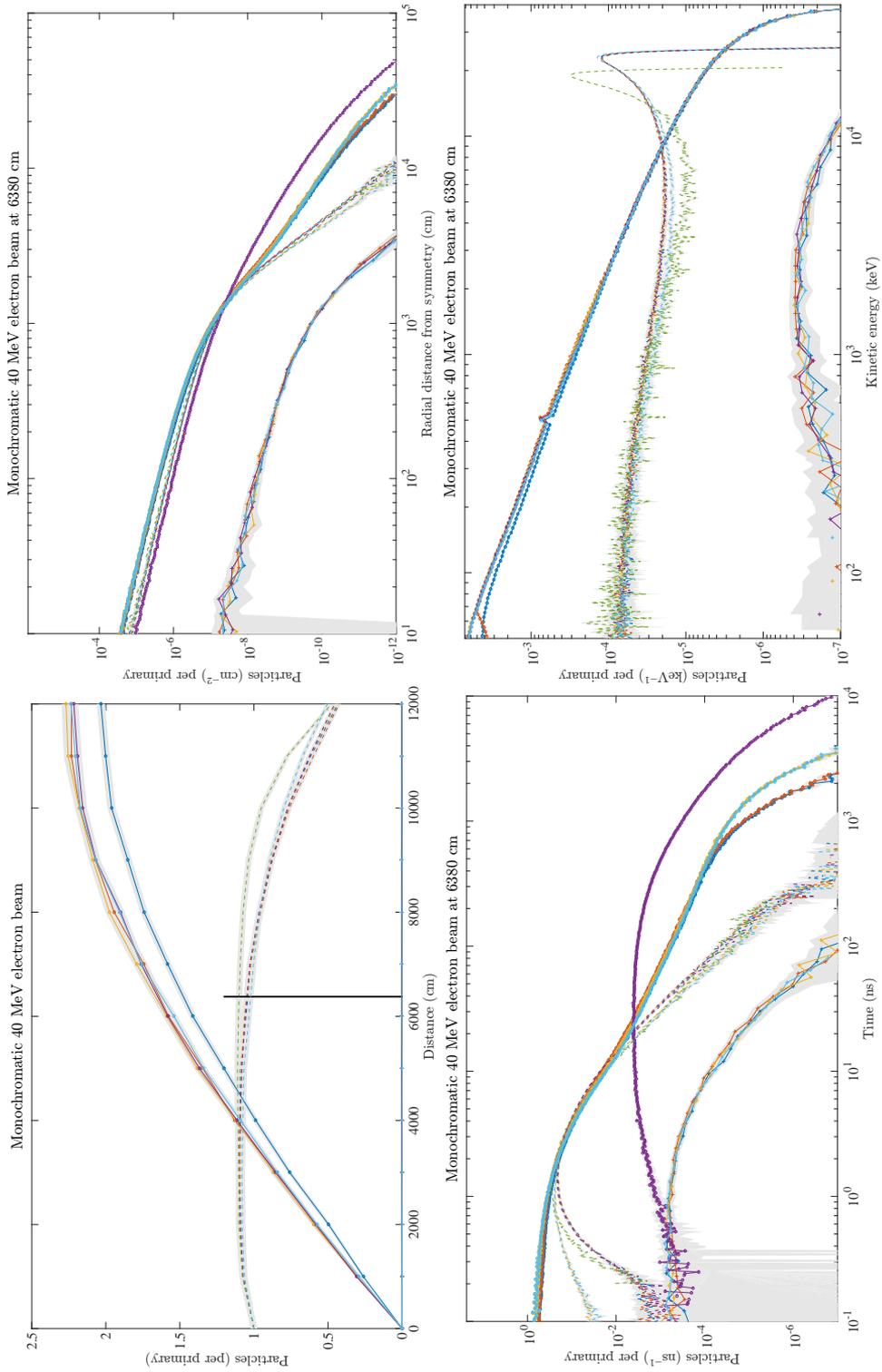
Figure 7.18: 40 MeV electron beam.
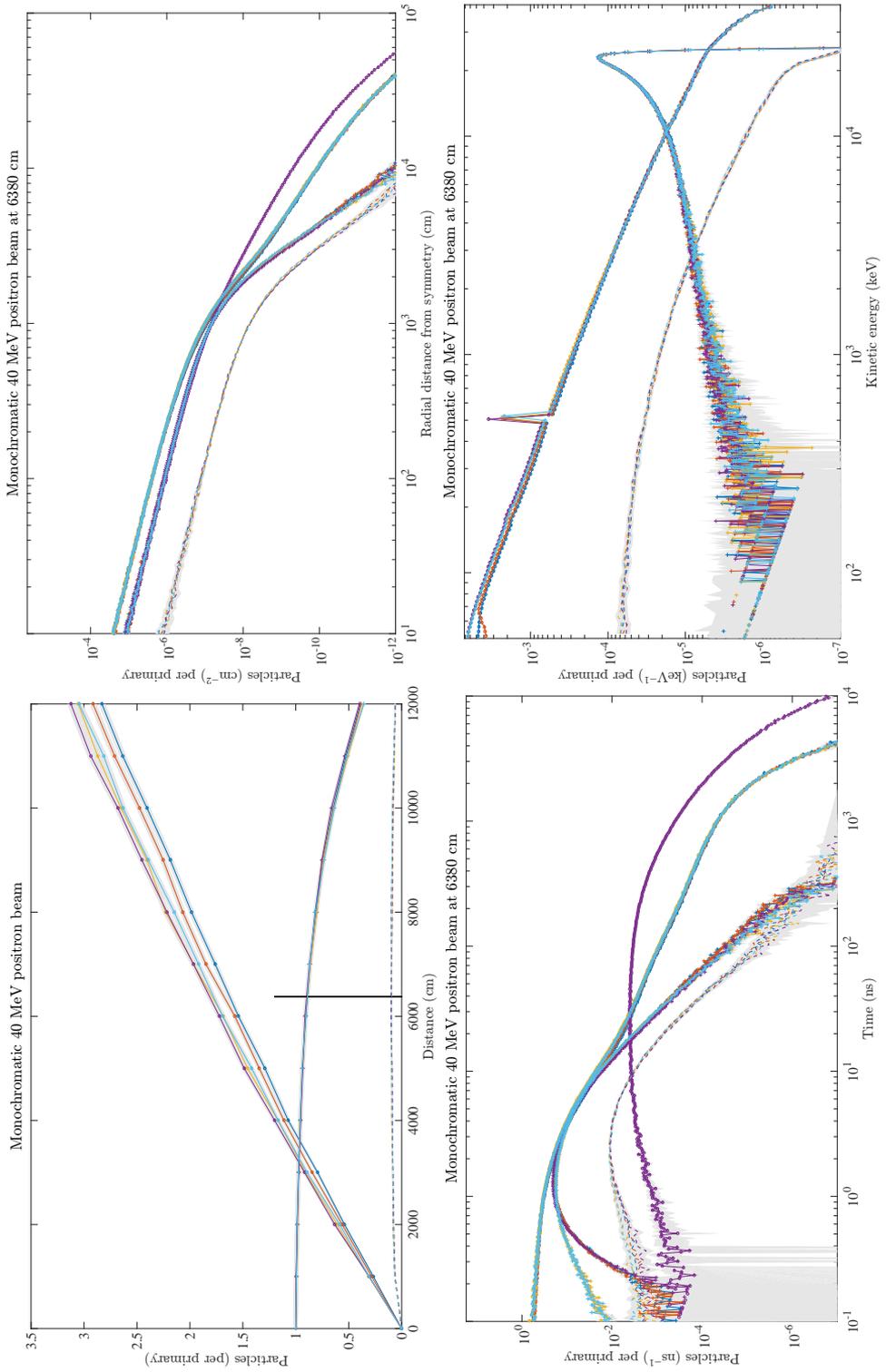
Figure 7.19: 40 MeV positron beam.
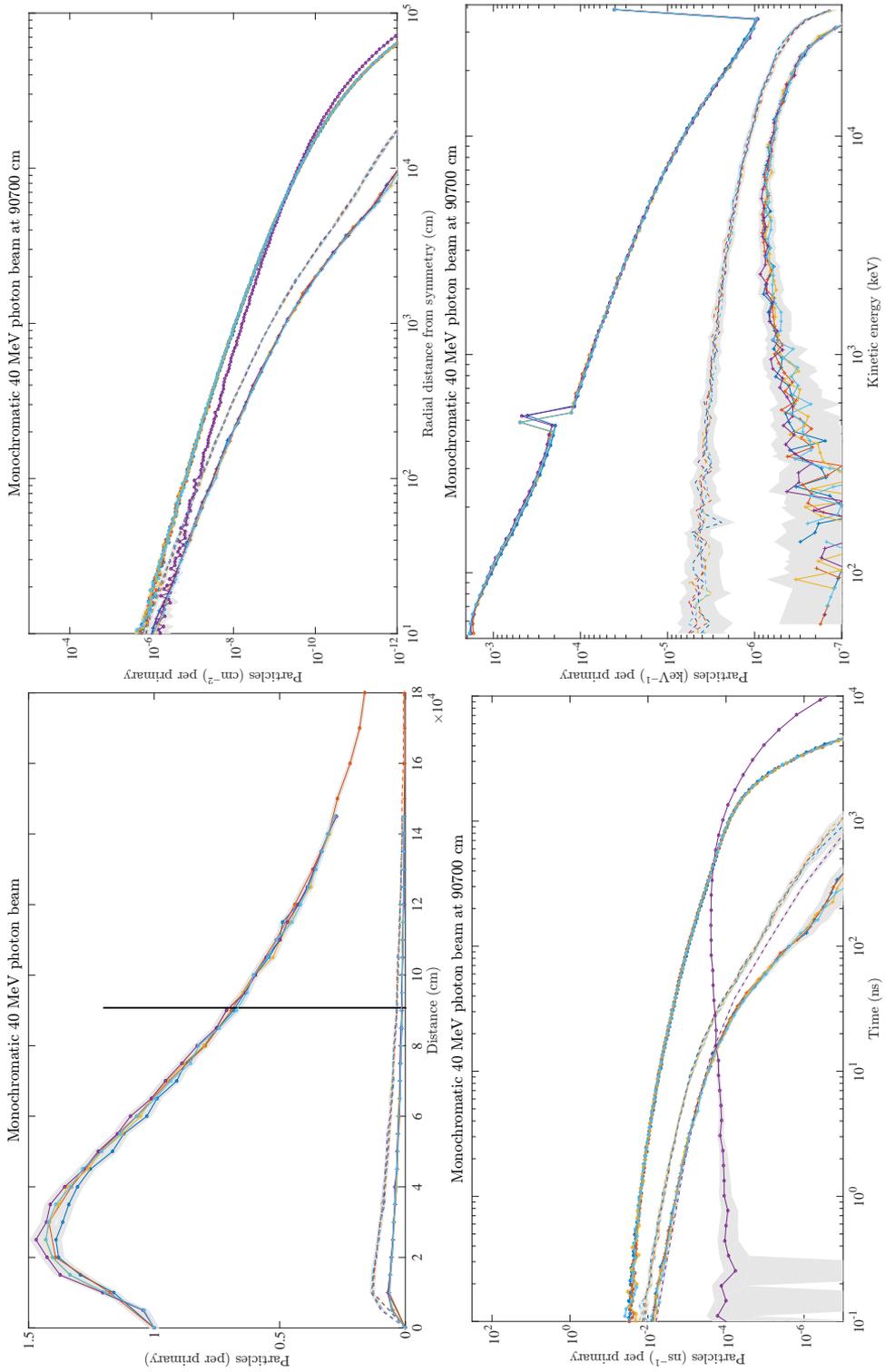
Figure 7.20: 40 MeV photon beam.

Chapter 8

Evaluation of Monte Carlo methods in high atmospheric physics II

The emerging field of High Energy Atmospheric Physics (HEAP) studies events producing high energy particles associated with thunderstorms, such as terrestrial gamma-ray flashes, electronpositron beams, thunderstorm ground enhancements and gamma-ray glows. Understanding these phenomena requires appropriate models, for the interaction of high energy electrons, positrons and photons with air. In our first study [115] (i.e. chapter 7), we investigated the results and performance of several Monte Carlo codes in the absence of electric fields. In this paper, we investigate the results of several codes used in the HEAP community when the effects of the electric fields are included, more precisely when its magnitude is high enough to produce Relativistic Runaway Electron avalanches (RREA).

First, we provide two simple set-ups that could be used by the community for future benchmark of codes that are not included in this study. The first one checks the probability to create a RREA from a given seed electron inside a given electric field, and the second one checks the full properties of the RREA at different records in time and space and for several electric field magnitudes. Then we proceed to use these two benchmarks on three codes used in the HEAP community: Geant4, GRRR, and REAM.

We found that different implementations of step-length limitations (null collisions) are responsible for significant differences in codes. For Geant4, appropriate step-length limitations are mandatory to get accurate results. We show that accurate RREA simulations can be achieved with an energy cut-off of 10 keV and higher (depending on the electric field magnitude), allowing for fast simulations. Precise modeling of the interactions of particles below 10 keV (e.g. by using more precise and slower models that include atomic shell effects) provides negligible differences.

8.1 Introduction

8.1.1 Phenomena and observations in high energy atmospheric physics

Wilson [116] had the conceptual awareness of the generation of high energy radiation from thunderclouds, 60 years in before recent theoretical publications and observational results [117]. Traditionally, thunderclouds have been studied by classical electromagnetism, but understanding the production and propagation of high energy radiation from thunderstorms falls in the context of relativistic High Energy Atmospheric Physics (HEAP). A review is provided by Dwyer et al. [62].

Different penetrating radiation types have been observed from thunderclouds, naturally categorized by duration. Microsecond-long burst of photons, which were first observed from space [53, 105], are known as Terrestrial Gamma-ray Flashes (TGF). TGFs produce bursts of electron positron pairs [54, 55, 100] that follow the geomagnetic field lines into space. Seconds to minutes or even hours long of penetrating X and gamma radiation have been observed on ground, from balloons and aircraft, by Kelley et al. [18], McCarthy and Parks [56], Eack et al. [57], Tsuchiya et al. [58], Adachi et al. [59], Chilingarian et al. [60, 61], Torii et al. [118], Kochkin et al. [119] (and references therein), which are called gamma-ray glows or thunderstorm ground enhancements. In between, on the millisecond duration, a new radiation mechanism has been introduced as TGF afterglows by Rutjes et al. [120] (i.e. chapter 13) and measured by Bowers et al. [121], Teruaki et al. [122], based on the intermediate neutrons produced by a TGF, creating a

prolonged and relocated signal. Emissions of lower energy particles (typically 100 keV) are seen in nanosecond pulsed discharges [74–77] and during the formation of long sparks [78–84] in high voltage and pulsed plasma technology.

Following the idea of Wilson [116], penetrating radiation is created by runaway electrons, which may further grow by the effect of Moller scattering in the form of so called relativistic run-away electron avalanches (RREA) [64, 85]. The difference in duration between TGFs and gamma-ray glows can be explained by two different scenarios to create run-away electrons, which traditionally are illustrated using the average energy-loss or friction curve. In this curve, there is a maximum at around $\varepsilon \approx 200$ eV, illustrating the scenario that for electric fields higher than the critical electric field, of maximally $E_c \approx 26$ MV/m at standard temperature and pressure (STP), thermal electrons can be accelerated into the run-away regime, described in the so-called Cold Runaway theory [65]. The actual critical electric field E_c may be significantly lower, as electrons could overcome the friction barrier due to their intrinsic random interactions [114, 123–125]. Cold Runaway could happen in the streamer phase [66, 68, 73, 114] or leader phase [69–72, 126] of a transient discharge, explaining the high energy electron seeding that will evolve to RREA that will produce a TGF gamma-rays as an effect of bremsstrahlung of the accelerated electrons. Alternatively, the relativistic feedback discharge model is also proposed to explain TGF production using large scale and high potential electric fields [88], where the RREA initial seeding may be provided by cosmic-ray secondaries, background radiation, or cold runaway [69, 127].

For fields significantly below the critical electric field E_c but above the runaway breakdown electric field of $E_{rb} = 0.28$ MV/m (STP), runaway behavior is still observed in detailed Monte Carlo studies [62]. At thundercloud altitudes, cosmic particles create energetic electrons that could runaway in patches of the thundercloud where the electric field satisfies this criterion. RREA are then formed if space permits and could be sustained with feedback of photons and positrons creating new avalanches [86–88]. Gamma-ray glows could be explained by this mechanism, as they are observed irrespectively of lighting or observed to be terminated by lightning [17, 18, 56, 119]. The fact that gamma-ray glows are not (necessarily) accompanied by classical discharges, results in the conclusion that the electric fields causing gamma-ray glows are usually also below classical breakdown. The classical breakdown field, of $E_k \approx 3.2$ MV/m at STP, is where low energy electrons (< 200 eV) exponentially grow in number as ionization overcomes attachment. In this work we focus on electric fields above runaway breakdown $E_{rb} = 0.28$ MV/m and below classical breakdown $E_k \approx 3.2$ MV/m to investigate RREA responsible for gamma-ray glows.

8.1.2   Theoretical understanding of RREA

In the energy regime of HEAP, the evolution of electrons is driven by electron impact ionization [115] (i.e. chapter 7), as this energy loss channel is much larger than the radiative (bremsstrahlung) energy loss (by a few orders of magnitude). However, this is only true for the average, and bremsstrahlung does have significant effect on the electron spectrum because of straggling [115]. For the electron impact ionization, straggling only occurs for thin targets, as the energy is much more unequally separated [115] (i.e. chapter 7). Almost all energy loss of ionization is going into producing secondary electrons of lower energy ($\epsilon_2 \ll 200$ eV), very frequently. For this reason it is fair to approximate that channel as continuous energy loss or friction because it does not alter the primary electron significantly.

In the case of electric fields above runaway breakdown, there is a possible mode where electrons, when considered as a population, keep on growing. Some individual electrons do not survive, but the ensemble grows exponentially as new electrons keep being generated. The production of secondaries, above a value much larger than the ionization threshold, can be approximated by the Møller cross section (see page 321 of [128]), which is the exact solution for a free-free electron-electron interaction:

$$
\begin{aligned}
\frac{\mathrm{d}\sigma_{\mathrm{M}}}{\mathrm{d}\Delta_2} \;=\; & Z\frac{2\pi r_e^2}{\gamma_1^2 - 1}\left[\frac{(\gamma_1 - 1)^2\gamma_1^2}{\Delta_2^2(\gamma_1 - 1 - \Delta_2)^2}\right. \\
& \left. - \frac{2\gamma_1^2 + 2\gamma_1 - 1}{\Delta_2(\gamma_1 - 1 - \Delta_2)} + 1\right],
\end{aligned}
\tag{8.1}
$$

where γ_1 is the Lorentz factor of the primary particle, $\Delta_2 = \gamma_2 - 1 = \epsilon_2/(m_e c^2)$ represents kinetic energy ϵ_2 of the secondary (in units of electron rest energy, with rest mass m_e), $r_e = \frac{1}{4\pi\epsilon_0}\frac{e^2}{m_e c^2} \approx 2.8 \times 10^{-15}$ m the classical electron radius and Z the number of electrons in the molecule. In the case $\Delta_2 \ll \gamma_1 - 1$ and $\Delta_2 \ll 1$, we see observe that the term $\propto 1/\Delta_2^2$ is dominating. Thus, we can write equation 8.1 as

$$
\frac{\mathrm{d}\sigma_{\mathrm{M}}}{\mathrm{d}\Delta_2} \approx Z\frac{2\pi r_e^2}{\beta_1^2}\frac{1}{\Delta_2^2},
\tag{8.2}
$$

with $v_1 = \beta_1 c$ the velocity of the primary particle. Integrating equation (8.2) for Δ_2 to the maximum energy ($\epsilon_1/2$) yields a production rate

$$
\sigma_{\mathrm{prod}} \approx Z\frac{2\pi r_e^2}{\beta_1^2}\frac{1}{\Delta_2} \propto \frac{1}{\epsilon_2},
\tag{8.3}
$$

using again $\epsilon_2 \ll \epsilon_1$. The remaining sensitivity of σ_{prod} to the primary particle is given by the factor β_1^2 which converges strongly to 1 as the mean energy of the primary electrons exceeds 1 MeV. In other words, as the mean energy of the electrons grows towards even more relativistic energies, the production rate σ_{prod} becomes independent of the energy spectrum.

8.1.3   One dimensional deterministic case

For illustrative purposes, we consider the one dimensional deterministic case, which results in an analytical solution of the electron energy spectrum. We make the system deterministic by assuming that the differential cross section is a delta-function at $\epsilon_2^{\min}$ and use $\lambda_{\mathrm{prod}} = \frac{1}{N\sigma_{\mathrm{prod}}}$ as the constant collision length, with N the air number density. In other words, every length λ_{prod} a secondary electron of energy $\epsilon_2^{\min}$ is produced. The derivation below is equivalent to what was presented by Dwyer et al. [62], Skeltved et al. [103], Celestin and Pasko [129] and references therein.

Consider a population of electrons in 1-dimension with space-coordinate z, a homogenous and constant electric field E above runaway breakdown and a friction force $F(\epsilon)$. The minimum energy $\epsilon_2^{\min}$ that can runaway is given by the requirement $F(\epsilon_2^{\min}) > E$, that is to say $\epsilon_2^{\min} = \mathrm{function}(F, E)$ is constant. Assuming that the mean energy of the ensemble is relativistic, results in a constant production rate $\lambda_{\mathrm{prod}} = \lambda_{\mathrm{prod}}(\epsilon_{\min})$. In space, the distribution f_e thus grows exponentially as

$$\frac{\partial f_e}{\partial z} = \frac{1}{\lambda_{\mathrm{prod}}} f_e. \tag{8.4}$$

The differential equation for energy depends on the net force,

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}z} = qE - F(\epsilon). \tag{8.5}$$

Solving for homogeneity means

$$\frac{\mathrm{d}f_e}{\mathrm{d}z} = \frac{\partial f_e}{\partial z} + \frac{\partial f_e}{\partial \epsilon}\frac{\mathrm{d}\epsilon}{\mathrm{d}z} = 0, \tag{8.6}$$

and using equation 8.4 and 8.5 results in,

$$\frac{\partial f_e}{\partial \epsilon} = -\frac{1}{\lambda_{\mathrm{prod}}(qE - F(\epsilon))} f_e. \tag{8.7}$$

For the largest part of the energy spectrum, specifically above 0.511 MeV and below 100 MeV, $F(\epsilon)$ is not sensitive to ϵ, see [115]. Only at around an electron energy $\epsilon \approx 100$ MeV, $F(\epsilon)$ starts increasing again because of bremsstrahlung. One may thus assume $F(\epsilon) \approx F$ constant, which yields that the RREA energy spectrum $f(\epsilon)$ at steady state is given by

$$f_e(\epsilon) = \frac{1}{\bar{\epsilon}} \exp\left(-\frac{\epsilon}{\bar{\epsilon}}\right) \tag{8.8}$$

with the exponential shape parameter and approximated average energy $\bar{\epsilon}(E)$ given by

$$\bar{\epsilon}(E) = \lambda_{\mathrm{prod}}(qE - F). \tag{8.9}$$

Equivalently, in terms of collision frequency $\nu_{\text{prod}} = \frac{\beta c}{\lambda_{\text{prod}}}$, equation 8.9 can be written as

$$\bar{\epsilon}(E) = \frac{\beta c}{\nu_{\text{prod}}}(qE - F), \tag{8.10}$$

with β the velocity v/c of the RREA avalanche front. For the 1-d case there is no momentum-loss or diffusion, so $\beta \approx 1$. Remark that λ_{prod} depends on $\epsilon_2^{\min} = \epsilon_2^{\min}(E)$, which is depenent on the electric field E as that determines the minimum electron energy that can go into runaway.

With this simple analysis we illustrate with equation 8.8 and 8.9, that the full RREA characteristics, such as the multiplication length λ_{prod} or the mean energy $\bar{\epsilon}$ are driven by processes determining $\epsilon_2^{\min}$. See for illustration figure 8.1.

8.1.4   Three dimensional stochastic case

In reality there are important differences to the one dimensional deterministic case described above, which we discuss qualitatively for understanding the Monte Carlo simulations evaluated in this study.

During collisions, electrons deviate from the path parallel to E. Therefore in general, electrons experience a reduced net electric field as the cosine function of the opening angle θ, which reduces the net force to $qE\cos(\theta) - F$ and thereby the mean energy $\bar{\epsilon}$ of equation 8.9. In reality of course θ changes along the path of the particle. Although the velocity remains still close to c (as the mean energy is still $>$ MeV), the RREA front velocity parallel to E is reduced again because of the opening angle as function of its cosine:

$$\beta_{\parallel} = \beta\cos(\theta), \tag{8.11}$$

which also reduces the mean energy $\bar{\epsilon}$. Or equivalently the avalanche scale length λ_{RREA} in 3-D is $\cos(\theta)\lambda_{\text{prod}}$. However most importantly, the momentum-loss of the lower energetic electrons results in a significant increase of $\epsilon_2^{\min}$, as it is much harder for electrons to runaway. The increase of $\epsilon_2^{\min}$ significantly increases λ_{prod} and thereby increases the characteristic mean energy $\bar{\epsilon}$. On the other hand, the stochasticity creates an interval of possible energies $\epsilon_2^{\min}$ that can run-away with a certain probability and for thin targets a straggling effect [115].

The effects discussed above prevent a straight forward analytical derivation of the RREA characteristics in 3 dimensions (as far as we know), but what remains is the important notion that the physics is completely driven by the production of electrons with intermediate energy. Intermediate electron energy in the sense that they are far above ionization threshold ($\gg 200$ eV) but much below relativistic energies ($\ll 1$ MeV). The parametrization of the electron energy spectrum, given by equation 8.9 turns out to be an accurate empirical fit, as it was already shown by Dwyer et al. [62], Skeltved et al. [103], Celestin and Pasko [129] and references therein. Nevertheless in these works $\lambda^{\min}(E)$, or
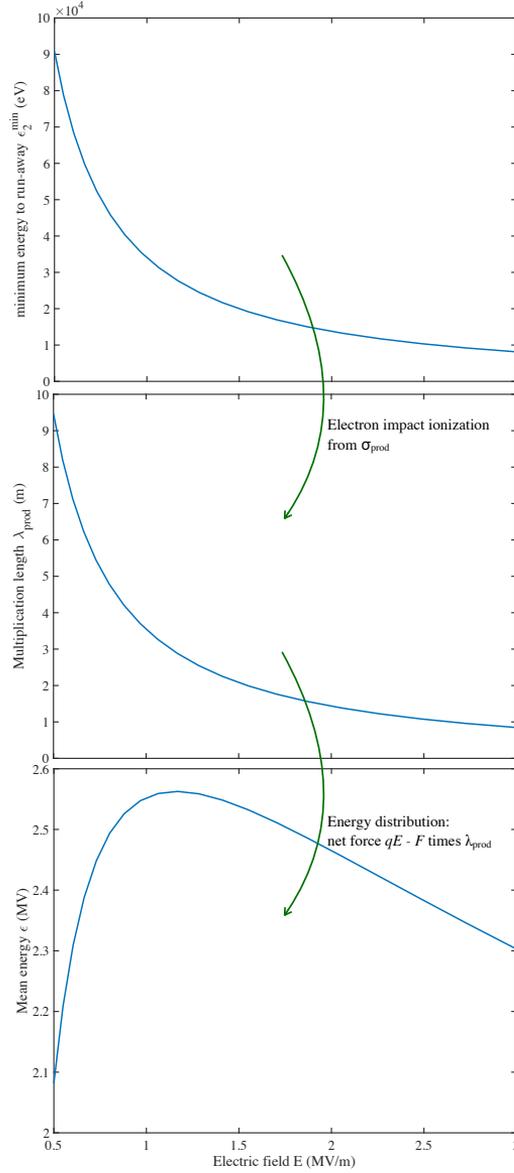
Figure 8.1: For the 1D-case, the mean energy as function of E (bottom panel) is derived from equation 8.9, a function of the production rate λ_{prod} (middle panel), from the cross section σ_{prod}, see equation 8.3. λ_{prod} (middle panel) is in turn a result of the minimum energy to run-away $\epsilon_2^{\min}$ (top panel) calculated from the requirement $F(\epsilon_2^{\min}) > E$ (i.e. the friction curve).

equivalently the velocity over collision frequency $\beta c / \nu_{\text{prod}}$, is fitted by numerical Monte Carlo studies and the final direct relation to $\epsilon_2^{\min}$ is not executed. Celestin and Pasko [129] calculated that $\nu^{\text{prod}}(E) \propto E$, which explains why $\bar{\epsilon}(E)$ must

saturate to a constant value. Celestin and Pasko [129] argue that $\epsilon_2^{\min}(E)$ is given by the deterministic friction curve F (as illustrated in figure 8.1), for which they use Bethe's formula and an integration of a more sophisticated electron impact ionization cross section (RBEB) including molecular effects, but that is only true in 1-D without stochastic fluctuations.

8.1.5   Content and order of present study

Here we continue the work of Rutjes et al. [115], now in the case of electric fields, and the results we will present can also be considered as a continuation of the work of Skeltved et al. [103]. We evaluate the main parameters governing RREAs and see how to optimize the low energy cutoff given the electric field.

During our study we found that reducing the low energy cutoff results in smaller integration steps which can also be achieved in different ways, resulting in a different interpretation of the conclusions of Skeltved et al. [103] and outlook of Rutjes et al. [115], that not the low energy cutoff is responsible for the differences but the average step-lengths the particles have to travel between collisions. Furthermore, as introduced in section 8.1.2, the RREA characteristics must be driven by physics around $\epsilon_2^{\min}$ and not by physics at energies below 1 keV.

To prove our new insights and to benchmark codes capable of computing RREA characteristics for further use, we first calculate the probability for an electron to accelerate into the runaway regime, which is practically the quantity $\epsilon_2^{\min}(E)$. For the definition of this probability see section 8.2.3. This study is similar to the works of [114, 123–125], but here with higher resolution, clearer definition and for multiple codes presented as one comparison. From the probability study is it directly clear that it is safe to choose the low energy cutoff ε_c higher than previously expected by Skeltved et al. [103], Rutjes et al. [115], even higher than $F(\varepsilon_c) > E$ given an electric field $E < E_k$. The probability that these particles below ε_c accelerate and participate in the penetrating radiation is negligible. In practice a value of $\varepsilon_c \approx 10$ keV can be used for any electric field below the classical breakdown field.

The benchmark set-ups, together with data and figures to compare with, are given in detail in the appendix.

The new insight presented is that step-length restrictions (i.e. the implementation of null collisions) are of major importance, which is described in section 8.2.2. Differences seen in the RREA characteristics between implementations are discussed in section 8.3.3. We conclude in in Sec. 8.4.

8.2   Model correctness and optimization

8.2.1   Computational models and physicslists

In HEAP, some researchers use general purpose Monte Carlo codes developed by large collaborations like Geant4 (used by Bowers et al. [121], Carlson et al. [130], Sarria et al. [131] and by Skeltved et al. [103]) or FLUKA (used by Dubinova et al. [26], Rutjes et al. [120]). On the other hand, to model, e.g., the radiation sources with self-consistent time dependent electric fields, other researchers develop custom made codes in small groups or as individuals, where the cross sections and numerical methods may come from already validated theory (e.g. Sarria et al. [24], Kohn et al. [72]).

Rutjes et al. [115] list in their section 1.3 the reasons why differences between codes occur and why it may be very complicated (if not impossible) to estimate and find differences studying the source files. Here we continue the work in [115] (i.e. chapter 7), to provide a comparison standard for the particle codes, as simple and informative as possible, by only considering their physical outputs. In this work the benchmark data is provided by the general-purpose code Geant4 (two different set-ups) and two custom-made codes GRRR and REAM.

Geant4

Geant4 is a software toolkit developed by the European Organization for Nuclear Research (CERN) and a worldwide collaboration [23, 132, 133]. We use the version 10.2.3. The electro-magnetic models can simulate the propagation of photons, electrons and positrons including all the relevant processes, and the effect of arbitrary electric and magnetic fields. Geant4 uses steps in distance, whereas REAM and GRRR use time steps. In the context of this study, three main different electro-magnetic cross-section sets are included, one based on analytical or semi-analytical models (e.g. the Moller cross section for ionization and Klein-Nishina cross section for Compton scattering), one based on the Livermore data set [109], and one based on the Penelope models [113]. Each of them can be implemented with a large number of different electro-magnetic parameters (binning of the cross section tables, energy thresholds, production cuts, maximum energies, multiple scattering factors, accuracy of the electro-magnetic field stepper, among others). Skeltved et al. [103] used two different physics list: LHEP and LBE. The first one, based on parametrization on measurement data and optimized for speed, was depreciated since the 10.0 version of the toolkit. The LBE physics list is based on the Livermore data, but it is not considered as the most accurate physics list in the Geant4 documentation. This last is given by the Option 4 physics list, that uses a mix of different models, and in particular the Penelope model for the the impact ionization of electrons. For this study, we will use two GEANT4 physics list options: Option 4 (referred to as O4 hereafter) that is the most accurate one according to the documentation, and

the Option 1 (referred to as O1 hereafter) that is less accurate and runs faster. In practice, O1 and O4 give very similar results for simulations without electric field and energies above 50 keV, as produced in our previous code comparison study [115].

Concerning how energy thresholds are handled, Geant4, by default, is following all primary particles down to zero energy. A primary particle is a defined as a particle with more energy than a threshold energy E_{th}^p. The O4 option follows this default behavior, but the O1 electro-magnetic model does not and proceeds to discrete energy loss above E_{th}^p, and continuous energy loss below. The default E_{th}^p is set to 990 eV and is not changed in this study. The LBE Physics list used in [103] uses a threshold down to 250 eV, and this parameter was thought to be responsible for a major change in the simulated steady-state RREA energy spectra. In this work, we will argue that this was a mis-interpretation, and that the main parameter governing the spectrum is the maximum allowed step ($\ell_{\max}$) of the tracking of the particle that undergoes friction from air and acceleration from the electric field. Note that this mis-interpretation was possible because reducing the value of E_{th}^p indirectly also reduces the average step that is used.

REAM

The REAM (Runaway Electron Avalanche Model) is a three dimensional Monte Carlo simulation of Relativistic Runaway Electron Avalanche (also refereed to as Runaway Breakdown), including electric and magnetic fields [85, 87, 134]. This code is inspired by earlier work by Lehtinen et al. [123] and takes accurately into account all the important interactions involving runaway electrons, including energy losses through ionization, atomic excitation and Moller scattering. A shielded-Coulomb potential is implemented in order to fully model elastic scattering, and it also includes the production of X/gamma-rays from radiation energy loss (bremsstrahlung) and the propagation of the photons, by including photoelectric absorption, Compton scattering and electron/positron pair production. The positron propagation is also simulated, including the generation of energetic seed electrons thought Bhabha scattering.

One technical detail (relevant for this study) is that REAM limits the time step so that the energy change within one time step cannot be more than 10%. The effect of reducing this factor down to 1% was tested and did not make any noticeable difference in the resulting spectra.

8.2.2   Null collisions

In Monte Carlo simulations, particles step, collide and interact with surrounding media by means of cross sections (and their derivatives). A particle simulation can be either *space-oriented* or *time-oriented*, if it steps in space or in time, respectively.

During steps however, charged particles can change energy when electric fields are present. To guarantee accuracy, energies should be updated frequently enough, by exponentially sampling with mean:

$$\bar{t}_{\text{step}} = \min_{\epsilon}\{(v(\epsilon)\sigma_t(\epsilon)N)^{-1}\}, \tag{8.12}$$

in time-oriented perspective or equivalently,

$$\bar{\ell}_{\text{step}} = \min_{\epsilon}\{(\sigma_t(\epsilon)N)^{-1}\}, \tag{8.13}$$

in space-oriented perspective, with v the velocity, σ_t the total cross section and N the density. Then at each updated location (and energy) the type of collision must be sampled, including the possibility of a so-called null-collision. Simply stated, null-collisions are collisions that do nothing, and only guarantee a correct update of particle energy, which is possible as the exponential distribution is memoryless.

In some cases, this would result in impractical too many collisions, so instead a different approach is used. The step-length ℓ at start location is sampled as

$$\ell(\epsilon) = -\log(\xi)\lambda(\epsilon), \tag{8.14}$$

where ξ is a uniform random variable between (0,1] and $\lambda(\epsilon) = (\sigma_t(\epsilon)N)^{-1}$ is the mean free path of the considered interactions, and the time step δt as

$$\delta t(\epsilon) = -\log(\xi)\tau(\epsilon), \tag{8.15}$$

where $\tau(\epsilon)$ is the inverse of the collision frequency $(v(\epsilon)\sigma_t(\epsilon)N)^{-1}$. To limit the error made in this approximation, null-collisions are used when the step length (or time) is above a threshold set by the user, with the so-called 'max step length' ℓ_{max}.

In space-oriented simulations null-collisions are generally used to cope with internal boundaries, usually of different materials and/ or different densities, and for scoring purposes. After each step the distance to the closest (internal) boundary d_{wall} is calculated and steps are only accepted if they are smaller than this distance. If a sampled step-length is larger than d_{wall}, the particle is just moved to the boundary (without a collision) and stored, enabling in the next sampling to use the cross section and density of the other material.

By construction, space-oriented simulations are thus not synchronous in time. Usually, a single particle is simulated over its entire lifetime before going to the next particle. This procedure makes it impossible to incorporate particle to particle interactions, such as a space charge electric field, because then information must by available synchronously. A big advantage of asynchronous simulations is that, besides the ability to include boundaries, particles step as far as is possible in the same material, minimizing the overhead due to null collisions. Particle codes used in this study or in [115] that are space-oriented are EGS5 [21], FLUKA [22], Geant4 [23] and MC-PEPTITA [24].

In time-oriented simulations null-collisions enable also synchronizability, where all particles are at the same moment in time after each user defined time interval. This is enforced by letting a particle step and collide up to the point that the remaining time t_{left} plus the sampled time δt to the next collision is larger than Δt, after which the particle is moved a time $\Delta t - t_{\text{left}}$ (without a collision) and is stored. Before continuing the same particle, first all other particles are processed a Δt. For self-consistent electrical discharge simulations space charge is most be incorporated. Particle codes that are time-oriented are GRRR.

Using max step length is an approximation, sensitive to the electric field

In contrast to the exact null collision implementation, i.e. sampling random steps with average by equation 8.12 or 8.13, using a length step limiter is an approximation with accuracy sensitive to the electric field. We explain why and what the expected error is.

The evolution of electrons in an electric field is essentially given by the net force, see equation 8.5. Moving and colliding particles with steps described by equations 8.14 and 8.15 is thus integrating equation 8.5 in a discrete manner. This numerical integration is essentially worse than first order, because besides updating the energy the mean free path (or collision time) is determined by the energy before the step. In other words, during the step the energy changes but this does not change the probability of collisions as that was set before the step.

Naturally, lowering the low energy cutoff ϵ_c, increases the integration interval of the differential cross sections and thereby the total cross section, lowering the mean free path, thus lowering step lengths and improving this integration. Not because lower energy physics is important or needed, but just because lower energy physics result in smaller steps. Smaller steps can also be achieved by maximizing steps, with the so-called 'max step length' ℓ_{max}.

For example in Geant4 the "max_step" ℓ_{max} option is available and defined as the maximum allowable step before rejection (it can be either specified in the Transportation Manager or in the Physics List). If a sampled random step ℓ, see e.g. equation 8.14, is larger than ℓ_{max} the particle is moved a distance ℓ_{max} using a null-collision and then sampled again.

To illustrate this, we calculate the relative error in mean free path for ionization of an electron in an electric field as function of max step length, see figure 8.2. The relative error in the mean free path is calculated by the relative difference of the mean free path before and after a step of length $\ell = \min(\lambda, \ell_{\text{max}})$. Visible in figure 8.2 is the saturation of the error due to the mean free path 8.2, i.e. the step length in this illustration does not exceed the mean free path. In reality step length could be randomly chosen, see equations 8.14 and 8.15, but on average it is limited by the mean free path (or collision frequency).
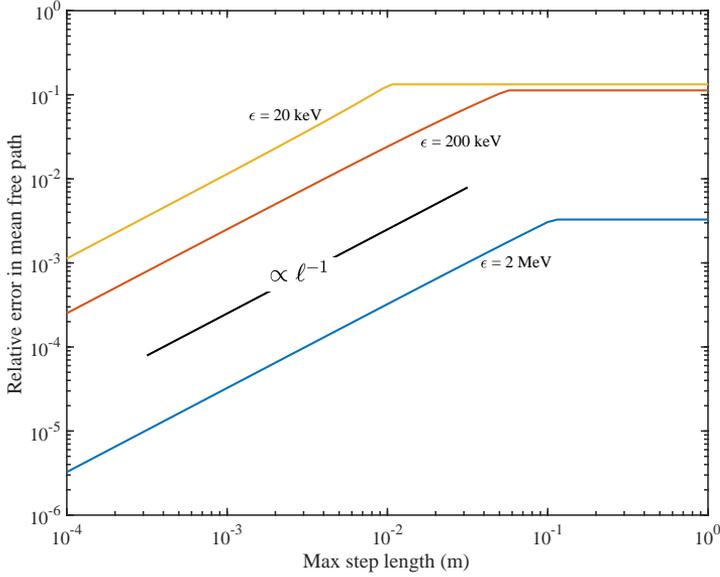
Figure 8.2: Relative error versus max step $\ell_{\max}$ for initial 20 keV, 200 keV and 2 MeV electron in an electric field of 1.2 MV/m. The error saturates when the mean free path becomes smaller than the max step, limiting the integration error by the physical step length.

8.2.3   Avalanche probability

As theoretically explained in section 8.1.3 and illustrated in figure 8.1, the final electron spectrum is essentially driven by the minimum energy $\epsilon_2^{\min}$ that can create a RREA. We calculated the probability that an electron accelerates into the run-away regime, given its initial energy and the electric field. We define this probability as the fraction of electrons that create a RREA, i.e. accelerates further than the minimum of the friction cure (i.e. > 1 MeV) and create an avalanche of new runaway electrons. To be precise, we define the probability as the fraction of electrons that created an avalanche of 20 electrons above 1 MeV, given an initial energy and electric field. The number 20 is arbitrary, to be well above 1 but small enough for computational reasons. For some initial conditions, we tested also 25 and 30 electrons which run slower but give the same probability.

In figure 8.3 the avalanche probability is given for Geant4 O4, with max step length setting of $\ell_{\max} = 1$ cm. Visible in figure 8.3 is the wide interval of electron energies that can run-away with some probability. In figure 8.4, we compare the 10%, 50% and 90% contour lines of Geant4 O4, Geant4 O1 and GRRR which turn out to be different. Here Geant4 O1/O4 compared to GRRR have different max length step settings, which can explain the difference. As test we calculated the probability as function of max step length for O4, for the configuration indicated by the purple cross (20 keV and 2 MV/m), see figure 8.5.

Although this configuration has zero probability at $\ell_{\max} = 1$ cm, the probability increases as the max step length decreases, which is consistent with the outcome of the GRRR simulation. In figure 8.5 it is visible that the RREA probability saturates for max step length setting below $\ell_{\max} < 10^{-4}$ m. For decreasing $\ell_{\max}$ below 0.5×10^{-4} m, however, the probability first decreases and then increases again. This is possibly because of roundoff errors, due to too many null collisions.

In addition, in figure 8.3 Geant4 O1 and O4 are very similar for energies above 50 keV, but quite different below 50 keV. The plausible reason is that O4 has more lower energy physics and thereby smaller mean free paths than O1, which below 50 keV limits the length step of O4 but not O1.
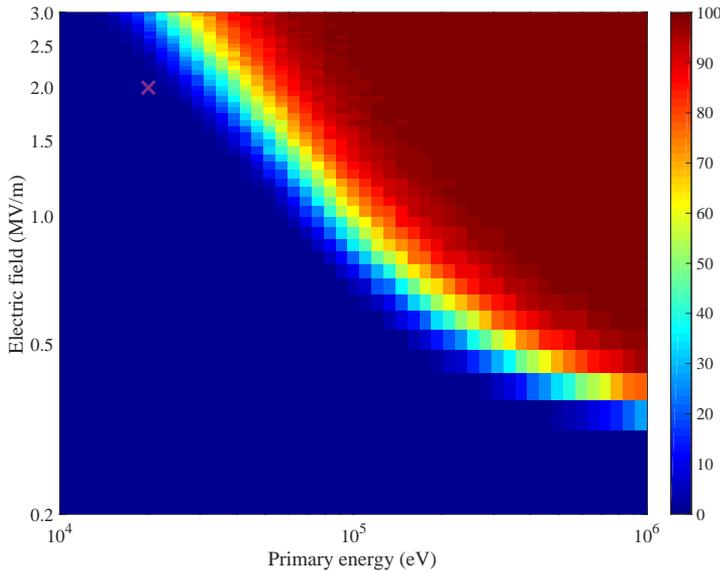
Figure 8.3: Avalanche probability versus the electron energy and electric field for O4 with max step setting of 1 cm. The purple cross highlights where we studied the behavior of the max step setting for O4 to the probability to run-away, see figure 8.5.

8.3   Relativistic runaway electron avalanche

For constant air density and electric field above runaway breakdown, but below classical breakdown, the relativistic runaway electron avalanche (RREA) converges to a steady state solution upon normalization of the total number of particles in the system. In section 8.1.2 we derived the reason of this behavior; the exponential growth of the secondaries dictates the spectrum.

First we show the avalanche speed and scales in space and time and the evolution to steady state. Secondly, we show the RREA characteristics in steady
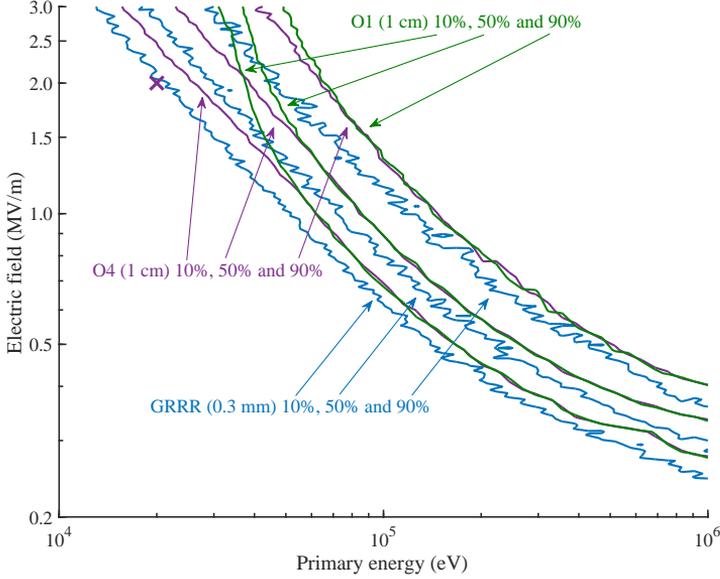
Figure 8.4: Avalanche probability comparison between O4 ($\ell_{\max} = 1$ cm), O1 ($\ell_{\max} = 1$ cm) and GRRR ($\ell_{\max} = 0.3$ mm), with three contour probabilities 10%, 50% and 90%. These contours are derived from the full probability scan, see figure 8.3 as example of O4. The purple cross highlights where we studied the behavior of the max step setting for O4 to the probability to run-away, see figure 8.5.

state.

8.3.1 Avalanche scale and velocity

The velocity of the avalanche parallel to the field is given in figure 8.6. We observe that GRRR is faster and O4 is slower than REAM/O1, although the differences and overall sensitivity towards the electric field E is small. The velocity β_z only changes from 0.83 to 0.9, see figure 8.6.

In figure 8.7 the calculated avalanche length and time scale are given and fitted by,

$$y = \frac{c_1}{x - c_2}, \tag{8.16}$$

where c_1 and c_2 are given in table 8.1. This empirical fit is motivated by, equation 8.9 and 8.10, derived for 1-D. As a result $c_1/c_2 \approx \beta_z c$, with $\beta_z \approx 0.88$, see figure 8.6. However, these fits remain empirical as they neglect the sensitivity of the mean energy and velocity to the electric field.
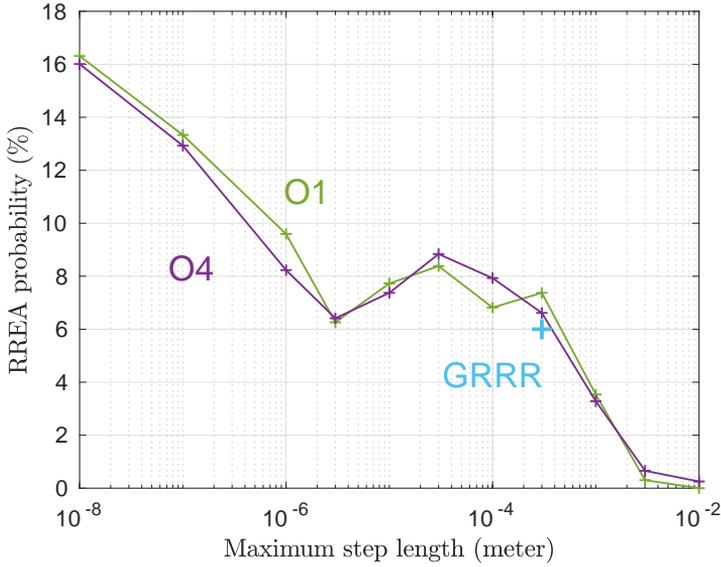
Figure 8.5: Avalanche probability versus max step length setting for O4, for electric field case $E = 2$ MV/m and $\epsilon = 20$ keV energy. This specific point is illustrated by a purple cross in figure 8.3 and 8.4. One point of GRRR (with 0.3 mm max step length) is added for comparison (blue cross).
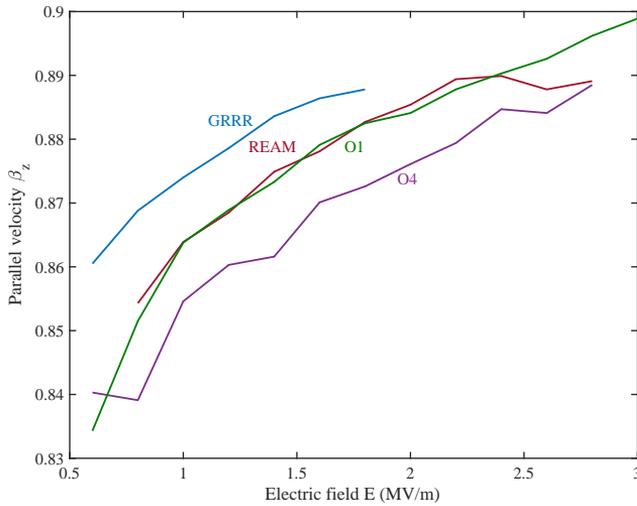
Figure 8.6: The parallel RREA velocity β_z in units of c as function of electric field E (in z direction) for several codes.

8.3.2   Evolution to steady state

All the RREA characteristic studies start with a monochromatic beam of 100 keV electrons (which are considered low compared to the reached steady-state mean
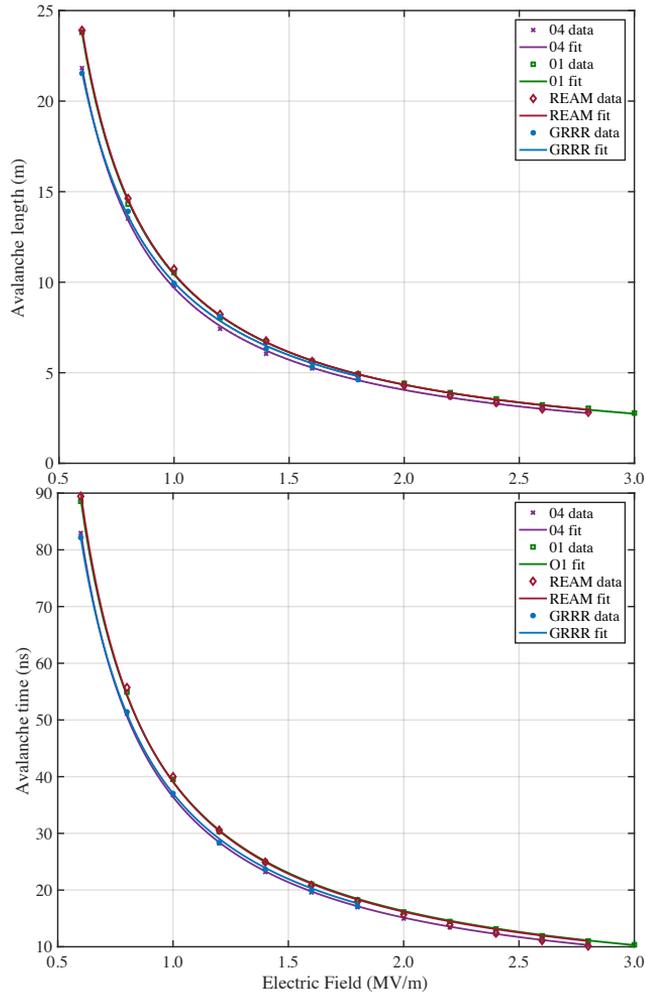
Figure 8.7: Avalanche multiplication length (top) and time (bottom) as function of ambient electric field, for each of the codes included in this study. See table 8.1 for fit parameters.

Table 8.1: Avalanche scale in space and time. For evaluated codes we fitted by equation 8.16. See figure 8.7 for results.

Code	Avalanche length		Avalanche time	
	c_1 (m MV/m)	c_2 (MV/m)	c_1 (ns MV/m)	c_2 (MV/m)
Geant O1	7.43	0.287	28.0	0.285
Geant O4	6.96	0.282	25.9	0.288
REAM	7.43	0.290	27.6	0.293
GRRR	7.41	0.257	27.0	0.272

energy of 6 to 9 MeV). We fitted the time evolution of the mean electron energy
by

$$\bar{\epsilon}(t) = a + b \times \exp(-t/T), \tag{8.17}$$

and defined the steady-state time to be equal to $T_{\mathrm{ss}} = 5T$, or five e-folding i.e.
converged to 99.3%, see figure 8.9.
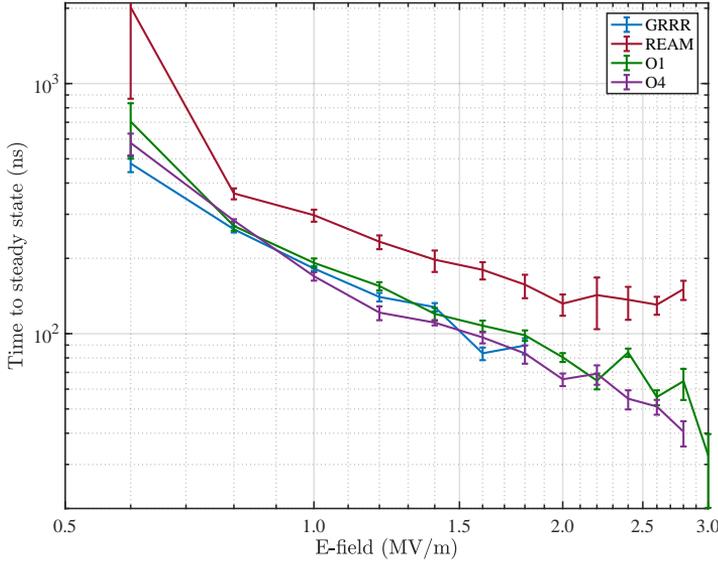
Figure 8.8: Time to steady state as function of ambient electric field, for each of
the codes included in this study. Time to steady state is defined as five e-folding
lengths, i.e. converged with in 99.3%, see equation 8.17.

In figure 8.7 we observe equivalent time to steady state for Geant O1, O4
and GRRR, in which the time correspondens with five e-folding lengths of time
scale seen in the bottom panel of figure 8.7. For REAM however, we observe a
significant longer time to steady state. The evolution to steady state is illustrated
for Geant4 O4, see figure 8.9.

8.3.3   Discussion RREA spectra

Electrons

At the location where the RREA is in steady state, we fit the energy spectrum
in a plane. For the empirical fit we use an exponential spectrum, motivated
by equation 8.8. The mean energy $\bar{\epsilon}$ of the exponential spectrum is calculated
for the several codes as function of electric field E, see figure 8.10. For Geant4
O1 and O4 the simulations and analysis were done twice, for a max step length
setting of $\ell_{\mathrm{max}} = 1$ cm and for $\ell_{\mathrm{max}} = 1$ mm. The final data was fitted again by
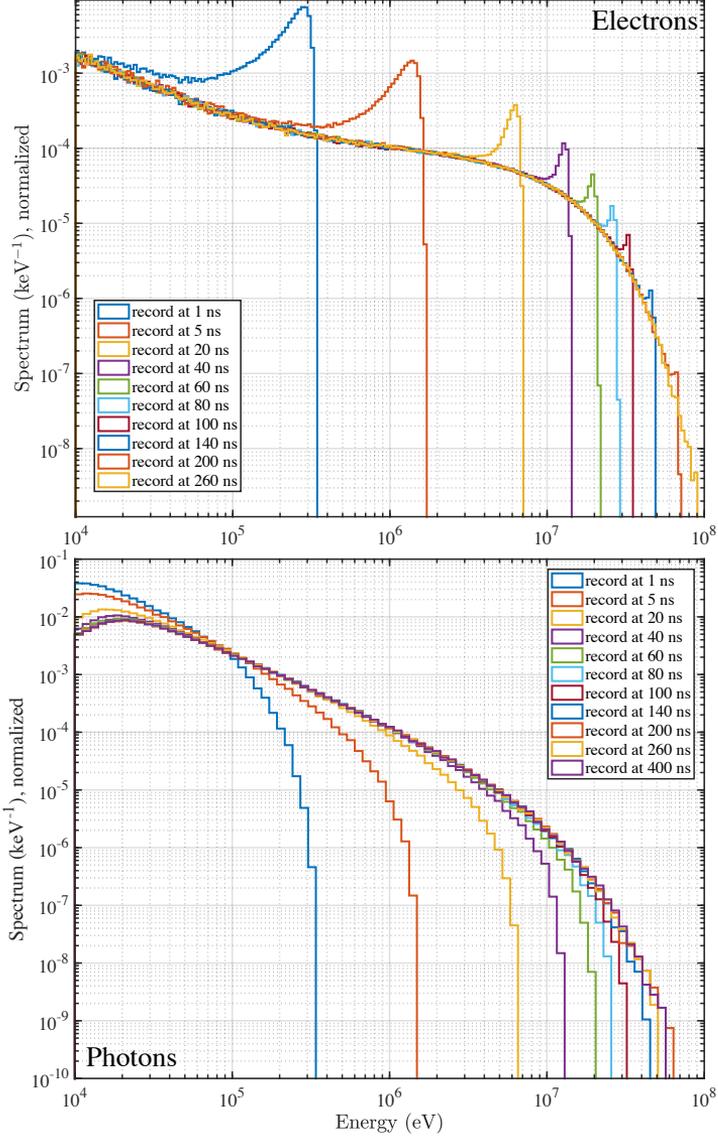
Figure 8.9: Time evolution of electron (top) and photon (bottom) spectrum for O4 with $\ell_{\max} = 1$ cm, and am electric field of 14×10^5 V/m

three parameters a_1, a_2 and a_3 with the ansatz

$$\bar{\epsilon}_{\mathrm{fit}}(E) = \lambda(E)(qE - F), \quad \lambda(E) = \beta c \left[a_1 \left(\frac{qE}{F} \right)^{a_2} + a_3 \right]^{-1}, \tag{8.18}$$

motivated by the fact that $\epsilon_2^{\min}$ is a power-law of E (see figure 8.3) and λ is a power-law of $\epsilon_2^{\min}$ (see equation 8.3), and we took $F = 0.28$ MV/m. The electric force is divided by F in $\lambda(E)$ so the parameter in Equation 8.18 is dimensionless.
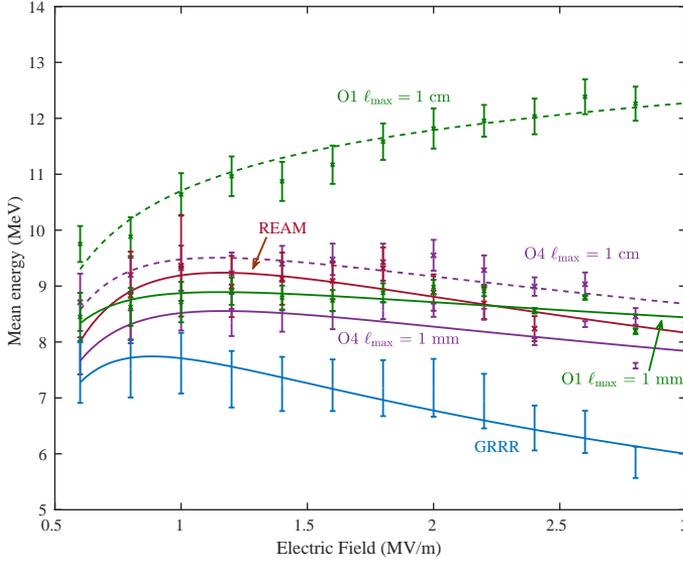
Figure 8.10: Fitted electron mean energies, including model.

β is chosen constant as 0.87 because the parallel RREA velocity does not change significantly over the range of electric fields, see Figure 8.6. The fits are in general agreement with the calculations of [70], where $\lambda(E)$ presents an approximately linear relation with the electric field.

Photons

In figure 8.11 the electron and photon spectrum is given for Geant O4 ($\ell_{\max} = 1$ mm), O4 ($\ell_{\max} = 10$ cm) and REAM for the 1.2 MV/m configuration, recorded at 77 m. We observe an overall small difference in spectrum, but significant at the

Table 8.2: Mean energy variation with electric field. For evaluated codes we fitted by equation 8.18, with $F = 0.28$ MV/m. See figure 8.10 for results.

Parameter / Code	a_1 [$10^6\mathrm{s}^{-1}$]	a_2	a_3 [$10^6\mathrm{s}^{-1}$]
GeantO1 ($l_{max} = 1$ mm)	5.96	1.14	-4.15
GeantO4 ($l_{max} = 1$ mm)	4.99	1.23	-1.85
GeantO1 ($l_{max} = 1$ cm)	6.96	0.929	-5.15
GeantO4 ($l_{max} = 1$ cm)	4.51	1.23	-1.79
REAM	3.82	1.3175	-3.25×10^{-3}
GRRR	6.87	1.1760	-5.63
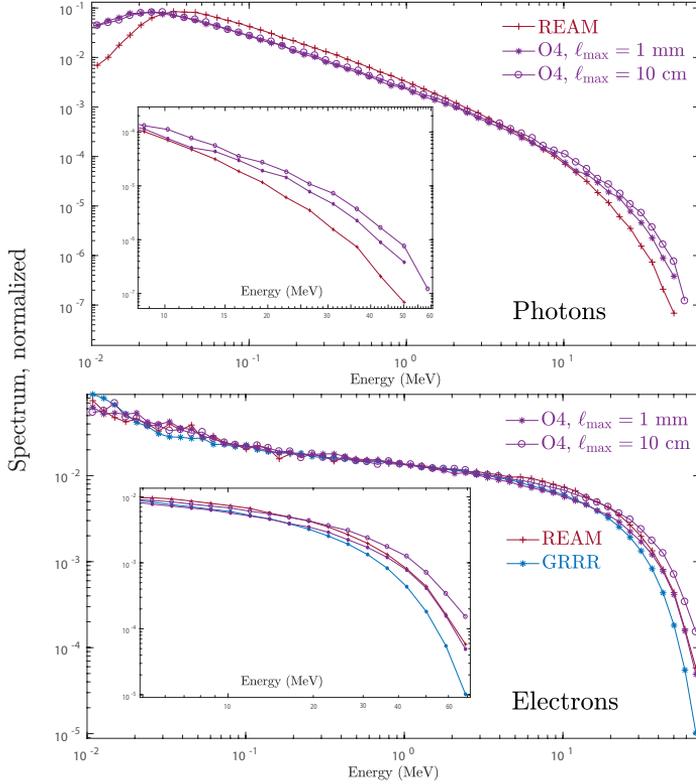
Figure 8.11: Photon and electron spectra of Geant O4 (10 cm), O4 (1 mm), REAM and GRRR (electron only) for the 1.2 MV/m field, and record at 77 m.

lowest and highest photon energies. Further investigation is needed to establish what is causes the observed differences.

8.4   Conclusion

We continued our first study [115] (i.e. chapter 7), now investigated performance of several Monte Carlo codes, in the presence of electric fields. We provide two simple set-ups that could be used by the community for future benchmark of codes that are not included in this study. These set-ups are described, together with data and figures to compare with, are given in detail in the attachment.

We found, surprisingly that not the low energy cutoff, but the implementation of null collisions result in significant differences between codes. Physics of RREA is determined by intermediate energy electrons between 10-100 keV; more precisely the minimum electron energy ($\epsilon_2^{\min}$) that can still create a RREA and this probability is very sensitive to the chosen step limiter.

We show that accurate RREA simulations can be achieved with an energy cut-off of 10 keV and higher (depending on the electric field magnitude), allowing for fast simulations. Precise modeling of the interactions of particles below 10 keV (e.g. by using more precise and slower models that include atomic shells effects) provides negligible differences.

8.A   Geant4 relative performance

Concerning the simulations aiming to fully characterize the Relativistic Runaway Electrons Avalanches (see section 3), we could not run them for all the electric fields with a maximum step lower than 1 millimeter, because of limitations in our available computation power.

Table 8.3 presents the relative computation times it takes to complete the simulation with an electric field magnitude of 1.2 MV/m, and 100 seed electrons with initial energy $\epsilon = 100$ keV, and a stop time (physical) of 233 nanoseconds. The fastest simulation uses Geant4 with the O1 physics list and $\ell_{\max} = 10$ cm and took 4.53 seconds to complete on one thread with the microprocessor we used. The simulations with the O4 physics list with $\ell_{\max} = 1$ mm requires about 400 times more computation time. To achieve it for the full range of electric fields we tested (in a reasonable amount of time), it required the use of the Norwegian FRAM computer cluster. The simulations with $\ell_{\max} = 0.1$ mm for all electric fields could not be achieved in a reasonable amount of time, even by using the computer cluster.

Model / $\ell_{\max}$	Option 1 (O1)		Option 4 (O4)	
10 cm	4.53 s	**1**	29.4 s	**6.49**
1 cm	51.9 s	**11.5**	123 s	**27.2**
1 mm	1004 s	**222**	1780 s	**393**
0.1 mm	9514 s	**2100**	16933 s	**3738**

Table 8.3: Computation time needed by different Geant4 configurations for the simulation of the same physical problem. The bold numbers are the times relative to the Geant4 O1 / 10 cm case.

8.B   Simulation Set-up

Medium composition

- The air is composed of 78.085% nitrogen, 20.95% oxygen and 0.965% argon,

- Absolute number density of 2.6881025 10^{25}m^3, equivalent to 1.293 kg.m^3. We always use an uniform air density.

Coordinate system and domain

- Cartesian coordinates x, y, z.

- Any field is always applied in the z (or -z) direction.

- Domain is a cylinder with radius (perpendicular to z) of 5 km and height (parallel to z) of 10 km, such that the middle of the bottom circle lies at (0,0,-5) and the middle of the top circle at (0,0,5).

- Delete particles if they fly out of the domain.

Initial seed

- Every simulation starts with 200 electrons of 100 keV.

- Particle start always in the center, i.e. (x,y,z) = (0,0,0), which their momentum directed upwards (positive z), parallel with the electric field such that they are accelerated by the field.

Energy threshold

- Take an energy threshold of 10 keV, meaning that all particles below 10 keV do not appear on the output.

- If your code is using friction for part of the spectrum, use 10 keV as the point to separate the part of friction and the part of explicit collisions.

Electric fields

- In this work we consider 13 electric fields: [6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30] x $10^5 \, V/m$.

- The field is oriented in the -z direction, such that the electrons are accelerated by the field upwards.

Time limit

- To avoid very large file size, a time limit must be set depending on the electric field.

- t_{stop} = [1078, 522, 351, 277, 233, 199, 179, 164, 153, 144, 135, 130, 124] nanoseconds; each value corresponding to an electric field (following the same order).

Output type 1 - evolution in time

Output all particles at 32 times : t = [14., 26., 39., 51., 64., 78., 90., 102., 116., 124., 135., 144., 153., 164., 179., 199., 215., 233., 262., 277., 290., 312., 351., 406., 464., 479., 522., 599., 719., 838., 958., 1078.] nanoseconds; regardless of position.

Output type 2 - evolution in space

Output all particles at 32 distances : z = [4., 7., 10., 14., 17., 21., 24., 27., 31., 33., 36., 38., 41., 44., 48., 53., 57., 62., 70., 74., 77., 83., 94., 108., 124., 128., 139., 160., 192., 224., 256., 288.] meters, regardless of time of arrival

Chapter 9

Adaptive selection of sampling points for uncertainty quantification

Simulations described in this thesis are in general computationally expensive. Ideally the simulation is performed on a large set of initial conditions, randomly distributed, to capture the propagation of uncertainty through the model. To execute this procedure more efficiently, we present a simple and robust strategy for the selection of sampling points in Uncertainty Quantification. The goal is to achieve the fastest possible convergence in the cumulative distribution function of a stochastic output of interest. We assume that the output of interest is the outcome of a computationally expensive nonlinear mapping of an input random variable, whose probability density function is known. We use a radial function basis to construct an accurate interpolant of the mapping. This strategy enables adding new sampling points one at a time, *adaptively*. This takes into full account the previous evaluations of the target non-linear function. We present comparisons with a stochastic collocation method based on the Clenshaw-Curtis quadrature rule, and with an adaptive method based on hierarchical surplus, showing that the new method often results in a large computational saving.

9.1   Introduction

We address one of the fundamental problems in Uncertainty Quantification (UQ): the mapping of the probability distribution of a random variable through a nonlinear function. Let us assume that we are concerned with a specific physical or engineering model which is computationally expensive. The model is defined by the map $g : \mathbb{R} \to \mathbb{R}$. It takes a parameter X as input, and produces an output Y, $Y = g(X)$. In this chapter we restrict ourselves to a proof-of-principle one-dimensional case. Let us assume that X is a random variable distributed with probability density function (pdf) P_X. The Uncertainty Quantification problem is the estimation of the pdf P_Y of the output variable Y, given P_X. Formally, the problem can be simply cast as a coordinate transformation and one easily obtains

$$P_Y(y) = \sum_{x \in \{x \mid g(x) = y\}} \frac{P_X(x)}{|\det J(x)|}, \tag{9.1}$$

where $J(x)$ is the Jacobian of $g(x)$. The sum over all x such that $g(x) = y$ takes in account the possibility that g may not be injective. If the function g is known exactly and invertible, Eq.(9.1) can be used straightforwardly to construct the pdf $P_Y(y)$, but this is of course not the case when the mapping g is computed via numerical simulations.

Several techniques have been studied in the last couple of decades to tackle this problem. Generally, the techniques can be divided in two categories: intrusive and non-intrusive [136–138]. Intrusive methods modify the original, *deterministic*, set of equations to account for the stochastic nature of the input (random) variables, hence eventually dealing with stochastic differential equations, and employing specific numerical techniques to solve them. Classical examples of intrusive methods are represented by Polynomial Chaos expansion [139–142], and stochastic Galerkin methods [143–146].

On the other hand, the philosophy behind non-intrusive methods is to make use of the deterministic version of the model (and the computer code that solves it) as a black-box, which returns one deterministic output for any given input. An arbitrary large number of solutions, obtained by sampling the input parameter space, can then be collected and analyzed in order to reconstruct the pdf $P_Y(y)$.

The paradigm of non-intrusive methods is perhaps best represented by Monte Carlo (MC) methods [147, 148]: one can construct an ensemble of input parameters $\{X_n \mid n = 1, \dots, N\}$ (N typically large) distributed according to the pdf $P_X(x)$, run the corresponding ensemble of simulations $g : X \to Y$, and process the outputs $\{Y_n \mid n = 1, \dots, N\}$. MC methods are probably the most robust of all the non-intrusive methods. Their main shortcoming is the slow convergence of the method, with a typical convergence rate proportional to $\sqrt{N}$. For many applications quasi-Monte Carlo (QMC) methods [147, 149] are now preferred to

MC methods, for their faster convergence rate. In QMC the pseudo-random generator of samples is replaced by more uniform distributions, obtained through so-called quasi-random generators [150, 151].

It is often said that MC and QMC do not suffer the 'curse of dimensionality'[152–154], in the sense that the convergence rate (but not the actual error!) is not affected by the dimension D of the input parameter space. Therefore, they represent the standard choice for large dimensional problems. On the other hand, when the dimension D is not very large, collocation methods [155–157] are usually more efficient.

Yet a different method that focuses on deriving a deterministic differential equation for cumulative distribution functions has been presented, e.g., in [158, 159]. This method is however not completely black-box.

Collocation methods recast an UQ problem as an interpolation problem. In collocation methods, the function $g(x)$ is sampled in a small (compared to the MC approach) number of points ('collocation points'), and an interpolant is constructed to obtain an approximation of g over the whole input parameter space, from which the pdf $P_Y(y)$ can be estimated.

The question then arises on how to effectively choose the collocation points. Recalling that every evaluation of the function g amounts to performing an expensive simulation, the challenge resides in obtaining an accurate approximation of P_Y with the least number of collocation points. Indeed, a very active area of research is represented by collocation methods that use sparse grids, so to avoid the computation of a full-rank tensorial product, particularly for model order reduction (see, e.g., [160–166]

As the name suggests, collocation methods are usually derived from classical quadrature rules [167–169].

The type of pdf P_X can guide the choice of the optimal quadrature rule to be used (i.e., Gauss-Hermite for a Gaussian probability, Gauss-Legendre for a uniform probability, etc. [155]). Furthermore, because quadratures are associated with polynomial interpolation, it becomes natural to define a global interpolant in terms of a Lagrange polynomial [170]. Also, choosing the collocation points as the abscissas of a given quadrature rule makes sense particularly if one is only interested in the evaluation of the statistical moments of the pdf (i.e., mean, variance, etc.) [171].

On the other hand, there are several applications where one is interested in the approximation of the full pdf P_Y. For instance, when g is narrowly peaked around two or more distinct values, its mean does not have any statistical meaning. In such cases one can wonder whether a standard collocation method based on quadrature rules still represents the optimal choice, in the sense of the computational cost to obtain a given accuracy.

From this perspective, a downside of collocation methods is that the collocation points are chosen a priori, without making use of the knowledge of $g(x)$ acquired at previous interpolation levels. For instance, the Clenshaw-Curtis

(CC) method uses a set of points that contains 'nested' subset, in order to re-use all the previous computations, when the number of collocation points is increased. However, since the abscissas are unevenly spaced and concentrated towards the edge of the domain (this is typical of all quadrature rules, in order to overcome the Runge phenomenon [170, 172]), it is likely that the majority of the performed simulations will not contribute significantly in achieving a better approximation of P_Y. Stated differently, one would like to employ a method where each new sampling point is chosen in such a way to result in the fastest convergence rate for the approximated P_Y, in contrast to a set of points defined a priori.

As a matter of fact, because the function g is unknown, a certain number of simulations will always be redundant, in the sense that they will contribute very little to the convergence of P_Y. The rationale for this work is to devise a method to minimize such a redundancy in the choice of sampling points while achieving fastest possible convergence of P_Y.

Clearly, this suggests to devise a strategy that chooses collocation points *adaptively*, making use of the knowledge of the interpolant of $g(x)$, which becomes more and more accurate as more points are added.

A well known adaptive sampling algorithm is based on the calculation of the so-called hierarchical surplus [163, 165, 173, 174, see e.g]. This is defined as the difference, between two levels of refinement, in the solution obtained by the interpolant. Although this algorithm is quite robust, and it is especially efficient in detecting discontinuities, it has the obvious drawback that it can be prematurely terminated, whenever the interpolant happens to exactly pass through the true solution on a point where the hierarchical surplus is calculated, no matter how inaccurate the interpolant is in close-by regions (see Figure 9.1 for an example).

The goal of this chapter is to describe an alternative strategy for the adaptive selection of sampling points. The objective in devising such strategy is to have a simple and robust set of rules for choosing the next sampling point. The chapter is concerned with a proof-of-principle demonstration of our new strategy, and we will focus here on one dimensional cases and on the case of uniform P_X only, postponing the generalization to multiple dimensions to future work. It is important to appreciate that the stated goal of this work is different from the traditional approach followed in the overwhelming majority of works that have presented sampling methods for UQ in the literature. Indeed, it is standard to focus on the convergence of the nonlinear unknown function $g(x)$, trying to minimize the interpolation error on $g(x)$, for a given number of sampling points. On the other hand, we will show that the convergence rates of $g(x)$ and of its cumulative distribution function can be quite different. Our new strategy is designed to achieve the fastest convergence on the latter quantity, which is ultimately the observable quantity of an experiment.

The chapter is organized as follows. In Section 2 we define the mathemati-
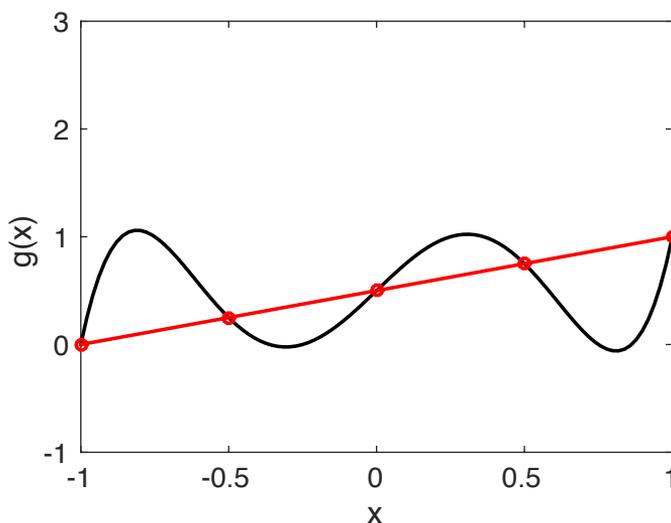
Figure 9.1: Example for which the algorithm based on hierarchical surplus fails. The function $g(x) = \frac{256}{30}x^5 - \frac{32}{3}x^3 + \frac{79}{30}x + \frac{1}{2}$ (in black) goes exactly through the red straight line at the points $x = -1, -0.5, 0, 0.5, 1$. Calculating the piecewise linear interpolant between two $(x = -1, 1)$, three $(x = -1, 0, 1)$, and five $(x = -1, -\frac{1}{2}, 0, \frac{1}{2}, 1)$ points would result in a null hierarchical surplus on these points.

cal methods used for the construction of the interpolant and show our adaptive strategy to choose a new collocation points. In Section 3 we present some numerical examples and comparisons with the Clenshaw-Curtis collocation method, and the adaptive method based on hierarchical surplus. Finally, we draw our conclusions in Section 4.

9.2   Mathematical methods

9.2.1   Clenshaw-Curtis (CC) quadrature rule

In Section 3, we compare our method with the CC method, which is the standard appropriate collocation method for a uniform P_X. Here, we recall the basic properties of CC, for completeness. The Clenshaw-Curtis (CC) quadrature rule uses the extrema of a Chebyshev polynomial (the so-called 'extrema plus endpoints' collocation points in [175]) as abscissas. They are particularly appealing to be used as collocation points in UQ, because a certain subset of them are nested. Specifically, they are defined, in the interval $[-1, 1]$ as:

$$x_i = -\cos\left(\frac{\pi(i-1)}{N-1}\right) \quad i = 1, \dots, N. \tag{9.2}$$

One can notice that the the set of $N = 2^w + 1$ points is fully contained in the set of $N = 2^{w+1} + 1$ points (with w an arbitrary integer, referred to as the level of the set). In practice this means that one can construct a nested sequence of collocation points with $N = 3, 5, 9, 17, 33, 65, 129, \ldots$, re-using all the previous evaluations of g.

Collocation points based on quadratures are optimal to calculate moments [1]:

$$\mu_Y^p = \int y^p P_Y(y) \mathrm{d}y = \int g(x)^p P_X(x) \mathrm{d}x, \qquad (9.3)$$

where we used the identity relation,

$$P_Y(y)\mathrm{d}y = P_X(x)\mathrm{d}x. \qquad (9.4)$$

It is known that integration by quadrature is very accurate (for smooth enough integrand), and the moments can be readily evaluated, without the need to construct an interpolant:

$$\mu_Y^p \simeq \sum_i w_i (g(x_i))^p, \qquad (9.5)$$

where the weights w_i can be computed with standard techniques (see, e.g. [171]). The interpolant for the CC method is the Lagrange polynomial.

9.2.2 Selection of collocation points based on hierarchical surplus

The hierarchical surplus algorithm is widely used for interpolation on sparse grids. It is generally defined as the difference between the value of an interpolant at the current and previous interpolation levels [163]:

$$\Delta^n = \tilde{g}^n - \tilde{g}^{n-1} \qquad (9.6)$$

The simplest algorithm prescribes a certain tolerance and looks for all the point at the new level where the hierarchical surplus is larger than the tolerance. The new sampling points (at the next level, $n + 1$) will be the neighbors (defined with a certain rule) of the points where this condition is met. In one-dimension, the algorithm is extremely simple because the neighbors are defined by only two points, that one can define in such a way that cells are always halved. In this work, we compare our new method with a slightly improved version of the hierarchical surplus algorithm. The reason is because we do not want our comparisons to be dependent on the choice of an arbitrary tolerance level, and we want to be able to add new points two at the time. Hence, we define a new interpolation level by adding only the two neighbors of the point with the largest hierarchical surplus. All the previous hierarchical

---

[1]Here p on the left-hand side is a label, such that μ^1 is the mean, μ^2 is the variance, and so on. On the right-hand side it is an exponent.

surpluses that have been calculated, but for which new points have not been added yet are kept. The pseudo-code of the algorithm follows. The interpolant is understood to be piece-wise linear interpolation, and the grid is $x \in [-1, 1]$.

Calculate the interpolant on the grid $x = \{-1, 0, 1\}$.
Define $x_h = \{-1/2, 1/2\}$ and add them on the grid
while *Not converged* **do**
> Calculate the interpolant on the new grid
> Calculate the hierarchical surplus on the last two entries of x_h and store them in the vector h_s
> Find the largest hierarchical surplus in h_s, remove it from h_s and remove the corresponding x from x_h
> Append the two neighbors to x_h and add them to the grid

end

Algorithm 1: Hierarchical surplus algorithm

9.2.3 Multiquadric biharmonic radial basis

We use a multiquadric biharmonic radial basis function (RBF) with respect to a set of points $\{x_i\}$, with $i = 1, \ldots, N$, defined as:

$$\Phi_i(x, c) = \sqrt{(x - x_i)^2 + c_i^2},\tag{9.7}$$

where c_i are free parameters (referred to as shape parameters). The function $g(x)$ is approximated by the interpolant $\widetilde{g}(x)$ defined as

$$\widetilde{g}(x) = \sum_{i=1}^{N} \lambda_i \Phi_i(x, c).\tag{9.8}$$

The weights λ_i are obtained by imposing that $g(x_i) = \widetilde{g}(x_i)$ for each sampling point in the set, namely the interpolation error is null at the sampling points. This results in solving a linear system for $\lambda = (\lambda_1, \ldots, \lambda_N)$ of the form $A\lambda^T = g(\mathbf{x})^T$, with A a real symmetric $N \times N$ matrix. We note that, by construction, the linear system will become more and more ill-conditioned with increasing N, for fixed values of c. This can be easily understood because when two points become closer and closer the corresponding two rows in the matrix A become less and less linearly independent. To overcome this problem one needs to decrease the corresponding values of c. In turns, this means that the interpolant $\widetilde{g}(x)$ will tend to a piece-wise linear interpolant for increasingly large N.

9.2.4 New adaptive selection of collocation points

We focus, as the main diagnostic of our method, on the cumulative distribution function (cdf) $C(y)$, which is defined as

$$C(y) = \int_{y_{\min}}^{y} P_Y(y)\mathrm{d}y,\tag{9.9}$$

where $y_{\min} = \min g(x)$. As it is well known, the interpretation of the cumulative distribution function is that, for a given value y^*, $C(y^*)$ is the probability that $g(x) \le y^*$. Of course, the cdf $C(y)$ contains all the statistical information needed to calculate any moment of the distribution, and can return the probability density function $P_Y(y)$, upon differentiation. Moreover, the cdf is always well defined between 0 and 1. The following two straightforward considerations will guide the design of our *adaptive selection strategy*. A first crucial point, already evident from Eq. (9.1), is whether or not $g(x)$ is bijective. When $g(x)$ is bijective this translates to the cdf $C(y)$ being continuous, while a non-bijective function $g(x)$ produces a cdf $C(y)$ which is discontinuous. It follows that intervals in x where $g(x)$ is constant (or nearly constant) will map into a single value $y = g(x)$ (or a very small interval in y) where the cdf will be discontinuous (or 'nearly' discontinuous). Secondly, an interval in x with a large first derivative of $g(x)$ will produce a nearly flat cdf $C(y)$. This is again clear by noticing that the Jacobian J in Eq. (9.1) ($dg(x)/dx$ in one dimension) is in the denominator, and therefore the corresponding $P_Y(y)$ will be very small, resulting in a flat cdf $C(y)$.

Loosely speaking one can then state that regions where $g(x)$ is flat will produce large jumps in the cdf $C(y)$ and, conversely, regions where the $g(x)$ has large jumps will map in to a nearly flat cdf $C(y)$. From this simple considerations one can appreciate how important it is to have an interpolant that accurately capture both regions with *very large* and *very small* first derivative of $g(x)$. Moreover, since the cdf $C(y)$ is an integrated quantity, interpolation errors committed around a given y will propagate in the cdf for all larger y values. For this reason, it is important to achieve a global convergence with interpolation errors that are of the same order of magnitude along the whole domain.

The adaptive section algorithm works as follows. We work in the interval $x \in [-1, 1]$ (every other interval where the support of $g(x)$ is defined can be rescaled to this interval). We denote with $\{x_i\}$ the sampling set which we assume is always sorted, such that $x_i < x_{i+1}$. We start with 3 points: $x_1 = -1$, $x_2 = 0$, $x_3 = 1$. For the robustness and the simplicity of the implementation we choose to select a new sampling point always at equal distance between two existing points. One can decide to limit the ratio between the largest and smallest distance between adjacent points: $r = \max\{d_i\}/\min\{d_i\}$ (with $i = 1, \ldots, N-1$), where d_i is the distance between the points x_{i+1} and x_i. This avoids to keep refining small intervals when large intervals might still be under-resolved, thus aiming for the above mentioned global convergence over the whole support. At each iteration we create a list of possible new points, by halving every interval, excluding the points that would increase the value of r above the maximum desired (note that r will always be a power of 2). We calculate the first derivative of $\widetilde{g}(x)$ at these points, and alternatively choose the point with largest/smallest derivative as the next sampling point. Notice that, by the definition of the interpolant, Eq. (9.8),

its first derivative can be calculated exactly as:

$$\frac{\mathrm{d}\widetilde{g}(x)}{\mathrm{d}x} = \sum_{i=1}^{N} \lambda_i \frac{\mathrm{d}\Phi_i(x,c)}{\mathrm{d}x} \qquad (9.10)$$

without having to recompute the weights λ_i. At each iteration the shape parameters c_i are defined at each points, as $c_i = 0.85 \cdot \min(d_{i-1}, d_i)$, i.e. they are linearly rescaled with the smallest distance between the point x_i and its neighbors. The pseudo-code of the algorithm follows.

while *Not converged* **do**
    $x_{\text{guess}} \leftarrow 0.5 \cdot (x_i + x_{i+1})$
    Exclude points in x_{guess} such that $r = \max\{d_i\}/\min\{d_i\} > R$
    Calculate $\widetilde{g}^n(x)'$ through (9.10) at $\{x_{\text{guess}}\}$
    Alternatively choose x_{guess} with largest/smallest values of $|\widetilde{g}^n(x)'|$ as new collocation point
    Calculate new weights λ_i
end

Algorithm 2: Adaptive selection of sampling points

9.3   Numerical examples

In this section we present and discuss four numerical examples where we apply our adaptive selection strategy. In this work we focus on a single input parameter and the case of constant probability $P_X = 1/2$ in the interval $x \in [-1, 1]$, and we compare our results against the Clenshaw-Curtis, and the hierarchical surplus methods. We denote with $\widetilde{g}^n(x)$ the interpolant obtained with a set of n points (hence the iterative procedure starts with $\widetilde{g}^3(x)$). A possible way to construct the cdf $C(y)$ from a given interpolant $\widetilde{g}^n(x)$ would be to generate a sample of points in the domain $[-1, 1]$, randomly distributed according to the pdf $P_X(x)$, collecting the corresponding values calculated through Eq. (9.8), and constructing their cdf. Because here we work with a constant $P_X(x)$, it is more efficient to simply define a uniform grid in the domain $[-1, 1]$ where to compute $\widetilde{g}^n(x)$. In the following we will use, in the evaluation of the cdf $C(y)$, a grid in y with $N_y = 10001$ points equally spaced in the interval $[\min \widetilde{g}^n(x), \max \widetilde{g}^n(x)]$, and a grid in x with $N_x = 1001$ points equally spaced in the interval $[-1, 1]$. We define the following errors:

$$\varepsilon_C = \frac{||C(\widetilde{g}^n(x)) - C(g(x))||_2}{\sqrt{N_y}} \qquad (9.11)$$

$$\varepsilon_g = \frac{||\widetilde{g}^n(x) - g(x)||_2}{\sqrt{N_x}} \qquad (9.12)$$

where $|| \cdot ||_2$ denotes the L_2 norm. It is important to realize that the accuracy of the numerically evaluated cdf $C(y)$ will always depend on the binning of y, i.e. the points at which the cdf is evaluated. As we will see in the following examples, the error ε_C saturates for large N, which thus is an artifact of the finite bin size. We emphasize that, differently from most of the previous literature, our strategy focuses on converging rapidly in ε_C, rather than in ε_g. Of course, a more accurate interpolant will always result in a more accurate cdf, however the relationship between a reduction in ε_g and a corresponding reduction in ε_C is not at all trivial. This is because the relation between $P_X(x)$ and $P_Y(y)$ is mediated by the Jacobian of $g(x)$, and it also involves the bijectivity of g. Finally, we study the convergence of the mean μ_Y, see equation 9.3, and the variance σ_Y^2, which is defined as

$$\sigma_Y^2 = \int_{-1}^{1} (\widetilde{g}(x) - \mu_Y)^2 P_X(x) \mathrm{d}x. \tag{9.13}$$

These will be calculated by quadrature for the CC methods, and with an integration via trapezoidal method for the adaptive methods.
We study two analytical test cases:

- Case 1: $g(x) = \arctan(10^3 x^3)$;

- Case 2: $g(x) = \frac{1}{(2+\sin(3\pi x))^2}$;

and two test cases where an analytical solution is not available, and the reference $g(x)$ will be calculated as an accurate numerical solution of a set of ordinary differential equations:

- Case 3: Lotka-Volterra model (predator-prey);

- Case 4: Van der Pol oscillator.

While Case 1 and 2 are more favorable to the CC method, because the functions are smooth and analytical, hence a polynomial interpolation is expected to produce accurate results, the latter two cases mimic applications of real interest, where the model does not produce analytical results, although $g(x)$ might still be smooth (at least piece-wise, in Case 4).

9.3.1   Case 1: $g(x) = \arctan(10^3 x^3)$

In this case $g(x)$ is a bijective function, with one point ($x = 0$) where the first derivative vanishes. Figure 9.2 shows the function $g(x)$ (top panel) and the corresponding cdf $C(y)$ (bottom panel), which in this case can be derived analytically. Hence, we use the analytical expression of cdf $C(y)$ to evaluate the error ε_C. The convergence of ε_C and ε_g is shown in Figure 9.3 (top and bottom panels, respectively). Here and in all the following figures blue squares denote
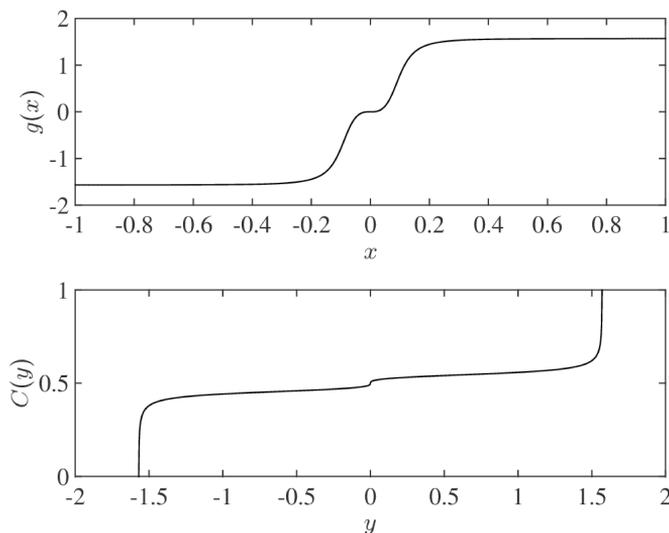
Figure 9.2: Case 1: $g(x) = \arctan(10^3 x^3)$. Top panel: $g(x)$; bottom panel: cdf $C(y)$.

the new adaptive selection method, red dots are for the CC methods, and black line is for the hierarchical surplus method. We have run the CC method only for $N = 3, 5, 9, 17, 33, 65, 129$ (i.e. the points at which the collocation points are nested), but for a better graphical visualization the red dots are connected with straight lines. One can notice that the error for the new adaptive method is consistently smaller than for the CC method. From the top panel, one can appreciate the saving in computer power that can be achieved with our new method. Although the difference with CC is not very large until $N = 17$, at $N = 33$ there is an order of magnitude difference between the two. It effectively means that in order to achieve the same error $\varepsilon_C \sim 10^{-5}$, the CC method would run at least twice the number of simulations. The importance of focusing on the convergence of the cdf, rather than on the interpolant, is clear in comparing our method with the hierarchical surplus method. For instance, for $N = 80$, the two methods have a comparable error ε_g, but our method has achieved almost an order of magnitude more accurate solution in $C(y)$. Effectively, this means that our method has sampled the new points less redundantly. In this case $g(x)$ is an anti-symmetric function with zero mean. Hence, any method that chooses sampling points symmetrically distributed around zero would produce the correct first moment μ_Y. We show in figure 9.4 the convergence of σ_Y^2, as the absolute value of the different with the exact value σ_{an}, in logarithmic scale. Blue, red, and black lines represent the new adaptive method, the CC, and the hierarchical surplus methods, respectively (where again for the CC,
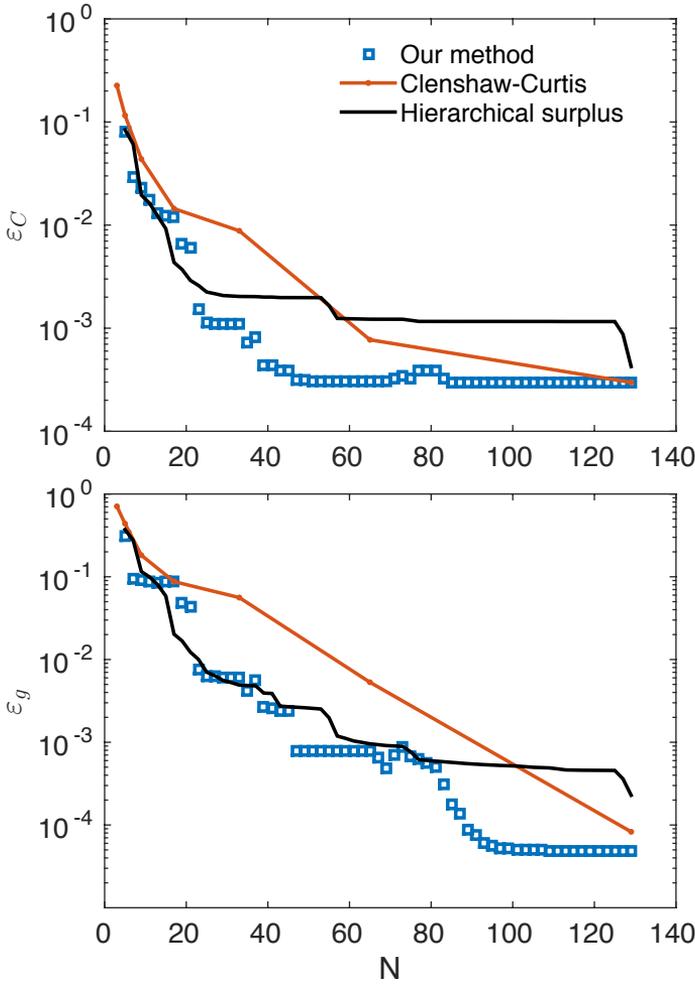
Figure 9.3: Case 1. Error ε_C (top) ε_g (bottom) as function of number of sampling points N. Blue squares: new adaptive selection method. Red dots: Clenshaw-Curtis. Black curve: adaptive method based on hierarchical surplus.

simulations are only performed where the red dots are shown). The exact value is $\sigma_{an}^2 = 2.102$. As we mentioned, the CC method is optimal to calculate moments, since it uses quadrature. Although in our method the error does not decrease monotonically, it is comparable with the result for CC.
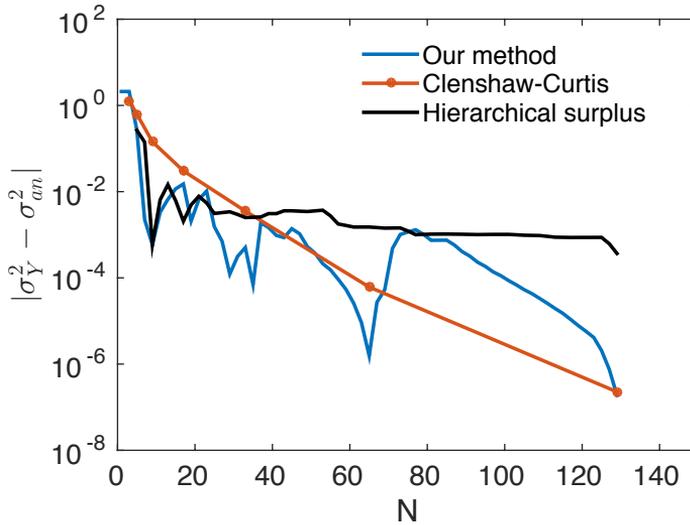
Figure 9.4: Case 1. Absolute error in the variance σ_Y^2 versus number of sampling points N. Blue: new adaptive selection method. Red: Clenshaw-Curtis. Black: adaptive method based on hierarchical surplus.

9.3.2 Case 2: $g(x) = \frac{1}{(2+\sin(3\pi x))^2}$

In this case the function $g(x)$ is periodic, and it presents, in the domain $x \in [-1, 1]$ three local minima ($y = 1/9$) and three local maxima ($y = 1$). The function and the cdf $C(y)$ are shown in Figure 9.5 (top and bottom panel, respectively). Figure 9.6 shows the error for this case (from now on the same format of Figure 9.3 will be used). The first consideration is that the hierarchical surplus method is the less accurate of the three. Second, ε_g is essentially the same for the CC and the new method, up to $N = 65$. For $N = 129$ the CC methods achieve a much accurate solution as compared to the new adaptive method, whose error has a much slower convergence. However, looking at the error in the cdf in top panel of Figure 9.6, the two methods are essentially equivalent. This example demonstrates that, in an UQ framework, the primary goal in constructing a good interpolant should not be to minimize the error of the interpolant with respect to the 'true' $g(x)$, but rather to achieve the fastest possible convergence on the cdf C_Y. Although, the two effects are intuitively correlated, they are not into a linear relationship. In other words, not all sample points in x count equally in minimizing ε_C. The convergence of μ_Y (exact value $\mu_{an} = 0.385$) and σ_Y^2 (exact value $\sigma_{an} = 0.087$) is shown in Figures 9.7 and 9.8, respectively. It is interesting to notice that our method presents errors that are always smaller than the CC method, although the errors degrade considerably in the regions between two CC points, where the two adaptive methods yield

comparable results.
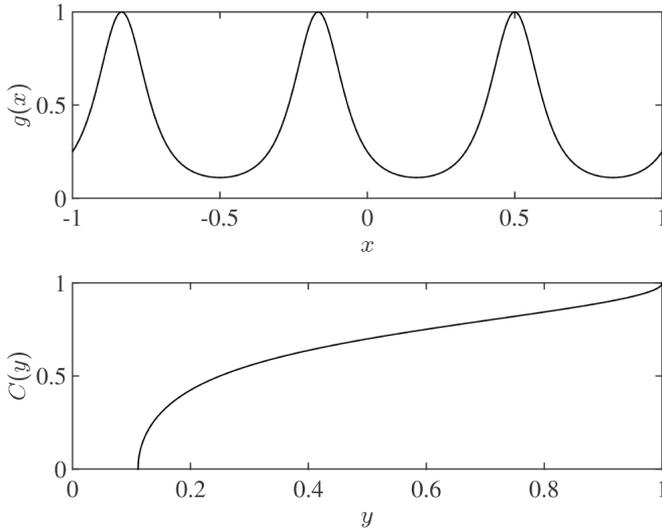
Figure 9.5: Case 2: $g(x) = \frac{1}{(2+\sin(3\pi x))^2}$. Top panel: $g(x)$; bottom panel: $C(y)$.

9.3.3 Case 3: Lotka-Volterra model (predator-prey)

The Lotka-Volterra model [176–178] is a well-studied model that exemplifies the interaction between two populations (predators and preys). This case is more realistic than Cases 1 and 2, as the solution of the model cannot be written in analytical form. As such, both the $g(x)$ and the cdf $C(y)$ used to compute the errors are calculated numerically. We use the following simple model:

$$\frac{\mathrm{d}h(t)}{\mathrm{d}t} = h(t) - (5x+6)h(t)l(t) \qquad (9.14)$$

$$\frac{\mathrm{d}l(t)}{\mathrm{d}t} = h(t)l(t) - l(t) \qquad (9.15)$$

where $h(t)$ and $l(t)$ denote the population size for each species (say, horses and lions) as function of time. The ODE is easily solved in MATLAB, with the `ode45` routine, with an absolute tolerance set equal to 10^{-8}. We use, as initial conditions, $h(t=0) = l(t=0) = 1$, and we solve the equations for $t \in [0, 10]$. Clearly, the solution of the model depends on the input parameter x. We define our test function $g(x)$ to be the result of the model for the l population at time $t = 10$:

$$g(x) = l(t = 10, x). \qquad (9.16)$$
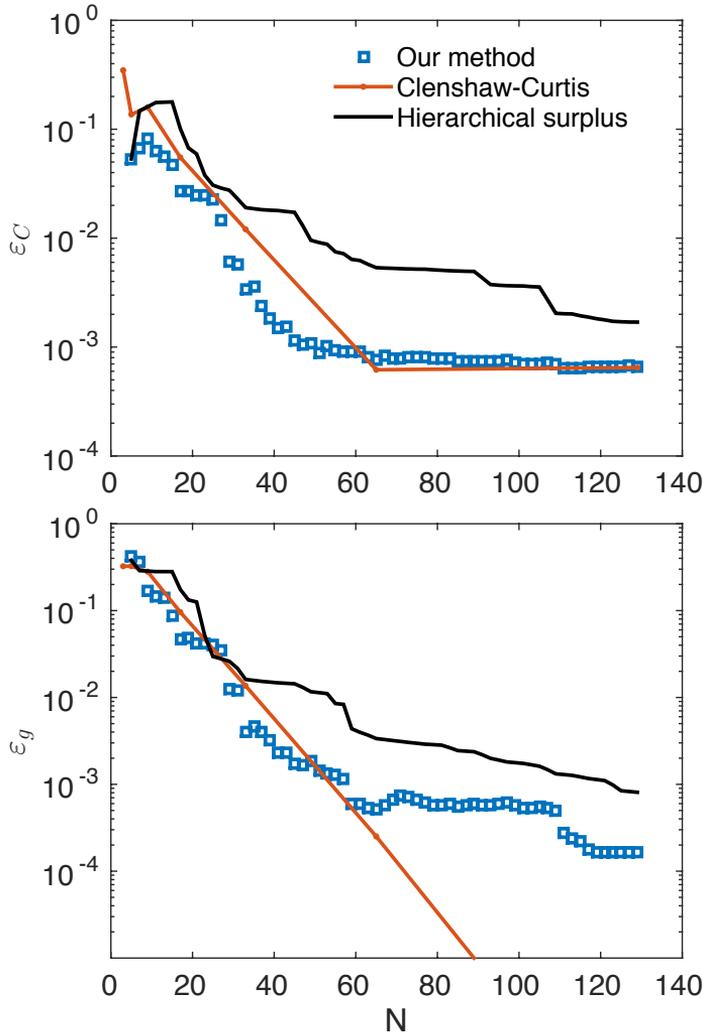
Figure 9.6: Case 2. Error ε_C (top) ε_g (bottom) as function of number of sampling points N. Blue squares: new adaptive selection method. Red dots: Clenshaw-Curtis. Black curve: adaptive method based on hierarchical surplus.

The resulting function $g(x)$, and the computed cdf $C(y)$ are shown in Figure 9.9 (top and bottom panel, respectively). We note that, although $g(x)$ cannot be expressed as an analytical function, it is still smooth, and hence it does not present particular difficulties in being approximated through a polynomial interpolant. Indeed the error ε_g undergoes a fast convergence both for the adaptive methods and for the CC method (Figure 9.10). Once again, the new adaptive method is much more powerful than the CC method in achieving a better convergence rate,
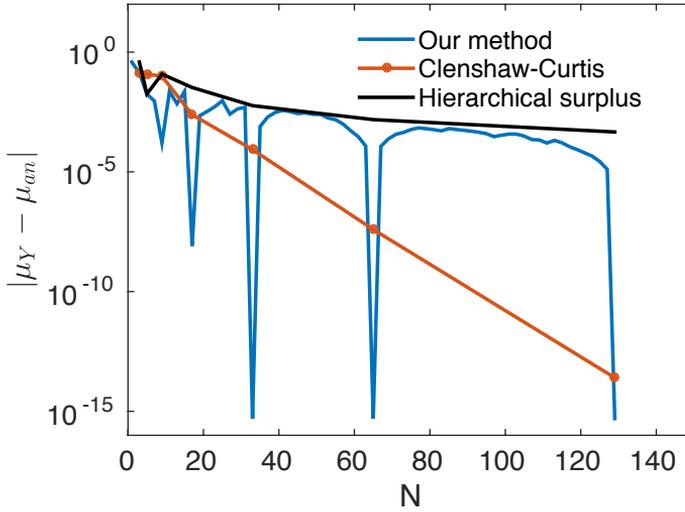
Figure 9.7: Case 2. Absolute error in the mean μ_Y versus number of sampling points N. Blue: new adaptive selection method. Red: Clenshaw-Curtis. Black: adaptive method based on hierarchical surplus.
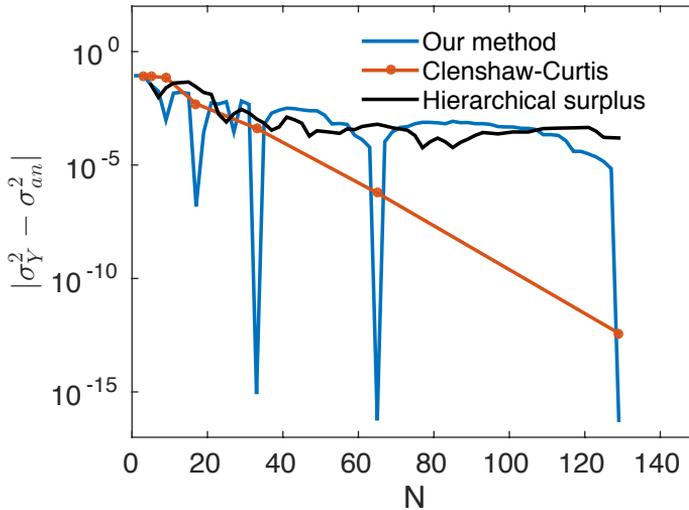
Figure 9.8: Case 2. Absolute error in the variance σ_Y^2 versus number of sampling points N. Blue: new adaptive selection method. Red: Clenshaw-Curtis. Black: adaptive method based on hierarchical surplus.

and thus saving computational power, while the hierarchical surplus method is the worst of the three. Convergence of μ_Y and σ_Y^2 are shown in Figures 9.11 and

9.12, respectively. Similar to previous cases, the CC presents a monotonic convergence, while this is not the case for the adaptive methods. Only for $N = 129$, the CC method yields much better results than the new method.
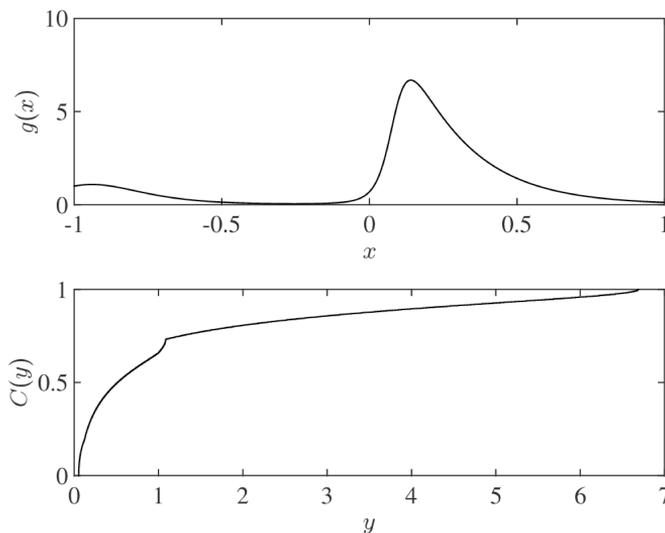
Figure 9.9: Case 3: Lotka-Volterra model. Top panel: $g(x)$; bottom panel: $C(y)$.

9.3.4 Case 4: Van der Pol oscillator

Our last example is the celebrated Van der Pol oscillator[142, 179–181], which has been extensively studied as a textbook case of a nonlinear dynamical system. In this respect this test case is very relevant to Uncertainty Quantification, since real systems often exhibit a high degree of nonlinearity. Similar to Case 3, we define our test function $g(x)$ as the output of a set of two ODEs, which we solve numerically with MATLAB. The model for the Van der Pol oscillator is:

$$\frac{dQ(t)}{dt} = V(t) \tag{9.17}$$

$$\frac{dV(t)}{dt} = (-50 + 100(x + 2))(1 - Q(t)^2)V(t) - Q(t). \tag{9.18}$$

The initial conditions are $Q(t = 0) = 2$, $V(t = 0) = 0$. The model is solved for time $t \in [0, 300]$, and the function $g(x)$ is defined as

$$g(x) = V(t = 300, x). \tag{9.19}$$

The so-called nonlinear damping parameter is rescaled such that for $x \in [-1, 1]$, it ranges between 50 and 250. The function $g(x)$ and the corresponding cdf
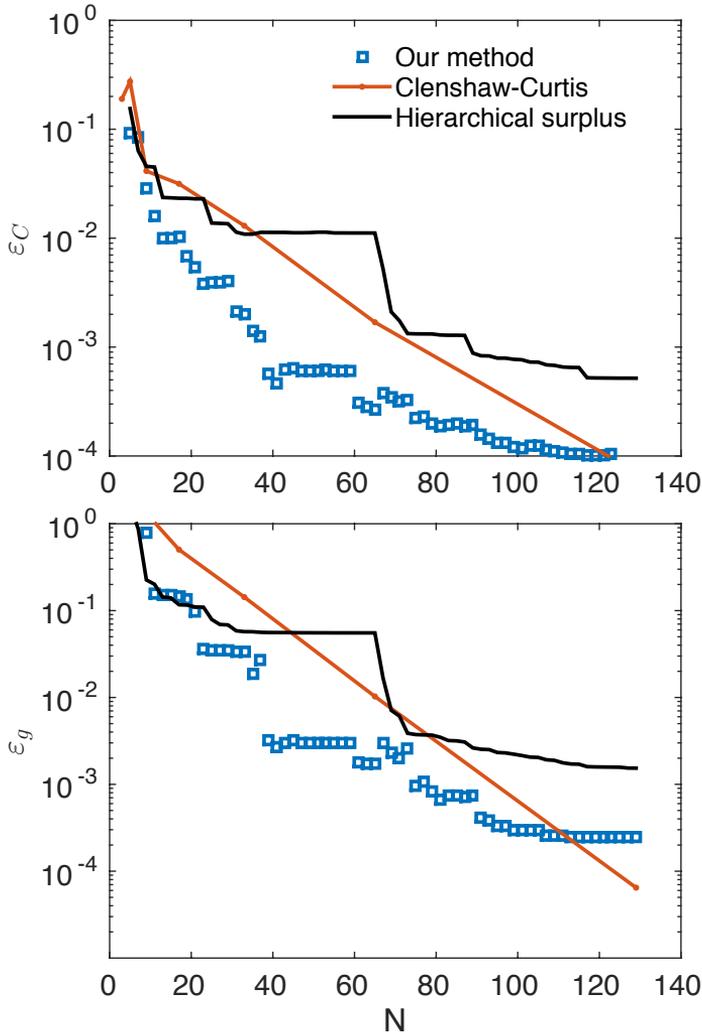
Figure 9.10: Case 3. Error ε_C (top) ε_g (bottom) as function of number of sampling points N. Blue squares: new adaptive selection method. Red dots: Clenshaw-Curtis. Black curve: adaptive method based on hierarchical surplus.

$C(y)$ are shown in Figure 9.13. This function is clearly much more challenging than the previous ones. It is divided in two branches, where it takes values $-2 \leq y \leq -1$ and $1 \leq y \leq 2$, and it presents discontinuities where it jumps from one branch to the other. Correspondingly, cdf $C(y)$ presents a flat plateau for $-1 \leq y \leq 1$, which is the major challenge for both methods. In figure 9.14 we show the errors ε_g and ε_C. The overall convergence rate of the CC and the new method is similar. For this case, the hierarchical surplus method yields a
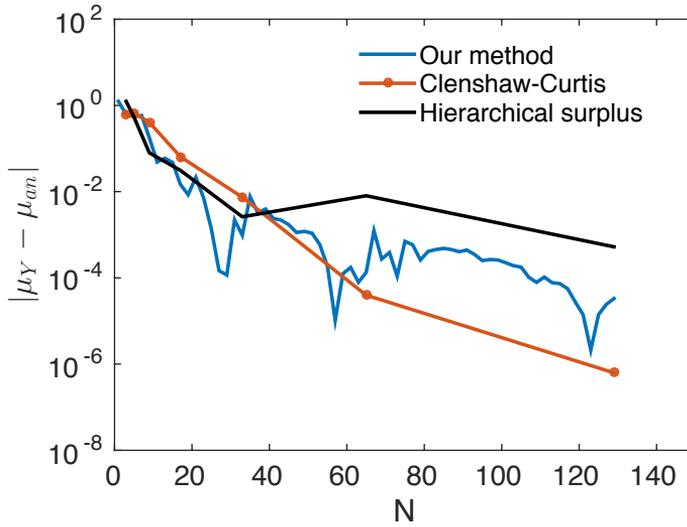
Figure 9.11: Case 3. Absolute error in the mean μ_Y versus number of sampling points N. Blue: new adaptive selection method. Red: Clenshaw-Curtis. Black: adaptive method based on hierarchical surplus.
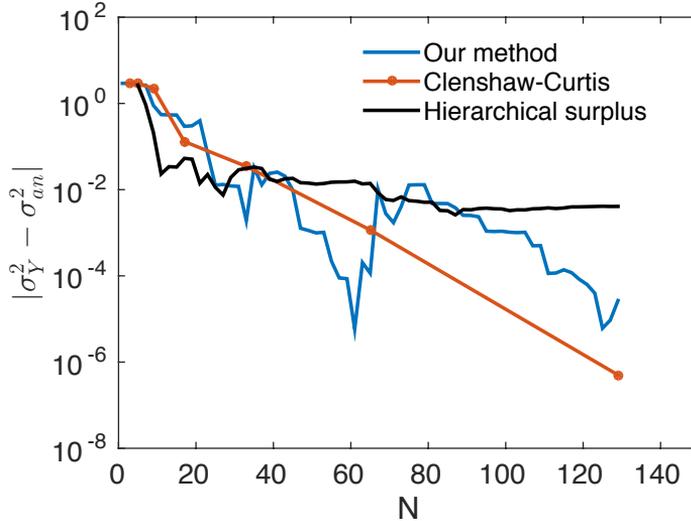
Figure 9.12: Case 3. Absolute error in the variance σ_Y^2 versus number of sampling points N. Blue: new adaptive selection method. Red: Clenshaw-Curtis. Black: adaptive method based on hierarchical surplus.

better convergence, but only for $N > 80$. As we commented before, the mean μ_Y has no statistical meaning in this case, because the output is divided into

two separate regions. The convergence for σ_Y^2 is presented in Figure 9.15.
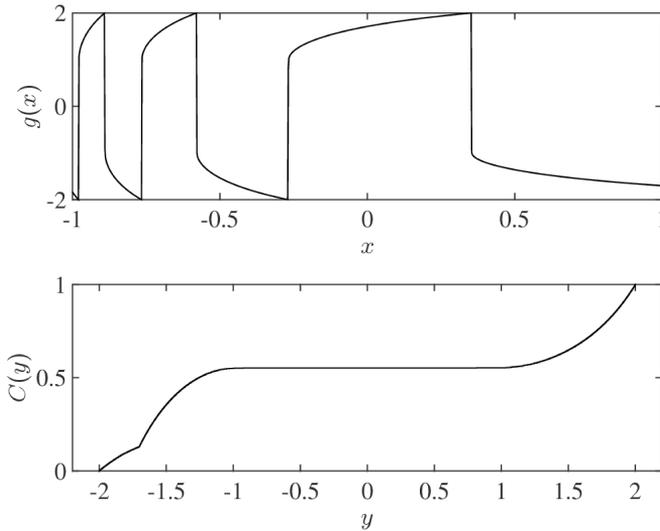
Figure 9.13: Case 4: Van der Pol oscillator. Top panel: $g(x)$; bottom panel: $C(y)$.

9.4   Conclusions and future work

We have presented a new adaptive algorithm for the selection of sampling points for non-intrusive stochastic collocation in Uncertainty Quantification (UQ). The main idea is to use a radial basis function as interpolant, and to refine the grid on points where the interpolant presents large and small first derivative.
In this work we have focused on 1D and uniform probability $P_X(x)$, and we have shown four test cases, encompassing analytical and non-analytical smooth functions, which are prototype of a very wide class of functions. In all cases the new adaptive method improved the efficiency of both the (non-adaptive) Clenshaw-Curtis collocation method, and of the adaptive algorithm based on the calculation of the hierarchical surplus (note that the method used in this chapter is a slight improvement of the classical algorithm). The strength of our method is the ability to select a new sampling point making full use of the interpolant resulting from all the previous evaluation of the function $g(x)$, thus seeking the most optimal convergence rate for the cdf $C(y)$. We have shown that there is no one-to-one correspondence between a reduction in the interpolation error ε_g and a reduction in the cdf error ε_C. For this reason, collocation methods that choose the distribution of sampling points a priori can perform poorly in attaining a fast convergence rate in ε_C, which is the main goal of UQ. Moreover,
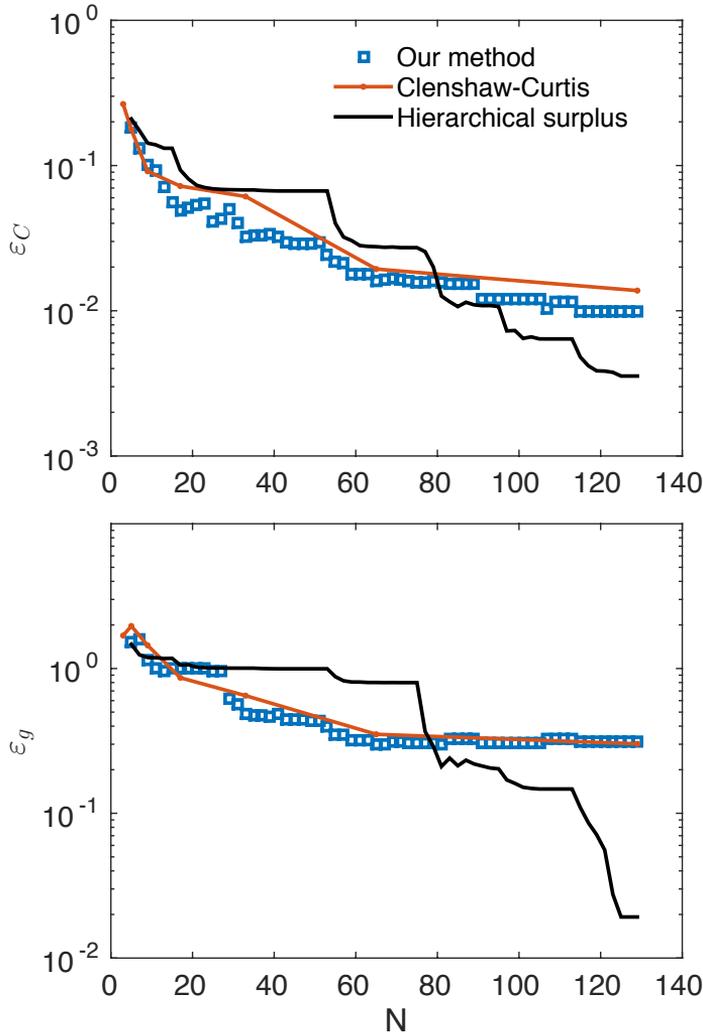
Figure 9.14: Case 4. Error ε_C (top) ε_g (bottom) as function of number of sampling points N. Blue squares: new adaptive selection method. Red dots: Clenshaw-Curtis. Black curve: adaptive method based on hierarchical surplus.

in order to maintain the nestedness of the collocation points the CC method requires larger and larger number of simulations (2^w moving from level w to level $w + 1$), which is in contrast with our new method where one can add one point at the time.

We envision many possible research directions to further investigate our method. The most obvious is to study multi-dimensional problems. We emphasize that the radial basis function is a mesh-free method and as such we anticipate that
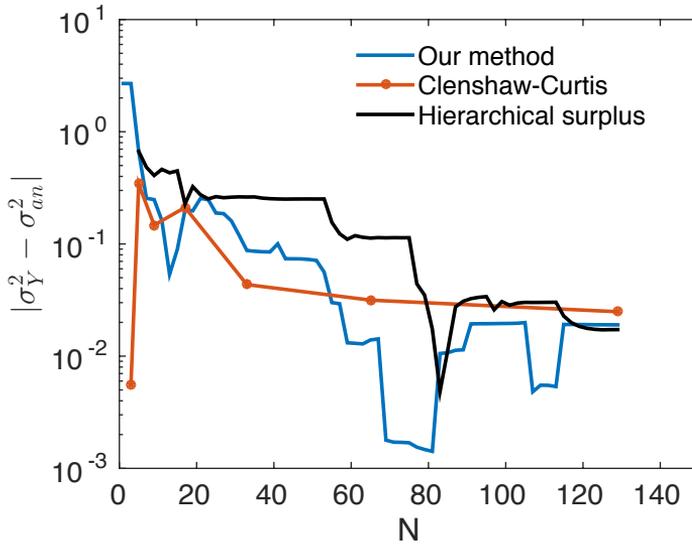
Figure 9.15: Case 4. Absolute error in the variance σ_Y^2 versus number of sampling points N. Blue: new adaptive selection method. Red: Clenshaw-Curtis. Black: adaptive method based on hierarchical surplus.

this will largely alleviate the curse of dimensionality that afflicts other collocation methods based on quadrature points (however, see [161] for methods related to the construction of sparse grids, which have the same aim). Moreover, it will be interesting to explore the versatility of RBF in what concerns the possibility of choosing an optimal shape parameter c [182]. Recent work [183, 184] investigated the role of the shape parameter c in interpolating discontinuous functions, which might be very relevant in the context of UQ, when the continuity of $g(x)$ cannot be assumed a priori. Finally, a very appealing research direction, would be to simultaneously exploit quasi-Monte Carlo and adaptive selection methods for extremely large dimension problems.

Part III

Lightning initiation problem

Chapter 10

Prediction of lighting inception by ice particles and extensive air showers

We derive that lightning can start if the electric field is 15% of the breakdown field, and if elongated ice particles of 6 cm length and 100 free electrons per cm^3 are present. This is one particular example set from a parameter range that we discuss as well. Our simulations include the permittivity $\epsilon(\omega)$ of ice. 100 free electrons per cm^3 exist at 5.5 km altitude in air showers created by cosmic particles of at least 5×10^{15} eV. If the electric field zone is 3 m high and 0.2 km^2 in the horizontal direction, at least one discharge per minute can be triggered. The size distribution of the ice particles is crucial for our argument; more detailed measurements would be desirable.

10.1   Introduction.

Lightning inception is the first out of the "top ten questions in lightning research" according to a recent review [185]. How can lightning start when the electric fields in thunderclouds are well below the classical breakdown field E_k [186] that is required for electron multiplication and ionization growth? And when the height of the high electric field zone is typically smaller than a kilometer?

It has been suggested already a few decades ago that a discharge could start in a lower electric field due to the relativistic run-away electron breakdown: cosmic particles could create ionization avalanches of relativistic particles when the electric field exceeds the threshold field $E_{RREA} \approx E_k/10$ for the formation of a relativistic run-away electron avalanche (RREA). However, as the RREA length is of the order of 300 m for a field of 2.8 kV/cm (at standard temperature and pressure) [187], the electric field needs to exceed E_{RREA} over heights of several km.

Another suggestion is that hydrometeors, i.e., airborne particles consisting of liquid or frozen water (droplets, snowflakes, graupel, hail etc.) could enhance the electric field locally in their neighborhood due to their high permittivity [2]. Experiments show how air discharges start from ice particles [10, 188]; however, the background electric fields are here as large as $0.3\ E_k$, and the free electrons needed to start the discharge are created through a radioactive source. In [34] an ice particle is modeled as an ionized patch of air. This model demonstrates the field enhancement around a real hydrometeor and the emergence of a discharge, but the electrons are trivially available from the ionized patch while a lack of free electrons is an essential issue in a thundercloud.

Free electrons in the high field region are needed to start a discharge. They are generated up to a few km altitude by the decay of radioactive elements emitted from the ground, and furthermore by solar energetic particles and by cosmic rays. However, within the troposphere these free electrons attach within tens of nanoseconds to oxygen molecules and form roughly 10^3 positive and negative ions per cm^3. In dry air, the electrons can detach again and start a discharge when the electric field exceeds E_k [189, 190]. But in humid air, the O_2^- ions attract water molecules within microseconds [191]. The electron detachment time from such ion-water clusters is of the order of micro- or even milliseconds [192], and it is negligible on the nanosecond time scale of the primary discharge evolution. Gurevich and Karashtin [193] suggested that the free electrons near a hydrometeor could be supplied by RREAs in air showers created by cosmic particles with energies between 10^{11} and 10^{12} eV. However, they do not elaborate whether a discharge would actually start — according to our analysis below it wouldn't — and their frequency of cosmic particles is 2 or 3 orders of magnitude smaller than in the Review of Particle Physics 2014 [194]. Furthermore, the frequency dependence of the dielectric permittivity $\epsilon(\omega)$ of ice has to be taken into account when calculating the field enhancement near a frozen hydrometeor

— it is 90 for static electric fields, but only 3 for fields changing on a nanosecond time scale.

10.1.1 Structure of the approach.

Whether lightning can be started by an extensive air shower hitting a hydrometeor, depends (i) on the distributions of hydrometeor sizes and shapes, (ii) on the distribution of electric fields in the thundercloud, and (iii) on the distribution and properties of extensive air showers created by high energy particles penetrating the atmosphere. Here we determine one set of parameters in this high-dimensional space that is likely to start lightning.

We start with analyzing the requirements on hydrometeor sizes and shapes, on background electron density and on background electric field to start a discharge, and we perform simulations showing that and how the discharge actually starts under these conditions. Then we analyze the energy of the galactic cosmic particles required to create the necessary density of free electrons. Finally we investigate the probability that the requirements on hydrometeors, electric fields and cosmic particles coincide.

The altitude for our calculations is 5.5 km, a typical altitude for lightning inception [195]. According to the International Standard Atmosphere, we assume $T = 250$ K, $p = 500$ mbar, and hence an air density $n = 0.6n_0$, where n_0 is at ground level. The transport and reaction coefficients (electron mobility and diffusion and effective Townsend coefficient including 2- and 3-body attachment) for an air discharge are calculated with BOLSIG+ [196] with Phelps database.

10.2 Requirements

10.2.1 Hydrometeor size and shape as a function of the background field

A frozen hydrometeor moving in a thundercloud electric field that changes on a millisecond time scale or more slowly will locally enhance the field due to its high dielectric permittivity $\epsilon = 90$. To start a self propagating streamer discharge, a free electron needs sufficiently many ionization lengths in this high field region. This ionization length as a function of the local electric field E is given by the inverse of the effective Townsend coefficient $\alpha_{\text{eff}}(E)$, which is basically the balance of electron impact ionization and electron attachment (hence $\alpha_{\text{eff}}(E_k) = 0$ defines the breakdown field E_k). The electron avalanche multiplication factor e^M along a given path is given by the Meek number $M = \int \alpha_{\text{eff}}(E)dz$ that is widely used in electrical engineering. In our simulations, a Meek number of 10 was sufficient, and we take that number as a benchmark.
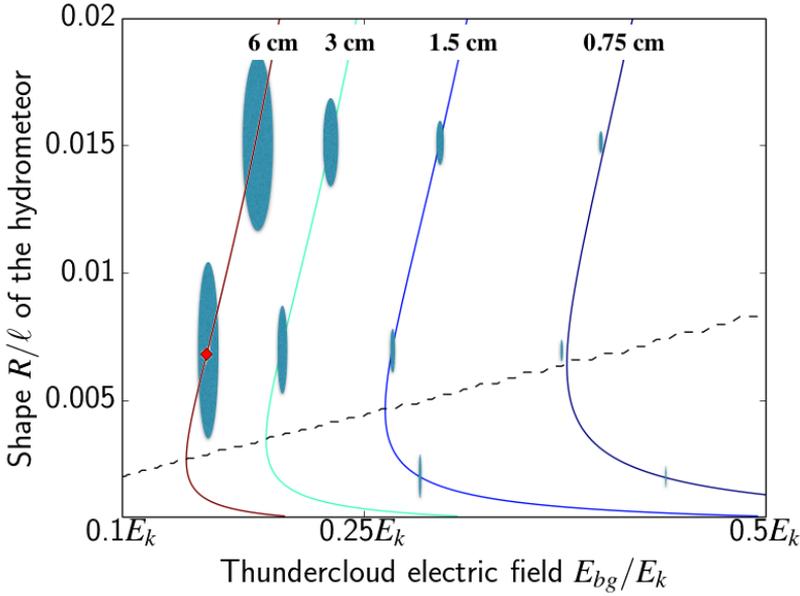
Figure 10.1: Conditions of discharge inception at 5.5 km altitude. The lines of constant hydrometeor length ℓ show when a discharge can start. The axes show the reduced thundercloud field E_{bg}/E_k and the shape parameter R/ℓ. The red diamond indicates the case in Fig. 10.2. The dashed curve indicates the optimal R/ℓ ratio for given E_{bg}/E_k. The ellipsoids are drawn to scale to illustrate the hydrometeor shape and length.

Hydrometeors appear in a large variety of shapes, yet their shape in the direction perpendicular to the thundercloud field does not contribute much to the field enhancement at their tip. Essentially the length of a hydrometeor ℓ and its radius of curvature R at the tip parallel to the field determine the electric field enhancement near the tip. Therefore, we approximate the hydrometeor as a prolate ellipsoid of revolution with length ℓ and radius of curvature R. (In this case the pre-discharge field can be calculated analytically [128].) For a Meek number of $M_0 = 10$, the hydrometeor length ℓ required to start a discharge is a function of the reduced background electric field E_{bg}/E_k and the R/ℓ ratio, and is given by $\ell = M_0 n_0 / n F(R/\ell, E_{bg}/E_k)$. The Meek number is calculated on the symmetry axis where the field is above the breakdown value E_k.

Fig. 10.1 shows as a result the conditions for a discharge to start at 5.5 km altitude. The contour lines indicate the lines of constant hydrometeor length ℓ as a function shape parameter R/ℓ and the reduced thundercloud field E_{bg}/E_k. The lines are altitude dependent, and only the length ℓ has to be rescaled. At 0 km, the lines are from left to right 3.4 cm, 1.7 cm, 0.85 cm and 0.43 cm. At 8 km, they are 7 cm, 3.6 cm, 1.8 cm and 0.95 cm, respectively. Lightning inception is possible either for large hydrometeors (the left part in Fig. 10.1) or

for high electric fields (the right part). In other words, to create a sufficiently large number of electron multiplications, a hydrometeor in a given thundercloud field not only has to be sufficiently sharp to significantly enhance the field, but also sufficiently long to enhance it in a large region. The optimal aspect ratio R/ℓ for given E_{bg}/E_k is indicated by the dashed curve. For a given length ℓ, hydrometeors sharper than optimal are less likely to initiate a discharge, even though they have a higher electric field directly at the tip.

Below we present the full analysis for one case indicated by the red diamond in Fig. 10.1. In our simulations, we chose the parameters to minimize the electric field within the parameter space, and therefore the size of the HM had to be on the extreme side in a thundercloud. Yet, such hydrometeors are observed with a density of roughly 0.1 m^{-3} [197].

10.2.2   Requirement on the density of free electrons.

The Meek number analysis can be applied if there is at least one free electron ahead of the positive end of the hydrometeor at such a distance that it can drift toward it within its time. For an air density of $n = 0.6n_0$, the effective electron life time (based on attachment and impact ionization times) is approximately 30 ns. By tracing the electron drift from the hydrometeor surface backward in time for half the life time, we found that they came from a volume with 1.5 mm radius and length, i.e., from a volume of 10 mm^3. This means that with a homogeneous density of 100 free electrons per cm^3, on average one electron will be available within the relevant volume to start the discharge

10.3   Simulation of actual discharge inception from the hydrometeor.

The discussion above suggests that a positive streamer (developing subsequently into a lightning leader) can start from an hydrometeor with a radius of curvature of $R = 0.4$ mm and 6 cm length in a thundercloud field of 2.7 kV/cm at 5.5 km altitude when the density of free electrons is initially at least 100 cm^{-3}. We now take these parameters as an input for our 3D cylindrically symmetric discharge model and investigate whether a streamer discharge actually forms and propagates.

The discharge is modeled with the classical diffusion-drift-reaction model of [198] with space charge effects, and with photoionization included as in [199].

The hydrometeor is modeled as a dielectric; the dielectric function $\epsilon(\omega)$ of ice depends on frequency ω, it is 90 for slow responses and 3 on the nanosecond time scale [47]. We assume that electrons, when reaching the hydrometeor, attach to the surface. No transport, reactions or secondary electron emission are assumed

on its surface.

The length of the simulation domain is 8.5 cm and its diameter is 4 cm, sufficiently much larger than the hydrometeor that the background field can be fixed by appropriate Dirichlet boundary conditions for the electric potential on the boundary. Together with the hydrometeor with dielectric constant 90, this fixes the stationary field. The discharge plasma develops its own electric field. Due to the superposition principle, this can be calculated from the charge distribution within the discharge with a constant potential on the outer boundary, and with a dielectric response of the hydrometeor with $\epsilon = 3$, and then added to the other field.

The equations are discretized on a static nonuniform grid. The grid is refined in the area where a streamer is expected to propagate. The size of the finest grid cells is 1 μm. Away from the area of streamer propagation grid cells quadratically increase in size up to 0.2 mm on the boundaries.

The result of our simulations is shown in Fig. 10.2. The left panel shows the electron density and the right panel the electric field strength after 46 ns. Clearly a streamer discharge with its strong field enhancement ahead of the tip has formed below the hydrometeor. The streamer incepts after about 25 ns, leaves the area of enhanced electric field and propagates due to its self-generated field enhancement into a region where the field is below the breakdown field. The average streamer velocity is about 10^5 m/s. The small ϵ of ice on the ns time scale substantially hinders the streamer propagation; if we would erroneously take $\epsilon = 90$ on all time scales, the streamer would propagate twice as fast. As the ice responds with a small ϵ to the rapid streamer evolution, the field penetrates into the tip of the hydrometeor, as the right panel of Fig. 10.2 shows.

10.4   Occurrence rate of required electron density

Without assuming a sufficiently high thunderstorm electric field of sufficient extension to form relativistic runaway electron avalanches as in [193, 200], energetic cosmic particles are able to produce enough free electrons in so called extensive air showers [194]. We will now calculate the occurrence rate of such events based on models developed for cosmic ray physics.

We focus our analysis on protons with energies between 5×10^{15} and 5×10^{16} eV. Our calculations described below show that below 5×10^{15} eV it is very unlikely that the core of an extensive air shower reaches a density of 100 thermal electrons per cm^3 at 5.5 km altitude, while above 5×10^{16} eV the electron density is always sufficient, but the occurrence rate decreases with the energy E of the cosmic particle as E^{-2}.

Cosmic protons with energies between 5×10^{15} and 5×10^{16} eV first interact with an air molecule at about 15 to 25 km altitude, which marks the begin of the
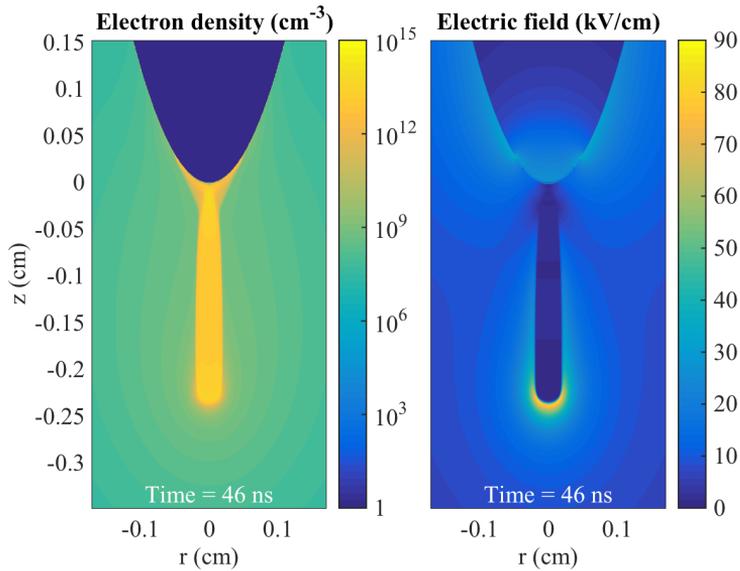
Figure 10.2: Streamer below an icy hydrometeor of length $\ell = 6$ cm and curvature radius $R = 0.4$ mm in a background field of $E_{bg}0.15$ E_k at 5.5 km altitude. Left: the electron density (green to yellow) with the hydrometeor (blue). Right: The electric field strength. Both figures zoom into $r \leq 2$ mm and -0.35 cm $\leq$ z ≤ 0.15 cm after 46 ns of simulation.

shower. The number of secondary particles in the shower then increases until it reaches a maximum between 4 and 8 km altitude, depending on the energy of the primary particle, on the inclination of the particle trajectory with respect to the vertical axis and on the altitude of first interaction. The shower develops downward with nearly the speed of light, and resembles a disc of high energetic particles that leaves a trail of non-relativistic particles behind. The shower has a narrow core with a very high particle concentration.

Extensive air showers can be simulated by the Monte Carlo program COR-SIKA [101] that follows particle movement and interactions explicitly down to an user defined energy threshold of at least 50 keV. However, we need the density of free electrons in the eV range that could start the discharge shown in Fig. 10.2. So we need to follow the particle generation and motion below CORSIKA's energy threshold. As the cross sections for impact ionization by electrons and positrons are four orders of magnitude larger than for photo-ionization, while the photon number is only a factor 20 larger, we only calculate the lower energy electrons generated by electrons and positrons. Furthermore, we use the fact that electrons and positrons of 1 MeV or less can not travel for more than 1 meter at 5.5 km altitude [72], and we only post-process the electrons and positrons within 3 meters from the core center to derive the electron density within the core. The particle density above this threshold is calculated using the restricted collisional stopping power [] divided by the net cost per ionization of about 20-30 eV (that depends on energy as elaborated in [72]), this strategy is in line with [200]. This combination of approaches determines the thermalized free electron density within the core of each shower.

We have simulated 297 showers with primary proton energies between 5×10^{15} and 5×10^{16} eV and with random inclinations, and we have determined the flux of electrons and positrons with energy above 1 MeV in the shower core at 5.5 km altitude. A typical output are 10 to 40 million particles with roughly 90 % photons, 4 % electrons and 4 % positrons. From these we determined the density of electrons with eV energy within the core of 1 meter radius according to the prescription above. Taking the measured differential cosmic ray flux [194] into account, the density of 100 thermal electrons per cm^3 in the core is reached with a frequency of at least 5 km^{-2} min^{-1} within our sample.

10.5   Probability of coincidence of large hydrometeor and extensive air shower inside the thundercloud field.

As sketched in Fig. 10.3, we require three phenomena to coincide in space and time: a thunderstorm electric field, a hydrometeor that is large and elongated enough, and an extensive air shower creating a sufficient density of thermal electrons.
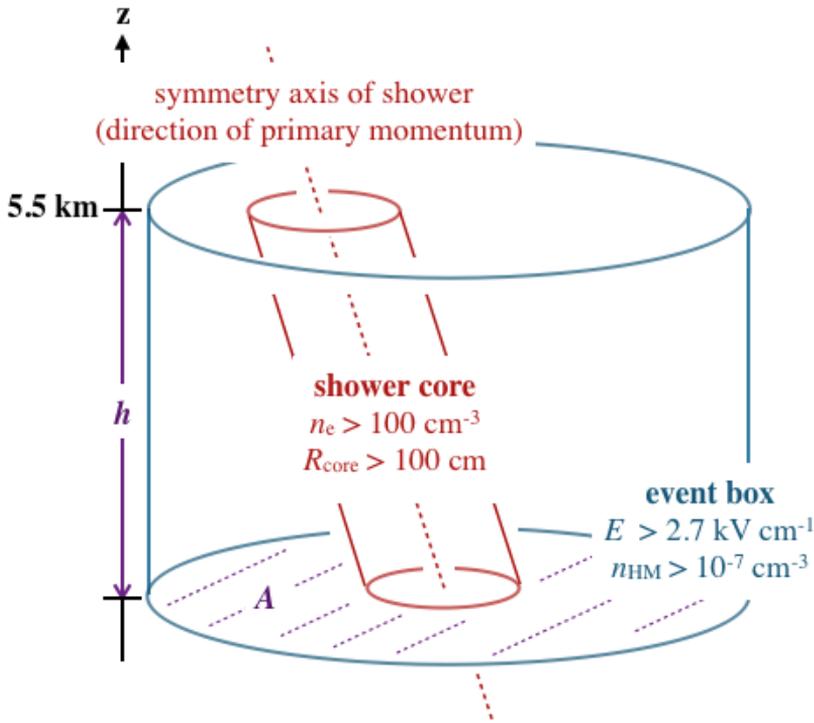
Figure 10.3: Sketch of the model ingredients; not to scale. In the event box (blue) at an altitude of 5.5 km we assume an electric field $E \geq 2.7$ kV cm^{-1}, a density $n_{\mathrm{HM}} \geq 10^{-7}$ cm^{-3} of hydrometeors of correct shape and size, and at least one extensive air shower creating a thermalized free electron density $n_e \geq 100$ cm^{-3} in a core (red) of radius ≥ 100 cm. The height h and horizontal area A of the event box are determined by shower distribution and inception frequency.

First, at least one hydrometeor of sufficient size and shape has to be inside the core of the air shower. For a core radius of 100 cm and a hydrometeor density of 10^{-7} cm^{-3}, a height of $h \approx 3$ m is sufficient. This limits the spatial height where the electric field has to exceed 2.7 kV cm^{-1} to 3 m as well.

Second, the air shower has to hit the horizontal area in the cloud where the electric field exceeds 2.7 kV cm^{-1}. With at least 5 sufficiently energetic showers per km^2 and per minute, we find at least one shower per second, if the horizontal high field area A is 12 km^2, or equivalently 1 shower per minute, if the area is 0.2 km^2.

10.6   Summary and discussion

We have found that elongated hydrometeors of centimeter size are required to start a discharge at 5.5 km altitude, if the background field is as low as 0.15 times the classical breakdown field, and our simulations with a realistic permittivity $\epsilon(\omega)$ of ice showed that a streamer discharge actually can emerge, if 100 free electrons per cm^3 are present. But due to electron attachment to oxygen and the formation of water clusters around these ions, free electrons are typically too rare for a discharge to start. However, a cosmic proton with energy above 5×10^{15} eV can create an extensive air shower whose core can provide the necessary electron density. These air showers appear with a frequency of 1 per 0.2 km^2 per minute, and hence can explain how lightning discharges can start in an undervolted region. According to this argument, lightning inception at higher altitudes – say 12 km – is less likely as the hydrometeor has to be larger and as the electron density in the extensive air shower is lower at these altitudes.

The strongest constraint in our analysis comes from the sizes of the hydrometeors which are little investigated within thunderclouds. Mason [201] suggested already in 1953 a correlation between graupel size and lightning inception — which should be studied further.

The height of the zone where the electric field has to exceed the runaway threshold is only of the size of meters, rather than of more than a kilometer as required for a relativistic runaway avalanche, and relativistic avalanching in an external electric field is not required in our model. Rather all free electrons are generated by the energy of the primary cosmic particle. The required height of the electric field zone is inversely proportional to the density distribution of large hydrometeors.

Finally, in our rare event analysis for the electron density in the shower core, we have focussed on a core radius of 1 m. Future analysis might reveal even more favorable core radii and related primary particle energies.

Chapter 11

Availability of extreme electron seeds in thunderclouds

As discussed in the last chapter, lightning needs free electrons to initiate. We use rare event analysis, from a large set of simulated extensive air showers, to calculate the availability of extreme electron seeds in thunderclouds. We present the available electron densities as function of altitude and rareness of the events. We observe that free electron densities at altitudes between 5 and 13 km do not exceed values of 5×10^3 cm^{-3} and then only in cores of centimeter scale. Above 6 km, the availability of extreme free electron densities decreases significantly with higher altitudes. Recent measurements [202] reveal streamers that must be triggered simultaneously, presumably by an extensive air shower event. The measurements show further that the streamers are laterally separated by more than tens of meters; so the presented results must be triggered by densities down to 1 cm^{-3}. Such low electron densities demand a stochastic approach for streamer initiation, as presented in chapter 6.

11.1 Introduction

Thunderclouds produce strong electric fields, which drives lightning. The initiation of lightning, however, is an active line of research, because the measured electric fields are too low for classical breakdown of the air [6, 203, 204]. The three important ingredients for starting a discharge in a thundercloud are known, as discussed in the last chapter 10: the thundercloud electric field, a hydrometeor (a droplet or ice particle) and free electrons. If the ingredients are combined in the correct way a preliminary discharge, called a streamer, could be initiated.

Dubinova et al. [26], Sadighi et al. [29], Babich et al. [30], Liu et al. [34], Babich et al. [205] and references therein, have studied, this system with computer simulations, identifying if and how a streamer would start from a hydrometeor, see also chapters 5, 6 and 10. Typically, assuming a certain sub breakdown thundercloud electric field, some type of large and or charged hydrometeor and an initial free electron seed are assumed. More free electrons increase the probability that the streamer starts, but in the humid environment of the thundercloud, free electrons are actually hard to find [26]. The crux is that although free electrons are repeatedly refreshed by background radiation, i.e., by cosmic rays and radioactive decay in a rate of the order of 10 cm^{-3}s^{-1} [206], they attach to electronegative molecules which in their turn act as condensation nuclei forming ion-water clusters, efficiently removing the electron. For example, electrons attach to oxygen on a timescale of the order of 50 ns [207] forming O_2^-, and in microseconds they form ion-water clusters $O_2^-(H_2O)_n$ [192]. As a direct consequence, the steady state of the free electron density is only 5×10^{-7} cm^{-3}, making a streamer initiation from a hydrometeor very unlikely. The only way to get a significant free electron seed is in a very energetic cosmic ray event, called an extensive air shower (EAS), as was introduced in this context by Dubinova et al. [26]. In that work it was shown that the free electron densities, produced in the core of an EAS, can be nine orders of magnitude larger than the steady state, but they occur only rarely. In this work we calculate the availability of extreme electron seeds in thunderclouds, by performing a rare event analysis of EAS events.

The quest for extreme electron seeds from EASs has recently become even more crucial, as the leading theory is now that lightning initiates from a volumetric system of streamers [202]. In this theory multiple streamers from multiple hydrometeors should initiate close together at the same moment in time. That is to say, synchronous on the nanosecond streamer timescale. The reasoning could be that a single streamer, with typical dimensions of only millimeters to tens of centimeters, would not be able by itself to evolve to a full-fledged lightning stroke. The theory is linked to the observations of narrow bipolar events and is called fast positive breakdown [202].

This study is organized as follows, we first introduce the physical system in Sec. 11.2, which are the extensive air showers producing free electrons. In
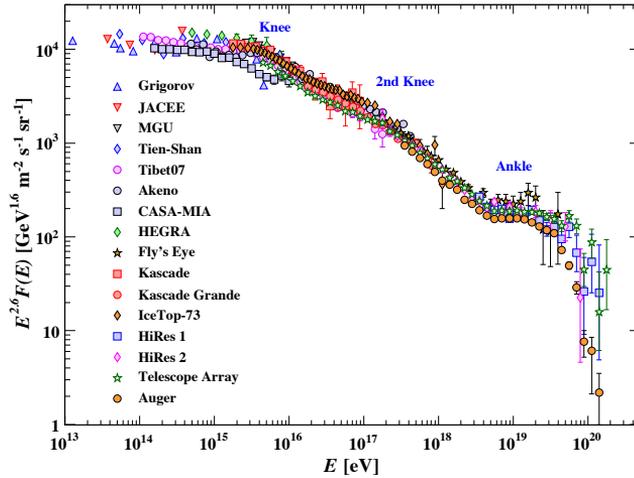
Figure 11.1: The all-particle spectrum as a function of energy E (eV) (energy-per-nucleus) from air shower measurements, Fig. 29.8 from review of Particle Data Group [208]. Mark that the spectrum is rescaled with a factor of $E^{2.6}$.

Sec. 11.3 we explain our methodology. After presenting the results in Sec. 11.4 we discuss in Sec. 11.5 their implications in context of lightning initiation. We conclude in Sec. 11.6 with an outlook on further studies related to this work.

11.2 Physical system

11.2.1 Cosmic ray induced extensive air showers

In Fig. 11.1 the flux spectrum of cosmic rays is given as function of energy, which are mainly protons as most of the universe consists of hydrogen. The flux (of a particular energy interval) should be interpreted as the average number of cosmic rays per unit of area, time and solid angle. Furthermore, the flux is uniformly distributed per solid angle (i.e. isotropic) and it follows a Poisson distribution in space and time. It is important to note that this is different from cosmic rays from the Sun (i.e. with energies below about 10^{11} eV) which are non-isotropic (the flux is strongest near the poles because of the magnetic field) and the flux changes in time (day and night cycle and anti-correlation with the solar-11y-cycle) [208]. The spectrum steepens around the energy of 10^{15} eV, called the *knee* and reflecting the fact that most cosmic accelerators in the galaxy have reached their maximum energy. From the so called *ankle*, at 3.1×10^{18} eV, extragalactic flux begins to dominate over the galactic flux [208].

Approximating the cosmic ray flux from Fig. 11.1, with energy from 10^{14} to 10^{15} eV as proportional to $E^{-2.7}$ and above 10^{15} eV to 10^{18} eV as proportional to E^{-3} we can quote typical fluxes per energy interval, see Tab. 11.1. In Sec. 11.3

Table 11.1: Flux per energy interval, defined as $f_i = \int_{10^i}^{10^{i+1}} F(E)\,\mathrm{d}E$. For the quoted flux we use Fig. 11.1 and approximate with 10^{13} to 10^{15} eV as proportional to $E^{-2.7}$ and 10^{15} eV to 10^{17} eV as proportional to E^{-3}. We use a solid angle of 4.13 sr, i.e. using an maximal inclination of 70 degrees.

Symbol	Interval (eV)	Flux ($\mathrm{km^{-2}\,s^{-1}}$)
f_{14}	$[10^{14}, 10^{15}]$	1.1×10^3
f_{15}	$[10^{15}, 10^{16}]$	39
f_{16}	$[10^{16}, 10^{17}]$	0.4
f_{17}	$[10^{17}, 10^{18}]$	4×10^{-3}

on methodology we return to these fluxes and explain which cosmic ray energies are taken in this study.

Once a high energy cosmic ray enters the atmosphere and collides with an air molecule it creates a cascade of secondary particles. Basically, every collision of a secondary particle with still sufficient kinetic energy is creating new secondary particles by converting energy into mass. This cascade will continue down to MeV energies; creating electrons and positrons which are the stable elementary particles with the minimum rest mass; this is called an extensive air shower (EAS). In Fig. 11.2 an illustration is given, where the top view illustrates that within a certain area A and time interval T of a cloud system multiple EAS will be created with variable primary energy, see also Tab. 11.1 for typical fluxes. Particles in the EAS have typically much larger kinetic energies than their rest mass, so they behave relativistically and move practically with the speed of light. In the side view of Fig. 11.2, the time integrated paths of the particles in the EAS are given as illustration. The EAS should be thought of as a thin pancake structure moving with almost the speed of light towards the ground (two snapshots are illustrated in green). The particle flux in this pancake is most dense around the symmetry axis of the shower, called the core of the EAS.

11.2.2   System decoupling

While the pancake of particles in the EAS moves through the air downwards, all particles interact with the air and loose energy, for example by ionization of air molecules. Sporadically they transfer significant energy to the bound electron promoting it to the EAS population, but most of the time they only free an electron and transferring up to a few tens of eV of energy, as can be seen in the left panel of Fig. 11.3 for electron impact ionization. Two important remarks should be made here. Firstly, as the energy loss per collision is so small compared to the kinetic energy of the particle in the EAS, it is fair to approximate it as a continuous energy loss per unit of length or friction for the EAS particle. The friction of an EAS electron producing secondary electrons below 50 keV is given in blue in the right panel of Fig. 11.3. Secondly, the ionization cross
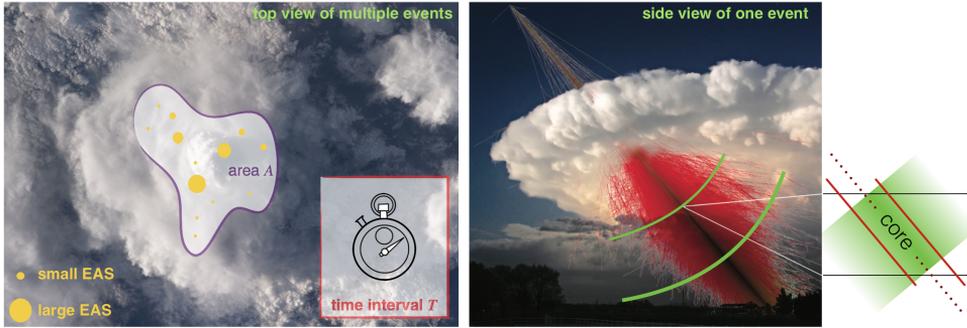
Figure 11.2: Illustration of the system. Left panel is the top view of a thundercloud, where within a certain area A and time interval T of interest multiple EAS events take place. Some small, some large, randomly distributed in time, location and size, but we assume that they do not interact or overlap. Right panel is the illustration of the side view of one single event. In red the time integrated paths of the particles of the extensive air shower. In green are two snapshots in time to illustrate that it is actually a pancake structure moving with almost the speed of light towards the ground. Around the axis (the momentum vector of the initial cosmic ray) particles are concentrated in a dense core.
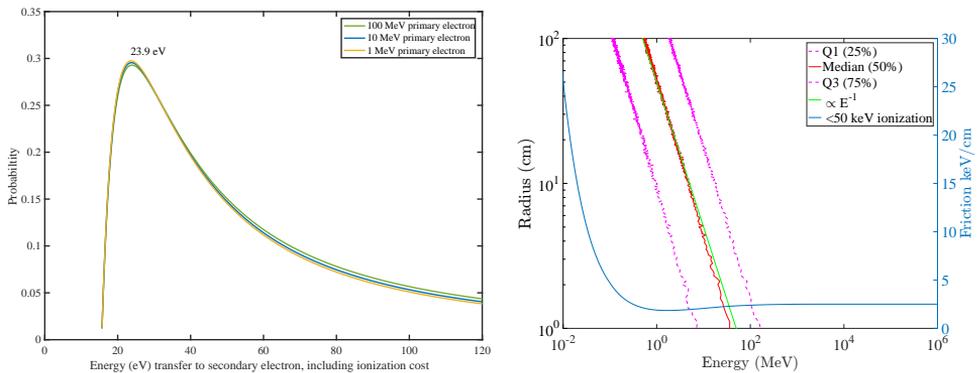
Figure 11.3: Left panel, secondary energy distribution for electron impact ionization for different initial energies, using the RBEB theory. Right panel, the typical (i.e. the boxplot: Q1, median and Q3) energy distribution of electrons and positrons in EASs show a typical pattern. The sub collisional friction, i.e. the energy into ionizations below 50 keV is for energies above 0.1 MeV independent of kinetic energy, mark the linear scale. The average friction used in this study is equal to 2.5 keV/cm. Friction data are from [108, 109] and displayed here for an air density of 1.293×10^{-3} g cm^{-3} corresponding to 1 bar and 273 K.

section to produce electrons say below 50 keV is not sensitive to the energy of
the EAS particle if it has enough energy. Equivalently, for all energies above
0.5 MeV the friction for producing electrons below 50 keV is nearly constant.
As a result, there is a strong decoupling in energy: between EAS particles and
free / seed electrons produced by the EAS. EAS particles are particles with
high energy (typically much larger than MeV) that behave relativistically. Free
electrons produced by the EAS are of order tens of eV, move very slowly and
behave classically. They are basically just left behind by the EAS, while to the
EAS moves practically with c. As is described in detail in the methodology in
Sec. 11.3, for computational reasons it is actually critical to decouple the system
into these two populations. To recapitulate,

$$\text{system decoupling} \begin{cases} \textbf{EAS particles}: \text{relativistic and moving practically with } c, \\ \textbf{free / seed electrons}: \text{left behind, nearly thermalized} \\ \qquad \text{and static w.r.t. the EAS.} \end{cases}$$

$$(11.1)$$

It turns out that only EAS electrons and positrons leave a significant trail of ther-
malized free electrons behind, even though the flux of EAS photons is roughly
ten times more (see Fig. 11.4). The reason is that the friction of electrons and
positrons is almost six orders of magnitude larger compared to photons. The
free electrons can be approximated as an EAS particle yield,

$$Y(h) = F(h)/W = 74 \text{ cm}^{-1} \left(\frac{n_{\text{air}}(h)}{n_{\text{air}}(0)} \right), \qquad (11.2)$$

as function of altitude h, using an average friction of 2.5 keV/cm (see Fig. 11.3)
at STP density $n_{\text{air}}(0)$. Note that the yield scales with air density $n_{\text{air}}(h)$, it
is lower at higher altitudes (h). In the yield the energy cost per free electron
is taken as $W = 34$ eV from Jesse and Sadauskis [209], Cole [210], Knoll [211].
This procedure is similar to the one in [212] which uses $W = 35$ eV/ion and in
the work of Dwyer and Babich [200] who use $W = 34$ eV/ion.

11.2.3   Randomness of EASs

In Fig. 11.4 panel A to C, the longitudinal evolution of the EAS is given for
five random showers per panel for a different primary energy. The EAS par-
ticle output as function of altitude is grouped in photons, electrons, positrons,
muons and hadrons. Mark the logarithmic scale, the three most significant EAS
particle types are photons, electrons and positrons. Electrons and positrons are
produced in a symmetric way by pair production, but with a small surplus of
electrons due to sporadic hard impact ionization, as already discussed above.
The longitudinal shower looks like a bell-shaped structure, where the altitude
of the shower maximum decreases on average with increasing primary energy,
because higher energy particles penetrate deeper into the atmosphere. For the

average EAS the maximum number of particles in the shower scales roughly linearly with primary energy. Nevertheless, there is a very large shower-to-shower fluctuation mainly due to the random inclinations and first one hundred or so collisions. Only thereafter self averaging will take place. As a consequence, for a given altitude the number of particles is thus not one-to-one related to primary energy. This is illustrated in panel D of Fig. 11.4 with two handpicked showers: even though the 9.5×10^{16} eV shower has more particles at its shower maximum than the 2×10^{15} eV shower at its shower maximum, above 8 km altitude the 2×10^{15} eV shower has more particles than the 9.5×10^{16} eV one.

In Fig. 11.5 an example of electron and positron positions passing at 6 km altitude of a 10^{16} eV shower is given in four different window scales. In panel D the cylindrical symmetry is clearly visible, but in panel A the number of particles in this event are too low to see any structure. The NKG function, named after Kamata and Nishimura [213] and Greisen [214], can be used for a simple approximation of the lateral electromagnetic particle distribution. The NKG function with respect to radius r from the core center at an altitude z is given by,

$$\frac{\mathrm{d}\mathcal{N}(z)}{r\mathrm{d}r\mathrm{d}\phi} = \mathcal{N}(z)\frac{1}{2\pi R_m^2}\frac{\Gamma(4.5-s)}{\Gamma(s)\Gamma(4.5-2s)}\left(\frac{r}{R_m}\right)^{s-2}\left(1+\frac{r}{R_m}\right)^{s-4.5}, \quad (11.3)$$

where Γ is the Euler function, $s = s(z)$ is the age parameter and R_m is the Moliere radius. EASs start typically with a proton, where for hadrons the NKG function can be extended, to include a radial dependence of the age parameter. This would be done by using the so-called lateral age $s_\perp(z,r)$, see Bourdeau et al. [215], Apel et al. [216] and reference therein. As we further describe in the methodology Sec. 11.3, we will approximate the lateral shape of the inner core (i.e. $r \leq 10$ m) just by using the single age parameter $s(z)$ of Eq. 11.3.

It is important to note that EAS happen at distinct places in space and time and thus should be considered as individual, independent and non-overlapping events. Returning back to the top view illustration of Fig. 11.2, given a certain area A of interest and time interval T there is a random number of high energy cosmic rays (see Tab. 11.1 for the expected number) that produce an EAS. They start at a random location and time, with a random energy, random inclination and random evolution through the atmosphere. Thus for a given altitude, the number of EAS particles to occur is random and so are the number of free electrons which are directly proportional to the number of EAS particles (see Sec. 11.2.2). We are interested in the number of free electrons that can occur in thunderclouds, specifically the extreme cases which are rare. Therefore we have to introduce a concept of rareness. We follow a similar strategy as in other fields in geophysics. For example, a 10-year-flood is the strength definition for a rare and extreme flood that has a 10 year expected interval between two events or better to say a 10% probability to occur within one year. A 100-year-flood is rarer and thus more extreme. EAS events happen in expected time intervals
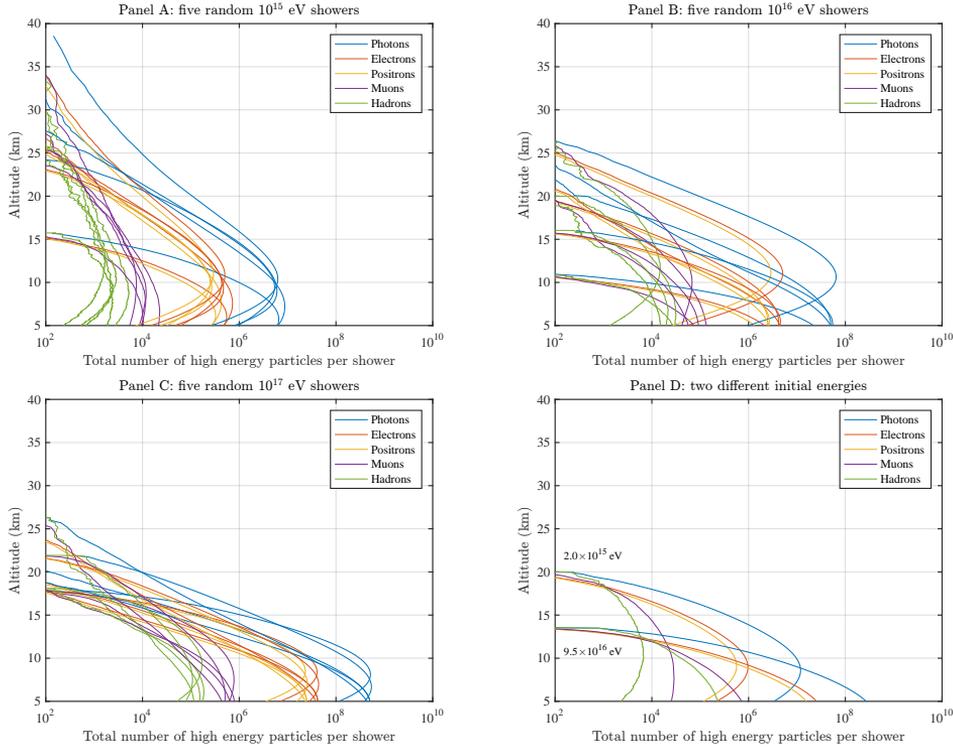
Figure 11.4: Number of passing particles as function of altitude for random EASs produced by cosmic rays with different primary energy. The shown EAS particle counts are above a threshold energy, which is 50 keV for photons, electrons and positrons, 30 MeV for muons and 300 MeV for hadrons. In panel A to panel C five random showers are given, for 10^{15} eV, 10^{16} eV and 10^{17} eV respectively. Here random means that only the initial energy of the proton is fixed, but the initial inclination and all collisions are sampled randomly. The large differences from shower-to-shower are mainly caused by the random inclination and the first one hundred or so collisions, thereafter self averaging will take place. In panel D two different energies are plotted to illustrate the huge shower-to-shower fluctuations and the importance that while higher initial energies are more likely to start at lower altitudes, they can produce less particles than EAS with lower initial energies at some altitudes (in this example above 8 km). The two showers in panel D are chosen just to illustrate this effect, not representing average behavior.
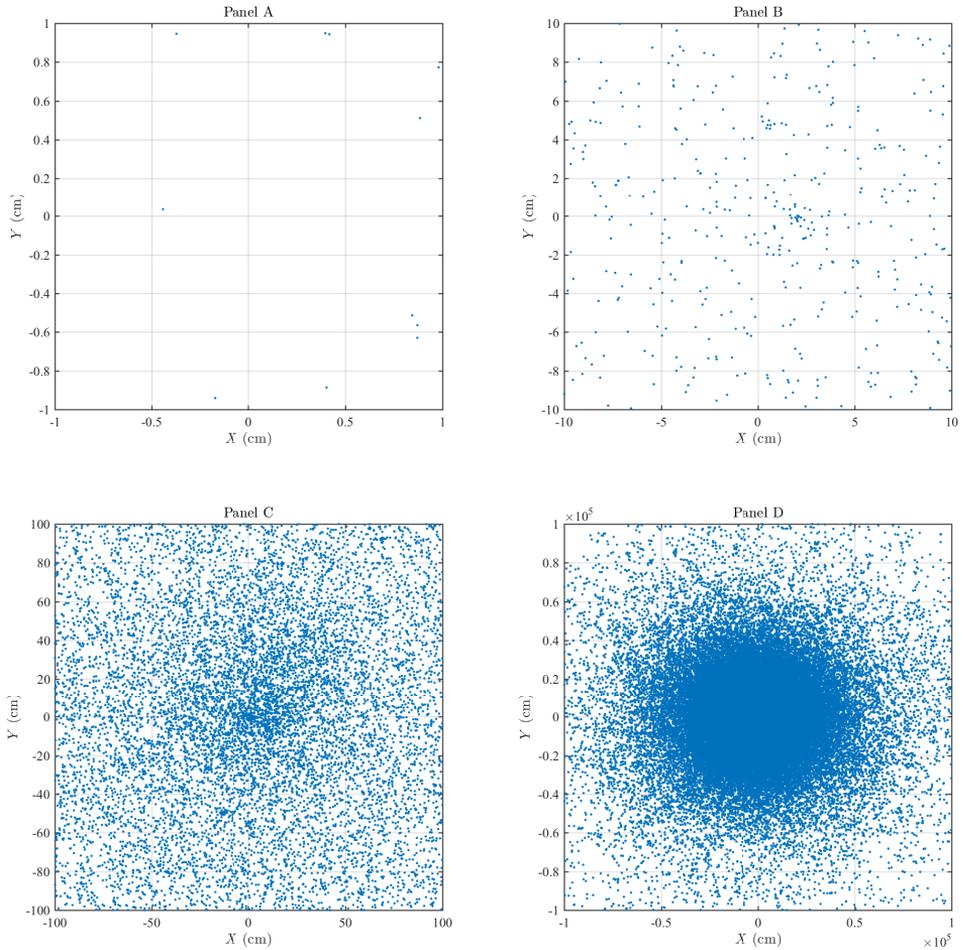
Figure 11.5: Top view of the electron and positron positions of a 10^{16} eV shower passing at 6 km altitude. The four panels A to D give the same data but with different zoom, i.e. axis: $\pm$ 1 cm, $\pm$ 10 cm, $\pm$ 1 m and $\pm$ 1 km respectively. The cylindrical symmetry is clearly visible in panel D, but for radii below 10 cm the total number of particles in the plot becomes to small to see any structure.

T assuming a certain area A and for the rareness only the product $A \times T$ is relevant. We use the notation of 'x-km^{2}s-EAS', where x is the product of $A \times T$. For example, a 4-km^{2}s-EAS is the strength definition for a rare and extreme EAS that has 25% probability to occur per one km^{2}s. We thus define,

x-km^{2}s-EAS: an extreme EAS strength with 1/x probability to happen per km^{2}s.
(11.4)

11.3　Methodology

11.3.1　Simulations

Simulations of the EASs are preformed with the Monte Carlo software package CORSIKA [101]. We use the high-energy hadronic interaction model QGSJET-II [217] and for the low-energy interaction (below TeV) we use FLUKA [218]. We do not use the thinning option (simulating less particles with a higher particle weight) and we follow particles to their lowest energy threshold recommended in the software, which is 50 keV for electrons, positrons and photons. We thus decouple the system (see Eq. 11.1) and calculate the free electrons as described in Sec. 11.2.2. To know the primary cosmic ray energy to study we have to approximate the system size. Following the study of Stolzenburg et al. [219], thunderstorms come in a variety of sizes and lightning intensities, small single-celled storms are generally about 10 km in diameter and their lifetime is typically 60-90 minutes up to supercell storms with a typical horizontal coverage of 250 km^2 and lifetimes of 2-5 hours. The system size in this study is defined as the assumed horizontal area A of high electric field times time interval T between two successive studied events, see for illustration Fig. 11.2. The median per-cell flash rate is of the order of 3 per minute and the median cell nearest neighbor range is 30 km [220]. Assuming that the high electric field areas in the cloud are roughly 1% of the single-cell cloud coverage, we approximate the system size as $\pi(5\,\text{km})^2 \times 1\% \times 20\text{s} \approx 15$ km^{2}s. Comparing this typical system size to cosmic ray fluxes, see Tab. 11.1, we focus our study on cosmic rays with energies in the range of 10^{15} eV to 10^{17} eV. This range is divided in 99 equal intervals of 10^{15} eV, meaning the first interval is $[1, 2] \times 10^{15}$ eV and the last interval (the 99th) is $[99, 100] \times 10^{15}$ eV. Within each interval the primary energy E is sampled randomly with a probability proportional to E^{-3}, meaning we chose the spectral slope of -3 as input. The inclination is sampled randomly, uniform per solid angle, up to a zenith angle of 70 degrees, bound by the choice of using a horizontally geometrical atmosphere (no curvature).

For each shower we first retrieve at altitudes of 5 km to 13 km, in steps of 1 km, the full output (particle ID, momentum, position and arrival time) of each particle passing the hypothetical detector at that altitude. We only

simulate down to 5 km for computational reasons, as in CORSIKA all particles are followed down to the lowest observational level. We post-process this detailed data by counting the numbers $\mathcal{N}_{R_1}$, $\mathcal{N}_{R_2}$ and $\mathcal{N}_{R_3}$ of electrons and positrons within a circle of radius $R_1 = 10\,\text{cm}$, $R_2 = 1\,\text{m}$ and $R_3 = 10\,\text{m}$. after which we delete the detailed output, because of storage reasons. The final output is thus time and energy integrated flux for electrons plus positrons. The choice for only post-processing electrons and positrons, independent of energy and arrival time is reasoned by the following considerations (some parts already mentioned in Sec. 11.2.2):

- The most abundant species in an EAS are: photons, electrons and positrons, in proportion of roughly: 100, 10 and 9, respectively. The ionization cross section, i.e. the probability to produce low energetic electrons, is proportional to roughly: 1, 10^6 and 10^6, respectively. This means that although photons are roughly ten times more abundant, their ability to produce low energy electrons is much lower and thus their contribution is negligible.

- For the energies considered in modeling EASs electrons and positrons produce almost an equal amount of low energy ionizations. Thus we can sum the contributions of electrons and positrons.

- For the energies considered in modeling EASs the amount of low energy ionizations is almost independent of energy, see friction curve Fig. 11.3. Therefore we can energy integrate the number of particles.

- The spread between the first particle to arrive at an altitude and the last 99% is of the order of subnanoseconds, well within the uncertainty introduced by the other approximations (such as the friction to electron yield, see the end of Sect. 11.2.2). Therefore we can time integrate the number of particles.

By integrating the NKG function (Eq.11.3) over the area of a circle with radius R we get,

$$\mathcal{N}_{\text{R}}(z) = \mathcal{N}(z)\frac{\Gamma(4.5 - s)}{\Gamma(1 + s)\Gamma(4.5 - 2s)}\left(\frac{r}{R_m}\right)^s {}_2F_1(4.5-s, s, 1+s, -R/R_m), \quad (11.5)$$

in which ${}_2F_1$ is the ordinary hypergeometric function, we can directly relate the recored output $(R_i, \mathcal{N}_{R_i})$ of the simulations to the three unknowns in 11.3. For all simulations we have numerically solved Eq. 11.5 for $(\mathcal{N}, s$ and $R_m)$ with high precision ($|\text{error}| < 10^{-12}$). Only for the showers that had an empty inner core count $\mathcal{N}_{R_1} = 0$ the unknowns where fitted with a least χ^2 fit.

We used 125 000 CPU hours to compute 5811 EASs and processed 150TB of binary data. The number of simulations per energy interval of 10^{15} eV is given in Fig. 11.6, where the blue line represents a distribution if all EASs where sampled as E^{-3}, we thus have preformed importance sampling, i.e. oversampling the tail of the distribution, for higher accuracy.
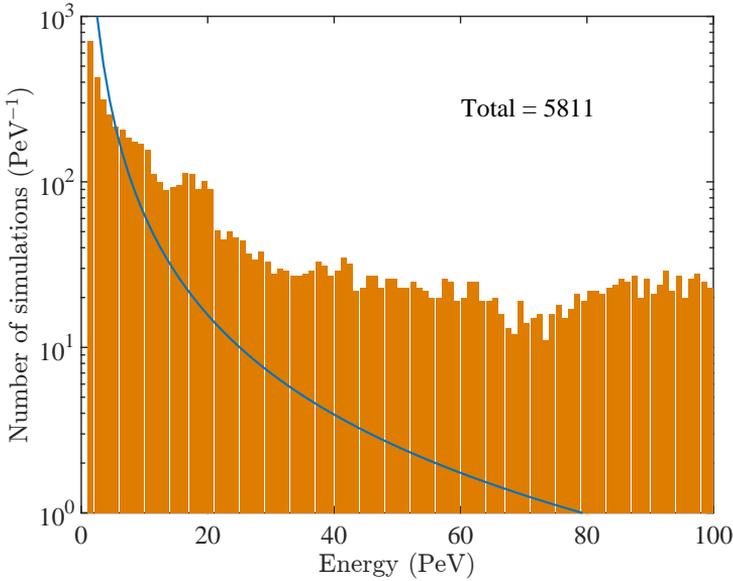
Figure 11.6: Overview of performed simulations. In total we used 125 000 CPU hours to compute 5811 EASs. The blue line represents the expected distribution if sampled as E^{-3}. We have performed importance sampling, i.e. oversampling the tail of the distribution, for higher accuracy.

11.3.2 Rare event analysis

In this subsection we derive an expression for the strength of a x-km^{2}s-EAS (see definition Eq. 11.4), based on the statistical data from the EAS simulations. This section is rather technical and uses notation common in stochastics.

In a given system size ($A \times T$, see illustration Fig. 11.2) and energy interval, the number of EASs K is a discrete random variable and Poisson distributed,

$$\mathbf{P}(K = k) = \frac{\bar{k}^k}{k!} \exp(-\bar{k}), \tag{11.6}$$

where $\bar{k}$ is the expectation value of K, given by the integration of the differential flux Fig. 11.1. See Tab. 11.1 for $\bar{k}$ for several energy intervals.

In a general sense, let the Y be a random output of interest of a single EAS. Now, let the M_k be the maximum of a set of k instances of Y, that is to say

$$M_k = \max\{y_1, \ldots, y_k\}. \tag{11.7}$$

This means that if one considers k EASs and evaluates the k random instances of Y, then there will be one maximum which is set to M_k. M_k thus depends on the number of instances k and on the probability distribution of Y. The larger k, the higher the probability that M_k is near the maximum of possible outcomes

of Y. If k is infinite, than M_k can practically only be the maximum possible value of Y with probability asymptotically going to 1.

As already said in Sec. 11.2.3 EASs happen at distinct places in space and time and thus should be considered as individual, independent and non-overlapping events. So all instances $\mathbf{P}(Y_i \leq y_i)$ are independent and identically distributed which yields,

$$\mathbf{P}(M_k \leq m) = \prod_{i=1}^{k} \mathbf{P}(Y_i \leq m) = \mathbf{P}(Y \leq m)^k, \qquad (11.8)$$

where $\mathbf{P}(M_k \leq m)$ and $\mathbf{P}(Y_i \leq m)$ denote now the cumulative distribution function.

For increasing k we find, by using the Central-Limit Theorem in the log-domain, that $\mathbf{P}(M_k)$ is converging to a lognormal distribution $\ln \mathcal{N}(\mu_k, \sigma_k^2)$, which is analytically,

$$\mathbf{P}(M_k = m) \approx \frac{1}{m \sigma_k \sqrt{2\pi}} \exp\left(-\frac{[\ln(m) - \mu_k]^2}{2\sigma_k^2}\right) \qquad (11.9)$$

and for the cumulative distribution function,

$$\mathbf{P}(M_k \leq m) \approx \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\ln(m) - \mu_k}{\sqrt{2}\sigma_k}\right), \qquad (11.10)$$

where erf() is the error function. Now for an arbitrary number of EASs, we perform the weighted sum over k and call the outcome maximum M. Combining Eq. 11.6, 11.8 and 11.10 thus yields,

$$\mathbf{P}(M \leq m \text{ given } \bar{k}) = \sum_{k=0}^{\infty} \mathbf{P}(K = k) \prod_{i=1}^{k} \mathbf{P}(Y_i \leq m), \qquad (11.11)$$

$$= \sum_{k=0}^{\infty} \frac{\bar{k}^k}{k!} \exp(-\bar{k}) \, \mathbf{P}(Y \leq m)^k, \qquad (11.12)$$

$$\approx \sum_{k=0}^{\infty} \frac{\bar{k}^k}{k!} \exp(-\bar{k}) \left[\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\ln(m) - \mu_k}{\sqrt{2}\sigma_k}\right)\right] (11.13)$$

$$\approx \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\ln(m) - \mu}{\sqrt{2}\sigma}\right). \qquad (11.14)$$

Eq. 11.13 is close to equality when $\bar{k}$ is large and in Eq. 11.14 we approximate the sum of lognormal distributions again as lognormal (Fenton-Wilkinson approximation).

In this work we consider the energy interval between 10^{15} eV to 10^{17} eV, resulting in a $\bar{k} = 39.4$ (Tab. 11.1) with basic system unit of km^{2}s. Now the

strength of x-km^{2}s-EAS in Eq. 11.4 is the value m where $\mathbf{P}(M \leq m$ given $\bar{k} = 39.4) = 1 - 1/x$ or $\mathbf{P}(M > m$ given $\bar{k} = 39.4) = 1/x$. The practical meaning of these statements will be more clear while presenting the results in Sec. 11.4 and discussing their implication towards lightning initiation in Sec. 11.5.

11.4 Results

We have simulated 5811 EASs (see Fig. 11.6) and recorded the number of electrons and positrons passing through three core sizes (10 cm, 1 m and 10 m) at nine observational altitude (5 to 13 km), as described in Sec. 11.3. As an example in Fig. 11.7 the variation of the raw data for core size 10 cm at 8 km is presented. We see the linear trend with energy, that is the number of EAS particles scales roughly linear with primary energy. However, the variation also increases and we see that the 25% largest recored 50 PeV showers have a denser core than the 25% smallest recored 100 PeV showers, reflecting the statement in Sec. 11.2.3 from panel D in Fig. 11.4 that not only primary energy is important: a full energy interval should be analyzed.
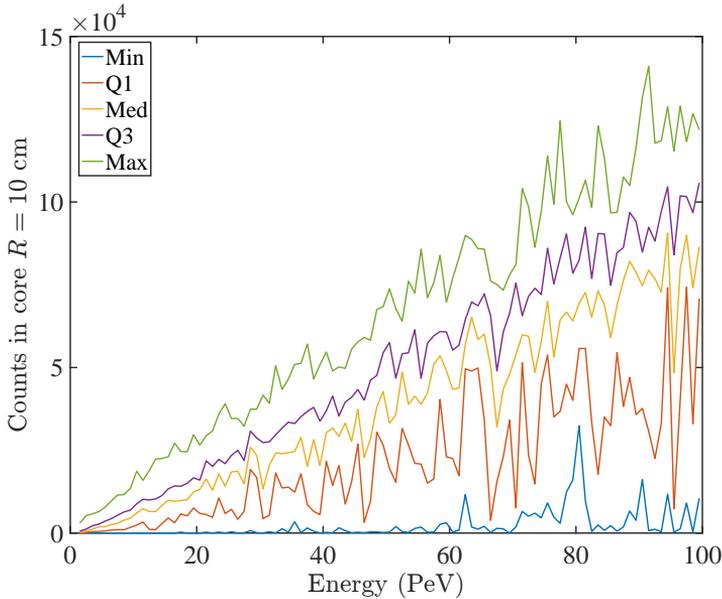
Figure 11.7: Example of variation of dataset recordings of electron and positron passing through a disc core of 10 cm at 8 km altitude. The five lines represent a boxplot: when sorted the minimum, first quartile at 25%, the median, the third quartile at 75% and the maximum value is given.

The first step in the post-processing is a fitting to the integrated NKG function (Eq. 11.5) for the three unknowns ($\mathcal{N}$, s and R_m). In Fig. 11.8 the distri-
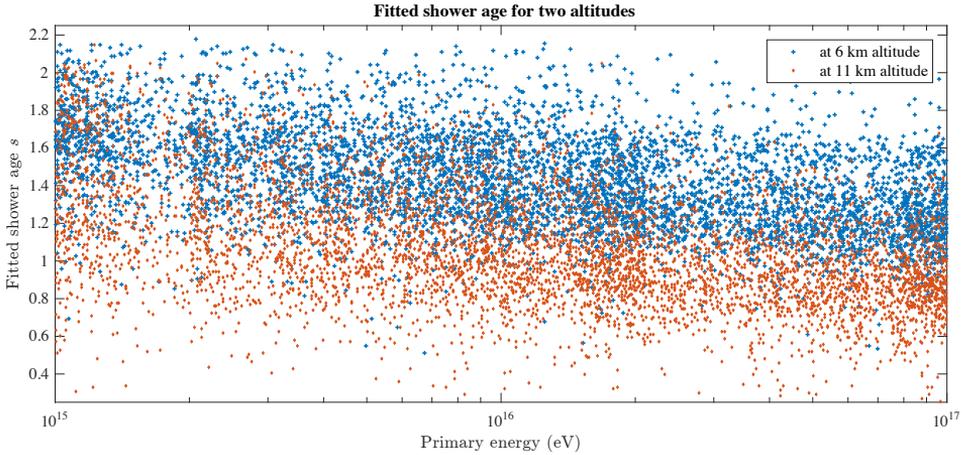
Figure 11.8: Fitted shower age at an altitude of 6 km (blue cross) and 11 km (red diamond).

bution of fitted shower ages is shown for 6 km and 11 km altitude as function of energy. The large variation at each energy displays the large shower-to-shower fluctuations, as introduced in Sec. 11.2.3. Two trends are visible. First, showers at 6 km altitude are older (have a larger age parameter) compared to 11 km altitude. This is natural because showers evolve downwards. Second, at the same altitude the showers tend to be younger (have a smaller age parameter) for higher primary energies. This is also expected as the higher the cosmic ray primary energy the deeper it can penetrate the atmosphere before starting the EAS.

The NKG function (Eq. 11.3) is used to approximate the lateral profile. There are multiple nonequivalent ways to sort a lateral profile in strength. One possibility is to look at the maximum density in the middle of the core of the shower. One other possibility is to look at the extension of the core having a minimum density in mind. Both perspectives are relevant for discharge inception in thunderclouds and are discussed in more detail in Sec. 11.5. Below the results in both perspectives are given.

11.4.1   Maximum expected density in the EAS core

In this subsection we take the fitted NKG density at 1 cm from the core as an output of interest. We compute the cumulative distribution function as described Sec. 11.3.2 and fit the result by a lognormal distribution, which is parameterized as a function of μ and σ. That is to say, we numerically fit the best μ and σ parameters of Eq. 11.14, to match the cumulative distribution function as close as possible. A lognormal distribution, looks like a normal distribution when plotted on logarithmic horizontal scale. μ and σ represent the 'mean' and the

Alt. (km)	μ	σ	median (cm^{-2})
5	3.2062 ± 0.0006	0.6732 ± 0.0004	24.7
6	3.4754 ± 0.0005	0.6144 ± 0.0004	32.3
7	3.6257 ± 0.0005	0.5591 ± 0.0003	37.5
8	3.6876 ± 0.0004	0.5389 ± 0.0003	39.9
9	3.6959 ± 0.0004	0.5175 ± 0.0003	40.3
10	3.6228 ± 0.0004	0.5063 ± 0.0003	37.4
11	3.4435 ± 0.0004	0.4904 ± 0.0003	31.3
12	3.2050 ± 0.0004	0.4748 ± 0.0003	24.7
13	2.9209 ± 0.0004	0.4750 ± 0.0003	18.6

Table 11.2: Fitted lognormal distribution (see Eq. 11.14) of the ionizing flux of electrons and positrons (≤ 50 keV) in the EAS at 1 cm from the center for each observational altitude and system size of 1 km^2s. Given confidence interval is 99%. Altitude 5, 9 and 13 km are plotted in Fig. 11.9.

'standard deviation' of the normal distribution when the horizontal axis is in logarithmic scale. In the real coordinate system we have the following identities:

$$\text{mean} = \exp\left(\mu + \sigma^2/2\right) \tag{11.15}$$

$$\text{median} = \exp\left(\mu\right) \tag{11.16}$$

$$\text{variance} = \left[\exp\left(\sigma^2\right) - 1\right]\exp\left(2\mu + \sigma^2\right) \tag{11.17}$$

Remark, that due to the long tail of the lognormal distribution, it is not symmetric and the mean is larger than the median.

For illustration in Fig. 11.9 the statistical data and their fits are shown for three altitudes. We observe that the ionizing flux at 9 km altitude is larger than at 6 km or 13 km altitude. This optimum occurs because two factors balance, see also discussion about panel D of Fig. 11.4. Parameters of the lognormal fits for all altitudes are given in Tab. 11.2. The lognormal parameter fits are very accurate, the given 99% confidence interval (error bar) is smaller than 0.08%.

The probability that a rare event takes place increases when the system size increases. Using the lognormal distributions from Tab. 11.2 we compute the strength of a x-km^2s-EAS (see definition Eq.11.4) as function of altitude, see Fig. 11.10. In the left panel the ionizing flux is given. The rarer the event the larger the strength, and for 23-km^2s-EAS an ionizing flux of 100 cm^{-2} (of electrons and positrons ≤ 50 keV) can be expected around 8 km altitude. We notice that the maximum lowers in altitude, because the larger the system size the more are significant high energy cosmic rays, which penetrate deeper, contribute. In the right panel the free electron density is presented, calculated by the ionizing flux multiplied by the average yield as function of altitude (see Eq. 11.2). Because the yield drops for lower air densities, that is higher altitudes, we see that the maximum of the free electron density is lower than the maximum in

the ionizing flux. From 23-km^{2}s-EAS a free electron density of 4000 cm^{-3} can be expected around 6 km altitude. We also notice the contour of 1000 cm^{-3} tends to converge to 13 km, thus stating that larger densities are not expected above 13 km altitude for any system size. This convergence is expected because to increase the density one needs to have a stronger EAS to occur, which more likely penetrates deeper into the atmosphere. Hence, large free electron densities are not expected at high altitudes (> 13 km).
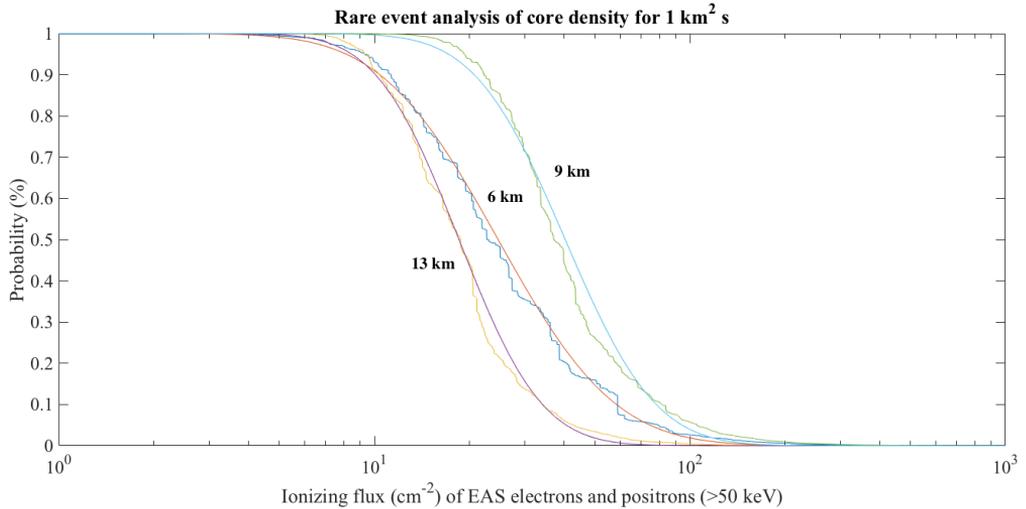
Figure 11.9: Numerical data and lognormal fits (Eq. 11.14) for the commutative distribution function for the ionizing flux of EAS particles in a system size of 1 km^{2}s.

11.4.2 Maximum extension of the EAS core

In this subsection we search for the maximum extension of the EAS core. As in Sec. 11.4.1 the maximum density at 1 cm was used to sort the strength of the EAS, now the output of interest is the core radius where the density is equal to a certain amount. In Tab. 11.4, 11.5 and 11.6 the fitted lognormal distributions for 1 km^{2}s are given for radii with 10, 100 and 1000 cm^{-3} density respectively. As in Sec 11.4.1 the fits are very accurate, the 99% confidence interval (error bar) is smaller than 0.05%. We notice that increasing the density by an order of magnitude the median radius decreases by an order of magnitude. We have plotted radius of the 100 cm^{-3} and 1000 cm^{-3} case in Fig. 11.11 as function of rareness of the event, see Eq. 11.4 for the x-km^{2}s-EAS definition. We see a clear trend that the EAS core is wider at lower altitudes. As it should, the 1 cm contour line of the right panel in Fig. 11.11 is identical to the 1000 cm^{-3} density contour line in the right panel of Fig. 11.10 because at this line both perspectives are equivalent. The core radius for high altitudes (say above 9 km) is almost not
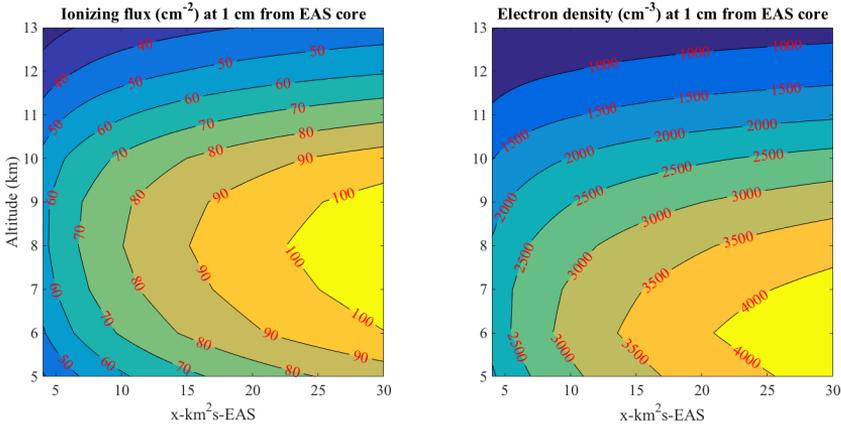
Figure 11.10: Analysis at 1 cm in horizontal direction from the EAS core as function of altitude and the rareness of the event (see Eq. 11.4 for the x-km^{2}s-EAS definition). On the left side the ionizing flux (cm^{-2}) is given, which is the result after post-processing the records of electrons and positrons above 50 keV passing the given altitude, as described in Sec. 11.3. On the right side the resulting thermalized coherent free electron density is presented, which is the left results multiplied by the average yield as function of altitude (see Eq. 11.2).

sensitive anymore to rarer events (contour lines converge to horizontal), implying that it not possible to get much bigger extensions than quoted even for rarer events. For every altitude above 5 km and up to 30-km^{2}s-EAS no shower is expected that has 1000 cm^{-3} and extends more than 6 cm.

11.5   Discussion

As said in the introduction, streamer initiation in thunderclouds require three ingredients [26]: the thundercloud electric field, a hydrometeor (a droplet or ice particle) and free electrons. In this work we focus only on the third ingredient and calculated the maximum number to expect in the thundercloud system. The first noticeable observation between our results and simulation works in the literature is that typically too many initial electrons are considered. In the simulation works of Liu et al. [34] and Sadighi et al. [29] a background density of 10^{11} and 2×10^9 cm^{-3} is taken respectively, while from Fig. 11.10 we see that the maximum density is not exceeding 5×10^3 cm^{-3} even for very large systems and that these very high densities only occur in very thin cores, see Fig. 11.11. Sadighi et al. [29] justify the large initial background density by a pre-existing corona discharge. They refer to the work of Sin'kevich and Dovgalyuk [221] where the number of corona discharges due to hydrometeor collisions are calculated. But the point is, that the initiation part of a corona discharge and

Alt. (km)	μ	σ	median (cm)
5	7.3696 ± 0.0002	0.2274 ± 0.0001	1587
6	7.3089 ± 0.0002	0.2252 ± 0.0001	1494
7	7.2055 ± 0.0002	0.2267 ± 0.0001	1347
8	7.0679 ± 0.0002	0.2314 ± 0.0001	1173
9	6.8929 ± 0.0002	0.2300 ± 0.0001	985.3
10	6.6858 ± 0.0002	0.2529 ± 0.0001	800.9
11	6.4158 ± 0.0002	0.2707 ± 0.0002	611.5
12	6.1017 ± 0.0002	0.2898 ± 0.0002	446.6
13	5.7092 ± 0.0003	0.3177 ± 0.0002	301.6

Table 11.3: Fitted lognormal distribution (see Eq. 11.14) radius where the density is equal to 1 cm^{-3} for system size of 1 km^{2}s. Given confidence interval is 99%.

Alt. (km)	μ	σ	median (cm)
5	5.9065 ± 0.0003	0.3801 ± 0.0002	367
6	5.8438 ± 0.0003	0.3725 ± 0.0002	345
7	5.7060 ± 0.0003	0.3724 ± 0.0002	301
8	5.4970 ± 0.0003	0.3786 ± 0.0002	244
9	5.2232 ± 0.0003	0.3944 ± 0.0002	186
10	4.8890 ± 0.0003	0.4168 ± 0.0002	133
11	4.4463 ± 0.0004	0.4465 ± 0.0003	85.3
12	3.9525 ± 0.0004	0.4742 ± 0.0003	52.1
13	3.4002 ± 0.0004	0.4823 ± 0.0003	30.0

Table 11.4: Fitted lognormal distribution (see Eq. 11.14) radius where the density is equal to 10 cm^{-3} for system size of 1 km^{2}s. Given confidence interval is 99%.

Alt. (km)	μ	σ	median (cm)
5	3.2254 ± 0.0006	0.6870 ± 0.0004	25.2
6	3.2403 ± 0.0005	0.6302 ± 0.0004	25.5
7	3.1436 ± 0.0005	0.5935 ± 0.0003	23.2
8	2.9617 ± 0.0004	0.5594 ± 0.0003	19.3
9	2.7048 ± 0.0004	0.5276 ± 0.0003	15.0
10	2.3881 ± 0.0004	0.5018 ± 0.0002	10.9
11	1.9877 ± 0.0004	0.4736 ± 0.0003	7.30
12	1.5670 ± 0.0004	0.4541 ± 0.0003	4.79
13	1.1136 ± 0.0004	0.4379 ± 0.0003	3.05

Table 11.5: Fitted lognormal distribution (see Eq. 11.14) radius where the density is equal to 100 cm^{-3} for system size of 1 km^{2}s. Given confidence interval is 99%.

Alt. (km)	μ	σ	median (cm)
5	0.1943 ± 0.0007	0.8441 ± 0.0005	1.21
6	0.3847 ± 0.0006	0.7149 ± 0.0004	1.47
7	0.4352 ± 0.0005	0.6258 ± 0.0004	1.54
8	0.3691 ± 0.0005	0.5713 ± 0.0003	1.44
9	0.2446 ± 0.0004	0.5280 ± 0.0003	1.28
10	0.0481 ± 0.0004	0.4946 ± 0.0003	1.05
11	-0.2255 ± 0.0004	0.4601 ± 0.0003	0.80
12	-0.5434 ± 0.0004	0.4289 ± 0.0002	0.58
13	-0.8766 ± 0.0003	0.4109 ± 0.0002	0.42

Table 11.6: Fitted lognormal distribution (see Eq. 11.14) radius where the density is equal to 1000 cm^{-3} for system size of 1 km^{2}s. Given confidence interval is 99%.

a streamer discharge are equivalent. Both need the three ingredients quoted above, including initial free electrons to start. Thus, arguing that the initial electrons come from another discharge is therefore not solving the requirement.

[30, 205] and [26] use an initial density in their fluid model, but argue that only a single electron was needed. They differ in approach of the required ionization integral to start a positive streamer near a hydrometeor, but that is outside of the scope of this work. Important here is that [30, 205] just postulate that there is a free electron available, either by background cosmic ray flux or by cloud charge fluctuations, but they do not give a calculation for it. In [26] a minimum density is calculated by requiring that a free electron must survive the path to the hydrometeor before it attaches which yields a minimum density of 100 cm^{-3}. They did a the statistical analysis of 297 EASs from 5×10^{16} eV to 5×10^{17} eV (without importance sampling), to conclude that a 12-km^{2}s-EAS (0.2 km^2 once per minute) could produce 100 cm^{-3} at a core radius of $R = 1$ m at 5.5 km altitude. Comparing this with the current and more extended work (see Fig. 11.11), we find that the core radius is more accurately[1] $R \approx 63$ cm.

Definitely at least one electron is needed, but more is always better, it softens the Meek criterion [26, 30, 205, 222]. In other words, a larger initial number of electrons will soften the requirement for the other two ingredients, the value of the thundercloud electric field or the hydrometeor (properties such as size, sharpness [26] and or charge [30, 205, 222]). From our results here however, there is not a lot of room for increasing the density much. From Fig. 11.10 and the right panel of Fig. 11.11 we see that a density of a few times 1000 cm^{-3} is possible, but only in a very limited core radius. To find the right hydrometeor there is maybe limiting the initiation probability more than the logarithm of the

---

[1]The consequence of the difference in the work of [26] of 1 m to 0.6 m core radius is only a change of height of the event box from 3 m to 8 m to get the example hydrometeor presented in the study.
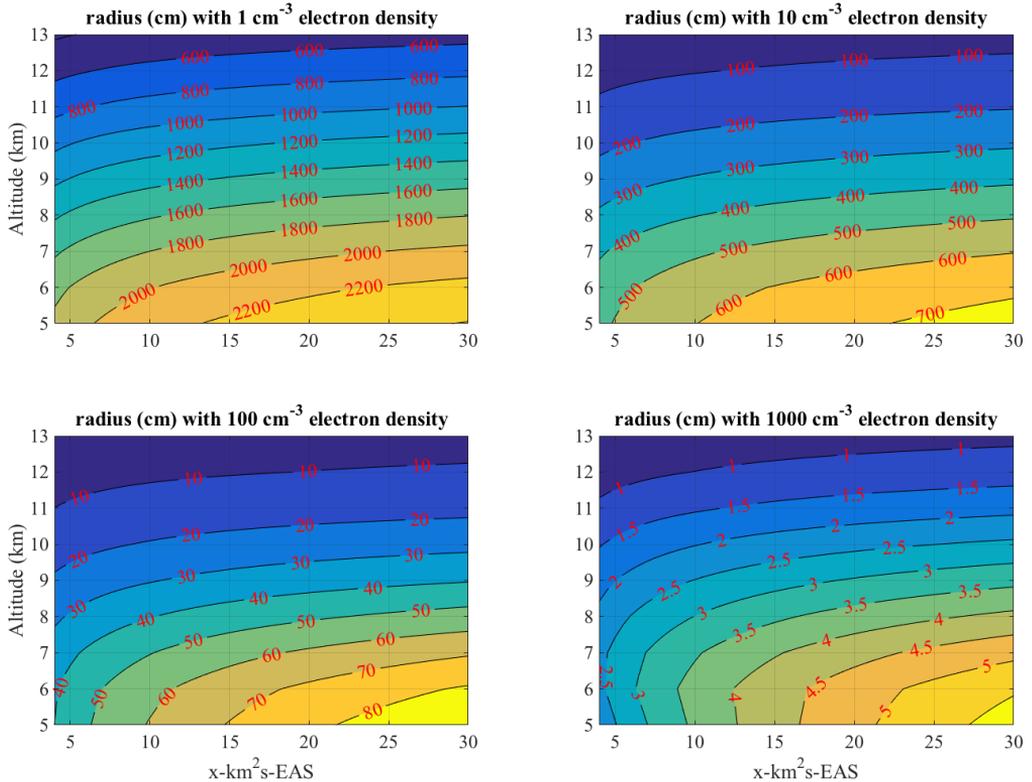
Figure 11.11: Analysis of EAS core radius for given density as function of altitude and the rareness of the event (see Eq. 11.4 for the x-km^{2}s-EAS definition). On panels the core radius is given where the thermalized free electron density is 1, 10, 100 and 1000 cm^{-3} respectively. We have post-processed the time integrated records of electrons and positrons above 50 keV passing the given altitude, as described in Sec. 11.3, and multiplied by the average yield as function of altitude (see Eq. 11.2).

density increase.

As indicated shortly already in the introduction, it would be more crucial to have enough free electrons over a large enough area at the same moment in time when going beyond initiation of only one streamer. In the hypothesis that a system of multiple streamers from multiple hydrometeors are necessary to explain the start of lightning, EAS core needs enough extension. This hypothesis was addressed by Rison et al. [202] from detailed observations of narrow bipolar events (NBE). These type of observations have actually been made before in other measurements where no lightning starts [223, 224], as an indication of attempted breakdown; interpreted as a system of streamers which is triggered, but not capable of developing to a full-fledged lightning stroke [202]. The observations of Rison et al. [202] show a limited vertical extent (less then or equal to

500 m) and fast group speed (typically 3 to 7×10^7 ms^{-1}), moving up and downwards. In a first naive hypothesis that an EAS would trigger all these streamers in a single causal relation, it would be expected that the pattern only moves downwards (possibly with some inclination) and roughly with the speed of light, contradicting the observations. And even though the initiation of a streamer after the EAS is passed is not instantaneous (electrons need to drift towards the hydrometeor), it than still only adds random noise to a trend going with c. The velocity measured in the system of streamers is actually more streamer like and independent of altitude (as streamer scaling-laws demand) [202]. So the full fast positive breakdown going up or down is not driven by an EAS, it is already an evolving structure of streamers [202]. The initiation of the system of streamers however does require an extreme EAS core density over a certain extent, as we explain in the next paragraph.

As is discussed in Rison et al. [202], to explain the strong spherics multiple streamers side-by-side are necessary and the observations show a relatively large lateral extent of around 50 to 100 meters. Examining closely NBE3 from [202] (see their supplementary Fig. 9), we see that the first two data points recorded span horizontally 129 m (and vertically 58 m) but the time between these events is only 3.8×10^{-7} s. They thus can not be causally related, they are outside of each other's light cone. We see this as an evidence that they must be triggered together in a simultaneous event, a fast event with a timescale of less than a microsecond. Even considering both end points of the error bars in the position (see supplementary Fig. 18 of NBE3 from [202]), the speed would be much faster than is typical for streamers and the propagation direction would then be perpendicular to the velocity which is set in after initiation. In other words, when the downwards propagation is probably streamer-like assumed to be driven by the electric field which is then vertical, the horizontal propagation at that altitude would be perpendicular to the field, which is not expected. We even think that all first four points of the NBE3 observation are triggered simultaneously by an EAS event, but this can not be proven from the data presented in [202]. For the other two NBEs presented in Rison et al. [202], the evidence is less convincing, as data points in the presented plots overlap, but for NBE1 it is still convincing. Studying closely NBE1 (see their supplementary Fig. 7), we see that the first two events are spanned vertically 87 m in only 1.8×10^{-7} s (it was not clear how much the azimuth angle was). Thus, they are also outside each other's light cone or at least much faster than streamers and must therefore be triggered by the same event, presumably from one extreme EAS event.

To trigger a system of streamers by an EAS event one should increase the event box, as was introduced in [26], such that multiple streamers are initiated. The top area A of the event box times the waiting time T defines the probability of a rare event (i.e. the availability of an x-km^{2}s-EAS, where x is the product of A and T), see also Fig. 11.2. From all NBE results presented in [202] we approximate that the EAS core radius must be at least of the order of tens of

meters. From Fig. 11.11 we find that the maximum expected density of such large core radii is only of the order of 1 cm^{-3}. The volume above breakdown near the hydrometeor to start the streamer, is less than a cubic millimeter. If the volume is considered that free electrons can drift towards the hydrometeor before attachment approximately 10 mm^3 is available [26]. Free electron densities as low as 1 cm^{-3} mean that only a small subset of hydrometeors can be triggered, demanding a stochastic approach for streamer initiation.

11.6 Conclusions and outlook

We have calculated the availability of an extreme free electron seed, which can be expected in a thunderclouds due to a rare cosmic ray induced extensive air shower event. The results are presented in Fig. 11.10 and 11.11, which give the maximal free electron density as function of altitude and rareness of the event. This initial free electron density can be used when studying streamer initiation in the context of starting lightning. In the leading theory that lightning starts by a volumetric system of streamers [202], it is even more crucial that enough free electrons are created simultaneously than was reported in [26] before.

The measurements of Rison et al. [202] reveal that the first few data points must be triggered simultaneously, as they lay outside each other's light cone or the information between the first few data points travels much faster than streamers do. We think that this is evidence that they are probably initiated simultaneously by one EAS event, as we discuss in Sec. 11.5. The measurements of Rison et al. [202] reveal further that the lateral width of the volumetric system of streamers is of order of tens of meters, which from our results yield initiation from electron densities as low as 1 cm^{-3}, demanding a stochastic approach for streamer initiation.

We are currently developing models to incorporate the stochastic streamer initiation, starting from only one electron, see chapter 6. In addition we are extending our models to simulate the initiation of multiple streamers from multiple hydrometeors side-by-side, where streamers could interact and connect to explain the further growth in the start of lightning. For this, cylindrically symmetric modeling, as was done in [26, 29, 30, 34, 205] and reference therein, does not suffice. For simulating multiple interacting streamers the only direction is full 3D, which can be done with the Afivo framework created by Teunissen and Ebert [27] Afivo is open source and publicly available at our Gitlab page (https://gitlab.com/MD-CWI-NL/afivo). Or by model reduction to discharge tree models [31].

Part IV

Neutrons from thunderclouds and lightning

Chapter 12

Modeling ground detected neutrons from thunderstorms

Neutron emissions with different durations have been observed on ground during thunderstorms. These neutrons can be produced by fast Terrestrial Gamma-ray Flashes correlated with lightning, or by Gamma-ray Glows lasting several minutes. In both cases the neutrons are produced through a photonuclear reaction with nuclei of air molecules, if gamma-rays in the energy range of 10 to 30 MeV are present, and are thus a diagnostic for studying high energy atmospheric physics. Here we present simulations of downward gamma-ray beams with energies between 10 and 30 MeV, separated into 4 energy intervals with 5 MeV of width, at varying starting altitudes. The individual signature at sea level of each energy interval allows the extrapolation of our results to an arbitrary source spectrum and altitude. Our results indicate that neutrons are created along all photons motion and the majority do not reach sea level farther than 0.5 km from the source axis. This last feature indicates the primary photons do not collide much with air molecules before creating a neutron, hence, most of neutrons dispersion is result of their own motion in the air. A second photon pulse at sea level, which recently have been predicted and measured.This second photon pulse is replaced in time relative to the primary one due to the slower neutron time scale, hence, it constitute an afterglow effect.

12.1   Introduction

12.1.1   Neutrons as footprint of thunderstorms

The connection between neutron research and thunderclouds originates from radiocarbon studies of tree rings by Libby and Lukens [225], see review by Rutjes and Ebert [226]. Later, [11] proved a relative increase of neutron emissions compared to the average cosmic background, correlated with thunderstorms. The first attempts to explain neutron production by lightning focused on fusion, and dominated literature for some decades, but was proven counter-wise by [227]. The fusion mechanism was impossible due to a lack of acceleration of the deuterium ions to reach the reaction threshold, and also because of the small fraction of deuterium ions in air. Even for electric fields 20-30 times the conventional breakdown, the neutron yield from fusion remains small in comparison to the original projections of Libby and Lukens [225]. Babich and Roussel-Dupré [227] and in more detail by Babich et al. [228], proved that the correct explanation of neutrons in thunderclouds, is the photonuclear reaction by high energetic photons between 10 MeV to 30 MeV. Photonuclear reaction is due the the so-called Giant Dipole Resonance (GDR) mechanism [229, 230]. Neutrons are thus diagnostic of high energy atmospheric phenomena, created in thunderclouds.

12.1.2   Neutron production through TGFs and gamma-ray glows

The photonuclear reaction mechanism requires a photon source generating gamma-rays with energies above 10 MeV which is the approximate binding energy of neutrons in atmospheric nuclei. More precisely, the binding energies are 10.55 MeV for nitrogen, 15.66 MeV for oxygen and 9.87 MeV for argon [231]. There are two known types of high energy atmospheric mechanisms that can produce gamma-rays above these thresholds, namely Terrestrial Gamma-ray Flashes (TGFs) and gamma-ray glows.

TGFs which are fast microsecond-long bursts of photons that were first observed from space [53, 105]; they can be accompanied by bursts of electron positron pairs [54, 55] and they correlate with leader propagation. Glows on the other hand, originate from relativistic runaway electron avalanches (RREA) in the field of the thundercloud and last much longer, for minutes or even hours; they have been observed on ground, from balloons and aircraft [56, 58–61, 118**?**]. The different properties of flashes and glows have been related to different physical mechanisms as explained in [62, 115, 120]. As lightning leaders produce TGFs, lightning is observed to terminate gamma-ray glows [17, 18, 56].

12.1.3   Neutron emissions from thunderstorms measured with different detectors

Three main types of neutron detectors were used: a thermal helium-based detector, a boron counter detector and the Neutron Monitors NM-64. The thermal detectors are most efficient for particles with energy of the order of 1 eV, the boron based detector has a peak of efficiency at 2.45 MeV, while the Neutron Monitors NM-64 are focused on higher energies and have a small efficiency below the MeV energy range (reaching 10 % for 3 MeV and 2 % for 0.5 MeV). All three detectors work on very distinct and narrow energy windows, which makes it difficult to analyze the whole spectrum (see [19, 232–238]).

Neutron emissions with different characteristics have been observed during thunderstorms. There are long-term neutron emissions, e.g., [236], lasting for up to 30 minutes at Yangbajing at 4.3 km altitude above sea level; they were not correlated with a lightning discharge, and they were detected with boron based detectors and NM-64.

Chilingarian et al. [237] also measured long neutron emissions during thunderstorms that were not correlated with lightning as they lasted much longer; they were observed at 3.2 km above sea level on Mt. Aragats with boron-based and NM-64 detectors. The authors have related these neutron emissions to gamma-ray glows (which they call Terrestrial Gamma-ray Enhancements) which are long lasting gamma-ray emissions during thunderstorms on the same time scale as the neutron emissions. Gurevich et al. [235] have used three sets of thermal neutron detectors with a summed area of approximately 127.23 cm^2 for observations at the Tien-Shan Mountain (3340 m above sea level). They also report long lasting neutron signals; in their measurement, they found a high value of neutron flux (3-5$\times10^{-2}$ cm^{-2}s^{-1}).

Meanwhile Martin and Alves [233] in Brazil, using a helium-based detector at 600 m above sea level and Toropov et al. [19] in Russia, using the Yakutsk cosmic ray spectrograph at 105 m above sea level, measured neutron pulses lasting from seconds to a couple of minutes that were correlated with lightning discharges.

Recently, [121] has measured a burst of neutrons following a lightning strike in a wind turbine at Japan. Their neutron detection was followed by a glow of gamma-ray photons which were presumably created by the neutrons themselves. Their detection system consists in three scintillation detectors and two additional electronics for noise control. The detectors' scintillation material are SmPl, LgPl and NaI(Tl). These neutrons were detected through a capture reaction between the emitted neutrons and the LgPl that releases photons with a characteristic energy of 2.223 MeV. The photons were measured on a millisecond timescale and with characteristic energy signature of neutron capture, consistent with the TGF afterglow predicted by [120].

The neutron detection depends on the type of gamma-ray source: from bursts of 690 neutrons per minute associated with short emissions from lightning [233]

to fluxes of $(3-5) \times 10^{-2}$ neutrons per $cm^{-2}s^{-1}$ from long term emissions associated with the thunderstorm time scale [235].

Associated with TGFs, initial estimations of [11] on the neutron production per lightning event reach from $10^8 - 10^9$, to $10^{12} - 10^{13}$ in more recent observations by [233]. Associated gamma-ray glows [236] estimate a time integrated flux of neutrons from the thundercloud arriving at the observatory of 1.4×10^4 m^{-2}. [236] used a solar-neutron telescope of 9 m^2 total area, which results in 1.2×10^5 neutrons at the observatory.

12.1.4   Neutrons propagation to ground

The created neutrons can interact with air through elastic scattering, inelastic scattering and capture. The dominant process is elastic scattering for most of the energy range (0-20 MeV) while the capture process is the most rare among the three [126]. But scattering only cools the neutrons down, while capture is the only loss process. Capture is the most efficient at low neutron energies. Therefore the gamma-rays formed on neutron capture in a nucleus have energies in the order of a few MeV due to the energy released on the capture. This is the mechanism of the TGF afterglow described by Rutjes et al. [120].

Considering the gamma-ray source to be the step of a negative lightning leader, each step would be a possible source at different altitudes. The stepped leader, considering high-speed optical observations, travels at an average speed of $2.7 - 6.2 \times 10^5$ m/s and the average interval between steps is on the order of 10 μs [239, 240]. Using these time scales we can estimate that the leader takes about 10 ms to travel from 3 km altitude to the ground and 1 ms from 300 m, performing several steps during its motion. But, since the steps are separated from each other on the time scale of 10 μs and the majority of the neutron pulse is in the order of seconds, one lightning leader does not explain well the experimental measurements with tens of minutes time scale, such as [236], but suits well in the measurements correlated with lightning discharge such as [19].

Therefore, the gamma-rays formed on neutron capture in a nucleus have energies in the order of a few MeV due to the energy released on the capture. This is the mechanism of the TGF afterglow, which was first theoretically predicted by [120] and first measured by [121] in a downwards directed TGF.

12.1.5   Earlier simulations

Neutron emissions were earlier simulated by [73, 130, 228]. Babich et al. [228] have reached a conclusion that there are 4.3×10^{-3} neutrons produced per photon and the neutron energy is in the range of 0-20 MeV both in the production and at the sea level altitudes. The neutrons diffuse to distances of the order of km from the source axis, when reaching ground.

[130] assumed an RREA spectrum to simulate an initial photon source and report neutron production and detection at sea level. Their results are in general agreement with [228] energy range and estimate production. Both works provided a time scale of a couple seconds to their neutron pulse and have concluded that the neutrons are produced along the photons motion through the atmosphere instead of in a point-like source, as it was used in the work of [233].

[73] simulated a negative lightning leader, following the ideas of previous simulation works such as [66, 69, 70, 241], in the moment of the step interacting with electrons as source of the bremsstrahlung photons that would produce neutrons. [73] have recorded neutrons, at their initial altitudes, with energy up to ~ 20 MeV.

[121] have performed TGF Monte Carlo simulations in the GEANT4 program in order to confirm the neutron signature in the LgPl detector. A comparison between the measured 1-10 MeV photon spectrum and the GEANT4 simulations showed general agreement in the characteristic energy release on the capture reaction, hence, confirming their observations.

[121] have used a initial photon spectrum produced by RREA in their simulations, following distributions of [62]. They've used the initial photon spectrum to simulate the creation of photoneutrons and, then, recorded the photoneutrons signature generated at the ground. Furthermore, this ground signature has been used as an input of a third simulation where they have filtered the particles that would reach the detector. They have design the simulation to infer the number of TGF gamma-rays from the observed count rate. [121] have also notice a longer signal of low energy photoneutrons in their simulations due to neutron thermalizing in the ground, not only in the air.

The previous works results are strictly attached to their initial source spectra. Therefore, they describe the production and detection of a specific source. Since we are using a small, flat photon source spectrum our results may be used to estimate the detection of an arbitrary source detection as it is exemplified later in this work.

From the energetic and temporal characteristics of the neutron signal, we have found that the neutrons source geometry must be spread through all photons motion to the detector because there are neutrons arriving sea level at the same as the photons, indicating that the neutrons were produced at low altitudes so they could have this time coincidence with the pulse of primary photons and also there are neutrons arriving at sea level after the pulse of photons from the source is gone, meaning that these neutrons were created at high altitudes or collide enough with air particles so their motion time to the ground is increased.

Since the photons can not have many interactions without leaving the energy range of photonuclear reaction, we estimate that most neutrons are created in a straight line during the photons motion to sea level; and the neutrons lateral spreading must be due to their collisions with air molecules. These neutron collisions with the air occur in a larger time scale than the time scale of the

source particles and generate a second, shifted in time, photon pulse constituting a TGF afterglow effect [120].

12.1.6   Questions and organization of this chapter

In this work, we split the relevant gamma-ray energies of 10 to 30 MeV into four intervals of 5 MeV width, so we investigate the intervals [10 MeV, 15 MeV], [15 MeV, 20 MeV], [20 MeV, 25 MeV], and [25 MeV, 30 MeV] separately. We assume that each energy interval is equally filled with gamma-rays, and that all rays are directed in the same downward direction. This approach allows us to construct neutron distributions for arbitrary gamma-ray sources, without further hypotheses on the gamma-ray source. This enables us to investigate arbitrary gamma-rays sources suggested in the literature that could produce ground detectable neutrons.

Since the neutrons can be created all along the photon trajectory and as the photon source can be located different altitudes, we have studied how different source altitudes influence the neutron distribution and spectrum at sea level.

We have analyzed the position, time and energy of the neutron and photon signal at the ground. Our results may be extrapolated to an arbitrary source altitude via equivalence of integrated density between the considered source position and the detection altitude; one may extrapolate our results for a different source energy spectrum by weighting individually the signature of each energy interval. A detailed description of the neutrons energy range at sea level is presented so further observations may focus in a efficient energy window. The particles temporal signature revealed a second photon pulse that is not directly related to the primary gamma-ray, instead, it may be result of collisions between secondary particles and the air molecules.

The chapter is organized as follows: The section 2 is dedicated to explain the simulation environment and initial conditions. The simulation results are presented in section 3; first, we have reported how the neutrons are distributed at sea level, secondly, we have presented in detail the detected particles energy spectra and finally a description of the neutrons and photons temporal pulse. Our analysis on the simulation results are presented on section 4. Further, we compare our results with previous works on section 5 where we exemplify how to extrapolate the results to a different source spectrum. Finally, we summarize our conclusions on section 6.

12.2   Method and setup of the simulations

12.2.1   FLUKA

We simulate the photons and neutrons in air by Monte Carlo simulation with the software package FLUKA [22], which has a complete physics list for air in

our energy regime [115]. A comparison between FLUKA and other Monte Carlo simulation codes (GEANT4, EGS5, GRRR, MC-PEPTITA) was performed by [115] in which a standard input was used to test the effects of the different approximations and methods of each program on the output. The authors have noticed a general agreement between the programs, but differences due to the straggling effect and to the implementation of energy friction were observable. The straggling effect is appropriately implemented in FLUKA.

FLUKA is a general purpose tool to simulate the particle interaction with matter; it covers a broad range of energies from a few keV to hundreds of TeV. The program does not provide explicit information on how it transports the particles, however it provides several options of how to detect the particles and define the media through which they travel. We implemented a user-modified routine to detect the particles.

The photon interactions implemented in FLUKA are [22]: pair production, Compton scattering, photoelectric effect, Rayleigh scattering, photon polarization, photonuclear reactions (GDR, Quasideuteron interactions, Vector Meson Dominance Model), and generation and transport of Cherenkov, scintillation and transition radiation. Since there is no energy gain in our simulations the energy of the particles will always be below the initial energy, 10-30 MeV in this case.

Because about 10 MeV are required to liberate neutrons in a photonuclear reaction, the neutrons in our simulation have energy below 20 MeV. In this energy range, FLUKA transports the neutrons with a multi-group algorithm using a cross-section library of 260 groups [22]. The angular probabilities are treated with a Legendre polynomial expansion of the actual scattering distribution. Further details on FLUKA physics can be found on the manual available at the official website http://www.fluka.org/fluka.php.

We use a low energy cutoff of 10 keV for photons and of 10^{-12} eV for neutrons which are the recommended minimum thresholds in FLUKA.

12.2.2   Setup of simulations

For this work, a cylindrical domain from sea level up to 19 km altitude and with a radius of 12 km was adopted. The atmospheric density profile was simulated by 76 layers of constant density with 250 m of height, consisting of 78.085 % N_2, 20.950 % O_2 and 0.965 % Ar. The atmosphere density in these layers was linearly interpolated from the density of the 1976 U.S. standard atmosphere, defined by the United States Committee on Extension to the Standard Atmosphere [242].

All simulations start with a downward directed beam of ten million photons distributed within a given energy interval and starting at a given altitude, and we detect photons and neutrons at sea level. We do not assume a specific photon spectrum, but allow for the construction of an arbitrary initial spectrum. For this purpose, we decompose the relevant energy regime into four ranges with a

uniform spectrum. These ranges are 10-15 MeV, 15-20 MeV, 20-25 MeV, and 25-30 MeV. The source altitudes range from 300 m to 6 km. At higher source altitudes, too few particles reach sea level. So in total we performed 36 different simulations, namely for 9 different source altitudes and for 4 different energy ranges.

12.3 Simulation Results: Photon and neutron distributions in space, time and energy at sea level

We have analyzed the photon and neutron distributions in space, time and energy at sea level as a function of the energy range of the source photons for different altitudes. We found that these distributions do not depend on the energy range of the source photons, besides a few minor differences which are discussed. Only the neutron production rate depends significantly on the photon energy; it will be discussed in the next section. In the present section, we therefore sum over the four energy intervals; so effectively we present results for $4 \cdot 10^7$ source photons, uniformly distributed over the energy interval of 10 to 30 MeV.

12.3.1 Neutron number and distribution at sea level as a function of source altitude

Table 12.1 shows the total number of neutrons at sea level and their spatial distribution on ground as a function of photon source altitude. As expected, the neutron number strongly decreases with increasing source altitude. On the other hand, the spatial spread on ground is fairly constant. Measured as the standard deviation from the beam axis, the spatial spread of the neutrons ranges form 0.29 to 0.45 km. This behavior can be understood based on the spatial distribution of neutron generation discussed later. In general, there is an approximately exponential decay of the neutron density on ground for all source altitudes.

12.3.2 Energy spectra of photons and neutrons at sea level

The energy spectra of photons and neutrons at sea level are presented in Figures **??** and 12.2 as a function of source altitude for the same uniform distribution of source photon energies between 10 and 30 MeV. The only dependence on the energy spectrum of the source photons lies in the fact that, e.g., photons in the energy range from 10 to 15 MeV can not produce 10 MeV neutrons and hence the produced neutron spectrum has this upper energy limit, but this initial neutron energy difference rapidly disappears due to cooling down of the neutrons.

Let us start with the photon energy spectra recorded at sea level, shown in Figure **??**. The source photon spectrum is an equal distribution between 10 and 30 MeV as indicated in the figure, it would correspond to a source altitude of

Table 12.1: Neutron number and distribution at sea level as a function of source altitude for a uniform distribution of source photon energies between 10 and 30 MeV.

Source altitude [km]	Standard deviation [km] of distance to the beam axis	Neutron number on ground for 4×10^7 primaries
0.3	0.291	51777
1.0	0.396	24913
1.5	0.417	10666
2.0	0.427	4666
2.5	0.445	2018
3.0	0.408	917
4.0	0.413	227
5.0	0.373	81
6.0	0.347	13

0 km. This initial spectrum is still clearly visible at source altitudes of 0.3 and 1 km, though attenuated, with a clear discontinuity to the secondary photon spectrum at lower energies. This discontinuity diminishes for higher source altitudes due to multiple interactions where photons loose energy or disappear, and the discontinuity disappears above 4 km. There is a second discontinuity visible in the spectra at 511 keV which is the positron annihilation line. When photon energies above 1.02 MeV are available, electron positron pairs can be produced, and the annihilation of the positrons creates the 511 keV line.

Figure 12.2 shows the energy spectra of the neutrons at sea level that are generated by the same photon sources. There are two regimes visible: a spectrum with an index of $\sim$ -0.8 between 0 and 10 keV and an exponential cutoff in MeV energy regime. The latter thus depend on the initial photon source interval as that determines the maximum neutron energy. However, the shapes of neutron spectra hardly depend on the source altitude, this is a consequence of the elastic scattering dominance in the set of neutron interactions. The neutron energy spectra below 10 keV becomes noisy (i.e., enters in the region of 1 to 10 counts per bin) with increasing source altitude, indicating that low energy neutrons could not reach the ground and thermalized at the air via inelastic scattering or capture. The whole neutron spectra are noisy for source altitudes of 4 km, 5 km and 6 km, meaning that our simulation did not had enough resolution for these altitudes, thus, the sources at these high altitudes need a larger number of primaries ($\gg 10^7$) to create ground detectable neutrons.

Table 12.2 summarizes the neutron number for different energy ranges and all source altitudes. The absence of detected neutrons with low energy (below 300 eV) for higher source altitudes indicate that neutrons created far from the ground lost all their energy (or got captured) during their motion in the atmosphere, therefore, the high energy detected neutrons must have been created
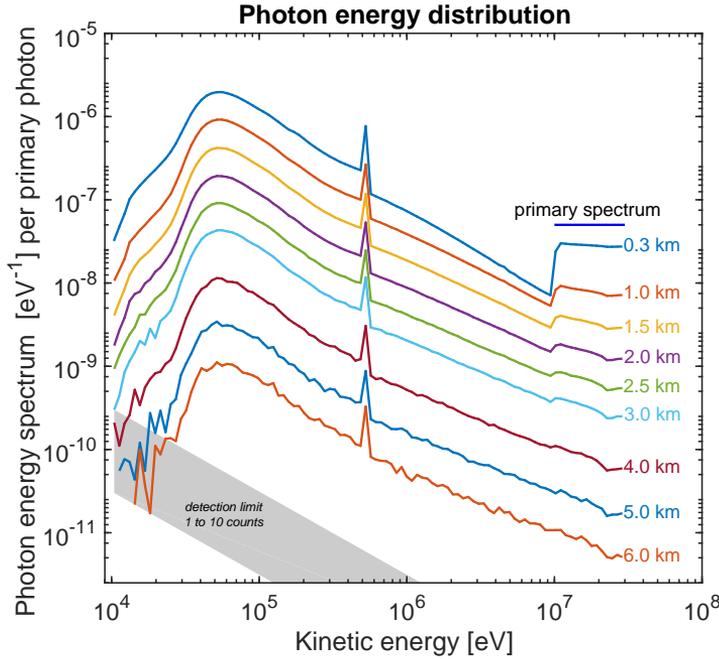
Figure 12.1: Photon energy spectra at the ground for different source altitudes. The gray shade indicates low statistics.

near the ground for all source altitudes.

12.3.3 Arrival time

We investigated the arrival time photons and neutrons, i.e. the number of particles arriving per primary per time (ns in the case of photons and μs for the neutrons) as function of time, that reached ground for different photon source altitudes. As in Section 12.3.2, it is clear to see that the shape of the particles temporal signal does not depend on the initial energy interval. And, again, the presented signals are a simple sum of the temporal signature produced by each one of the four energy intervals.

The photon temporal signal at the sea level, shown in Figure 12.3 reveals two separated pulses for every source altitude. The first one with 0.1 ms duration and the second with $\sim$ 500 ms duration, the total photon signal has a 0.5 s duration in the presented simulations. For every source altitude, the first photon reaches the ground within the time of a photon traveling straight down to the ground without any collision. The first pulse is identified as the source photons and secondary photons created due source photons interactions with the air. The second pulse is a result of late particles interaction with the air. More results on this effect are published in other work [120].

The number of neutrons arriving per unit of time is presented in Figure 12.4.
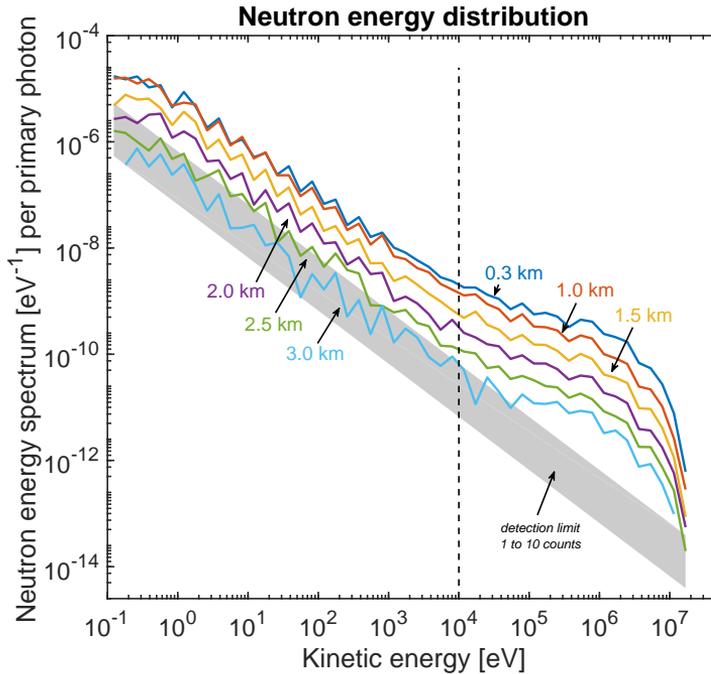
Figure 12.2: Neutron energy spectra at the ground for different source altitudes. The gray shade indicates low statistics. The vertical line marks the energy range at 10 keV.

It also has a total duration of 0.5 s for the given number of primaries and it is separated in two parts, before and after 0.1 ms. The first neutron reaches the ground close to the time of the first photon.

12.4   Analysis

Considering the first photon to reach sea level without collision, from a source altitude at 300 m, it will arrive straight at the detectors within 1 μs (considering the light speed as 3×10^8 m/s). On the other hand, if the most energetic neutron created (20 MeV, which is a non-relativistic particle) moved straight to the ground without an interaction, it would travel within the same time interval only 60 m. Since we have neutrons arriving at ground on ≈ 1 μs, as shown in Figure 12.4, these neutrons must be created along the photons motion near the ground and without many interactions with air particles. Therefore, the less energetic detected neutrons must have been created with higher energy and either were created near source altitudes and lose energy colliding through their motion or they were created near sea level but collide with air molecules enough to lose most of their energy during a motion parallel to the detector.

Thus, the ground distribution shown in Table 12.1 is a result from the neu-

Table 12.2: Neutron number at sea level for a downward directed beam of 4×10^7 photons at different source altitudes and in different neutron energy intervals (in eV).

Source altitude	(0, 0.025]	(0.025, 1]	(1, 300]	(0.3,10^6]	(1, 2×10^7]
0.3 km	17	489	1976	20966	28329
1.0 km	13	466	1709	11708	11017
1.5 km	7	223	796	5019	4621
2.0 km	2	109	374	2216	1965
2.5 km	0	42	166	942	868
3.0 km	0	21	55	441	400
4.0 km	0	3	12	107	105
5.0 km	0	0	4	43	34
6.0 km	0	1	0	7	5

trons diffusion in a random walk manner through the air from each origin point. Therefore, the neutron source is interpreted as a cone of photons, illustrated in Figure 12.5, in which the neutrons are created widely spread and start their random motion contributing to the ground ground distribution.

The photon energy range was separated in 4 intervals within the GDR energy range: 10-15, 15-20, 20-25, 25-30 MeV. Figure 12.6 shows the contribution of each energy range for the neutron detection for different source altitudes. The neutron spectra are created mostly by photons with energy between 20-30 MeV, as the photonuclear cross section of nitrogen peaks at 23 MeV [243], while the 10-15 MeV photon energy range has only significant signals in the MC size used in this study for low source altitudes, this may be seen in Figure 12.6 in which is presented the total neutron count per primary as a function of photon source altitude.

The detection for all intervals follow approximately the same behavior as a function of source altitude, as shown in Figure 12.6. Therefore, each energy interval labeled i contribute with N_{pi} neutrons per primary to the total signal per primary N_t, according to the Equation 12.1; where h is the source altitude, β_i is the constant of decay and I_i is the extrapolated value of the curve for h = 0. The factors a_i in Equation 12.2 are weight factors to account for how much each energy interval contribute to the signal. All a_i are considered as 1 in our simulations, this means that our results are the simple sum of each energy interval products. In order to approximate the detection of an arbitrary initial photon spectrum, one may adjust the factors a_i.

$$N_{pi}(h) = I_i e^{-\beta_i h} \tag{12.1}$$

$$N_t(h) = \Sigma_i a_i N_{pi}(h) \tag{12.2}$$

The β_i and I_i values are summarized in Table 12.3. Each β_i was calculated
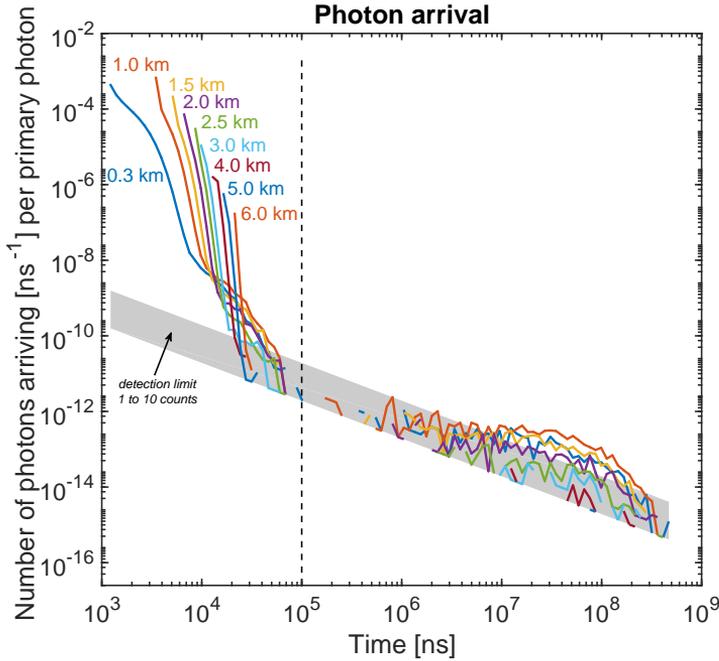
Figure 12.3: Photon temporal signal at the ground for all simulated source altitudes, as indicated by the legend. The vertical line marks the end of the primary pulse at 0.1 ms. The gray shade indicates low statistics.

interpolating the non zero extreme values for each curve, i.e., using the values of detected neutrons per primary for a source altitude at 300 m and the highest source altitude with neutron counts; the highest source altitude vary from one curve to another due to increasing air layer thickness that unable the neutrons to reach sea level.

The photon spectrum is generally fitted by Equation 12.3 [62], resembling a steady state RREA photon spectrum. The amount of photons with energy above the photonuclear reaction threshold increases with the exponential cutoff (E_{th}). Figure 12.7 shows the neutron production as function of this exponential cutoff following the strategy described previously. Here the weight factors a_i are

Table 12.3:   Values for analytic interpolated curve for each energy range detection decay with source altitude.

Energy Range [MeV]	β [km^{-1}]	I per primary
10-15	1.77	1.022×10^{-4}
15-20	1.34	8.323×10^{-4}
20-25	1.50	4.088×10^{-3}
25-30	1.36	2.944×10^{-3}
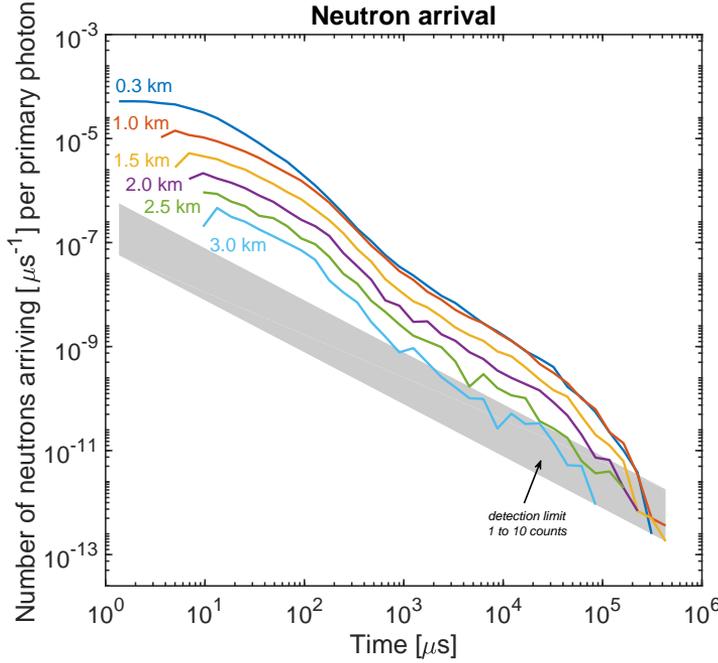
Figure 12.4: Neutron temporal signal at the ground for different source altitudes indicated by the legend. The gray shade indicates low statistics.

normalized by the total amount of photons between 10 and 30 MeV, i.e., the Equation 12.3 integrated over 10 and 30.

$$f(E) = \frac{1}{E} e^{\frac{-E}{E_{th}}} \tag{12.3}$$

All neutrons that do not reach sea level will eventually be captured again, accompanied with a release of a high energy photon, which will produce the delayed photon signal seen in the second part of the temporal photon signal. Thus, there is a neutron induced photon signature at the ground that can be used to more easily detect a neutron emission during measurements on thunderstorms due to the delay between the first and the second photon pulse. Due to a temporal correlation between the neutron signal and the second photon pulse, which can be seen with the spectra present in Figure 12.8 that are the same spectra present in Section 12.3.2 but separated in time intervals of before and after 0.1 ms, we assume that the neutrons are the source of those photons.

The majority of the photons are within the first photon pulse (before 0.1 ms) , while the secondary pulse is composed by a, relative to the spectra, small photon number with energy between 10 keV and some MeV, this feature is shown in Figure 12.8 a and b. The rate of photons present in the second pulse for the photons present in the first pulse is, for the same source altitude, $\sim 10^{-5}$ which is in agreement with the hypothesis of the neutrons being their source; since,

considering the [228] rate of neutrons created by photons (4.3×10^{-3}), we would have $\sim 10^4$ neutrons created by the our 10^7 primary photons. This last number of neutrons is enough to generate our detection and the secondary photon pulse, see [120] which identifies this secondary photon pulse as a TGF afterglow.

Meanwhile, the neutron spectra before 0.1 ms are composed by the more energetic neutrons that are were created near the ground and interact little with the air. The neutron spectra after 0.1 ms are composed by neutrons with energy lower than 1 MeV, as can be seen in the right column of Figure 12.8, since they were created at higher altitudes and lost energy in collisions with the air and created the second photon pulse.

The time scale separation on the neutron spectra also distinguish them on two different populations. The signature present on the first 0.1 ms of the discharge will be dominated by MeV energies and, hence, the detectors focused on this energy regime (such as the NM-64) will record this part of the spectra. Meanwhile, the detectors such as the Helium based thermal neutron counter is more likely to register counts after 0.1 ms. Therefore, these two different time scales must be taken into consideration on time correlated detections in order to see correctly what phenomena can cause the measurements.

The simulated results can be generalized to any detection altitude through equivalence on the integrated density between the source and the sea level. Due to the ambient described on Section 12.2, the integrated density can be calculated following the Equation 12.4.

$$I_d = \Sigma_i^{n(z)} N_i dz \qquad (12.4)$$

Where I_d is the integrated density; N_i is the density value linearly interpolated from the [242]; dz is the constant air layer thickness of 250 m; z is altitude and n is the final air layer index to be summed. This equivalence is justify by the same number of air molecules available to particles interactions for equal values of integrated density. In this manner, measurements with the detector at high ground altitude such as 3 km can estimate the source altitude based on their detections.

12.5 Conclusions and outlook

We simulated neutron footprints from gamma-ray point sources, of between 10 MeV and 30 MeV initial energy, differentiated between four initial energy intervals of 5 MeV width. Clearly we see that, even though the photons start as point source, neutrons do not originate from a point source, as they are produced along the kilometers long paths of the high energy (primary) photons. The later limits simple inverse approximations of the amount of photons needed to explain observed fluxes, as was done in reviewed literature. The current provided fluxes can be used to make better estimations

We observed that spectra shapes from the four different energy intervals do not change significantly, besides the number of neutrons produced. In other words, as the number of neutrons produced is so sensitive to the gamma-ray energy, the neutron yield depends on the gamma-ray initial spectrum. We calculated the number of neutrons as function of the characteristic (mean) energy cutoff of a steady state RREA for several altitudes, see figure 12.7. This dependence can be used to calculate the expected number of photons, e.g. in a gamma-ray glow, when the characteristic (mean) energy cutoff can be approximated based for example on its sensitivity of electric field. Or in the other direction, knowing the total neutron to photon ratio results in the characteristic (mean) energy cutoff (assuming RREA) and by that a diagnostic for the ambient electric field.
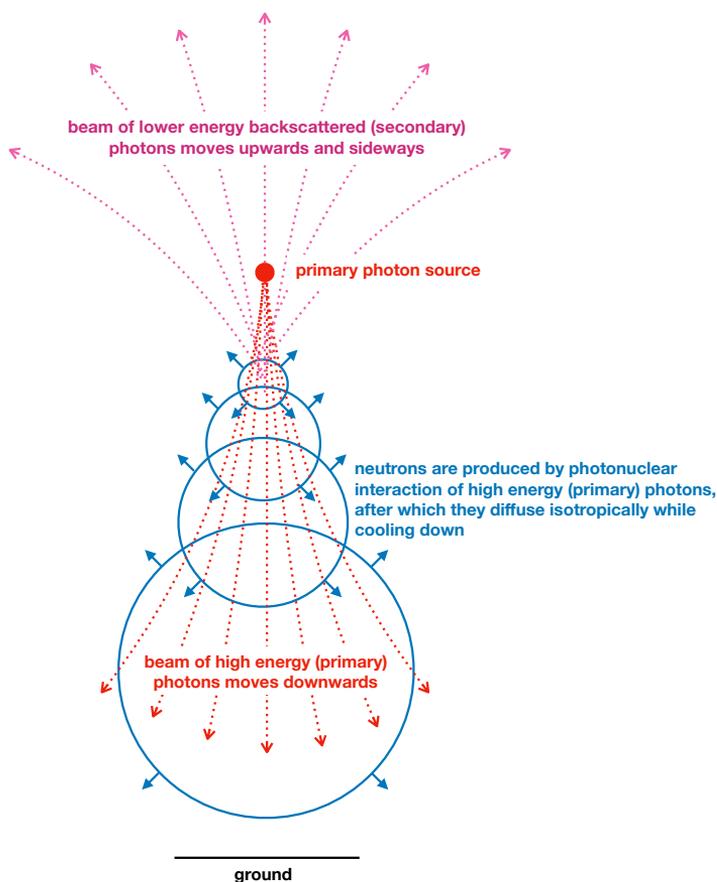
Figure 12.5: Illustration of neutron source geometry due to a point-like photon emission, initially directed downwards. Beam of high energy (primary) photons (in red) diffuse slightly, but remain beamed downwards. In the beam of high energy (primary) photons, individual photons move in straight lines and along their path have the probability to create neutrons by photonuclear interaction. Created neutrons diffuse isotropically while they cool down, illustrated in blue. The beam of lower energy (secondary) photons can move upwards or sideways, but do not have enough energy (anymore) to create neutrons.
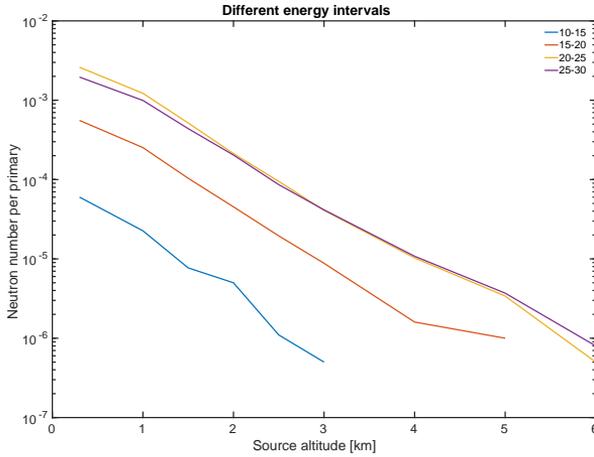
Figure 12.6: Neutron number per primary at the ground as function of source altitude for the four different energy range, indicated by the legends on the figure (10-15 MeV; 15-20 MeV; 20-25 MeV; 25-30 MeV).
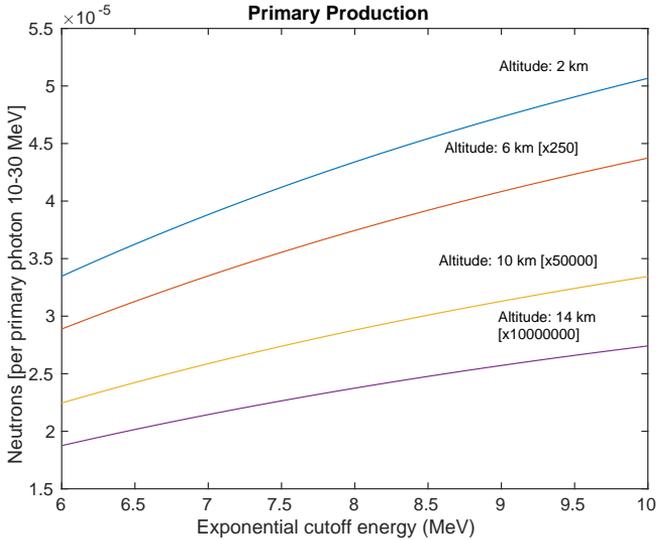
Figure 12.7: Neutron number per primary at the ground as function of initial photon spectra exponential cutoff with different source altitude.
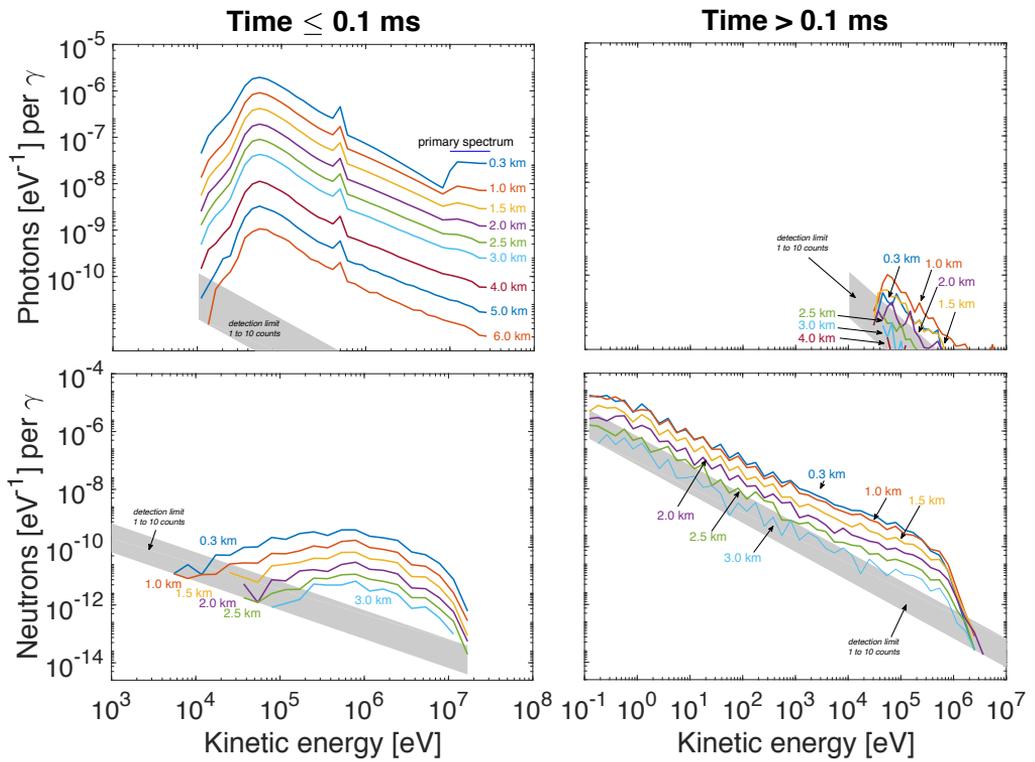
Figure 12.8: Energy distributions of photons (top row) and neutrons (bottom raw), equivalent as figures 12.1 and 12.2, but now differentiated between arriving before (left column) and after 0.1 ms (right column).

Chapter 13

TGF afterglows: a new radiation mechanism from thunderstorms

Thunderstorms are known to create terrestrial gamma-ray flashes (TGFs) which are microsecond-long bursts created by runaway of thermal electrons from propagating lightning leaders, as well as gamma-ray glows that possibly are created by relativistic runaway electron avalanches (RREA) that can last for minutes or more and are sometimes terminated by a discharge. In this work we predict a new intermediate thunderstorm radiation mechanism, which we call TGF afterglow, as it is caused by the capture of photonuclear neutrons produced by a TGF. TGF afterglows are milliseconds to seconds long; this duration is caused by the thermalization time of the intermediate neutrons. TGF afterglows indicate that the primary TGF has produced photons in the energy range of 10 - 30 MeV; they are nondirectional in contrast to the primary TGF. Gurevich *et al* might have reported TGF afterglows in 2011.

13.1 Introduction

Thunderstorms emit energetic radiation of different types. Best known are Terrestrial Gamma-ray Flashes (TGFs) which are microsecond-long bursts of photons that were first observed from space [53, 244]; they can be accompanied by bursts of electron positron pairs [54, 55]. On the other hand, gamma-ray glows last much longer, for minutes or even hours; they have been observed on ground, from balloons and aircraft [56–61, 118]. Chilingarian *et al.* call them thunderstorm ground enhancements, which refers to the fact that the detector is located on ground.

The different properties of flashes and glows have been related to different physical mechanisms. TGFs originate from cold runaway [65] where thermal electrons accelerate to tens of MeV in the strong electric fields of a propagating leader discharge. TGFs appear in bursts that last for microseconds to milliseconds with a temporal distribution sketched in Fig. 13.1; they correlate with leader propagation. Researchers have investigated how the streamer phase [66–68, 245] or the leader phase [69–73] could accelerate electrons to energies that could explain the gamma-rays as an effect of bremsstrahlung. Experimentally cold runaway has been found in pulsed discharges [74–77] and during the formation of meter long laboratory sparks [78–84, 246]. Glows on the other hand would originate from relativistic runaway electron avalanches (RREA) [64, 85], with feedback of photons and positrons creating new avalanches [86–88]; they evolve on the timescale of seconds to minutes and even hours, as sketched in Fig. 13.1 as well. Whereas lightning leaders produce TGFs, lightning is observed to terminate gamma-ray glows [17, 18, 56].
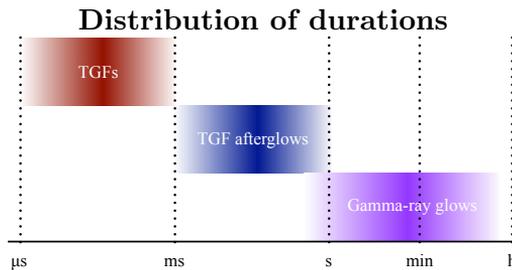
Figure 13.1: Sketch of the distribution of durations of TGFs, TGF afterglows and gamma-ray glows.

Here we predict a new intermediate thunderstorm radiation mechanism, which we call TGF afterglow, that evolves on the timescale of milliseconds to seconds, as also sketched in Fig. 13.1. In short, when photons in the TGF are energetic enough to release neutrons from air molecules by a photonuclear reaction, the neutrons have initial energies of tens of MeV and slowly cool down through collisions with nuclei of air molecules (as neutrons have no electric charge). Dur-

ing thermalization they can be captured again by nuclei and sometimes with the release of a high energy photon, hence in those cases reverting the photonuclear reaction.

That thunderstorms produce neutrons is observed [11, 19, 232, 234, 235, 237, 238, 247, 248]; and the relevant generation channels have been identified [14, 227, 249–251] as photonuclear reactions $\gamma + {}^{14}\text{N} \rightarrow \text{n} + {}^{13}\text{N}$, $\gamma + {}^{16}\text{O} \rightarrow \text{n} + {}^{15}\text{O}$ and $\gamma + {}^{40}\text{Ar} \rightarrow \text{n} + {}^{39}\text{Ar}$, with threshold energies of $\epsilon_N = 10.55$ MeV, $\epsilon_O = 15.7$ MeV and $\epsilon_{Ar} = 9.55$ MeV, respectively [231]. The photonuclear cross section is maximal for photons of roughly 23 MeV, creating neutrons of roughly 13 MeV; for a further discussion of the energy spectrum of the neutrons we refer to Babich et al. [228]. Electrodisintegration reactions (where electrons react with nuclei) could contribute to neutron generation as well, but their contribution is negligible [14]. The simulations by Köhn and Ebert [73], Carlson et al. [130], Babich et al. [252, 253], Drozdov et al. [254] have focussed on neutron production from TGFs, with the number of neutrons produced by a typical TGF varying from 10^{12} neutrons by Carlson et al. [130] to 10^{15} neutrons by Babich et al. [252, 253]. This is mainly due to different assumptions of the total number of photons and their spectrum, or of the initial electrons that create the photons by bremsstrahlung. These studies focus on the neutron emission, and we will return to them in Sec. 13.2. The present study addresses for the first time the prolonged and relocated gamma-ray glow generated by the nuclear capture of the neutrons during their thermalization.

Gurevich et al. [255] have recently observed gamma-ray emissions lasting 100 to 600 ms during lightning activity, with some inner temporal structures with durations that are too long for a TGF, which on ground maximally lasts a few hundreds of microseconds. These observations might be the first measurement of TGF afterglows. We will return to these observations in Sec. 13.3 to illustrate how TGF afterglows would qualitatively appear in measurements.

13.2   Simulations

13.2.1   Setup of simulations

Here we present two simulations, made with the general purpose Monte Carlo code FLUKA [www.fluka.org] [22, 112], which performs very well in the energy regime relevant for TGFs [115], and which has state-of-the-art neutron transport and interactions [112]. We simulate in air (78.085% N_2, 20.95% O_2 and 0.965% Ar) with the altitude dependent density profile given by the 'US Standard Atmosphere (1976)' (by the U.S. Committee on Extension to the Standard Atmosphere).

We perform both simulations within a cylindrical section of the atmosphere extending from ground up to 18 km altitude, with a radius of 12 km. Within FLUKA, this volume is partitioned into 72 horizontal slabs of 250 m thickness

in altitude. Every slab is filled with a homogenous air density determined by the air density of the 'US Standard Atmosphere (1976)' at the bottom of each slab, resulting in an exponential density profile starting from 1.225 kg m^{-3} in the lowermost slab up to 0.1216 kg m^{-3} in the uppermost slab. The temperature, however, is constant and equal to 293K everywhere. Each horizontal slab interface acts as an infinitely thin virtual detector, which detects any passing particle. The output that we record at these interfaces is: the particle type and its kinetic energy, its position in the interface plane, and the time of passing.

The output allows to calculate directly the average flux through the interfaces within a time bin. The time bins are equally spaced on logarithmic scale, with edges given by 10^p s, where p ranges from -9 to 2 in steps of 0.1. The particle density at the interfaces is approximated by dividing the flux by the velocity; more precisely, for the neutrons we add the inverse square root of the kinetic energy of all particles passing the interface within a time bin with an appropriate factor ($\sqrt{m_n/2}$). This approximation of the density is within the accuracy level of other parameters. This time averaged density is presented in different manners in Figures 13.2 and 13.3. The top panel shows these averaged densities integrated over the horizontal interfaces as a function of the discrete altitudes of these interfaces. The middle panel shows the average density as a function of radius at a given altitude. The bottom panel shows the total particle numbers in the system. Discretization artifacts in these figures are due to the discreteness of the interface altitudes and the time bins.

After the primary electron acceleration in a discharge, electrons, photons, neutrons and other TGF products move independently through the atmosphere colliding only with neutral gas molecules, hence the further evolution is linear. We therefore always start with 10^8 particles, and we can get higher particle numbers by multiplying initial, intermediate and final states by the same number. The number 10^8 is chosen as a compromise between statistical accuracy and computational demands.

As already discussed, electrons can gain high energies near leader discharges, and these electron energies are converted into photons by bremsstrahlung. A recent study by ?] of 46 TGFs constrains the average number of electrons with energies above 1 MeV to approximately 2×10^{18}, with a range from 4×10^{16} to 3×10^{19}, for source altitudes above 10 km. According to Briggs [244] photon energies can reach up to tens of MeV. We here concentrate on the photons with energies between 10 and 30 MeV as they can create neutrons by a photonuclear interaction. Gjesteland et al. [256] analyze three TGFs and estimate that the number of photons with energy above 1 MeM varies between 10^{17} and 10^{20} under the assumption that the TGFs have started at 8 km altitude.

13.2.2   TGF afterglow generated by the primary TGF

Our first simulation assumes that the TGF is at 8 km altitude and directed downward. It starts with 10^8 photons with uniformly distributed energies between 10 and 30 MeV. Using the results of Gjesteland et al. [256], and assuming that 1% of the photons with energy above 1 MeV have an energy above 10 MeV, we should actually consider $10^{16\pm1}$ photons above 10 MeV rather than 10^8. But as the evolution outside the TGF source is linear, we can take this into account by multiplying the result of the evolution of 10^8 photons by a factor $10^{8\pm1}$.

Fig. 13.2 shows the evolution as function of the logarithm of time. Photons are included only if their energy exceeds 10 keV. The presented quantities are defined in Sec. 13.2.1. The top panel of Fig. 13.2, viewed from left to right, shows first the light cone of the developing TGF as non-filled red to yellow contours. Photons moving upward have been backscattered or they are secondary, which implies that they have lost a significant amount of energy. Therefore only the primary photons (that move downward), will be energetic enough to produce neutrons; hence the neutron cloud appears only at lower altitudes in this configuration. The mean free path of the photonuclear reaciton scales with density as $\ell = \ell_0 \frac{n_0}{n}$. For the integrated density (starting at 8 km downward), the mean free path of the photonuclear cross section equals 5 km, consistent with Fig. 13.2.

When the neutrons are just created, their typical energy is of the order of 13 MeV (the energy of the maximum of the photonuclear cross section minus the neutron binding energy in nitrogen nuclei); then the neutrons diffuse isotropically and cool down (the neutron energy is given in the bottom panel of Fig. 13.2). While cooling down, the intermediate neutrons do create some photons by inelastic scattering, visible in the top panel at around 3 km, where the TGF envelope extends longer in time than at other altitudes, but after 10^{-4} s the secondary photons produced by inelastic scattering have energies below 10 keV and are thus not shown. The time for neutron thermalization scales as $t = t_0 \frac{n_0}{n}$. We see in the bottom panel of Fig. 13.2 that around 3 km altitude the intermediate neutrons take 0.5 s to thermalize. Neutrons can (at any energy) be captured again, but the cross section for neutron capture increases for decreasing energy as $\sigma_{\text{capture}} \propto 1/\sqrt{E_{\text{neutron}}} \propto 1/v_{\text{neutron}}$, according to the so-called $1/v$-law (see chapter II of [257]). Because of the $1/v$-law, the rate k_{capt} of neutron capture and hence of photon production in the TGF afterglow is constant for constant air density, as $k_{\text{capt}} = v_{\text{neutron}}\sigma_{\text{capture}}n_{\text{air}} \propto \frac{n}{n_0}$. Actually, the most significant capture pathway is not producing a high energy photon, but of radiocarbon (i.e. $n + {}^{14}\text{N} \rightarrow {}^{14}\text{C} + \text{p}$). The cross section for this reaction is $\sigma_{\text{capt}} = 1.8 \times 10^{-28}$ m^2 [?] at thermal velocities (0.025 eV, 2200 m/s), yielding a neutron capture rate of 15.8 s$^{-1}\frac{n}{n_0}$. The TGF afterglow timescale is thus

$$T_{\text{afterglow}} = 1/k_{\text{capt}} \approx 0.063\,\text{s} \exp\left(\frac{h}{7\,\text{km}}\right), \tag{13.1}$$

Evolution of TGF afterglow at 3 km
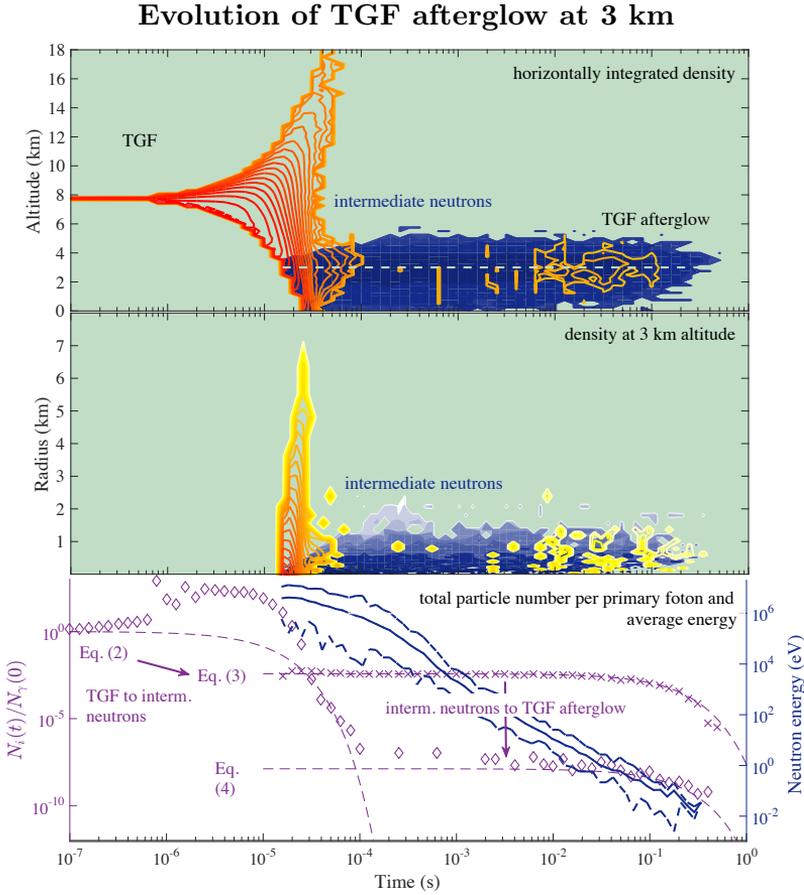
Figure 13.2: Evolution of the TGF afterglow generated by the primary TGF as a function of the logarithm of time. Top and middle panel are contour figures of the photon and neutron density (see definitions in Sec. 13.2.1), on a logarithmic scale; contours represent half a decade (i.e. a factor of $10^{0.5}$). The contour lines (red, yellow to white) are photons above 10 keV, the filled contours (blue to white) are neutrons. In the top panel the density is horizontally integrated. The middle panel gives the density profile as a function of radius at 3 km altitude, the density is averaged over rings around the symmetry axis. The bottom panel shows two quantities: on the left y-axis in purple the total particle number $N_i(t)$ of photons (diamonds) and neutrons (crosses), per initial photon $N_\gamma(0)$, with their approximations given by equations (13.2), (13.3) and (13.4); on the right y-axis in blue the average neutron energy is drawn as a solid line, together with the minimal and the maximal neutron energy as dashed lines.

if one assumes an exponential air density profile with a scale height of 7 km.

The bottom panel in Fig. 13.2 shows the total number of photons and neutrons. This number is the domain integrated time-averaged density, as explained

in Sec. 13.2.1. The evolution can be explained in a simple way, with three species and four rates, where we for convenience neglect the altitude (i.e., air density) dependence of the reaction rates (as all frequencies scale as $f = f_0 \frac{n}{n_0}$). The first reaction is the absorption of high energy photons,

$$N_\gamma(t) \approx N_\gamma(0) \exp\left[-k_{\text{ph-absorp}}t\right] \tag{13.2}$$

with the photon absorption rate $k_{\text{ph-absorb}} = \mu c \approx 2 \times 10^5$ s^{-1} at STP, (where μ is the photon attenuation coefficient, for a discussion see Rutjes et al. [115]). The loss due to the production of neutrons, i.e. the photonuclear reaction $k_{\text{ph-nuc}} = c\sigma_{\text{ph-nuc}}n_{\text{air}} \approx 8 \times 10^2$ s^{-1}, can be neglected in Eq. 13.2 as $k_{\text{ph-absorp}} \gg k_{\text{ph-nuc}}$. In Fig. 13.2 one sees that the photon number $N_\gamma(t)$ (displayed as diamonds) first increases, as the TGF beam creates also secondary photons, which are counted in the simulation, but Eq. 13.2 approximates only the number of high energy photons (with energies say $\gtrsim 1$ MeV), see further discussion by Rutjes et al. [115].

The photonuclear cross section for nitrogen and for photons between 10 and 30 MeV ranges from 1 mb to a peak value of 14 mb at photon energy of 23 MeV [243]. For the approximation of $k_{\text{ph-nuc}}$ above, we took the average photonuclear cross section of nitrogen $\sigma_{\text{ph-nuc}} \approx 2$ mb. The number of neutrons per TGF photon (between 10 MeV and 30 MeV) can then be approximated as $\frac{k_{\text{ph-nuc}}}{k_{\text{ph-absorp}}} \approx 4 \times 10^{-3}$ (consistent with the result of 4.3×10^{-3} by Babich et al. [228]). One may assume that all neutrons are generated - as this is limited by the photon absorption timescale $k_{\text{ph-absorp}}^{-1} \approx 5$ μs - before they start to disappear by capture, which happens with a rate of $k_{\text{capt}} \approx (80\,\text{ms})^{-1}$ at 3 km altitude.

As already mentioned above k_{capt} does not depend on energy, but only on altitude. For the number of intermediate neutrons this yields

$$N_n(t) \approx N_\gamma(0)\frac{k_{\text{ph-nuc}}}{k_{\text{ph-absorp}}} \exp\left[-k_{\text{capt}}t\right] \quad \text{for } t \gg k_{\text{ph-absorp}}^{-1}. \tag{13.3}$$

This equation is consistent with our simulated neutron number, indicated with crosses in the bottom panel of Fig. 13.2. For the gamma-radiation of the TGF afterglow we need to use the number of neutrons and the reaction rate from the most significant pathway producing high energy photons, i.e. $n + {}^{14}\text{N} \rightarrow {}^{15}\text{N} + \gamma$, which happens with a rate of $k_{\text{n-ph}} = 0.7$ s^{-1} as the cross section equals 7.98×10^{-30} m^2 [?] at thermal velocities (0.025 eV, 2200 m/s). Together this results in

$$N_{\gamma-\text{TGF afterglow}}(t) \approx \frac{k_{\text{n-ph}}}{k_{\text{ph-absorp}}}N_n(t) \approx N_\gamma(0)\frac{k_{\text{ph-nuc}}k_{\text{n-ph}}}{k_{\text{ph-absorp}}^2} \exp\left[-k_{\text{capt}}t\right] \quad \text{for } t \gg k_{\text{ph-abs}}^{-1}$$

$$\tag{13.4}$$

where $\frac{k_{\mathrm{ph-nuc}}k_{\mathrm{ph-nuc}}}{k_{\mathrm{ph-absorp}}^2} \approx 1.3 \times 10^{-8}$, consistent with our simulated photon numbers, indicated with diamonds in the bottom panel of Fig. 13.2.

13.2.3 TGF afterglow generated by neutrons (for better statistics)

The number of simulated photons in the TGF afterglow in Fig. 13.2 is limited, as we started the Monte Carlo simulation with 10^8 primary photons, and as the conversion rate from photon to neutron and consecutively from neutron back to photon is low. To achieve better statistics, our second simulation starts directly with 10^8 neutrons at an altitude of 10 km. As photons with energies between between 10 and 30 MeV are converted into neutrons with a probability of about 4×10^{-3} according to our calculations and to Babich et al. [228], we have to multiply our simulation results for particle numbers now with a factor of $4 \times 10^{5\pm1}$ to simulate a TGF with $10^{16\pm1}$ primary photons in the required energy range.

The 10^8 neutrons of our simulation initially all have the most probable energy of 13 MeV, and they are directed downwards, but they rapidly transit to isotropic diffusion. Fig. 13.3 presents the evolution of neutrons and photons in a similar manner as Fig. 13.2, but now focused on the TGF afterglow after 1 ms. Apart from the better statistics of neutron to photon conversion, there are major differences to the earlier simulation. As the air density n_{air} is 2.2 times lower, the neutrons cool down 2.2 times more slowly, and they spread 2.2 times more widely, hence the TGF afterglow is much more extended in space and duration. At the altitude of 10 km it lasts for more than 1 s, as the rate constant k_{capt} in Eq. (13.3) is now $k_{\mathrm{capt}} \approx 5 \ \mathrm{s}^{-1}$.

The statistics of Fig. 13.3 are much better than those of Fig. 13.2, but unfortunately the simulated box (see Sec. 13.2.1) was too small to keep all particles. The top panel clearly indicates that many particles leave the system at its upper border at 18 km altitude. Therefore the normalization rates of Eqs. (13.2-13.4) do not apply in the same fashion, so we rescaled them to fit the data. The decay rate of particles, however, with a rate constant of 5/s at 10 km altitude represents a good fit.

13.2.4 The predicted detector signal

One question is whether the predicted TGF afterglow will be measurable above the cosmic background radiation. Figures 13.2 and 13.3 show that it will be hard to detect a TGF afterglow at sea level, if the neutrons are created above 3 km. We have calculated the predicted detector signal of the TGF afterglows for the simulation of Fig. 13.2 at 3 km altitude, and for the simulation of Fig. 13.3 at 10 km altitude. The detector is the one of Gurevich et al. [255] with an area of 475 cm^2, and we used a temporal bin size of 200 μs as in their published plots.
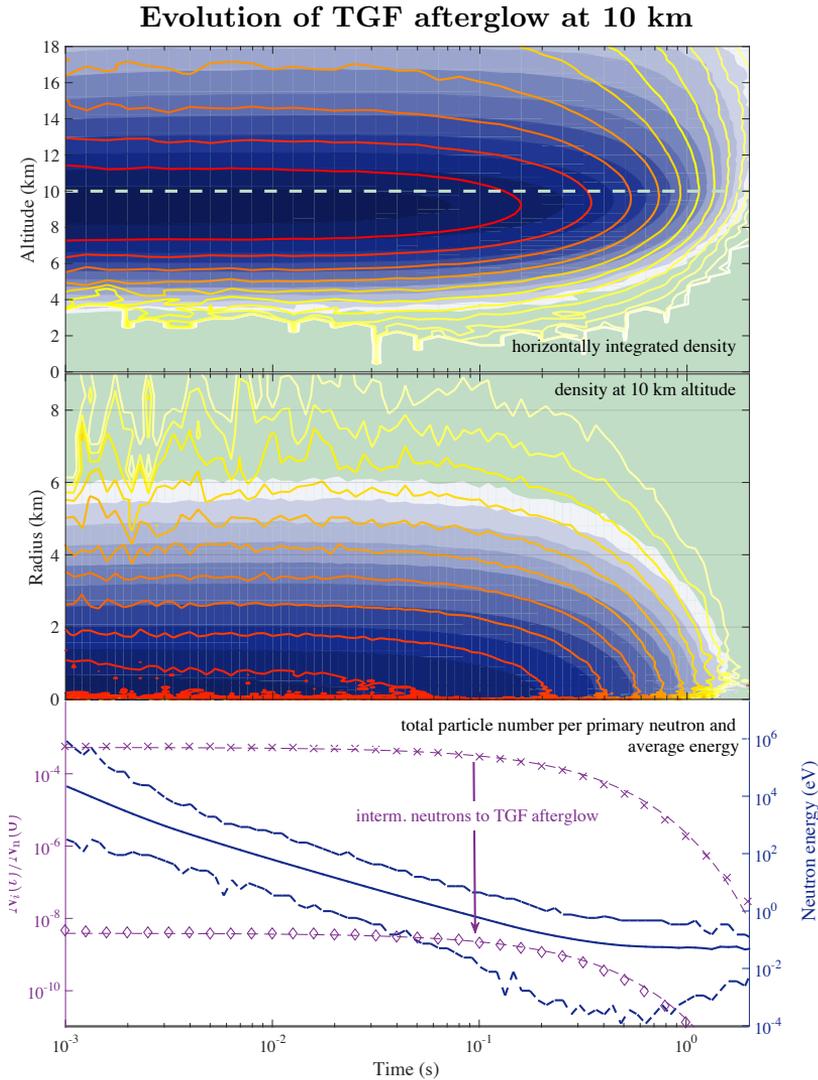
Figure 13.3: The same as in Fig. 13.2, but now for the TGF afterglow started from a neutron source at 10 km directed downwards. In this figure time is plotted only from 10^{-3} s on, focussing on the TGF afterglow. The bottom panel does not represent the total particle number as some escaped of the system at the upper boundary at 18 km, see text. The decay rates, i.e., the fits of the purple dashed lines, are the same as in Fig. 13.2, adapted to the lower air density (at 10 km compared to 3 km).

We assume that it is hit by 2 cm^{-2}s^{-1} or 9 cm^{-2}s^{-1} cosmic background photons with energy above our threshold of 10 keV at 3 or 10 km altitude, based on Bazilevskaya et al. [258]. This Poisson-distributed background is added to the signal in Fig. 13.4.

Simulated and measured detector signals
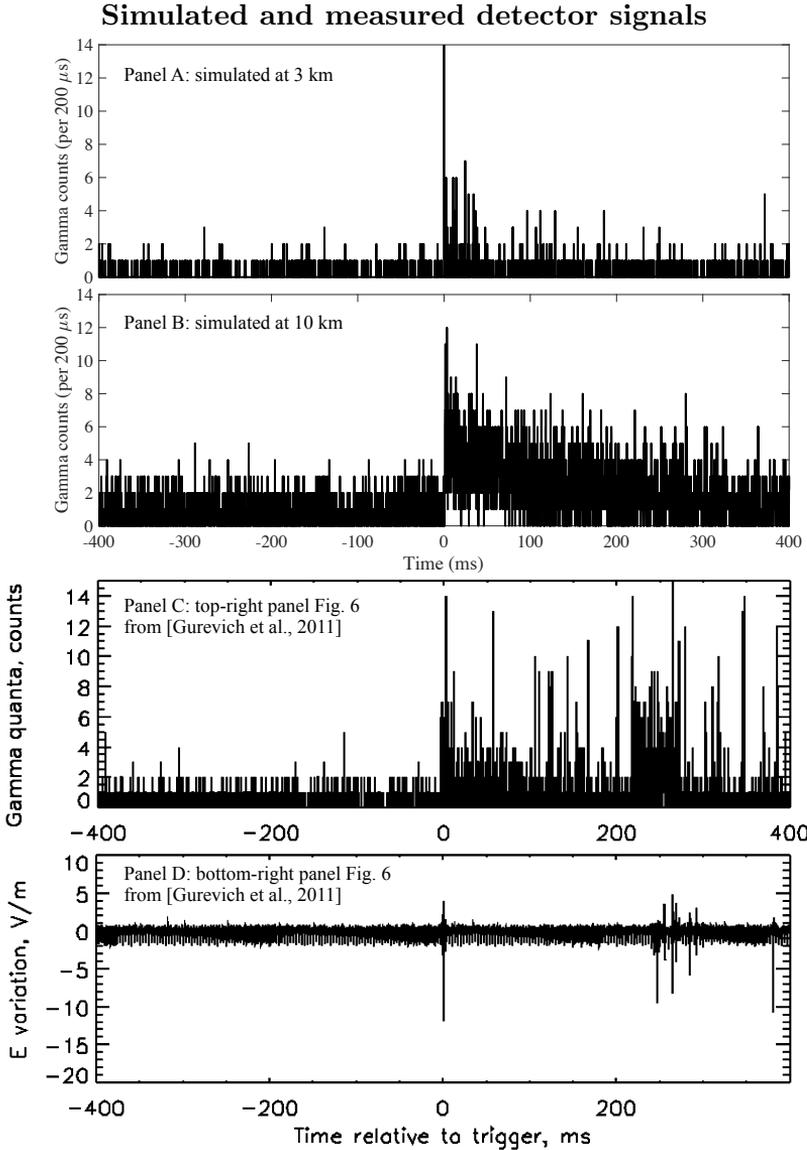
Figure 13.4: Panel A and B: simulated counts of gamma-radiation, from the simulation presented in Fig. 13.2 and Fig. 13.3. Panel C and D are taken from [255], in which it is denoted there as event "6". Panel A to C are gamma-ray counts per 200 μs interval on a detector of 475 cm^2, at 3, 10 and 3.8 km respectively. Panel D gives the measured fast electric field variation (20 μs sampling rate measured by the capacity sensor, see for more details [255]).

As discussed above, the particle numbers of our simulation are orders of magnitude lower than in a real TGF. The statistics of our simulation are corrected to 10^{16} initial photons between 10 and 30 MeV. Obviously, the TGF afterglow can

clearly be detected above the cosmic background radiation. The signal would be even more conspicuous for TGFs containing 10^{17} or 10^{18} photon, above 10 MeV. It is important to remark that photons decay with a rate of $k_{\mathrm{ph-absorp}} = (5\,\mu\mathrm{s})^{-1}$ at STP, or with $(7\,\mu\mathrm{s})^{-1}$ and $(15\,\mu\mathrm{s})^{-1}$ at 3 km and 10 km, respectively. Thus the TGF signal from the simulation of Fig. 13.2 is only visible as a point at $t = 0$ ms in Fig. 13.4, and the duration of the TGF afterglow is just the life-time of the neutrons, as explained with Eq. (13.4).

13.2.5   Summary of predictions for TGF afterglows

We have predicted a new thunderstorm radiation mechanism, the TGF afterglow. It is formed by the photonuclear production of neutrons by the TGF, neutron propagation and cooling and the inverse reaction that creates gamma-rays again. TGF afterglows are thus a signature of gamma-rays above 10 to 30 MeV. A TGF afterglow can be distinguished from TGFs or gamma-ray glows by the following criteria:

1. **Duration:** A TGF lasts not longer than 200 μs, or possibly 600 μs depending on the interpretation of some observations as one or several flashes. A gamma-ray glow lasts for seconds or more, see Fig. 13.1. A TGF afterglow lasts for 60 to 600 ms depending on the atmospheric altitudes crossed by the intermediate neutrons acting as their source, see Eq. (13.1). See also Fig. 13.1 for illustration.

2. **Signal shape:** Neutron and photon signal appear suddenly and decay in time, compared to the photon and neutron signal in a gamma-ray glow which first swells and then decays.

3. **Correlation with fast field changes:** TGF afterglows are created by TGFs which are triggered by leader propagation and related to fast electric field changes. Gamma-ray glows are seen before a discharge and can be terminated by one.

4. **Photon isotropy:** The photons of a TGF afterglow are fairly isotropic, in contrast the beams produced either in a TGF of in a gamma-ray glow by the beamed motion of electrons and their beamed gamma-ray emission by bremsstrahlung).

5. **Energy range:** The photon energy does not exceed the photonnuclear energy of $\epsilon_N \approx 10$ MeV for nitrogen, compared to many tens of MeV in a gamma-ray glow or TGF.).

13.3   Possible observations and outlook

As already mentioned in the introduction, Gurevich et al. [255] have reported gamma-ray emissions lasting for 0.1 to 0.6 s. Clearly, the duration is significantly longer than any TGF detected or simulated, which should disappear within a millisecond, but the signals reported by Gurevich et al. [255] are many orders longer. They occurred during the full duration of an atmospheric discharge at the Tien-Shan Cosmic Ray station at 3.3 to 3.9 km altitude, within the Tien-Shan mountains that reach up to almost 7.5 km altitude. Gurevich et al. [255] found that the temporal distribution of gamma-radiation intensity in a burst is quite non-uniform, with some time structures on the scale of ms strongly correlated with an electric field change during the discharge. Based on duration only, the measurements fall in the regime of TGF afterglows, see Fig. 13.1 for illustration.

To illustrate how TGF afterglows would qualitatively appear in measurements, we added one event from [255] to Fig. 13.4, see panel C and D. The measured gamma-ray counts in panel C appear suddenly at $t = 0$ ms, simultaneous with a fast field variation given in panel D, after which it decays in time. This measured structure in panel C, from 0 ms to 200 ms, shows similarities to our simulated TGF afterglow at 3 km. But, the observations are probably not produced by the specific TGF that we simulated (a TGF starting at 8 km and directed vertically downwards), there could be other scenarios (different altitude, orientation, opening angle and photon spectrum), in addition also the number of photons per TGF varies by an order of magnitude.

The measurements of Gurevich et al. [255] show also structures that would not fit in the description of a TGF afterglow. Namely, structures that first swell and then decay, centered around one or multiple fast field variations. An example of such a structure is also seen in panel C and D, between 200 ms to 300 ms. We speculate that it fits in the description of a gamma-ray glow, but a transient one with a much shorter life-time than typically measured. It could be the result of field development by previous partial discharges, producing a transient patch of air with an electric field above runaway breakdown, until the patch itself is discharged by a leader.

There may be more candidates of gamma-ray observations from thunderstorms which are actually TGF afterglows. We have summarized discriminators in Sec. 13.2.5 to search for TGF afterglows and we invite other researches to look for their signatures in their millisecond-timescale gamma-ray measurements.

Part V

Conclusions, references and Dutch summary

Chapter 14

Conclusions and outlook

14.1 Conclusions

This thesis has two research topics, 'the problem of lightning initiation' and 'high energy radiation from thunderstorms and lightning'. For both topics several articles / chapters have been written of predictive nature as well as on 'Model development, reduction and evaluation' applicable to more general problems than previously mentioned. The main body of this thesis has been organized into three distinct parts. Below, the main conclusions on all these parts are summarized, with references to relevant chapters.

Model development, reduction and evaluation

Chapter 5: Discharge inception from a positive tip occurs due to the accumulation of positive ions near the surface, where the historical Raether-Meek-criterion for streamer inception is not appropriate. When the correct frequency dependent permittivity is used, the streamer discharge develops with much lower velocity and stays much thinner.

Chapter 6: Starting a discharge with a low number of electrons demands a stochastic approach for streamer initiation. We combined existing software tools and extended them to incorporate dielectrics. We present preliminary results of discharge inception from an ice needle and charged droplets. We observe that discharges start with a significant jitter and usually off symmetry axis, limiting the use of cylindrically symmetric deterministic fluid simulations. Ideally one starts with a particle code to simulate the probability of inception and the realistic initial avalanche to streamer phase, and thereafter one continues with a fluid model for computational efficiency.

Chapter 7: Benchmark tests for codes in high energy atmospheric physics are constructed and used to evaluate custom-made codes and general purpose codes made by large collaborations, for the scenario without electric or

magnetic fields. We discuss important differences between the results of the different codes and provide plausible explanations, mainly caused by different implementations of the so-called low energy cutoff and the resulting straggling effect.

Chapter 8: For scenarios below classical breakdown and above runaway breakdown, we started benchmarking important relativistic runaway electron avalanche (RREA) characteristics. We observe and prove that not the low energy cutoff, but the maximal step length setting significantly changes the accuracy of the simulation, which changes the perspective in modeling RREA significantly compared to what was recently stated in literature. As a direct consequence we show that it is safe to increase the low energy cutoff by a few orders of magnitude, to improve computational efficiency and statistical accuracy for the same computation time. We derived the relative error as function of the maximal step length and present its sensitivity to the mean energy of the RREA.

Chapter 9: We present a simple and robust strategy for the selection of sampling points in Uncertainty Quantification. The goal is to achieve the fastest possible convergence in the cumulative distribution function of a stochastic output of interest. We present comparisons with a stochastic collocation method based on the Clenshaw-Curtis quadrature rule, and with an adaptive method based on hierarchical surplus, showing that the new method often results in a large computational saving.

Lightning initiation problem

Chapter 10: We prove that a streamer can start from an ice particle in a thundercloud electric field that is significantly below classical breakdown, with realistic initial conditions and dielectric properties of ice. We derive the interplay between hydrometer shape and thundercloud electric field strength, which shows that ellipsoidally shaped hydrometeors have an optimal aspect ratio for fixed length. Furthermore, we calculated by Monte Carlo study of extensive air showers the frequency to obtain at least 100 cm^{-3} initial free electrons to start the discharge and perform for a particular example the full time-dependent self-consistent streamer simulation, to prove that a streamer can start. The particular example configuration simulated is expected to occur with a frequency of once per minute.

Chapter 11: Recent observations predict that lightning starts by a volumetric system of streamers, which seems to be triggered simultaneously by an external event. A possible explanation is drawn from the hypothesis that a single extensive air shower initiates multiple independent streamers from hydrometeors in a coherent way. We thoroughly studied the probability

distribution of relevant extensive air showers and derived, with rare event analyses, the expected availability of extreme electron seeds. We observe that free electron densities at altitudes between 5 and 13 km do not exceed values of 5×10^3 cm^3, and then only in cores of centimeter scale. In addition we observe that above 6 km, the availability of extreme free electron densities decreases significantly with increasing altitudes. To match the width of the system of streamers seen in measurements, streamers must be triggered by electron densities down to 1 cm^3. Such low electron densities demand a stochastic approach for streamer initiation, as discussed in chapter 6.

Neutrons from thunderclouds and lightning

Chapter 12: Neutron emissions with different durations have been observed on ground during thunderstorms. These neutrons can be produced by fast Terrestrial Gamma-ray Flashes correlated with lightning, or by Gamma-ray Glows lasting several minutes. In both cases the neutrons are produced through a photonuclear reaction with nuclei of air molecules, if gamma-rays in the energy range of 10 to 30 MeV are present, and are thus a diagnostic for studying high energy atmospheric physics. We present simulations of downward gamma-ray beams with energies between 10 and 30 MeV, separated into 4 energy intervals with 5 MeV of width, at varying source altitudes. The individual signature at sea level of each energy interval allows the extrapolation of our results to an arbitrary source spectrum and altitude. Our results indicate that neutrons are created along all photon trajectories and the majority reaches sea level within a radius of around 0.5 km from the source axis.

Chapter 13: Thunderstorms are known to create terrestrial gamma-ray flashes (TGF) which are microsecond-long bursts created by runaway of thermal electrons from propagating lightning leaders, as well as gamma-ray glows that possibly are created by relativistic runaway electron avalanches that can last for minutes or more and are sometimes terminated by a discharge. We predict a new intermediate thunderstorm radiation mechanism, which we call TGF afterglow, as it is caused by the capture of photonuclear neutrons produced by a TGF. TGF afterglows are milliseconds to seconds long; this duration is caused by the thermalization time of the intermediate neutrons. TGF afterglows indicate that the primary TGF has produced photons in the energy range of 10 - 30 MeV; they are nondirectional in contrast to the primary TGF. Gurevich *et al* might have reported TGF afterglows in 2011.

14.2   Outlook

Lightning is observed to start as a volumetric system of streamers [202], potentially triggered by an external extensive air shower. Combining the probabilities of starting a streamer from a hydrometeor, the occurrence of a hydrometeor and the occurrence of an electron seed that triggers the discharge, would give insight in how these volumetric systems of streamers could form. For this, accurate hydrometeor size and shape distributions in thunderclouds are extremely desired, but up to now not available.

A next step in the understanding of lightning initiation is how the initial streamers further develop and interact with each other and with (other) hydrometeors. For example: *How far can a streamer propagate in a given sub-breakdown electric field? How does a hydrometeor influence path and evolution of a streamer?*. A direct challenge is to incorporate a physically accurate and computationally efficient way to simulate the interaction of the plasma with a dielectric surface (the hydrometeors). The latter is not only interesting for the understanding of lightning, but it is also valuable for the high voltage technology, for example to prevent surface flashovers.

Neutrons observed from thunderclouds and lightning are produced by photonuclear interactions with activation energy of 10 MeV [14, 227, 249–251], where the photons are created by bremsstrahlung of energetic electrons assumingely accelerated in electrostatic fields. However, recently neutrons were detected from laboratory discharges while the potential did not exceed 1 MV energy [259, 260]. Thus these observations suggest that other types of electron acceleration must play a role and may also be important for the acceleration of electrons near stepping lightning leaders. So: *How do the TGF source electrons get accelerated?* and strongly connected: *How do leaders step?* Recent findings, described in chapter 8, provide a new perspective on combining models for high and low energy electrons as the so-called low energy cutoff can be set orders of magnitude larger than expected. It enables a separation in energy scale, which could be used to construct a hybrid model where low energy electrons are modeled as a fluid and high energy electrons as discrete particles.

Bibliography

[1] Thomas C. Marshall and Maribeth Stolzenburg. Voltages inside and just above thunderstorms. *Journal of Geophysical Research: Atmospheres*, 106 (D5):4757–4768, 2001. doi:10.1029/2000JD900640.

[2] J. A. Crabb and J. Latham. Corona from colliding drops as a possible mechanism for the triggering of lightning. *Quarterly Journal of the Royal Meteorological Society*, 100(424):191, apr 1974. doi:10.1002/qj.49710042406.

[3] R. Gunn. Electric Field Intensity Inside of Natural Clouds. *Journal of Applied Physics*, 19:481, 1948. doi:10.1063/1.1698159.

[4] William P Winn and CB Moore. Electric field measurements in thunderclouds using instrumented rockets. *Journal of Geophysical Research*, 76 (21):5003–5017, 1971. doi:10.1029/JC076i021p05003.

[5] T. C. Marshall, M. Stolzenburg, and W. D. Rust. Electric field measurements above mesoscale convective systems. *Journal of Geophysical Research*, 101:6979, 1996. doi:10.1029/95JD03764.

[6] Maribeth Stolzenburg, Thomas C. Marshall, W. David Rust, Eric Bruning, Donald R. MacGorman, and Timothy Hamlin. Electric field values observed near lightning flash initiations. *Geophysical Research Letters*, 34 (4):L04804, 2007. doi:10.1029/2006GL028777. L04804.

[7] L. B. Loeb. The mechanisms of stepped and dart leaders in cloud-to-ground lightning strokes. *Journal of Geophysical Research*, 71:4711, 1966. doi:10.1029/JZ071i020p04711.

[8] R. F. Griffiths and J. Latham. Electrical corona from ice hydrometeors. *Quarterly Journal of the Royal Meteorological Society*, 100(424):163–180, apr 1974. doi:10.1002/qj.49710042404.

[9] RF Griffiths. The initiation of corona discharges from charged ice particles in a strong electric field. *Journal of Electrostatics*, 1(1):3–13, 1975.

[10] Danyal Petersen, Matthew Bailey, John Hallett, and William H Beasley. Laboratory investigation of positive streamer discharges from simulated ice hydrometeors. *Quarterly Journal of the Royal Meteorological Society*, 132(615):263–273, 2006.

[11] GN Shah, H Razdan, CL Bhat, and QM Ali. Neutron generation in lightning bolts. *Nature*, 313:773, 1985. doi:10.1038/313773a0.

[12] D. A. Petersen and W. H. Beasley. High-speed video observations of a natural negative stepped leader and subsequent dart-stepped leader. *Journal of Geophysical Research: Atmospheres*, 118:12, 2013. doi:10.1002/2013JD019910.

[13] Nikolai Østgaard, KH Albrecthsen, Thomas Gjesteland, and A Collier. A new population of terrestrial gamma-ray flashes in the rhessi data. *Geophysical Research Letters*, 42(24), 2015.

[14] L. P. Babich, E. I. Bochkov, I. M. Kutsyk, and H. K. Rassoul. Analysis of fundamental interactions capable of producing neutrons in thunderstorms. *Phys. Rev. D*, 89:093010, May 2014. doi:10.1103/PhysRevD.89.093010.

[15] A. Chilingarian. Thunderstorm Ground Enhancements (TGEs) - New High-Energy Phenomenon Originated in the Terrestrial Atmosphere. *Journal of Physics Conference Series*, 409(1):012019, 2013. doi:10.1088/1742-6596/409/1/012019.

[16] A. Chilingarian, N. Bostanjyan, T. Karapetyan, and L. Vanyan. Neutron production during thunderstorms. *Journal of Physics Conference Series*, 409(1):012216, 2013. doi:10.1088/1742-6596/409/1/012216.

[17] A Chilingarian, G Hovsepyan, G Khanikyanc, A Reymers, and S Soghomonyan. Lightning origination and thunderstorm ground enhancements terminated by the lightning flash. *EPL (Europhysics Letters)*, 110 (4):49001, 2015.

[18] Nicole A Kelley, David M Smith, Joseph R Dwyer, Michael Splitt, Steven Lazarus, Forest Martinez-McKinney, Bryna Hazelton, Brian Grefenstette, Alexander Lowell, and Hamid K Rassoul. Relativistic electron avalanches as a thunderstorm discharge competing with lightning. *Nature communications*, 6, 2015.

[19] AA Toropov, VI Kozlov, VA Mullayarov, and SA Starodubtsev. Experimental observations of strengthening the neutron flux during negative lightning discharges of thunderclouds with tripolar configuration. *Journal of Atmospheric and Solar-Terrestrial Physics*, 94:13–18, 2013.

[20] JP Santos, F Parente, and Yong-Ki Kim. Cross sections for k-shell ioniza-
tion of atoms by electron impact. *Journal of Physics B: Atomic, Molecular
and Optical Physics*, 36(21):4211, 2003.

[21] Hideo Hirayama, Yoshihito Namito, Walter R Nelson, Alex F Bielajew,
Scott J Wilderman, and U Michigan. The egs5 code system. Technical
report, United States. Department of Energy, 2005.

[22] Alfredo Ferrari, Paola R Sala, Alberto Fasso, and Johannes Ranft. Fluka:
A multi-particle transport code (program version 2005). Technical report,
2005.

[23] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth.*,
A506:250–303, 2003. doi:10.1016/S0168-9002(03)01368-8.

[24] D. Sarria, P.-L. Blelly, and F. Forme. Mc-peptita: a monte carlo model
for photon, electron and positron tracking in terrestrial atmosphere. ap-
plication for a terrestrial gamma-ray flash. *Journal of Geophysical Re-
search: Space Physics*, pages n/a–n/a, 2015. doi:10.1002/2014JA020695.
2014JA020695.

[25] Jannis Teunissen and Ute Ebert. 3d pic-mcc simulations of discharge incep-
tion around a sharp anode in nitrogen/oxygen mixtures. *Plasma Sources
Science and Technology*, 25(4):044005, 2016.

[26] Anna Dubinova, Casper Rutjes, Ute Ebert, Stijn Buitink, Olaf Scholten,
and Gia Thi Ngoc Trinh. Prediction of lightning inception by large ice par-
ticles and extensive air showers. *Physical Review Letters*, 115(1):015002,
Jun 2015. doi:10.1103/PhysRevLett.115.015002.

[27] Jannis Teunissen and Ute Ebert. Afivo: a framework for quadtree/octree
amr with shared-memory parallelization and geometric multigrid methods.
arXiv preprint arXiv:1701.04329, 2017.

[28] Jannis Teunissen and Ute Ebert. Simulating streamer discharges in 3d
with the parallel adaptive afivo framework. *Journal of Physics D: Applied
Physics*, 50(47):474001, oct 2017. doi:10.1088/1361-6463/aa8faf.

[29] Samaneh Sadighi, Ningyu Liu, Joseph R. Dwyer, and Hamid K. Ras-
soul. Streamer formation and branching from model hydrometeors in
subbreakdown conditions inside thunderclouds. *Journal of Geophysical Re-
search: Atmospheres*, 120(9):3660–3678, 2015. doi:10.1002/2014JD022724.
2014JD022724.

[30] L. P. Babich, E. I. Bochkov, I. M. Kutsyk, T. Neubert, and O. Chan-
rion. Positive streamer initiation from raindrops in thundercloud fields.

Journal of Geophysical Research: Atmospheres, 121(11):6393–6403, 2016. doi:10.1002/2016JD024901. 2016JD024901.

[31] A. Luque and U. Ebert. Growing discharge trees with self-consistent charge transport: the collective dynamics of streamers. *New Journal of Physics*, 16(1):013039, 2014. doi:10.1088/1367-2630/16/1/013039.

[32] Tim Schultz, Martin Pfeiffer, and Christian M. Franck. Optical investigation methods for determining the impact of rain drops on hvdc corona. *Journal of Electrostatics*, 77(Supplement C):13 – 20, 2015. doi:https://doi.org/10.1016/j.elstat.2015.06.007.

[33] M. Becerra, V. Cooray, S. Soula, and S. Chauzy. Effect of the space charge layer created by corona at ground level on the inception of upward lightning leaders from tall towers. *Journal of Geophysical Research: Atmospheres*, 112:D12205, 2007. doi:10.1029/2006JD008308.

[34] Ningyu Liu, Burcu Kosar, Samaneh Sadighi, Joseph R. Dwyer, and Hamid K. Rassoul. Formation of streamer discharges from an isolated ionization column at subbreakdown conditions. *Physical Review Letters*, 109(2):025002, Jul 2012. doi:10.1103/PhysRevLett.109.025002.

[35] L.P. Babich, E.I. Bochkov, and T. Neubert. The role of charged ice hydrometeors in lightning initiation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 154(Supplement C):43 – 46, 2017. doi:10.1016/j.jastp.2016.12.010.

[36] J. M. Meek. A Theory of Spark Discharge. *Physical Review*, 57:722, 1940. doi:10.1103/PhysRev.57.722.

[37] H. Raether. Die Entwicklung der Elektronenlawine in den Funkenkanal. *Zeitschrift fuer Physik*, 112(7-8):464–489, jul 1939. doi:10.1007/BF01340229.

[38] Leonard B Loeb and John M Meek. The mechanism of spark discharge in air at atmospheric pressure. I. *Journal of Applied Physics*, 11(6):438–447, 1940.

[39] Leonard B Loeb and John M Meek. The mechanism of spark discharge in air at atmospheric pressure. II. *Journal of Applied Physics*, 11(7):459–474, 1940.

[40] Heinz Raether. Electron avalanches and breakdown in gases. 1964.

[41] John Kuffel and Peter Kuffel. *High voltage engineering fundamentals*. Newnes, 2000.

[42] C. Montijn and U. Ebert. Diffusion correction to the Raether Meek criterion for the avalanche-to-streamer transition. *Journal of Physics D: Applied Physics*, 39:2979, 2006. doi:10.1088/0022-3727/39/14/017.

[43] A Dubinova, D Trienekens, U Ebert, S Nijdam, and T Christen. Pulsed positive discharges in air at moderate pressures near a dielectric rod. *Plasma Sources Science and Technology*, 25(5):055021, 2016. URL `http://stacks.iop.org/0963-0252/25/i=5/a=055021`.

[44] Anna Alexandrovna Dubinova. *Modeling of streamer discharges near dielectrics*. TU Eindhoven, 2016.

[45] Xu-Dong Liu, Ronald P. Fedkiw, and Myungjoo Kang. A boundary condition capturing method for poisson's equation on irregular domains. *Journal of Computational Physics*, 160(1):151 – 178, 2000. doi:https://doi.org/10.1006/jcph.2000.6444.

[46] S. Celestin, Z. Bonaventura, B. Zeghondy, A. Bourdon, and P. Ségur. The use of the ghost fluid method for Poisson's equation to simulate streamer propagation in point-to-plane and point-to-point geometries. *Journal of Physics D: Applied Physics*, 42(6):065203, 2009. doi:10.1088/0022-3727/42/6/065203.

[47] V. G. Artemov and A. A. Volkov. Water and ice dielectric spectra scaling at 0c. *Ferroelectrics*, 466(1):158–165, August 2014. doi:10.1080/00150193.2014.895216.

[48] Shuji Fujita, Takeshi Matsuoka, Toshihiro Ishida, Kenichi Matsuoka, and Shinji Mae. A summary of the complex dielectric permittivity of ice in the megahertz range and its applications for radar sounding of polar ice sheets. In *Physics of ice core records*, pages 185–212. Hokkaido University Press, 2000.

[49] U. Ebert, W. van Saarloos, and C. Caroli. Propagation and Structure of Planar Streamer Fronts. In *eprint arXiv:patt-sol/9702006*, page 2006, 1997.

[50] Jannis Teunissen and Ute Ebert. Controlling the weights of simulation particles: adaptive particle management using k-d trees. *Journal of Computational Physics*, 259(Supplement C):318 – 330, 2014. doi:https://doi.org/10.1016/j.jcp.2013.12.005.

[51] N. Liu, J. R. Dwyer, and H. K. Rassoul. Effects of pressure and humidity on positive corona inception from thundercloud hydrometeors. *J. Atm. Sol.-Terr. Phys.*, 80:179, 2012. doi:10.1016/j.jastp.2012.01.012.

[52] Kimmo Kärkkäinen et al. *On the finite-difference modelling of electromagnetic problems in structured lattices.* Helsinki University of Technology, 2003.

[53] G. J. Fishman, P. N. Bhat, R. Mallozzi, J. M. Horack, T. Koshut, C. Kouveliotou, G. N. Pendleton, C. A. Meegan, R. B. Wilson, W. S. Paciesas, S. J. Goodman, and H. J. Christian. Discovery of Intense Gamma-Ray Flashes of Atmospheric Origin. *Science*, 264:1313–1316, May 1994. doi:10.1126/science.264.5163.1313.

[54] J. R. Dwyer, B. W. Grefenstette, and D. M. Smith. High-energy electron beams launched into space by thunderstorms. *Geophysical Research Letters*, 35(2):L02815, 2008. doi:10.1029/2007GL032430.

[55] Michael S Briggs, Valerie Connaughton, Colleen Wilson-Hodge, Robert D Preece, Gerald J Fishman, R Marc Kippen, PN Bhat, William S Paciesas, Vandiver L Chaplin, Charles A Meegan, et al. Electron-positron beams from terrestrial lightning observed with fermi gbm. *Geophysical research letters*, 38(2):L02808, January 2011. doi:10.1029/2010GL046259.

[56] M McCarthy and GK Parks. Further observations of x-rays inside thunderstorms. *Geophysical Research Letters*, 12(6):393–396, 1985. doi:10.1029/GL012i006p00393.

[57] K. B. Eack, W. H. Beasley, W. D. Rust, T. C. Marshall, and M. Stolzenburg. Initial results from simultaneous observation of x-rays and electric fields in a thunderstorm. *Journal of Geophysical Research: Atmospheres*, 101(D23):29637–29640, 1996. doi:10.1029/96JD01705.

[58] H Tsuchiya, T Enoto, S Yamada, T Yuasa, M Kawaharada, T Kitaguchi, M Kokubun, H Kato, M Okano, S Nakamura, et al. Detection of high-energy gamma rays from winter thunderclouds. *Physical review letters*, 99 (16):165002, 2007. doi:10.1103/PhysRevLett.99.165002.

[59] Toru Adachi, Yukihiro Takahashi, Hiroyo Ohya, Fuminori Tsuchiya, Kozo Yamashita, Mamoru Yamamoto, and Hiroyuki Hashiguchi. Monitoring of lightning activity in southeast asia: Scientific objectives and strategies. *Kyoto Working Papers on Area Studies: G-COE Series*, 2008.

[60] A Chilingarian, A Daryan, K Arakelyan, A Hovhannisyan, B Mailyan, L Melkumyan, G Hovsepyan, S Chilingaryan, A Reymers, and L Vanyan. Ground-based observations of thunderstorm-correlated fluxes of high-energy electrons, gamma rays, and neutrons. *Physical Review D*, 82(4): 043009, 2010. doi:10.1103/PhysRevD.82.043009.

[61] Ashot Chilingarian, Gagik Hovsepyan, and Armen Hovhannisyan. Particle bursts from thunderclouds: Natural particle accelerators above our heads. *Physical review D*, 83(6):062001, 2011. doi:10.1103/PhysRevD.83.062001.

[62] J. R. Dwyer, D. M. Smith, and S. A. Cummer. High-Energy Atmospheric Physics: Terrestrial Gamma-Ray Flashes and Related Phenomena. *Space Science Review*, 173:133–196, November 2012. doi:10.1007/s11214-012-9894-0.

[63] CTR Wilson. The electric field of a thundercloud and some of its effects. *Proceedings of the Physical Society of London*, 37(1):32D, 1924.

[64] AV Gurevich, GM Milikh, and R Roussel-Dupre. Runaway electron mechanism of air breakdown and preconditioning during a thunderstorm. *Physics Letters A*, 165(5):463–468, June 1992. doi:10.1016/0375-9601(92)90348-P.

[65] AV Gurevich. On the theory of runaway electrons. *Sov. Phys. JETP*, 12 (5):904–912, 1961.

[66] G. D. Moss, V. P. Pasko, N. Liu, and G. Veronis. Monte Carlo model for analysis of thermal runaway electrons in streamer tips in transient luminous events and streamer zones of lightning leaders. *Journal of Geophysical Research (Space Physics)*, 111:A02307, February 2006. doi:10.1029/2005JA011350.

[67] D. W. Liu, F. Iza, and M. G. Kong. Electron avalanches and diffused γ-mode in radio-frequency capacitively coupled atmospheric-pressure microplasmas. *Applied Physics Letters*, 95(3):031501, 2009. doi:10.1063/1.3186073.

[68] O. Chanrion and T. Neubert. Production of runaway electrons by negative streamer discharges. *J. Geophys. Res. (Space Phys)*, 115:A00E32, 2010. doi:10.1029/2009JA014774.

[69] Sebastien Celestin and Victor P. Pasko. Energy and fluxes of thermal runaway electrons produced by exponential growth of streamers during the stepping of lightning leaders and in transient luminous events. *Journal of Geophysical Research: Space Physics*, 116(A3):A03315, mar 2011. doi:10.1029/2010JA016260.

[70] S. Celestin, W. Xu, and V. P. Pasko. Terrestrial gamma ray flashes with energies up to 100 MeV produced by nonequilibrium acceleration of electrons in lightning. *Journal of Geophysical Research (Space Physics)*, 117 (16):A05315, May 2012. doi:10.1029/2012JA017535. A05315.

[71] O. Chanrion, Z. Bonaventura, D. Çinar, A. Bourdon, and T. Neubert. Runaway electrons from a "beam-bulk" model of streamer: application to TGFs. *Environmental Research Letters*, 9(5):055003, May 2014. doi:10.1088/1748-9326/9/5/055003.

[72] C Kohn, U Ebert, and A Mangiarotti. The importance of electron–electron bremsstrahlung for terrestrial gamma-ray flashes, electron beams and electron–positron beams. *Journal of Physics D: Applied Physics*, 47 (25):252001, 2014. URL `http://stacks.iop.org/0022-3727/47/i=25/a=252001`.

[73] Christoph Köhn and Ute Ebert. Calculation of beams of positrons, neutrons, and protons associated with terrestrial gamma ray flashes. *Journal of Geophysical Research: Atmospheres*, 120(4):1620–1635, 2015.

[74] Yu L Stankevich and VG Kalinin. Fast electrons and x radiation during the initial stages of an impulse spark discharge in airr. *Dokl. Akad. Nauk SSSR, 177: 72-3 (Nov.-Dec. 1967).*, 1967.

[75] I. D. Kostyrya, V. F. Tarasenko, A. N. Tkachev, and S. I. Yakovlenko. X-ray radiation due to nanosecond volume discharges in air under atmospheric pressure. *Journal of Technical Physics*, 51(3):356, 2006. doi:10.1134/S1063784206030108.

[76] Viktor F Tarasenko, Evgeny Kh Baksht, Aleksandr G Burachenko, Igor D Kostyrya, Mikhail I Lomaev, and Dmitry V Rybka. Generation of supershort avalanche electron beams and formation of diffuse discharges in different gases at high pressure. *Plasma Devices and Operations*, 16(4): 267–298, 2008.

[77] Tao Shao, Cheng Zhang, Zheng Niu, Ping Yan, Victor F Tarasenko, Evgenii Kh Baksht, Alexander G Burahenko, and Yuliya V Shut'ko. Diffuse discharge, runaway electron, and x-ray in atmospheric pressure air in an inhomogeneous electrical field in repetitive pulsed modes. *Applied Physics Letters*, 98(2):021503, 2011. doi:10.1063/1.3540504.

[78] RC Noggle, EP Krider, and JR Wayland. A search for x rays from helium and air discharges at atmospheric pressure. *Journal of Applied Physics*, 39(10):4746–4748, 1968.

[79] JR Dwyer, Z Saleh, HK Rassoul, D Concha, Mahbubur Rahman, Vernon Cooray, J Jerauld, MA Uman, and VA Rakov. A study of x-ray emission from laboratory sparks in air at atmospheric pressure. *Journal of Geophysical Research: Atmospheres*, 113(D23):D23207, 2008. doi:10.1029/2008JD010315.

[80] AG Rep'ev and PB Repin. Spatiotemporal parameters of the x-ray radiation from a diffuse atmospheric-pressure discharge. *Technical Physics*, 53 (1):73–80, 2008. doi:10.1134/S1063784208010143.

[81] V. Cooray, L. Arevalo, M. Rahman, J. Dwyer, and H. Rassoul. On the possible origin of X-rays in long laboratory sparks. *J. Atm. Sol.-Terr. Phys.*, 71(17):1890, 2009. doi:10.1016/j.jastp.2009.07.010.

[82] PO Kochkin, CV Nguyen, APJ Van Deursen, and U Ebert. Experimental study of hard x-rays emitted from metre-scale positive discharges in air. *Journal of Physics D: Applied Physics*, 45(42):425202, 2012. doi:10.1088/0022-3727/45/42/425202.

[83] P. O. Kochkin, A. P. J. van Deursen, and U. Ebert. Experimental study of the spatio-temporal development of metre-scale negative discharge in air. *Journal of Physics D: Applied Physics*, 47(14):145203, 2014. doi:10.1088/0022-3727/47/14/145203.

[84] Pavlo Kochkin, Christoph Köhn, Ute Ebert, and Lex van Deursen. Analyzing x-ray emissions from meter-scale negative discharges in ambient air. *Plasma Sources Science and Technology*, 25(4):044002, 2016.

[85] J. R. Dwyer. A fundamental limit on electric fields in air. *Geophysical Research Letters*, 30(20):2055, 2003. doi:10.1029/2003GL017781.

[86] LP Babich, EN Donskoy, IM Kutsyk, and RA Roussel-Dupré. The feedback mechanism of runaway air breakdown. *Geophysical research letters*, 32(9):L09809, 2005. doi:10.1029/2004GL021744.

[87] J. R. Dwyer. Relativistic breakdown in planetary atmospheres. *Physics of Plasmas*, 14(4):042901, April 2007. doi:10.1063/1.2709652.

[88] J. R. Dwyer. The relativistic feedback discharge model of terrestrial gamma ray flashes. *J. Geophys. Res. (Space Phys)*, 117(A2):A02308, February 2012. doi:10.1029/2011JA017160.

[89] A. Luque. Relativistic Runaway Ionization Fronts. *Physical Review Letters*, 112(4):045003, 2014. doi:10.1103/PhysRevLett.112.045003.

[90] P Schellart, TNG Trinh, S Buitink, A Corstanje, JE Enriquez, H Falcke, JR Hörandel, A Nelles, JP Rachen, L Rossetto, et al. Probing atmospheric electric fields in thunderstorms through radio emission from cosmic-ray-induced air showers. *Physical review letters*, 114(16):165001, 2015.

[91] TNG Trinh, O Scholten, S Buitink, AM van den Berg, A Corstanje, U Ebert, JE Enriquez, H Falcke, JR Hörandel, C Köhn, et al. Influence of atmospheric electric fields on the radio emission from extensive air showers. *Physical Review D*, 93(2):023003, 2016.

[92] Pedro Andreo. Monte carlo techniques in medical radiation physics. *Physics in medicine and biology*, 36(7):861, 1991.

[93] J Sempau, A Sanchez-Reyes, F Salvat, H Oulad ben Tahar, SB Jiang, and JM Fernández-Varea. Monte carlo simulation of electron beams from an accelerator head using penelope. *Physics in medicine and biology*, 46(4): 1163, 2001.

[94] J-F Carrier, L Archambault, L Beaulieu, and R Roy. Validation of geant4, an object-oriented monte carlo toolkit, for simulations in medical physics. *Medical physics*, 31(3):484–492, 2004.

[95] S Pancheshnyi, S Biagi, MC Bordage, GJM Hagelaar, WL Morgan, AV Phelps, and LC Pitchford. The lxcat project: Electron scattering cross sections and swarm parameters for low temperature plasma modeling. *Chemical Physics*, 398:148–153, 2012.

[96] Nikolai Østgaard, Thomas Gjesteland, J Stadsnes, PH Connell, and B Carlson. Production altitude and time delays of the terrestrial gamma flashes: Revisiting the burst and transient source experiment spectra. *Journal of Geophysical Research: Space Physics*, 113(A2):A02307, February 2008. doi:10.1029/2007JA012618.

[97] J. R. Dwyer and D. M. Smith. A comparison between Monte Carlo simulations of runaway breakdown and terrestrial gamma-ray flash observations. *Geophysical Research Letters*, 32:L22804, November 2005. doi:10.1029/2005GL023848.

[98] B. E. Carlson, N. G. Lehtinen, and U. S. Inan. Constraints on terrestrial gamma ray flash production from satellite observation. *Geophysical Research Letters*, 34:L08809, April 2007. doi:10.1029/2006GL029229.

[99] B. J. Hazelton, B. W. Grefenstette, D. M. Smith, J. R. Dwyer, X.-M. Shao, S. A. Cummer, T. Chronis, E. H. Lay, and R. H. Holzworth. Spectral dependence of terrestrial gamma-ray flashes on source distance. *Geophysical Research Letters*, 36(1):L01108, jan 2009. doi:10.1029/2008gl035906.

[100] D. Sarria, P.-L. Blelly, M. S. Briggs, and F. Forme. Studying the time histogram of a terrestrial electron beam detected from the opposite hemisphere of its associated tgf. *Journal of Geophysical Research: Space Physics*, pages n/a–n/a, 2016. doi:10.1002/2015JA021881.

[101] Dieter Heck, G Schatz, J Knapp, T Thouw, and JN Capdevielle. Corsika: A monte carlo code to simulate extensive air showers. Technical report, 1998.

[102] B. E. Carlson, T. Gjesteland, and N. Østgaard. Terrestrial gamma-ray flash electron beam geometry, fluence, and detection frequency. *Journal of Geophysical Research (Space Physics)*, 116(15):A11217, November 2011. doi:10.1029/2011JA016812.

[103] Alexander Broberg Skeltved, Nikolai Østgaard, Brant Carlson, Thomas Gjesteland, and Sebastien Celestin. Modeling the relativistic runaway electron avalanche and the feedback mechanism with geant4. *Journal of Geophysical Research: Space Physics*, 119(11):9174–9191, nov 2014. doi:10.1002/2014ja020504.

[104] Tullio Basaglia, ZW Bell, PV Dressendorfer, A Larkin, and MG Pia. Writing software or writing scientific articles? In *Nuclear Science Symposium Conference Record, 2007. NSS'07. IEEE*, volume 1, pages 219–226. IEEE, 2007.

[105] M. S. Briggs, G. J. Fishman, V. Connaughton, P. N. Bhat, W. S. Paciesas, R. D. Preece, C. Wilson-Hodge, V. L. Chaplin, R. M. Kippen, A. von Kienlin, C. A. Meegan, E. Bissaldi, J. R. Dwyer, D. M. Smith, R. H. Holzworth, J. E. Grove, and A. Chekhtman. First results on terrestrial gamma ray flashes from the Fermi Gamma-ray Burst Monitor. *Journal of Geophysical Research (Space Physics)*, 115:A07323, July 2010. doi:10.1029/2009JA015242.

[106] Martino Marisaldi, Fabio Fuschino, Carlotta Pittori, Francesco Verrecchia, Paolo Giommi, Marco Tavani, Stefano Dietrich, Colin Price, Andrea Argan, Claudio Labanti, et al. The first agile low-energy (¡ 30 mev) terrestrial gamma-ray flashes catalog. In *EGU General Assembly Conference Abstracts*, volume 16, 2014.

[107] Savely G Karshenboim. Precision study of positronium: Testing bound state qed theory. *International Journal of Modern Physics A*, 19(23): 3879–3896, 2004.

[108] Dermott E. Cullen, John H. Hubbell, and Lynn Kissel. Epdl97: the evaluated photon data library, '97 version. *Accessed*, 1997.

[109] S. T. Perkins, D. E. Cullen, and S. M. Seltzer. Tables and graphs of electron-interaction cross sections from 10 eV to 100 GeV derived from the LLNL Evaluated Electron Data Library (EEDL), Z = 1 to 100. Technical report, November 1991.

[110] Hans Bethe and Walter Heitler. On the stopping of fast particles and on the creation of positive electrons. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 146(856): 83–112, 1934.

[111] Walter R Nelson, Hideo Hirayama, and David WO Rogers. Egs4 code system. Technical report, Stanford Linear Accelerator Center, Menlo Park, CA (USA), 1985.

[112] TT Böhlen, F Cerutti, MPW Chin, A Fassò, A Ferrari, PG Ortega, A Mairani, PR Sala, G Smirnov, and V Vlachoudis. The fluka code: developments and challenges for high energy and medical applications. *Nuclear Data Sheets*, 120:211–214, 2014.

[113] F. Salvat, J. M. Fernández-Varea, and Josep Sempau. *PENELOPE-2011: A Code System for Monte Carlo Simulation of Electron and Photon Transport.* 2011.

[114] Chao Li, Ute Ebert, and Willem Hundsdorfer. 3d hybrid computations for streamer discharges and production of runaway electrons. *Journal of Physics D: Applied Physics*, 42(20):202003, 2009. doi:10.1088/0022-3727/42/20/202003.

[115] Casper Rutjes, David Sarria, Alexander Broberg Skeltved, Alejandro Luque, Gabriel Diniz, Nikolai Østgaard, and Ute Ebert. Evaluation of monte carlo tools for high energy atmospheric physics. *Geoscientific Model Development*, 9(11):3961, 2016. doi:10.5194/gmd-9-3961-2016.

[116] C. T. R. Wilson. The Acceleration of β-particles in Strong Electric Fields such as those of Thunderclouds. *Proceedings of the Cambridge Philosophical Society*, 22:534, 1925. doi:10.1017/S0305004100003236.

[117] E. R. Williams. Origin and context of C. T. R. Wilson's ideas on electron runaway in thunderclouds. *Journal of Geophysical Research (Space Physics)*, 115:A00E50, August 2010. doi:10.1029/2009JA014581.

[118] Tatsuo Torii, Minoru Takeishi, and Teruo Hosono. Observation of gamma-ray dose increase associated with winter thunderstorm and lightning activity. *Journal of Geophysical Research: Atmospheres*, 107(D17):4324, 2002. doi:10.1029/2001JD000938.

[119] P. Kochkin, A. P. J. van Deursen, M. Marisaldi, A. Ursi, A. I. de Boer, M. Bardet, C. Allasia, J.-F. Boissin, F. Flourens, and N. Østgaard. In-flight observation of gamma-ray glows by ildas. *Journal of Geophysical Research: Atmospheres*, 2017. doi:10.1002/2017JD027405.

[120] C. Rutjes, G. Diniz, I. S. Ferreira, and U. Ebert. Tgf afterglows: a new radiation mechanism from thunderstorms. *Geophysical Research Letters*, oct 2017. doi:10.1002/2017GL075552.

[121] Gregory S Bowers, David Miles Smith, GF Martinez-McKinney, M Kamogawa, SA Cummer, JR Dwyer, D Wang, M Stock, and Z Kawasaki. Gamma ray signatures of neutrons from a terrestrial gamma ray flash. *Geophysical Research Letters*, 2017.

[122] Enoto Teruaki, Yuuki Wada, et al. Photonuclear reactions triggered by lightning discharge. *Nature*, 551(481), 2017. doi:10.1038/nature24630.

[123] N. G. Lehtinen, T. F. Bell, and U. S. Inan. Monte Carlo simulation of runaway MeV electron breakdown with application to red sprites and terrestrial gamma ray flashes. *jgr*, 104:24699–24712, November 1999. doi:10.1029/1999JA900335.

[124] Chang Liu, Dylan P Brennan, Amitava Bhattacharjee, and Allen H Boozer. Adjoint fokker-planck equation and runaway electron dynamics. *Physics of Plasmas*, 23(1):010702, 2016.

[125] Olivier Chanrion, Z Bonaventura, Anne Bourdon, and Torsten Neubert. Influence of the angular scattering of electrons on the runaway threshold in air. *Plasma Physics and Controlled Fusion*, 58(4):044001, 2016.

[126] Christoph Köhn, Gabriel Diniz, and Muhsin N Harakeh. Production mechanisms of leptons, photons, and hadrons and their possible feedback close to lightning leaders. *Journal of Geophysical Research: Atmospheres*, 122 (2):1365–1383, 2017.

[127] J. R. Dwyer. Source mechanisms of terrestrial gamma-ray flashes. *Journal of Geophysical Research: Atmospheres*, 113(D10):n/a–n/a, 2008. doi:10.1029/2007JD009248. D10103.

[128] Lev Davidovich Landau, JS Bell, MJ Kearsley, LP Pitaevskii, EM Lifshitz, and JB Sykes. *Electrodynamics of continuous media*, volume 8. elsevier, 2013.

[129] S. Celestin and V. P. Pasko. Soft collisions in relativistic runaway electron avalanches. *Journal of Physics D: Applied Physics*, 43:315206, 2010. doi:10.1088/0022-3727/43/31/315206.

[130] BE Carlson, Nikolai G Lehtinen, and Umran S Inan. Neutron production in terrestrial gamma ray flashes. *Journal of Geophysical Research: Space Physics*, 115(A4):A00E19, 2010. doi:10.1029/2009JA014696.

[131] D. Sarria, F. Lebrun, P.-L. Blelly, R. Chipaux, P. Laurent, J.-A. Sauvaud, L. Prech, P. Devoto, D. Pailot, J.-P. Baronick, and M. Lindsey-Clark. Taranis xgre and idee detection capability of terrestrial gamma-ray flashes and associated electron beams. *Geoscientific Instrumentation, Methods and Data Systems*, 6(2):239–256, 2017. doi:10.5194/gi-6-239-2017.

[132] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce Dubois, M. Asai, G. Barrand, R. Capra, S. Chauvie, R. Chytracek, G. A. P. Cirrone, G. Cooperman, G. Cosmo, G. Cuttone, G. G. Daquino, M. Donszelmann, M. Dressel, G. Folger, F. Foppiano, J. Generowicz, V. Grichine, S. Guatelli, P. Gumplinger, A. Heikkinen, I. Hrivnacova, A. Howard, S. Incerti, V. Ivanchenko, T. Johnson, F. Jones, T. Koi, R. Kokoulin, M. Kossov, H. Kurashige, V. Lara, S. Larsson, F. Lei, O. Link, F. Longo, M. Maire, A. Mantero, B. Mascialino, I. McLaren, P. Mendez Lorenzo, K. Minamimoto, K. Murakami, P. Nieminen, L. Pandola, S. Parlati, L. Peralta, J. Perl, A. Pfeiffer, M. G. Pia, A. Ribon, P. Rodrigues, G. Russo, S. Sadilov, G. Santin, T. Sasaki, D. Smith, N. Starkov, S. Tanaka, E. Tcherniaev, B. Tome, A. Trindade, P. Truscott, L. Urban, M. Verderi, A. Walkden, J. P. Wellisch, D. C. Williams, D. Wright, and H. Yoshida. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1):270–278, Feb 2006. doi:10.1109/TNS.2006.869826.

[133] J. Allison, K. Amako, J. Apostolakis, et al. Recent developments in geant4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 835:186 – 225, 2016. doi:https://doi.org/10.1016/j.nima.2016.06.125.

[134] E. S. Cramer, J. R. Dwyer, and H. K. Rassoul. Magnetic field modification to the relativistic runaway electron avalanche length. *Journal of Geophysical Research (Space Physics)*, 121:11, November 2016. doi:10.1002/2016JA022891.

[135] Enrico Camporeale, Ashutosh Agnihotri, and Casper Rutjes. Adaptive selection of sampling points for uncertainty quantification. *International Journal for Uncertainty Quantification*, 7(4), 2017.

[136] Dongbin Xiu. Fast numerical methods for stochastic computations: a review. *Communications in computational physics*, 5(2-4):242–272, 2009.

[137] Michael Eldred and John Burkardt. Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification, jan 2009.

[138] G Onorato, GJA Loeven, G Ghorbaniasl, H Bijl, and C Lacor. Comparison of intrusive and non-intrusive polynomial chaos methods for cfd applications in aeronautics. In *V European Conference on Computational Fluid Dynamics ECCOMAS, Lisbon, Portugal*, pages 14–17, 2010.

[139] Dongbin Xiu and George Em Karniadakis. The wiener–askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing*, 24(2):619–644, 2002.

[140] Thierry Crestaux, Olivier Le Maı̇tre, and Jean-Marc Martinez. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7):1161–1172, 2009.

[141] Kanali Togawa, Andrea Benigni, and Antonello Monti. Advantages and challenges of non-intrusive polynomial chaos theory. In *Proceedings of the 2011 Grand Challenges on Modeling and Simulation Conference*, pages 30–35. Society for Modeling & Simulation International, 2011.

[142] Michael Scott Eldred. Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design. *AIAA Paper*, 2274(2009):37, 2009.

[143] Ivo Babuska, Raúl Tempone, and Georgios E Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM Journal on Numerical Analysis*, 42(2):800–825, 2004.

[144] Mircea Grigoriu. *Stochastic systems: uncertainty quantification and propagation.* Springer Science & Business Media, 2012.

[145] Dongbin Xiu. *Numerical methods for stochastic computations: a spectral method approach.* Princeton university press, 2010.

[146] Olivier Le Maître and Omar M Knio. *Spectral methods for uncertainty quantification: with applications to computational fluid dynamics.* Springer Science & Business Media, 2010.

[147] Russel E Caflisch. Monte carlo and quasi-monte carlo methods. *Acta numerica*, 7:1–49, 1998.

[148] Malvin H Kalos and Paula A Whitlock. *Monte carlo methods.* John Wiley & Sons, 2008.

[149] Pierre L'Ecuyer and Art B Owen. *Monte Carlo and Quasi-Monte Carlo Methods 2008.* Springer, 2009.

[150] Harald Niederreiter. Quasi-monte carlo methods and pseudo-random numbers. *Bulletin of the American Mathematical Society*, 84(6):957–1041, 1978.

[151] Malvin H Kalos and Paula A Whitlock. Pseudorandom numbers. *Monte Carlo Methods, Second Edition*, pages 179–197, 2008.

[152] Richard Bellman. *Dynamic programming.* Courier Corporation, 2013.

[153] Piotr Indyk and Rajeev Motwani. Approximate nearest neighbors: towards removing the curse of dimensionality. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 604–613. ACM, 1998.

[154] Frances Y Kuo and Ian H Sloan. Lifting the curse of dimensionality. *Notices of the AMS*, 52(11):1320–1328, 2005.

[155] Dongbin Xiu and Jan S Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM Journal on Scientific Computing*, 27(3):1118–1139, 2005.

[156] Jasmine Foo and George Em Karniadakis. Multi-element probabilistic collocation method in high dimensions. *Journal of Computational Physics*, 229(5):1536–1557, 2010.

[157] Ivo Babuška, Fabio Nobile, and Raul Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.

[158] Peng Wang and Daniel M Tartakovsky. Uncertainty quantification in kinematic-wave models. *Journal of computational Physics*, 231(23):7868–7880, 2012.

[159] P. Wang, D. M. Tartakovsky, K. D. Jarman, and A. M. Tartakovsky. Cdf solutions of buckley–leverett equation with uncertain parameters. *Multiscale Modeling & Simulation*, 11(1):118–133, jan 2013. doi:10.1137/120865574.

[160] Sergey Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. In *Soviet Math. Dokl.*, volume 4, pages 240–243, 1963.

[161] Baskar Ganapathysubramanian and Nicholas Zabaras. Sparse grid collocation schemes for stochastic natural convection problems. *Journal of Computational Physics*, 225(1):652–685, jul 2007. doi:10.1016/j.jcp.2006.12.014.

[162] Marcel Bieri, Roman Andreev, and Christoph Schwab. Sparse tensor discretization of elliptic spdes. *SIAM Journal on Scientific Computing*, 31 (6):4281–4304, 2009.

[163] Xiang Ma and Nicholas Zabaras. An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. *Journal of Computational Physics*, 228(8):3084–3113, 2009.

[164] Fabio Nobile, Raúl Tempone, and Clayton G Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.

[165] John D Jakeman and Stephen G Roberts. Local and dimension adaptive stochastic collocation for uncertainty quantification. In *Sparse grids and applications*, pages 181–203. Springer, 2012.

[166] Ngoc-Hien Nguyen, Karen Willcox, and Boo Cheong Khoo. Model order reduction for stochastic optimal control. In *ASME 2012 11th Biennial Conference on Engineering Systems Design and Analysis*, pages 2–4. American Society of Mechanical Engineers, 2012.

[167] Dirk P Laurie. Computation of gauss-type quadrature formulas. *Journal of Computational and Applied Mathematics*, 127(1):201–217, 2001.

[168] Walter Gautschi. On the construction of gaussian quadrature rules from modified moments. *Mathematics of Computation*, 24(110):245–260, 1970.

[169] Jörg Waldvogel. Fast construction of the fejér and clenshaw–curtis quadrature rules. *BIT Numerical Mathematics*, 46(1):195–202, 2006.

[170] Jean-Paul Berrut and Lloyd N Trefethen. Barycentric lagrange interpolation. *Siam Review*, 46(3):501–517, 2004.

[171] Lloyd N Trefethen. *Spectral methods in MATLAB*. SIAM, 2000.

[172] James F Epperson. On the runge example. *Amer. Math. Monthly*, 94(4): 329–341, 1987.

[173] Xiang Ma and Nicholas Zabaras. An efficient bayesian inference approach to inverse problems based on an adaptive sparse grid collocation method. *Inverse Problems*, 25(3):035013, 2009.

[174] Jeroen AS Witteveen and Gianluca Iaccarino. Refinement criteria for simplex stochastic collocation with local extremum diminishing robustness. *SIAM Journal on Scientific Computing*, 34(3):A1522–A1543, 2012.

[175] John P Boyd. *Chebyshev and Fourier spectral methods*. Courier Corporation, 2001.

[176] Fred Brauer, Carlos Castillo-Chavez, and Carlos Castillo-Chavez. *Mathematical models in population biology and epidemiology*, volume 40. Springer, 2001.

[177] M Henry H Stevens. Lotka–volterra interspecific competition. In *A Primer of Ecology with R*, pages 135–159. Springer, 2009.

[178] Peter J Wangersky. Lotka-volterra population models. *Annual Review of Ecology and Systematics*, 9(1):189–218, 1978.

[179] Mohammad B Dadfar, James Geer, and Carl M Andersen. Perturbation analysis of the limit cycle of the free van der pol equation. *SIAM Journal on Applied Mathematics*, 44(5):881–895, 1984.

[180] Ruoshi Yuan, Xinan Wang, Yian Ma, Bo Yuan, and Ping Ao. Exploring a noisy van der pol type oscillator with a stochastic approach. *Physical Review E*, 87(6):062109, 2013.

[181] Liping Liu, Earl H Dowell, and Kenneth C Hall. A novel harmonic balance analysis for the van der pol oscillator. *International Journal of Non-Linear Mechanics*, 42(1):2–12, 2007.

[182] Gregory E Fasshauer and Jack G Zhang. On choosing "optimal" shape parameters for rbf approximation. *Numerical Algorithms*, 45(1):345–368, 2007.

[183] Jae-Hun Jung and Vincent R Durante. An iterative adaptive multiquadric radial basis function method for the detection of local jump discontinuities. *Applied Numerical Mathematics*, 59(7):1449–1466, 2009.

[184] JG Wang and GR Liu. On the optimal shape parameters of radial basis functions used for 2-d meshless methods. *Computer methods in applied mechanics and engineering*, 191(23):2611–2630, 2002.

[185] Joseph R Dwyer and Martin A Uman. The physics of lightning. *Physics Reports*, 534(4):147–241, 2014.

[186] Maribeth Stolzenburg and Thomas C Marshall. *Electric field and charge structure in lightning-producing clouds*. Springer, 2009.

[187] L. M. Coleman and J. R. Dwyer. Propagation speed of runaway electron avalanches. *Geophysical Research Letters*, 33:L11810, June 2006. doi:10.1029/2006GL025863.

[188] Danyal Petersen, Matthew Bailey, John Hallett, and William Beasley. Laboratory investigation of corona initiation by ice crystals and its importance to lightning. *Quarterly Journal of the Royal Meteorological Society*, 141 (689):1283–1293, 2015.

[189] A. B. Sun, J. Teunissen, and U. Ebert. Why isolated streamer discharges hardly exist above the breakdown field in atmospheric air. *Geophysical Research Letters*, 40(10):2417, 2013. doi:10.1002/grl.50457.

[190] Jannis Teunissen, Anbang Sun, and Ute Ebert. A time scale for electrical screening in pulsed gas discharges. *Journal of Physics D: Applied Physics*, 47(36):365203, 2014. doi:10.1088/0022-3727/47/36/365203.

[191] ML Huertas, J Fontan, and J Gonzalez. Evolution times of tropospheric negative ions. *Atmospheric Environment (1967)*, 12(12):2351–2362, 1978.

[192] I Gallimberti. The mechanism of the long spark formation. *Journal de Physique Colloques*, 40(C7):193, 1979. doi:10.1051/jphyscol:19797440.

[193] A. V. Gurevich and A. N. Karashtin. Runaway breakdown and hydrometeors in lightning initiation. *Physical Review Letters*, 110:185005, May 2013. doi:10.1103/PhysRevLett.110.185005.

[194] Keith A Olive, Particle Data Group, et al. Review of particle physics. *Chinese physics C*, 38(9):090001, 2014.

[195] Earle Williams. Charge structure and geographical variation of thunderclouds. *The Lightning Flash; Cooray, V., Ed.; The Institution of Engineering and Technology: London, UK*, pages 1–14, 2014.

[196] G. J. M. Hagelaar and L. C. Pitchford. Solving the Boltzmann equation to obtain electron transport coefficients and rate coefficients for fluid models. *Plasma Sour. Sci. Technol.*, 14:722, 2005. doi:10.1088/0963-0252/14/4/011.

[197] Dennis Lamb and Johannes Verlinde. *Physics and chemistry of clouds.* Cambridge University Press, 2011.

[198] G. Wormeester, S. Pancheshnyi, A. Luque, S. Nijdam, and U. Ebert. Probing photo-ionization: simulations of positive streamers in varying N_2 : O_2-mixtures. *Journal of Physics D: Applied Physics*, 43:505201, 2010. doi:10.1088/0022-3727/43/50/505201.

[199] A. Luque, U. Ebert, C. Montijn, and W. Hundsdorfer. Photoionization in negative streamers: Fast computations and two propagation modes. *Applied Physics Letters*, 90(8):081501, 2007. doi:10.1063/1.2435934.

[200] J. R. Dwyer and L. P. Babich. Low-energy electron production by relativistic runaway electron avalanches in air. *J. Geophys. Res. (Space Phys)*, 116(A9):A09301, 2011. doi:10.1029/2011JA016494. A09301.

[201] B. J. Mason. On the generation of charge associated with graupel formation in thunderstorms. *Quarterly Journal of the Royal Meteorological Society*, 79(342):501–509, 1953. doi:10.1002/qj.49707934206.

[202] William Rison, Paul R Krehbiel, Michael G Stock, Harald E Edens, Xuan-Min Shao, Ronald J Thomas, Mark A Stanley, and Yang Zhang. Observations of narrow bipolar events reveal how lightning is initiated in thunderstorms. *Nature communications*, 7, 2016. doi:10.1038/ncomms10721.

[203] Thomas C. Marshall, Michael P. McCarthy, and W. David Rust. Electric field magnitudes and lightning initiation in thunderstorms. *Journal of Geophysical Research: Atmospheres*, 100(D4):7097–7103, 1995. doi:10.1029/95JD00020.

[204] T. C. Marshall, M. Stolzenburg, C. R. Maggio, L. M. Coleman, P. R. Krehbiel, T. Hamlin, R. J. Thomas, and W. Rison. Observed electric fields associated with lightning initiation. *Geophysical Research Letters*, 32(3):L03813, 2005. doi:10.1029/2004GL021802. L03813.

[205] L.P. Babich, E.I. Bochkov, and T. Neubert. The role of charged ice hydrometeors in lightning initiation. *Journal of Atmospheric and Solar-Terrestrial Physics*, 154:43 – 46, February 2017. doi:https://doi.org/10.1016/j.jastp.2016.12.010.

[206] Ilya G. Usoskin, Laurent Desorgher, Peter Velinov, Marisa Storini, Erwin O. Flückiger, Rolf Bütikofer, and Gennady A. Kovaltsov. Ionization of the earth's atmosphere by solar and galactic cosmic rays. *Acta Geophysica*, 57(1):88–101, Mar 2009. doi:10.2478/s11600-008-0019-9.

[207] Sergey Pancheshnyi. Effective ionization rate in nitrogen–oxygen mixtures. *Journal of Physics D: Applied Physics*, 46(15):155201, 2013. doi:10.1088/0022-3727/46/15/155201.

[208] C Patrignani, P Richardson, OV Zenin, R-Y Zhu, A Vogt, S Pagan Griso, L Garren, DE Groom, M Karliner, DM Asner, et al. Review of particle physics, 2016-2017. *Chinese Physics C*, 40:100001, 2016.

[209] William P. Jesse and John Sadauskis. Ionization in pure gases and the average energy to make an ion pair for alpha and beta particles. *Physical Review*, 97:1668–1670, Mar 1955. doi:10.1103/PhysRev.97.1668.

[210] Arthur Cole. Absorption of 20-ev to 50,000-ev electron beams in air and plastic. *Radiation Research*, 38(1):7–33, 1969. doi:10.2307/3572707.

[211] Glenn F Knoll. *Radiation detection and measurement.* John Wiley & Sons, 2010.

[212] Ilya G. Usoskin and Gennady A. Kovaltsov. Cosmic ray induced ionization in the atmosphere: Full modeling and practical applications. *Journal of Geophysical Research: Atmospheres*, 111(D21):n/a–n/a, 2006. doi:10.1029/2006JD007150. D21206.

[213] Koichi Kamata and Jun Nishimura. The lateral and the angular structure functions of electron showers. *Progress of Theoretical Physics Supplement*, 6:93, 1958. doi:10.1143/PTPS.6.93.

[214] Kenneth Greisen. Cosmic ray showers. *Annual Review of Nuclear Science*, 10(1):63–108, 1960. doi:10.1146/annurev.ns.10.120160.000431.

[215] M F Bourdeau, J N Capdevielle, and J Procureur. The lateral age parameter in extensive air showers. *Journal of Physics G: Nuclear Physics*, 6(7): 901, 1980. doi:10.1088/0305-4616/6/7/013.

[216] W.D. Apel, A.F. Badea, K. Bekk, A. Bercuci, J. Blmer, H. Bozdog, I.M. Brancus, A. Chilingarian, K. Daumiller, P. Doll, R. Engel, J. Engler, H.J. Gils, R. Glasstetter, A. Haungs, D. Heck, J.R. Hrandel, K.-H. Kampert, H.O. Klages, G. Maier, H.J. Mathes, H.J. Mayer, J. Milke, M. Mller, R. Obenland, J. Oehlschlger, S. Ostapchenko, M. Petcu, H. Rebel, A. Risse, M. Risse, M. Roth, G. Schatz, H. Schieler, H. Ulrich, J. van Buren, A. Vardanyan, A. Weindl, J. Wochele, and J. Zabierowski. Comparison of measured and simulated lateral distributions for electrons and muons with kascade. *Astroparticle Physics*, 24(6):467 – 483, 2006. doi:https://doi.org/10.1016/j.astropartphys.2005.10.001.

[217] S Ostapchenko. Qgsjet-ii: towards reliable description of very high energy hadronic interactions. *Nuclear Physics B-Proceedings Supplements*, 151 (1):143–146, 2006. doi:10.1016/j.nuclphysbps.2005.07.026.

[218] Giuseppe Battistoni, F Cerutti, A Fasso, A Ferrari, S Muraro, J Ranft, S Roesler, and PR Sala. The fluka code: Description and benchmarking. In *Hadronic Shower Simulation Workshop(AIP Conference Proceedings Volume 896)*, volume 896, pages 31–49, 2007. doi:10.1063/1.2720455.

[219] M. Stolzenburg, W. D. Rust, and T. C. Marshall. Electrical structure in thunderstorm convective regions: 3. Synthesis. *Journal of Geophysical Research*, 103:14, 1998. doi:10.1029/97JD03545.

[220] Dennis J. Boccippio, Steven J. Goodman, and Stan Heckman. Regional differences in tropical lightning distributions. *Journal of Applied Meteorology*, 39(12):2231–2248, 2000. doi:10.1175/1520-0450(2001)040<2231:RDITLD>2.0.CO;2.

[221] A. A. Sin'kevich and Yu. A. Dovgalyuk. Corona discharge in clouds. *Radiophysics and Quantum Electronics*, 56(11):818–828, Apr 2014. doi:10.1007/s11141-014-9484-y.

[222] Qiheng Cai, Jaroslav Jánský, and Victor P. Pasko. Initiation of positive streamer corona in low thundercloud fields. *Geophysical Research Letters*, 44(11):5758–5765, 2017. doi:10.1002/2017GL073107. 2017GL073107.

[223] Eric Defer, Patrice Blanchet, Claire Théry, Pierre Laroche, James E. Dye, Martin Venticinque, and Kenneth L. Cummins. Lightning activ-

ity for the july 10, 1996, storm during the stratosphere-troposphere experiment: Radiation, aerosol, and ozone-a (sterao-a) experiment. *Journal of Geophysical Research: Atmospheres*, 106(D10):10151–10172, 2001. doi:10.1029/2000JD900849.

[224] Eric C. Bruning, W. David Rust, Donald R. MacGorman, Michael I. Biggerstaff, and Terry J. Schuur. Formation of charge structures in a supercell. *Monthly Weather Review*, 138(10):3740–3761, 2010. doi:10.1175/2010MWR3160.1.

[225] L. M. Libby and H. R. Lukens. Production of radiocarbon in tree rings by lightning bolts. *Journal of Geophysical Research*, 78:5902, 1973. doi:10.1029/JB078i026p05902.

[226] Casper Rutjes and Ute Ebert. A natural neutron source. *Physics World*, 2017.

[227] Leonid P Babich and Robert A Roussel-Dupré. Origin of neutron flux increases observed in correlation with lightning. *Journal of Geophysical Research: Atmospheres*, 112(D13):D13303, 2007. doi:10.1029/2006JD008340.

[228] Leonid P Babich, Evgenii I Bochkov, Igor M Kutsyk, and Robert A Roussel-Dupré. Localization of the source of terrestrial neutron bursts detected in thunderstorm atmosphere. *Journal of Geophysical Research: Space Physics*, 115(A5):A00E28, 2010. doi:10.1029/2009JA014750.

[229] GC Baldwin and GS Klaiber. Photo-fission in heavy elements. *Physical Review*, 71(1):3, 1947.

[230] AV Varlamov, VV Varlamov, DS Rudenko, and ME Stepanov. Atlas of giant dipole resonances. *Parameters and Graphs of Photonuclear Reaction Cross Sections. INDC (NDS)-394, IAEA NDS, Vienna, Austria*, pages 1–311, 1999.

[231] Samuel S Dietrich and Barry L Berman. Atlas of photoneutron cross sections obtained with monoenergetic photons. *Atomic Data and Nuclear Data Tables*, 38(2):199–338, 1988.

[232] A Shyam and TC Kaushik. Observation of neutron bursts associated with atmospheric lightning discharge. *Journal of Geophysical Research: Space Physics*, 104(A4):6867–6869, 1999. doi:10.1029/98JA02683.

[233] I. M. Martin and M. A. Alves. Observation of a possible neutron burst associated with a lightning discharge? *J. Geophys. Res. (Space Phys)*, 115:A00E11, 2010. doi:10.1029/2009JA014498.

[234] Sergei Anatol'evich Starodubtsev, VI Kozlov, AA Toropov, VA Mulla-yarov, Vladislav Georgievich Grigor'ev, and AV Moiseev. First experimental observations of neutron bursts under thunderstorm clouds near sea level. *Soviet Journal of Experimental and Theoretical Physics Letters*, 96 (3):188–191, 2012. doi:10.1134/S0021364012150106.

[235] AV Gurevich, VP Antonova, AP Chubenko, AN Karashtin, GG Mitko, MO Ptitsyn, VA Ryabov, AL Shepetov, Yu V Shlyugaev, LI Vildanova, et al. Strong flux of low-energy neutrons produced by thunderstorms. *Physical review letters*, 108(12):125001, 2012. doi:10.1103/PhysRevLett.108.125001.

[236] H. Tsuchiya, K. Hibino, K. Kawata, N. Hotta, N. Tateyama, M. Ohnishi, M. Takita, D. Chen, J. Huang, M. Miyasaka, I. Kondo, E. Takahashi, S. Shimoda, Y. Yamada, H. Lu, J. L. Zhang, X. X. Yu, Y. H. Tan, S. M. Nie, K. Munakata, T. Enoto, and K. Makishima. Observation of thundercloud-related gamma rays and neutrons in Tibet. *Phys. Rev. D*, 85(9):092006, 2012. doi:10.1103/PhysRevD.85.092006.

[237] Ashot Chilingarian, Nikolaj Bostanjyan, and L Vanyan. Neutron bursts associated with thunderstorms. *Physical review D*, 85(8):085017, 2012. doi:10.1103/PhysRevD.85.085017.

[238] VI Kozlov, VA Mullayarov, SA Starodubtsev, and AA Toropov. Neutron bursts during cloud-to-ground discharges of lightning. *Bulletin of the Russian Academy of Sciences: Physics*, 77(5):584–586, 2013.

[239] J. D. Hill, M. A. Uman, and D. M. Jordan. High-speed video observations of a lightning stepped leader. *Journal of Geophysical Research: Atmospheres*, 116:D16117, 2011. doi:10.1029/2011JD015818.

[240] Leandro ZS Campos, Marcelo MF Saba, Tom A Warner, Osmar Pinto, E Philip Krider, and Richard E Orville. High-speed video observations of natural cloud-to-ground lightning leaders–a statistical analysis. *Atmospheric research*, 135:285–305, 2014.

[241] Wei Xu, Sebastien Celestin, and Victor P. Pasko. Source altitudes of terrestrial gamma-ray flashes produced by lightning leaders. *Geophysical Research Letters*, 39(8):L08801, 2012. doi:10.1029/2012GL051351.

[242] US Coesa. Standard atmosphere, 1976. *US Government Printing Office, Washington, DC*, 1976.

[243] P Oblozinskỳ. Handbook of photonuclear data for applications: Cross sections and spectra. *International Atomic Energy Association report IAEA-TECDOC-1178, Vienna, Austria*, 2000.

[244] M. S. Briggs. Fermi GBM Observations of Terrestrial Gamma-ray Flashes. In V. Florinski, J. Heerikhuisen, G. P. Zank, and D. L. Gallagher, editors, *American Institute of Physics Conference Series*, volume 1366 of *American Institute of Physics Conference Series*, page 47, 2011. doi:10.1063/1.3625588.

[245] C Khn, O Chanrion, and T Neubert. The influence of bremsstrahlung on electric discharge streamers in n 2 , o 2 gas mixtures. *Plasma Sources Science and Technology*, 26(1):015006, 2017. URL `http://stacks.iop.org/0963-0252/26/i=1/a=015006`.

[246] Cung Vuong Nguyen, Alexander PJ van Deursen, and Ute Ebert. Multiple x-ray bursts from long discharges in air. *Journal of Physics D: Applied Physics*, 41(23):234012, 2008. doi:10.1088/0022-3727/41/23/234012.

[247] LS Bratolyubova-Tsulukidze, EA Grachev, OR Grigoryan, VE Kunitsyn, BM Kuzhevskij, DS Lysakov, O Yu Nechaev, and ME Usanova. Thunderstorms as the probable reason of high background neutron fluxes at l¡ 1.2. *Advances in Space Research*, 34(8):1815–1818, 2004.

[248] AV Gurevich, VP Antonova, AP Chubenko, AN Karashtin, ON Kryakunova, V Yu Lutsenko, GG Mitko, VV Piskal, MO Ptitsyn, VA Ryabov, et al. The time structure of neutron emission during atmospheric discharge. *Atmospheric Research*, 164:339–346, 2015.

[249] Robert L. Fleischer, J. A. Plumer, and K. Crouch. Are neutrons generated by lightning? *Journal of Geophysical Research*, 79(33):5013–5017, 1974. doi:10.1029/JC079i033p05013.

[250] Leonid Petrovich Babich. Generation of neutrons in giant upward atmospheric discharges. *JETP letters*, 84(6):285–288, 2006.

[251] L. P. Babich, A. Y. Kudryavtsev, M. L. Kudryavtseva, and I. M. Kutsyk. Neutron generation by gigantic upward atmospheric discharges. *Doklady Earth Sciences*, 415:885, 2007. doi:10.1134/S1028334X07060128.

[252] Leonid Petrovich Babich, A Yu Kudryavtsev, ML Kudryavtseva, and IM Kutsyk. Terrestrial gamma-ray flashes and neutron pulses from direct simulations of gigantic upward atmospheric discharge. *JETP Letters*, 85(10):483–487, 2007. doi:10.1134/S0021364007100037.

[253] LP Babich, A Yu Kudryavtsev, ML Kudryavtseva, and IM Kutsyk. Atmospheric gamma-ray and neutron flashes. *Journal of Experimental and Theoretical Physics*, 106(1):65–76, 2008. doi:10.1134/S1063776108010056.

[254] A. Drozdov, A. Grigoriev, and Y. Malyshkin. Assessment of thunderstorm neutron radiation environment at altitudes of aviation flights. *J. Geophys. Res. (Space Phys)*, 118:947, 2013. doi:10.1029/2012JA018302.

[255] AV Gurevich, AP Chubenko, AN Karashtin, GG Mitko, AS Naumov, MO Ptitsyn, VA Ryabov, AL Shepetov, Yu V Shlyugaev, LI Vildanova, et al. Gamma-ray emission from thunderstorm discharges. *Physics Letters A*, 375(15):1619–1625, 2011.

[256] Thomas Gjesteland, Nikolai Østgaard, Sante Laviola, MM Miglietta, Enrico Arnone, Martino Marisaldi, Fabio Fuschino, Andrew B Collier, F Fabró, and J Montanya. Observation of intrinsically bright terrestrial gamma ray flashes from the mediterranean basin. *Journal of Geophysical Research: Atmospheres*, 120(23), 2015.

[257] J. M. Blatt and V. F. Weisskopf. *Theoretical Nuclear Physics*. Springer-Verlag New York, 1979.

[258] G. A. Bazilevskaya, I. G. Usoskin, E. O. Flückiger, R. G. Harrison, L. Desorgher, R. Bütikofer, M. B. Krainev, V. S. Makhmutov, Y. I. Stozhkov, A. K. Svirzhevskaya, N. S. Svirzhevsky, and G. A. Kovaltsov. Cosmic Ray Induced Ion Production in the Atmosphere. *Space Science Review*, 137: 149, 2008. doi:10.1007/s11214-008-9339-y.

[259] A. V. Agafonov, A. V. Bagulya, O. D. Dalkarov, M. A. Negodaev, A. V. Oginov, A. S. Rusetskiy, V. A. Ryabov, and K. V. Shpakov. Observation of neutron bursts produced by laboratory high-voltage atmospheric discharge. *Phys. Rev. Lett.*, 111:115003, Sep 2013. doi:10.1103/PhysRevLett.111.115003.

[260] A V Agafonov, V A Bogachenkov, A P Chubenko, A V Oginov, A A Rodionov, A S Rusetskiy, V A Ryabov, A L Shepetov, and K V Shpakov. Observation of hard radiations in a laboratory atmospheric high-voltage discharge. *Journal of Physics D: Applied Physics*, 50(16):165202, 2017. URL http://stacks.iop.org/0022-3727/50/i=16/a=165202.

Curriculum Vitae

After my studies of Mathematics (BSc), Physics (BSc) and two-year research master Astrophysics, I started the traineeship Eerst de Klas (EDK); a two-year leadership program at de Baak. EDK provided the opportunity to further develop my interpersonal skills by combining training in the education and corporate sector. During the traineeship, I worked as a teacher and completed a master in Teaching Physics for Pre-Higher Education cum laude (Dutch: eerstegraads lesbevoegdheid). In the corporate sector, I participated in masterclasses and completed two major intern projects at Royal Schiphol Group and Royal Ahold.

In September 2013, I started a four-year physics and computer science PhD investigating lightning, that resulted in the dissertation in front of you, in the Multiscale Dynamics group of prof. dr. Ute Ebert at Centrum Wiskunde & Informatica (CWI). CWI is the Dutch national institute for mathematics and computer science. During my PhD I joined the Central Works Council (COR) right in the middle of the major organisational change of NWO, the Netherlands Organisation for Scientific Research. I was elected chair in 2017, for the last half year of my term. I received a PhD contract extension for these extra duties. In addition, in 2014, I co founded doorkennen; a company offering workshops, lectures and private courses.

Publications in international journals

1. Probing atmospheric electric fields in thunderstorms through radio emission from cosmic-ray induced air showers, P. Schellart, T.N.G. Trinh, S. Buitink, A. Corstanje, J.E. Enriquez, H. Falcke, J.R. Hörandel, A. Nelles, J.P. Rachen, L. Rossetto, O. Scholten, S. ter Veen, S. Thoudam, U. Ebert, C. Köhn, C. Rutjes, and the LOFAR collaboration, Phys Rev. Lett. 114, 165001 [5 pages] (2015).

2. Prediction of lightning inception by large ice particles and extensive air showers, A. Dubinova, C. Rutjes, U. Ebert, S. Buitink, O. Scholten, and T.N.G. Trinh, Phys. Rev. Lett. 115, 015002 [5 pages] (2015).

3. Influence of Atmospheric Electric Fields on Radio Emission from Cosmic-Ray Induced Air Showers, T.N.G. Trinh, O. Scholten, S. Buitink, A.M. van den Berg, A. Corstanje, U. Ebert, J.E. Enriquez, H. Falcke, J.R. Hörandel, C. Köhn, A. Nelles, J.P. Rachen, C. Rutjes, P. Schellart, S. Thoudam, K.D. de Vries, S. ter Veen, Phys. Rev. D 93, 023003 [17 pages] (2016).

4. Evaluation of Monte Carlo tools for high energy atmospheric physics, C. Rutjes, D. Sarria, A.B. Skeltved, A. Luque, G. Diniz, N. Ostgaard, U. Ebert, Geoscientific Model Development, 9, 3961-3974, 2016.

5. Thunderstorm electric fields probed by extensive air showers through their polarized radio emission, T.N.G. Trinh, O. Scholten, A. Bonardi, S. Buitink, A. Corstanje, U. Ebert, J.E. Enriquez, H. Falcke, J.R. Hörandel, B.M. Hare, P. Mitra, K. Mulrey, A. Nelles, J.P. Rachen, L. Rosetto, C. Rutjes, P. Schellart, S. Thoudam, S. ter Veen, T. Winchen, Phys. Rev. D 95, 083004 [10 pages] (2017).

6. Adaptive selection of sampling points for uncertainty quantification. E. Camporeale., A. Agnihotri, C. Rutjes. Int J Uncertain Quantif 7, issue 4 [22 pages] (2017).

7. TGF afterglows: a new radiation mechanism from thunderstorms. C. Rutjes, G. Diniz, I.S. Ferreira, U. Ebert. Geophys. Res. Lett. 44 2017GL075552 [11 pages] (2017).

Proceedings for international conferences

1. How do air showers of energetic cosmic particles interact with droplets in thunderclouds? [1 page], C. Rutjes, A. Dubinova, U. Ebert. Proceedings of COST TD1208 Annual meeting, Electrical discharges with liquids for future applications, Lisboa, March 2014.

2. A positive streamer discharge interaction with water droplets [1 page], A. Dubinova, C. Rutjes, U. Ebert. Proceedings of COST TD1208 Annual meeting, Electrical discharges with liquids for future applications, Lisboa, March 2014.

3. Proceedings of 2nd TEA-IS summer school, Collioure, France, June 2014: (1) Ionization due to extensive air showers in humid air [1 page], C. Rutjes, U. Ebert. (2) Effects of atmospheric electric fields on radio emission from air showers [2 pages], T. N. G. Trinh, O. Scholten, A.M. van den Berg, S. Buitink, U. Ebert, C. Rutjes, C. Köhn. (3) Positive streamer discharge inception from dielectrics [1 page], A. Dubinova, C. Rutjes, U. Ebert.

4. Proceedings of the 32nd Int. Conf. on Phenomena in Ionized Gases (32nd ICPIG) in Iasi, Rumania, July 2015: Lightning inception from ice particles and extensive air showers [2 pages], by A. Dubinova, C. Rutjes, U. Ebert, S. Buitink, O. Scholten, T.N.G. Trinh.

5. Proceedings of the 34th International Cosmic Ray Conference ICRC2015, The Hague, The Netherlands, Aug. 2015: Probing atmospheric electric fields in thunderstorms through radio emission from extensive air showers [8 pages], by T.N.G. Trinh, P. Schellart, S. Buitink, A. Corstanje, J.E. Enriquez, H. Falcke, J.R. Hörandel, A. Nelles, J.P. Rachen, L. Rossetto, O. Scholten, S. ter Veen, S. Thoudam, A.M. van den Berg, U. Ebert, C. Rutjes, Proceedings of Science (ICRC2015) 284.

6. Proceedings of the 9th ICRP, held together with the 68th Gaseous Electronics Conference GEC 2015 , Honolulu, Hawaii, Oct. 2015: Streamer discharge inception in a sub-breakdown electric field from a dielectric body with a frequency dependent dielectric permittivity [2 pages], by Anna Dubinova, Casper Rutjes and Ute Ebert.

7. Precision study of radio emission from air showers at LOFAR, O. Scholten, A. Bonardi, S. Buitink, A. Corstanje, U. Ebert, H. Falcke, J. Hörandel, P. Mitra, K. Mulrey, A. Nelles, J. Rachen, L. Rossetto, C. Rutjes, P. Schellart, S. Thoudam, G. Trinh, S. ter Veen, T. Winchen, Europhys. J. Web of Conferences 136, 02012 [4 pages] (2017).

8. Circular polarization of radio emission from air showers in thunderstorm conditions, T.N.G. Trinh, O. Scholten, A. Bonardi, S. Buitink, A. Corstanje, U. Ebert, J. E. Enriquez, H. Falcke, J.R. Hörandel, P. Mitra, K. Mulrey, A. Nelles, S. Thoudam, J.P. Rachen, L. Rossetto, C. Rutjes, P. Schellart, S. ter Veen, T. Winchen, Europhys. J. Web of Conferences 135, 03002 [4 pages] (2017).

9. Lightning Imaging with LOFAR, O. Scholten, Q.D. Hasankiadeh, F. Hendriks, D. Sardjan, G. Trinh, A. Bonard, S. Buitink, A. Corstanje, U. Ebert, H. Falcke, J. Hörandel, A. Nelles, J. Rachen, L. Rossetto, C. Rutjes, P. Schellart, S. Thoudam, S. ter Veen, T. Winchen, K. Mulrey, P. Mitra, J. Muller, R. Dina, W. Frieswijk, Europhys. J. Web of Conferences 135, 03003 [4 pages] (2017).

Publications for a general audience

- A natural neutron source. C. Rutjes, U. Ebert, Physics World Focus on: Neutron Science, October 2017, pp. 13-15.

- Onweer krijgt hulp uit de ruimte (Dutch). Dagblad Trouw, 20 juli 2017.

- Onweer: hoe veilig ben je in huis? (Dutch). Radar avrotros, June 2017.

- Hoe blijf je veilig tijdens onweer? (Dutch). `zorgnu.avrotros.nl`, June 2017

- Cosmic Lightning. Dutch Journal of Physics — DJoP 2016-1, Januari 2016.

- Kosmisch onweer (Dutch). Nederlands Tijdschrift voor Natuurkunde 82, Nummer 1 (januari), 22-26, 2016.

- Hoe sterrenexplosies helpen om onweer te begrijpen (Dutch). FOM Expres, juli 2015, jaargang 28, nummer 3, p. 18.

- Start of Lightning Explained: Hail and Cosmic Particles. ERCIM News Number 102, juli 2015.

- TU/e-professor helpt ontstaan van bliksem aantonen (Dutch). Technische Universiteit Eindhoven. Juli 2015.

- Mystery of how lightning forms. Daily Mail, Juli 2015.

- In het hart van de donderwolk (Dutch). NRC handelsblad, 1 mei 2015. NRC NEXT, 4 mei 2015.

- Cosmic rays reveal the secrets of thunderstorms. Nature — News, 23 april 2015.

Conference Talks

- Realistic 3D Particle Modelling of Discharge Inception near Ice Particles and other Dielectric Objects. The International Conference on Phenomena in Ionized Gases (ICPIG), Estoril, Portugal, July 2017. C. Rutjes, J. Teunisssen, Ute Ebert.

- Streamer inception from hydrometeors as a stochastic process with a particle-based model. EGU General Assembly, Vienna, Austria, April 2017. [invited] C. Rutjes, A. Dubinova, U. Ebert, J. Teunissen, S. Buitink, O. Scholten, T. N. G. Trinh.

- Simulating Streamer Inception From Hydrometeors as a Stochastic Process with a Particle-Based Model. AGU Fall Meeting, San Francisco, USA, December 2016. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh.

- Expected maximal electron densies from extended air showers for breakdown phenomena in thunderstorms. TEA-IS conference, Copenhagen, Denmark, June 2016. C. Rutjes, E. Camporeale, U. Ebert, S. Buitink, O. Scholten, G.T.N. Trinh.

- The origin of lightning. Scientific meeting CWI, Amsterdam, The Netherlands, May 2016. [invited] C. Rutjes.

- Avalanche-to-streamer transition near hydrometeors in thunderstorms. EGU General Assembly, Vienna, Austria, April 2016. C. Rutjes, E. Camporeale, U. Ebert, S. Buitink, O. Scholten, G.T.N. Trinh, J.A.S. Witteveen.

- Prediction of lightning inception by large ice particles and extensive air showers. Physics@FOM, Veldhoven, The Netherlands, January 2016. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh

- Cosmic Lightning. Astroparticules et Cosmologie, University Paris VII Diderot, Paris, France, October 2015. [invited] C. Rutjes.

- Modeling Discharge Inception in Thunderstorms. EGU General Assembly, Vienna, Austria, April 2015. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh.

- Modeling Discharge Inception in Thunderstorms. AGU Fall Meeting, San Francisco, USA, December 2014. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh.

- Free electrons from extensive air showers including interactions with hydrometeors. Workshop TEA-IS, Leiden, The Netherlands, October 2014. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh.

Conference Posters

- Evaluation of Monte Carlo tools for high energy atmospheric physics. AGU Fall Meeting, San Francisco, USA, December 2016. C. Rutjes, D. Sarria, A.B. Skeltved, A. Luque, G. Diniz, N. stgaard, U. Ebert.

- Evaluation of Monte Carlo tools for high energy atmospheric physics. TEA-IS conference, Copenhagen, Denmark, June 2016. C. Rutjes, D. Sarria, A.B. Skeltved, A. Luque, G. Diniz, N. stgaard, U. Ebert.

- Extreme atmospheric electron densities created by extensive air showers. EGU General Assembly, Vienna, Austria, April 2016. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh.

- Nanosecond enhancement of the atmospheric electron density by extensive air showers. 28th edition of the Symposium series on Plasma Physics and Radiation Technology, Lunteren, The Netherlands, March 2016. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh.

- Nanosecond enhancement of the atmospheric electron density by extensive air showers. AGU Fall Meeting, San Francisco, USA, December 2015. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh.

- Lightning inception from ice particles. Physics@FOM, Veldhoven, The Netherlands, January 2015. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh.

- Lightning inception from ice particles. International Low Temperature Plasma School, Bad Honnef, Germany, October 2015. C. Rutjes, A. Dubinova, U. Ebert, S. Buitink, O. Scholten, T. N. G. Trinh.

- Ionization due to extensive air showers in humid air. Thunderstorm Effects on the Atmosphere-Ionosphere System 2014 (TEA-IS), Collioure, France, June 2014. C. Rutjes, U. Ebert.

- How do air showers of energetic cosmic particles interact with droplets in thunderclouds? COST TD1208 Annual meeting: Electrical discharges with liquids for future applications, Lisboa, Portugal, March 2014. C. Rutjes, A. Dubinova, U. Ebert.

Acknowledgments

There are many people I would like to acknowledge who contributed to my dissertation and personal journey during my PhD. There are not enough words to describe my gratitude for the many different ways in which my family, friends and colleagues have supported me, thank you all.

Special thanks to Ashutosh Agnihotri, Gabriel Diniz, Anna Dubinova, David Sarria and Jannis Teunissen, for all our vivid discussions and fun about the PhD work and life itself. I will forever cherish these moments!

Special gratitude to my daily supervisor and promotor prof. dr. Ute Ebert. Ute, bedankt voor al jouw hulp, expertise, passie en gedrevenheid. Je deur stond altijd open, voor welk discussiepunt dan ook.

And last and foremost, to Laura Rutjes-Weurding, my wife, you know :)

Summary

This thesis has two research topics, 'the problem of lightning initiation' and 'high energy radiation from thunderstorms and lightning'.

Lightning is an electric discharge that develops in different phases. The last phase, called the return stroke, is observed as a flash and heard as thunder. In the first phase preliminary discharges try to start, which is far from trivial and sometimes referred to as the lightning initiation problem. Roughly half of this thesis is devoted to contributions to the lightning initiation problem. We prove that a streamer can start from an ice particle in a thundercloud electric field that is significantly below classical breakdown, with realistic initial conditions and dielectric properties of ice. We derive the interplay between hydrometer shape and thundercloud electric field strength, which shows that ellipsoidally shaped hydrometeors have an optimal aspect ratio for fixed length [chapter 10]. Recent observations predict that lightning starts by a volumetric system of streamers, which seems to be triggered simultaneously by an external event. A possible explanation is drawn from the hypothesis that a single extensive air shower initiates multiple independent streamers from hydrometeors in a coherent way. We thoroughly studied the probability distribution of relevant extensive air showers and derived with rare event analyses the expected availability of extreme electron seeds [chapter 11].

After the first initiation phase, lightning develops further and can produce flashes of high energy radiation, in so-called terrestrial gamma-ray flashes (TGFs), including the emission of neutrons. But also much longer lasting glows of high energy radiation have been observed from thunderclouds, called gamma-ray glows, including again the emission of neutrons. The presence of neutrons is stressed, as that indicates that the photon energy must have exceeded 10 MeV. Roughly half of this thesis is devoted to contributions to the understanding of high energy radiation from thunderstorms and lightning, and, in particular, of neutron emission and propagation. We present simulations of downward gamma-ray beams with energies between 10 and 30 MeV, separated into 4 energy intervals with 5 MeV of width, at varying source altitudes. The individual signature at sea level of each energy interval allows the extrapolation of our results to an arbitrary source spectrum and altitude [chapter 12]. Furthermore, these neutrons interact again with the air. We predict a new intermediate thunderstorm radia-

tion mechanism, which we call TGF afterglow, as it is caused by the capture of photonuclear neutrons produced by a TGF. TGF afterglows are milliseconds to seconds long; this duration is caused by the thermalization time of the intermediate neutrons [chapter 13].

The main body of this thesis has been organized into three distinct parts. The first part - not yet summarized - covers model development, reduction and evaluation. The chapters of the first part are created for predictions already summarized above, but apply to more general fields of study. In chapter 5, we revise the criterion for discharge inception, in particular for the avalanche-to-positive-streamer transition near a dielectric surface. In chapter 6, we report about a particle model to study the initial phase, specifically the stochastic nature of positive streamer inception near dielectrics and other surfaces. In chapter 7 and 8, we evaluate, compare and benchmark Monte Carlo methods in high atmospheric physics. And in chapter 9, we present a simple and robust strategy for the selection of sampling points to achieve the fastest possible convergence in the cumulative distribution function of a stochastic output of interest, particularly useful for computationally expensive models, such as the simulations in this thesis.

Nederlandse samenvatting (Dutch summary)

Het werk beschreven in dit proefschrift omvat twee onderzoeksonderwerpen, te weten: het starten van bliksem en energierijke straling afkomst van bliksem en onweerswolken. Binnen beide onderwerpen zijn er bijdragen geleverd aan zowel (algemene) methode ontwikkeling als ook (onderwerp specifieke) voorspellingen. Bijdragen van dit proefschrift zijn daarom onderverdeeld in drie delen. Het eerste deel gaat over modelontwikkeling, modelevaluatie en modelreductie. Het tweede deel gaat over het ontstaan van bliksem en het derde deel over straling afkomstig van bliksem en onweerswolken. Deze drie delen worden hieronder samengevat.

[Deel 1, hoofdstuk 5 t/m 9] Modelontwikkeling, modelevaluatie en modelreductie

Om de fenomenen beschreven in dit proefschrift (zoals bliksem) te bestuderen zijn modellen en simulaties gebruikt. Een model (of simulatie) is altijd een benadering van de werkelijkheid, waarbij steeds een keuze moet worden gemaakt tussen nauwkeurigheid en werkbaarheid. De reden is dat nauwkeurigere modellen vaak meer rekenstappen vereisen en het bij te veel rekenstappen te lang duurt om het model door te rekenen.

Van bliksem is nog veel niet bekend. Om te begrijpen hoe bliksem start zijn goede modellen nodig die bijvoorbeeld de invloed van ijsdeeltjes of regendruppels meenemen, maar het effect is een kostbare simulatie. Uit de studie, beschreven in hoofdstuk 5, blijkt dat een historisch veelgebruikte criterium een verkeerd beeld geeft in de buurt van objecten. Geladen deeltjes kunnen zich namelijk ophopen in de buurt van een object en zo ook bij lagere spanning een ontlading starten. Dit is een belangrijk gegeven voor het starten van bliksem, maar ook voor het voorspellen van het doorbranden van apparaten binnen de hoogspanningstechniek.

Hoofdstuk 6 rapporteert over de doorontwikkeling van een geavanceerde code om de volledige start van een ontlading te simuleren in de buurt van een object. In de code worden individuele elektronen gevolgd, omdat deze individueel een zeer willekeurig gedrag vertonen wat resulteert in een kans dat een ontlading

start. De code kan gebruikt worden om de start van bliksem te simuleren maar ook in algemenere toepassingen binnen bijvoorbeeld de hoogspanningstechniek.

Hoofdstuk 7 is een vergelijkingsstudie van softwarecodes die gebruikt kunnen worden om straling afkomstig van bliksem en onweerswolken te simuleren, met de aanname dat elektrische velden geen rol spelen. Deze aanname is geldig voor het stuk dat de straling aflegt vanaf de bron naar de observator, maar is vooral gemaakt om een vergelijkingsbasis te creren waarop de codes tenminste overeen moeten komen. Het hoofdstuk beschrijft standaarden om een goed vergelijk op te zetten en voerd deze daarna uit op zes verschillende softwarepakketten. Wat blijkt is dat er verschillende keuzes zijn gemaakt in de implementatie van de benodigde benaderingen wat resulteerde in afwijkende uitkomsten. Deze keuzes worden besproken en er wordt geadviseerd hoe het beste de simulatie moet worden opgezet om afwijkingen te voorkomen.

Hoofdstuk 8 is een vervolg van hoofdstuk 7 waarbij nu wordt gekeken hoe de codes presteren in de aanwezigheid van elektrische velden. De studie beschreven in hoofdstuk 8 presenteert een belangrijk verschil bij het opzetten van simulaties dan tot nu toe was bekend. Door de andere keuzes te maken kan een grotere nauwkeurigheid worden bereikt met minder rekenstappen wat een kortere rekentijd oplevert. Ook geeft het de mogelijkheid om uitgebreidere simulaties toe doen.

In hoofdstuk 9 wordt een strategie gepresenteerd die het mogelijk maakt om adaptief de juiste begincondities te bepalen om door te rekenen in een simulatie om zodoende een minder aantal simulaties uit te hoeven voeren voor dezelfde nauwkeurigheid. De nieuwe strategie is vergeleken met twee veel gebruikte strategien en presteert over het algemeen beter.

[Deel 2, hoofdstuk 10 en 11] Het ontstaan van bliksem en de invloed van kosmische deeltjes

Bliksem is een elektrische ontlading en voltrekt zich verschillende fasen. De laatste fase is te zien als flits en te horen als een donder. De eerste fase, waarbij de elektrische ontlading start, is beschreven in hoofdstuk 8. Voor een elektrische ontlading is het nodig dat het elektrische veld hoog genoeg is en dat er vrije elektronen aanwezig zijn. In een onweerswolk blijkt het elektrische veld in eerste instantie bij lange na niet hoog genoeg is om te starten, maar maximaal 20% van de benodigde waarde, die de doorslagspanning wordt genoemd. Druppels en ijsdeeltjes daarentegen, die van nature aanwezig zijn in een onweerswolk, kunnen het veld focusseren (versterken) tot boven de doorslagspanning. Echter, dit versterkte veld is dan alleen zeer lokaal aanwezig, in volumes van nog geen kubieke millimeter. In dit klein volume moeten er vrije elektronen aanwezig zijn om de ontlading te starten. Vrije elektronen worden steeds opnieuw aangemaakt door achtergrondstraling, maar blijken in de onweerswolk zeer schaars te zijn omdat de hoge luchtvochtigheid zorgt dat ze worden opgenomen door zeer kleine con-

densdruppels. Deze schaarste van vrije elektronen, in combinatie met de kleine volumes waar ze aanwezig moeten zijn, geeft een extra barrière om de bliksem te starten. Naast de gemiddelde achtergrondstraling kunnen zeer energetische kosmische deeltjes, afkomstig van ver buiten ons zonnestelsel, elektronenlawines in onze atmosfeer creëren en zorgen voor de vrije elektronen voor het starten van de ontlading bij de druppels en of ijsdeeltjes.

In hoofdstuk 10 wordt door middel van simulaties aangetoond, dat ontladingen in onweerswolken kunnen ontstaan bij 15% van de doorslagspanning, door rekening te houden met de aanwezigheid van ijsdeeltjes en vrije elektronen gemaakt door kosmische deeltjes.

In hoofdstuk 11 is verder uitgebreid berekend welke hoeveelheden vrije elektronen kunnen worden verwacht door deze zeer energetische kosmische deeltjes.

[Deel 3, hoofdstuk 12 en 13] Straling van onweerswolken en bliksem

Onweerswolken blijken ioniserende straling te produceren die grofweg te verdelen is in twee categorien, gebaseerd op de duur van de straling en de correlatie met bliksem. De eerste categorie zijn zeer kortdurende intense flitsen van straling, gecorreleerd met de ontwikkeling van bliksemkanalen. Dit fenomeen wordt uiteindelijk meestal waargenomen als gammastraling en daarom aardse gammaflisten genoemd (in het Engels terrestrial gamma-ray flashes). De tweede categorie duurt vele malen langer en is juist niet gecorreleerd met bliksemkanalen, maar ontstaat in de onweerwolk zelf. Ook dit fenomeen wordt uiteindelijk meestal waargenomen als gammastraling en wordt daarom een gamma-gloed genoemd (in het Engels gamma-ray glow).

Echter beide fenomenen beginnen in de vorm van elektronen die in sterke elektrische velden worden versneld tot zeer hoge energien. Bij aardse gammaflisten gebeurt deze versnelling tijdens de ontwikkeling van de bliksemkanalen en bij de gamma-gloed versnellen de elektronen door de elektrische krachten van de onweerswolk zelf. De versnelde elektronen produceren bij botsingen de gamma straling die relatief vaak gedetecteerd wordt omdat gammastraling veel verder rijkt. Maar het gaat verder, de gammastraling in beide categorien blijkt nog zo sterk te zijn dat deze ook interacties aangaat en neutronen losmaakt van de kernen van luchtmoleculen. Deze neutronen zijn ook gemeten en zijn bijzonder omdat voor het losmaken tenminste 10 MeV energie nodig is. Hiermee geven ze inzicht in de hoge energien die de eerste elektronen verkrijgen.

In hoofdstuk 12 wordt het neutronen signaal bestudeerd ten gevolge van verschillende gammabronnen. De profielen gepresenteerd in het hoofdstuk kunnen gebruikt worden om aardse gammaflisten en/of de gammagloed beter te diagnosticeren.

Het blijkt dat de neutronen op hun beurt weer gammastraling maken als ze worden ingevangen door andere moleculen en zo maken ze eigenlijk een derde categorie gammastraling.

In hoofdstuk 13 wordt geschreven dat aardse gammaflisten een gamma-nagloed (in het Engels TGF afterglow) geeft van gammastraling ten gevolge van de nucleaire interacties van de neutronen.