UNIVERSITY OF CALIFORNIA,
IRVINE

Design and Performance of a Smartphone-based Cosmic Ray Observatory

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Jeffrey R. Swaney, Jr.

Dissertation Committee:
Professor Daniel O. Whiteson, Chair
Professor Steven W. Barwick
Professor David P. Kirkby

2021

DEDICATION

To Benny and Isaac

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Years ago, I was told that unlike a typical Ph.D., with a clearly defined path to completion, CRAYFIS was more like the Wild Wild West. True, the field was wide-open, and the circuitous route to arrive at this thesis often went through uncharted territory. But the analogy has proven correct in more ways than one. The setbacks I encountered were seldom the setbacks I expected, and many of these, by necessity, were solo expeditions. For that matter, there certainly was not much money to go around. Most importantly though, this frontier required a village for survival to be possible. Here, I am grateful for the opportunity to thank those who made this journey possible, both through regular acts of encouragement, and for the occasional rescue when I was left high and dry.

First, I would like to thank the friends and family without whose support I would have abandoned this venture long ago. Most especially, I want to thank my wife, Kriztia, for enduring the endless uncertainties of this project with me, my mother-in-law Marisela for helping us to raise our two boys during these challenging times, and my parents for their constant willingness to encourage and to commiserate. Additionally, I am indebted to many others—non-scientists—who were able to participate in this crowdsourced experiment in more ways than I could have expected: to my father-in-law, Mauricio for allowing me to build a lab in his garage; to my brother-in-law JohnPaul, whose skill with a 3D-printer brought many of these ideas to life; to Marcel, the closest this project has had to a friend in high places; and to the many friends, especially Doug, who have collectively (and sometimes individually) brought this app all over the world.

I also would like to thank the professors and fellow graduate students who have helped to bring this thesis into fruition: my advisor, Daniel, for being understanding of the many non-academic challenges of the past five years—of which a global pandemic was the *least* inconvenient; to Mike, an invaluable source of feedback and a wellspring of ideas; and to Chase, Maxim, Eric, and Christian.

Lastly, I would like to give my gratitude to God our Creator, whose marvellous design has, from the beginning, set humankind's gaze towards the stars.

Work supported by the Fermi National Accelerator Laboratory, managed and operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes.

CURRICULUM VITAE

Jeffrey R. Swaney, Jr.

EDUCATION

Doctor of Philosophy in Physics **2021**
University of California, Irvine *Irvine, CA*

Bachelor of Arts in Physics and Mathematics **2014**
Vanderbilt University *Nashville, TN*
Magna cum laude, Highest Honors in Physics

HONORS AND AWARDS

Chancellor's Fellowship **2014**
University of California, Irvine

Underwood Memorial Award **2014**
Vanderbilt University

Phi Beta Kappa (Junior Inductee) **2013**
Alpha Chapter of Tennessee

McMinn Scholarship in the Natural Sciences **2012**
Vanderbilt University

Ernest A. Jones Scholarship **2012**
Vanderbilt University

TEACHING EXPERIENCE

Graduate Teaching Assistant **2014–2019**
University of California, Irvine *Irvine, CA*

SAT/ACT Tutor **2016–2018**
Powerful Prep, Inc. *Irvine, CA*

Teaching Assistant, Modern Physics Lab **2013**
Vanderbilt University *Nashville, TN*

Tutor, Physics Help Desk **2013**
Vanderbilt University *Nashville, TN*

RESEARCH EXPERIENCE

Astroparticle Physics | CRAYFIS **2016–2021**
Advisor: Daniel Whiteson

Cosmology Theory **2013–2015**
Advisor: Robert Scherrer

PUBLICATIONS

J. Swaney, M. Mulhearn, and D. Whiteson, *Simulated Response of Smartphone Sensors to Cosmic Ray Secondaries.* (2021). [in preparation]

J. Swaney, C. Shimmin, and D. Whiteson, *Data Acquisition System for a Distributed Smartphone Cosmic Ray Observatory.* (2021). arXiv:2108.04803 [astro-ph.IM]

J. Swaney, M. Mulhearn, C. Pratt, and D. Whiteson, *Measurement of Smartphone Sensor Efficiency to Cosmic Ray Muons.* (2021). arXiv:2107.06332 [astro-ph.IM]

J. Swaney and R.J. Scherrer, *Quadratic approximation for quintessence with arbitrary initial conditions.* Phys. Rev. D91, 123525. (2015). arXiv:1406.6026 [astro-ph.CO]

ABSTRACT OF THE DISSERTATION

Design and Performance of a Smartphone-based Cosmic Ray Observatory

By

Jeffrey R. Swaney, Jr.

Doctor of Philosophy in Physics

University of California, Irvine, 2021

Professor Daniel O. Whiteson, Chair

This dissertation examines the feasibility of appropriating the global network of smartphones as an observatory for ultra-high-energy cosmic rays. An application and cloud-based data-acquisition system are first proposed for such an observatory, in which heterogenous devices self-calibrate and server feedback optimizes individual device triggers. Methods of monitoring and manually controlling this network are also examined. Detection efficiencies for cosmic muons, MeV-scale gamma rays, and 120 GeV protons are then measured for this trigger, with which a Monte Carlo pixel model is fine-tuned. Extending this model to cosmic ray showers, the effective area of the observatory for super-GZK primaries is shown to equal that of the Pierre Auger Observatory with a global participation rate under 0.1%, far below initial estimates presented in Ref. [83]. A vastly improved sensitivity to photons and more precise modeling of electrons are primarily responsible for this discrepancy, though more refined models of the combinatorial background may yet lead to more stringent requirements.

Prologue

An Unconventional Cosmic Ray Detector

"Crayfis: the app that turns your phone into a cosmic ray detector. No joke."

These words, emblazoned in 34-pt font across the main page of `crayfis.io`, are typically the science enthusiast's first experience of the project. By its very nature, a crowdsourced experiment such as Crayfis requires this sort of enthusiastic tone—a message which was amplified on the usual popular-science media outlets in 2014 [1–6]. Such an introduction, however, may also leave a negative impression of popular-science clickbait, and the scientific community has often responded to Crayfis as such: as a clever idea that excites the public, but is lacking in nuance and is probably unable to deliver on its lofty promises. Indeed, the claim that smartphones may play a role in the future of cosmic ray physics requires some explanation. Such is the objective of this Prologue: to make this claim intelligible in the simplest possible terms before a more thorough review is attempted in Chapter 1.

Though the concept of Crayfis, a cosmic ray detector built with a network of smartphones, is novel to a degree, conducting citizen science with smartphones is certainly not unprecedented. In fields ranging from orinthology [7] to hydrology [8], smartphone apps predating Crayfis have been used to crowdsource fieldwork, par-

ticularly in remote areas. Other apps have enlisted users as computers, solving classification problems that, at least at the time, was beyond the capabilities of AI [9]. Even the concept of using smartphones as a distributed detector array predated CRAYFIS slightly, though in the context of meteorology using smartphone barometers [10]. Rather, CRAYFIS's novelty lies in extending this concept to the smartphone camera as a particle detector. The billions of smartphones scattered throughout the world had the necessary components—a silicon pixel detector (camera), location services, and Wi-Fi—to be employed as a global observatory. The only missing element was the software.

CRAYFIS became a bold idea in need of an application. Clearly, smartphones would not be able to outcompete existing scientific detectors in their own domains—this experiment would need to leverage its unique worldwide scale in a way that conventional experiments could not. The search for ultra-high-energy cosmic rays (UHECRs) fit this description perfectly: the incredibly small flux of these particles—1 per km^2 per century—necessitates detector arrays of enormous size. The Pierre Auger observatory in Argentina, for example, spans roughly the area of the state of Rhode Island. Naturally, CRAYFIS could far exceed this benchmark: Los Angeles county alone quadruples this surface area. The experiment's success, however, depended on the density of willing users and how well the typical smartphone CMOS can function as a cosmic ray detector.

Achieving a usable detection efficiency requires an app that fully utilizes each smartphone's capabilities. To a greater extent than silicon detectors used in high-energy physics, the hardware requires creative software to work properly for this purpose, as the standard camera pipeline was not designed with a cosmic ray trigger in mind. Hence, there has been an especially wide gap between theory and practice for this unconventional experiment, and after 6 years of development, a design has emerged with great promise for reconstructing UHECR showers. This challenge was also concurrently undertaken by DECO [11–13] and others [14] with a focus on public outreach and later by the CREDO collaboration [15], with encouraging preliminary

results in particle classification and phenomenology of exotic multi-shower events.

This thesis will explore these dimensions of the developing Crayfis experiment. In Part I, the theoretical dimensions of the experiment will be examined: first the science of UHECR showers and how a grid of smartphones might hope to detect them, followed by the physics and computer science underlying smartphone cameras. Part II will then delve into the infrastructure of Crayfis: a set of efficiency measurements with the initial beta version of the app will demonstrate the need for an improved trigger, and modifications to the Android app and backend will be discussed. In Part III, these advances will be used to calculate particle detection efficiencies, which will then be compared to simulation and applied to existing shower reconstruction models. These results, after 6 years of progress, will demonstrate a significantly more capable Crayfis. Now, a better response can be formulated to the skeptics' perpetual question: isn't this an unserious experiment meant to facilitate public outreach? As a functional cosmic ray observatory, isn't this wishful thinking? A joke?

Not in the least.

Part I

Theoretical Background

Chapter 1

The Search for UHECRs

Cosmic rays were discovered in 1911-12, so named as to distinguish them from the equally mysterious "X-rays" emitted from certain minerals [16]. At the time, it would have been hard to predict which of these new phenomena would still contain unsolved mysteries over 100 years later. Nonetheless, though our knowledge of cosmic rays is yet incomplete, much has been learned of their origins and composition over the past century.

While some advances were made in the 1920s and early 1930s—the discovery of the muon in 1936 as a constituent of cosmic rays at sea level, for instance—our knowledge of cosmic rays was accelerated by two breakthroughs: the development of Geiger counters and coincidence circuits, and the resulting discovery of extensive air showers in 1938 [17]. By observing the distribution of particles produced in these showers, Pierre Auger and his contemporaries extended the tail of the cosmic ray energy spectrum from 10 GeV to 10^{15} eV; in the subsequent 80 years, more sophisticated detector arrays have added another 5 orders of magnitude to this figure. While the spectrum, composition, and origins of cosmic rays below 10^{18} eV are now fairly well understood [18], at the extreme end of the tail, a number of open questions remain. This chapter will briefly review what is known about UHECRs—defined as those with energies above 10^{18} eV—and the observatories currently studying them. The theoretical capabilities of a new Crayfis observatory will then be positioned in this context.

1.1 Extensive air showers

Extensive air showers (EASs), distributed cascades of secondaries from cosmic rays scattering off the Earth's atmosphere, are the bedrock of our knowledge of UHECRs. While the cosmic ray spectrum has been probed by other means—balloon- and satellite-borne experiments, for example, have measured cosmic primaries with great precision below $10^{15}\,\text{eV}$ [19–30, 32]—the incredibly low flux of UHECRs (1 per km^2 per century for $E > 10^{20}\,\text{eV}$) makes an expansive observatory practically necessary. Distributed detectors covering vast areas, such as those to be discussed in Section 1.2, leverage the scales of these showers to counterbalance their minuscule flux. The fundamental challenge in such an experiment is then to accurately reconstruct these showers and glimpse the primary particles initiating them.

Reconstructing a cosmic ray primary—its energy, mass composition, and angle—from the observables of the EAS is naturally the purview of simulation. The software package CORSIKA [33] is the standard tool for bridging this gap; nonetheless, a number of lesser-known simulations are employed for specialized analyses, and exhibit strong agreement despite their slightly different approaches [34,35]. Yet, this approach suffers from a fundamental weakness: any EAS reconstruction will suffer systematic bias from the hadronic models employed in the simulation.

For UHECRs, the effective models are all extrapolated: the highest energy currently probed by the LHC is 13 TeV, corresponding to a fixed-target energy of $10^{17}\,\text{eV}$, well below the highest-energy cosmic rays. Though this systematic uncertainty limits UHECR experiments, put conversely, the hadronic models tuned by the LHC are instrumental in more accurate shower reconstruction [36, 37]; to cite a prominent example, such fine-tuning demonstrated that the observed "knee" in the power law spectrum at $\sim 10^{15.5}\,\text{eV}$ is intrinsic to the cosmic ray primaries themselves, rather than an artifact of a deficient hadronic model [38–41]. The upcoming FASER experiment may likewise further enhance such shower simulations through more precise muon and neutrino cross sections in the forward direction [42]. However, the reverse can also be true to a limited degree: in principle, reconstruction of EASs can be used to probe

energy scales up to 100 TeV [43, 44].

1.2 Detection methods

In currently operating observatories, EASs are profiled by two main classes of detectors: surface detectors (SDs) and fluorescence detectors (FDs). The former, through hundreds of evenly-spaced detector stations, measure the lateral density of secondary particles at the earth's surface. The latter use much smaller arrays of telescopes to measure the isotropic nitrogen fluorescence produced by a shower, thereby serving as a calorimeter with access to a shower's longitudinal development profile. The two currently-operating UHECR observatories, the Telescope Array (TA) and the Pierre Auger Observatory, use both of these methods in conjunction, discussed below.

1.2.1 The Pierre Auger Observatory

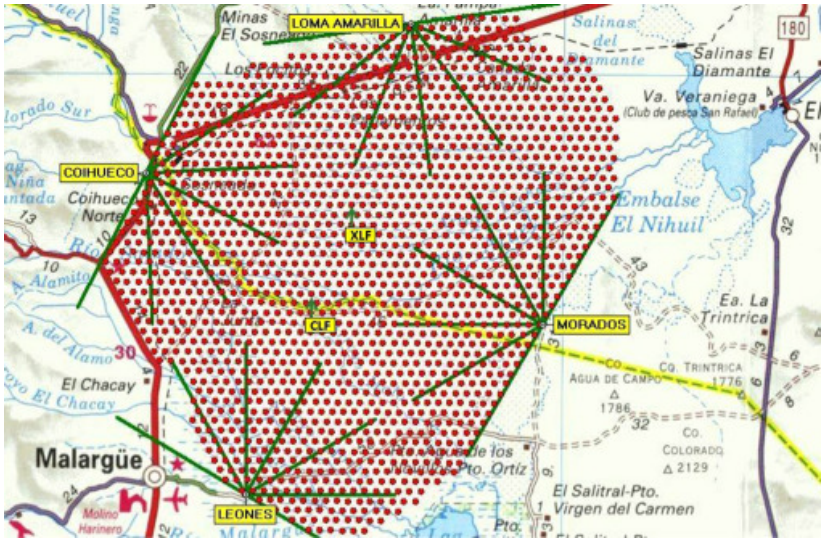

Figure 1.1: A map of the Pierre Auger Observatory, with 1660 SD elements marked in red, and the four FD enclosures, each with 6 telescopes occupying a 30° field of view, labelled in yellow. The green lines mark the divisions between telescope views. From Ref. [45].

The Pierre Auger observatory is located in Mendoza, Argentina, and utilizes a hybrid SD-FD approach over a 3000 km^2 surface area; see Figure 1.1. The SD is composed of 1600 water Cherenkov detectors. Each station consists of a tank 3.6 m

in diameter filled with 12 000 L of ultra-pure water, with three photomultiplier tubes (PMTs) detecting Cherenkov radiation from relativistic charged particles [45]. These detectors are also sensitive to high-energy photons produced in an EAS, as these convert to electron-positron pairs in the water volume. This array operates continuously, and has a perfect detection efficiency for UHECRs above 3×10^{18} eV.

Four stations on the perimeter of the observatory constitute the FD, each with 6 telescopes spanning 30° vertically and horizontally [45]. As mentioned above, integrating the energy deposited along the longitudinal axis provides a calorimetric measurement of an EAS, and the atmospheric slant depth at which the energy deposition is greatest ($X_{\max}$), a useful indicator of mass composition, is also readily available. However, the FD can only operate on dark nights with good weather, restricting its duty cycle to 15% [45].

The virtue of Auger, however, lies in hybrid design: the two detectors working in conjunction can achieve precision and accuracy well beyond that of each individual detector. As a second view of the same EAS events with independent uncertainties, the FD is partly intended to train the continuously-operating SD. By providing essential calorimetric data, the FD can set the energy scale of events seen by the SD, constraining the systematic bias from extrapolated hadronic models [45]. Likewise, the longitudinal profiles of the FD, when combined with the SD's lateral profiles at ground level, allow for a more precise view of a shower's development in both dimensions. For a detailed analysis of the resolution this attains, see Ref. [46].

1.2.2 The Telescope Array

The Telescope Array, located near Delta, Utah, uses the same hybrid SD-FD approach as Auger, but with important differences in the detectors themselves. As shown in Figure 1.2, the SD is composed of 507 SD stations on a square grid, spaced 1.2 km apart with a total surface area of 700 km^2 [47]. Unlike Auger, these detectors are plastic scintillation counters, so chosen as to be sensitive not just to muons, but also the electromagnetic component of the showers ($e^{\pm}$ and γ), which constitute 90% of the

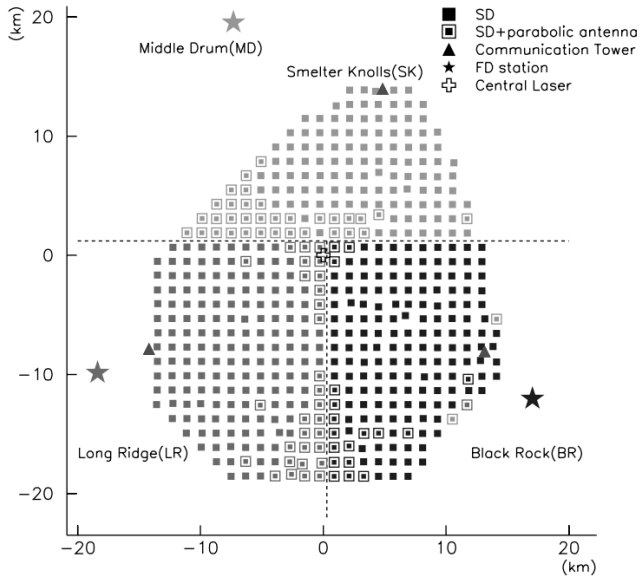

Figure 1.2: A map of the Telescope Array. The 507 SD stations are positioned on a rectangular grid and divided into three subarrays, each controlled by a communication tower (shown as triangles). The three FD stations, shown here as stars, are located slightly beyond the perimeter on nearby hills. From Ref. [47].

energy deposited by 10^{20} eV primaries [47]. Consequently, this array achieves 100% efficiency for primaries with energies above 10^{19} eV. The FD consists of three stations surrounding the SD, each containing between 12 and 14 telescopes and spanning a total field of view of 30° vertically and 108° horizontally [48]. Like Auger, the duty cycle of the FD is quite low—on the order of 10% [49]. However, the FD at TA has a slight advantage in terms of stereo-fluorescence measurements (combining the profiles of two FD stations) for energies below $10^{18.5}$ eV [50, 51]. In order to increase the rate of data collection, a larger, more sparse extension (TA×4) is planned, quadrupling the effective area [52].

1.3 Current advances from TA and Auger

Fifteen years into Auger's operation and twelve years into TA's, substantial progress has been made in answering the lingering questions about UHECRs. Below, several of these are discussed in brief.

1.3.1 Energy spectrum and mass composition

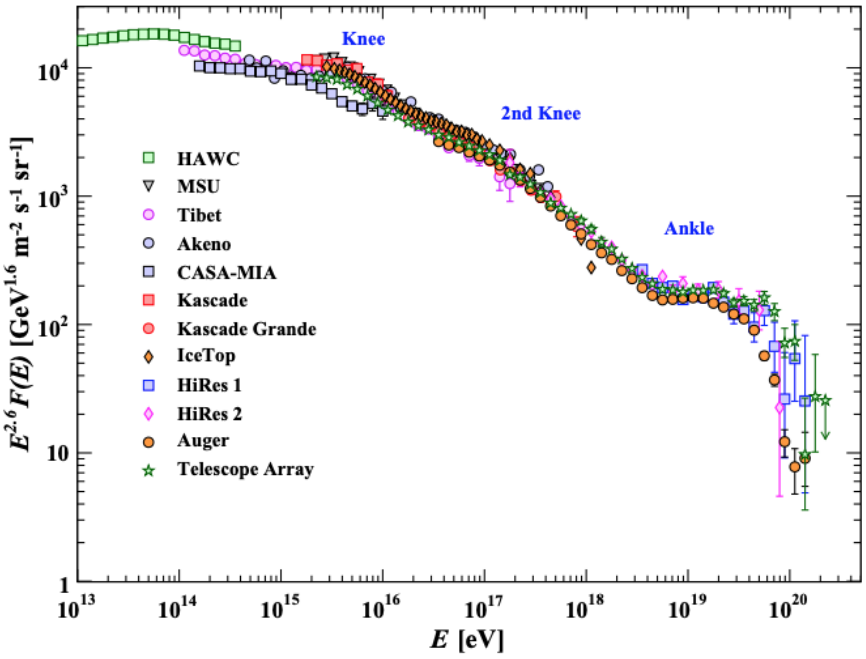

Figure 1.3: The all-particle spectrum from EAS reconstruction, compiled by [53] from [54–64]. The vertical axis has been multiplied by $E^{2.6}$ to better display the features of this steep power-law spectrum.

While the astounding energy scale of UHECRs was the initial impetus behind the study of cosmic rays, the shape of energy spectrum is itself rich with significance. As Figure 1.3 shows, this spectrum is well-approximated by a piecewise power law which steepens twice—first at the "knee" at $10^{15.5}$ eV and again at 10^{17} eV—then flattens at the $10^{18.5}$ eV "ankle" before steeply declining near $10^{19.5}$ eV. For each of these features, two possibilities exist: either they reflect an astrophysical process or a deficiency in our hadronic models. For an example of the latter, it was speculated that the knee could be the result of "invisible particles," such as gravitons, light supersymmetric particles, or neutrinos, consuming part of the primary's energy without fluorescing [65]—a hypothesis later excluded by the LHC [37].

Assuming the hadronic models are not at fault, there are reasonable explanations for the notable features at UHECR energy scales: the ankle and final cutoff. The latter, in fact, was hypothesized long before it was confirmed by TA and Auger. The Greisen-Zatsepin-Kuzmin (GZK) effect predicted that protons above 5×10^{19} eV would interact

with photons in the cosmic microwave background (CMB), producing pions through the Δ resonance [66, 67]; hence, any super-GZK cosmic rays are either from nearby sources ($< 100\,\mathrm{Mpc}$) or are composed of heavier nuclei. Indeed, results from Auger have suggested a trend towards heavier nuclei past the ankle [68], as shown in Figure 1.4, though this conclusion rests on measurements of $X_{\max}$ with ample systematic uncertainty. These heavy nuclei, however, are not exempted from such suppression: similar theoretical cutoffs have been shown to constrain heavier elements from He to Fe, though at slightly higher energies [69].

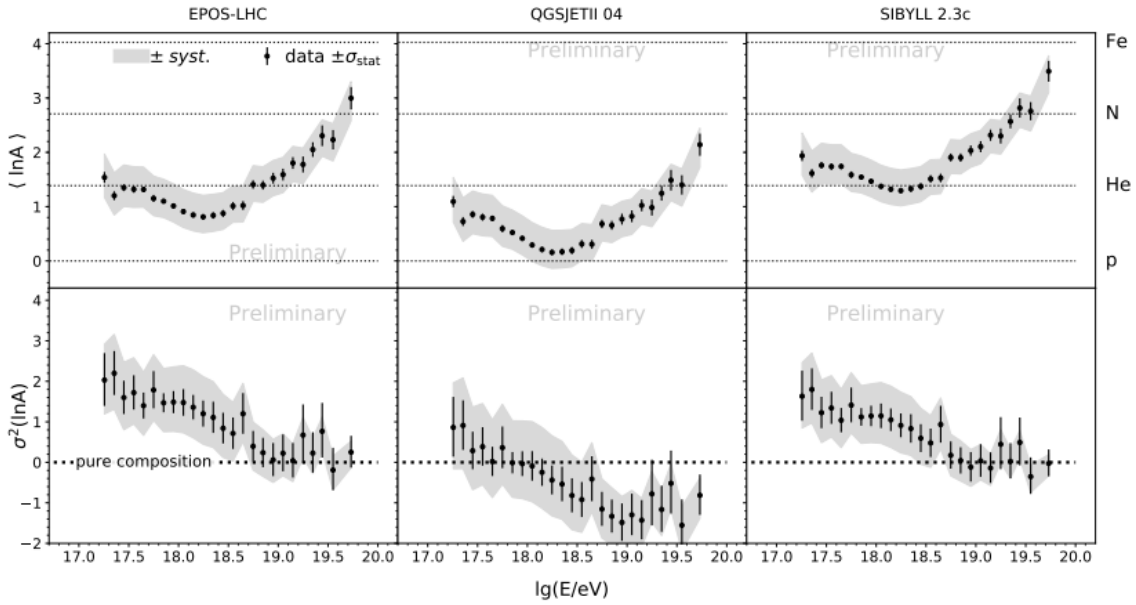

Figure 1.4: Preliminary measurements of atomic mass A vs. energy of primary cosmic rays at Auger. Top plots measure mean atomic mass at each energy, while bottom plots measure variance of $\ln A$. The three columns correspond to different hadronic models used in simulation. From Ref. [67].

Suppression of the super-GZK proton flux is clearly evidenced in measurements from both Auger and TA, as shown in Figure 1.5a; however, these two experiments yield inconsistent spectral shapes for this suppression. The spectra for TA and Auger can be made to agree through the ankle (Fig. 1.5b) by globally rescaling the energies by 5.2%, reflecting systematic uncertainties in their respective FDs; sampling only from the common declination band of the two observatories further extends this agreement to the GZK cutoff [70]. Disagreements in the shape of this suppression, however, require non-linear shifts in energies above $10^{19}\,\mathrm{eV}$ to resolve, at a rate of 10% per

decade [71, 72]. To date, no explanation has been found for proposed nonlinearities of this scale.

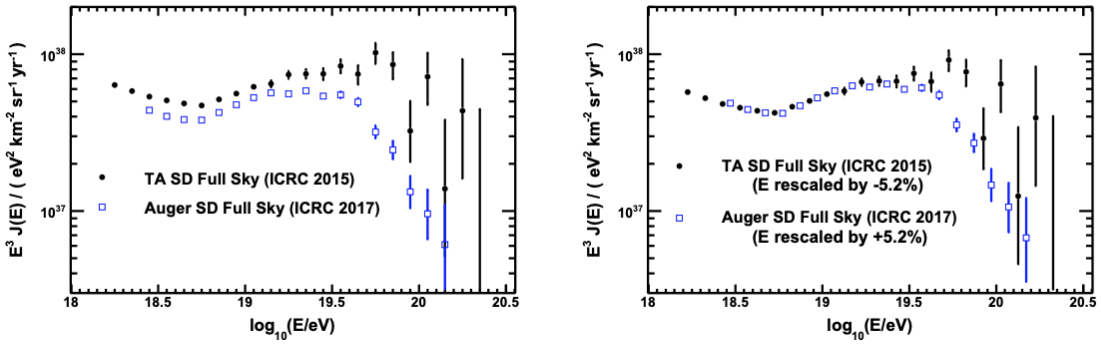

Figure 1.5: A comparison of the UHECR energy spectra for Auger and TA, where the vertical axis is multiplied by E^3 to better display changes to the power law. Left: the proper energy scales of TA and Auger are used. Right: Auger is rescaled by +5.2% and TA by -5.2% to achieve optimal agreement up to the ankle. From Ref. [69].

Meanwhile, the ankle is typically explained by one of two astrophysical processes. It is often attributed to an extragalactic population of protons overtaking the galactic flux, which at this energy scale is dominated by heavier nuclei due to galactic confinement [73]. This hypothesis is supported by the mass composition in Figure 1.4, in which the spectrum becomes lighter around the ankle before rapidly increasing in mass at the GZK limit. Another explanation is proposed in Ref. [74], where extragalactic protons lose energy to CMB photopion interactions (as in the GZK effect), accumulating at the ankle. Ultimately, a more precise mass composition will distinguish between these.

1.3.2 Anisotropy

Though the energy spectrum and mass composition are both rich with astrophysical phenomena, these only touch upon what is perhaps the most interesting question raised by UHECRs: what kind of process is capable of accelerating particles to 10^{20} eV? To more directly pursue this question, a number of anisotropy searches have been performed, on scales both large (through multipole expansion) and small (through correlation with possible sources).

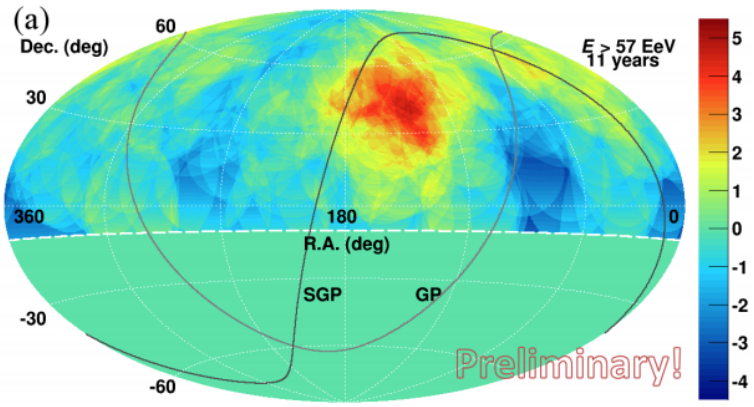

Figure 1.6: A preliminary significance map of directional UHECR flux, for energies above 5.7×10^{19} eV, from the Telescope Array (2008-2019). The color axis is in units of 1σ, and the galactic and supergalactic planes are shown as solid lines. From Ref. [82].

Small-scale anisotropy searches have seen moderate success: starburst galaxies (SBGs) and, to a lesser extent, active galactic nuclei (AGNs) have been implicated as a source of UHECRs by Auger. Centaurus A, one of the closest AGNs (< 4 Mpc), was proposed as a possible UHECR source as early as 1996 [75], and Auger has observed an increasingly significant excess throughout its operation [76–80], most recently at 3.9σ [80]. However, more general tests of local anisotropy have been performed [79,80], using catalogs of nearby sources from *Fermi*-LAT. Among these, the greatest deviation from isotropy was found in SBGs, reaching a significance of 4.5σ [80].

On an intermediate scale, TA detected a "hotspot" of UHECRs found near Ursa Major in 2014, shown in Figure 1.6. This excess of UHECRs above 5.7×10^{19} eV, roughly 20° in width, amounted to a 3.4σ significance above isotropy [81]; however, it was far less pronounced in the following six years, decreasing the 11-year significance to 2.9σ [82]. As this event is in the northern sky, it is unfortunately out of Auger's field of view for an independent confirmation.

1.4 CRAYFIS: a novel observatory

TA and Auger have greatly refined our understanding of UHECRs, yet an abundance of questions—of which those discussed above are a small sample—still remain. To contribute to this body of knowledge, CRAYFIS proposes to utilize the existing grid of

smartphones, which lie dormant for most of the night, as an SD requiring no expensive infrastructure or dedicated land use. In order to utilize this pre-existing grid, the only necessary components are an app to manage the data-acquisition (DAQ) and a server to receive the data over Wi-Fi. The architecture of the experiment, much improved through years of beta-testing, is discussed in Part II. Here, its theoretical performance relative to TA and Auger is considered, as well as global phenomena that Crayfis might be able to observe.

1.4.1 Shower reconstruction

As an SD, Crayfis's sensitivity can be evaluated separately for each constituent particle of EASs. While cosmic ray showers produce a wide variety of particle species, the the vast majority of the flux at the earth's surface is composed of muons, electrons and positrons, photons, neutrons, and neutrinos—the charged hadron flux falls sharply with atmospheric depth. Of these, neutrons and neutrinos stand a poor chance of detection in a thin layer of silicon. Sensitivity to the three remaining sectors is parametrized by the *detector cross sections* $A_i\epsilon_i$, signifying the rate of particle detections divided by the flux, for $i = \{\mu, \gamma, e\}$. For historical reasons, the notation is broken into per-particle efficiencies ϵ_i and acceptances A_i, though these are less precisely defined, as will be discussed in Chapter 6. Combined with the local density of users, the detector cross sections for muons, photons, and electrons determine Crayfis's sensitivity to EASs at a given energy and angle of incidence.

Following Ref. [83], the lateral density of the muonic and electromagnetic components of an EAS is approximately:

$$\rho_i(N_i, r, s) = \frac{N_i}{2\pi r_M^2}\left(\frac{r}{r_M}\right)^{s-2}\left(1+\frac{r}{r_M}\right)^{s-4.5}\left(\frac{\Gamma(4.5-s)}{\Gamma(s)\Gamma(4.5-2s)}\right) \tag{1.1}$$

where s is the shower age (scaled from 0 at first interaction to 1 at the shower maximum), r_M is the Moliére radius, and N_i is the total number of particles of species i in the shower, determined by CORSIKA. In the presence of an air shower with total

particle counts $\vec{N}$, a phone a distance r from the shower's center at the earth's surface will trigger, on average,

$$\lambda(\vec{N}, r, s) = \eta + \sum_i A_i \epsilon_i \rho_i(N_i, r, s) \tag{1.2}$$

times, where η is rate of noise per frame, including both electronic noise and background radiation. Such a phone will register no candidates with probability $P_0(\vec{N}, r, s) = e^{-\lambda(\vec{N},r,s)}$ and at least one candidate with probability $P_1(\vec{N}, r, s) = 1 - P_0(\vec{N}, r, s)$. Then, for the simplest shower reconstruction (i.e. neglecting trigger multiplicities), the likelihood can be constructed as:

$$L(E, \vec{x_0}, \theta, \phi) = \prod_i P_1(\vec{x_i} - \vec{x_0}) \prod_j P_0(\vec{x_j} - \vec{x_0}) \tag{1.3}$$

from which the a Bayesian estimate of the shower parameters can be obtained.

To translate this into a reconstruction efficiency, a limit of at least five phones with coincident triggers is imposed, based on initial estimates of the combinatorial background. As Figure 1.7 makes explicit, at a fixed primary energy and noise rate η, Crayfis's resolving power is dependent upon two unknowns: the detector cross sections $A_i \epsilon_i$ and the density of active users, each of which can directly compensate for deficiencies in the other. Figure 1.8 converts this shower efficiency into an effective area (that is, the integral of local per-shower efficiency across the detector) with Auger as a benchmark. This quantity is a more direct measurement of a detector's ability to accrue significance; under the more optimistic scenarios, Crayfis could at least rival Auger at detecting UHECRs near the GZK limit.

However, the curves in Figure 1.7 ignore the potentially significant contributions from electrons and positrons. Though the $e^{\pm}$ flux may indeed be attenuated by the overburden in a typical home, the GeV-scale $e^{\pm}$ flux from UHECR showers exceeds that of muons by nearly an order of magnitude [83]. At this energy scale, the electron range in carbon (a major component of the wood and asphalt comprising these structures) is nearly a meter [84]; consequently, Crayfis's $e^{\pm}$ sensitivity may be quite significant.

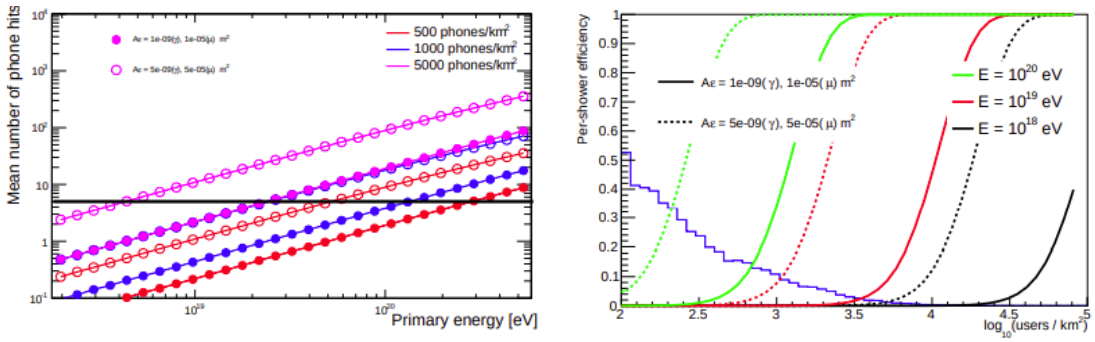

Figure 1.7: Left: the mean number of phone triggers per shower is plotted against primary energy, $A_i\epsilon_i$ for muons and photons, and user density. Right: the shower reconstruction efficiency of CRAYFIS under the same six choices of parameters. A histogram of global population density is superimposed in blue. From Ref. [83].

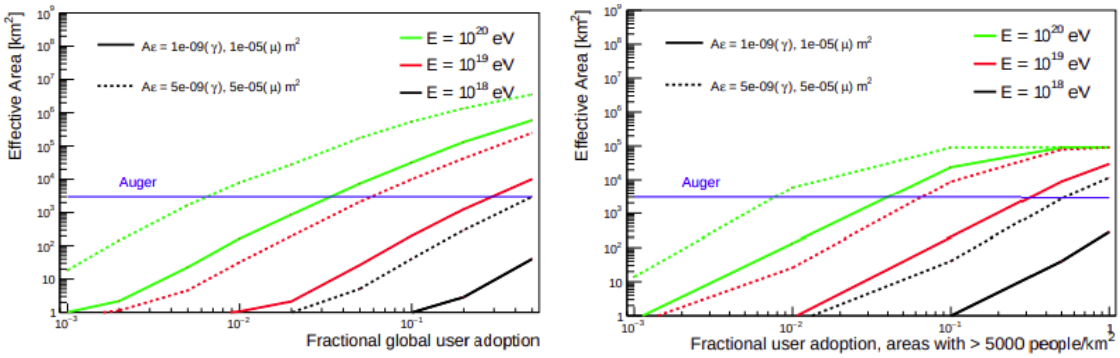

Figure 1.8: The effective area of the CRAYFIS observatory, determined by the same parameters as in Figure 1.7. The left plot quantifies the global effective area, while the right examines that of high-density metros, in which more targeted user adoption campaigns are possible. From Ref. [83].

.

1.4.2 Sensitivity to global phenomena

While reconstruction of individual UHECR showers, the primary focus of TA and Auger, was the motivation behind the Crayfis observatory, Crayfis's massive scale presents an opportunity for detecting hypothesized global phenomena beyond these observatories' capabilities. One well-known example, the Gerasimova-Zatsepin (GZ) effect [85], has been studied in detail in relation to Crayfis [86]. In this mechanism, a UHECR nucleus is split by solar photons, producing two daughter nuclei which the solar magnetic field then separates. If both daughter nuclei arrive at the earth, nearly-simultaneous showers separated by thousands of kilometers would result.

Such an event has never been observed to date, and with the initial efficiency efficiency estimates from Ref. [83], Crayfis would be able to distinguish such events only under unrealistic user adoption scenarios, or if the rate of such events was much higher than predicted. The approximate time to accruing significance for the GZ effect is shown in Figure 1.9. However, even a more realistic shower reconstruction efficiency may prove sufficient for detecting a UHECR "burst", in which a wavefront of UHECR particles and nuclei arrive simultaneously. A number of mechanisms have been proposed for such events, including relativistic dust-grains [87–90], pair-production from ultra-high-energy gamma rays [91–95], super-GZK neutrinos [96–99], extra-dimensions and localized gravity [100], and a collection of exotic particle scenarios [101–114]. The phenomenology of such scenarios will be left to later work.

As this last section has shown, the potential of Crayfis, which extends to phenomena undetectable by conventional cosmic ray observatories, is strongly constrained by the per-particle cross sections $A_i\epsilon_i$ of the individual detectors. While greater user adoption can compensate for a limited efficiency, the number of devices is still finite. To approach this pivotal question, the design of the detectors—the CMOS sensors within smartphone cameras—must first be explored.

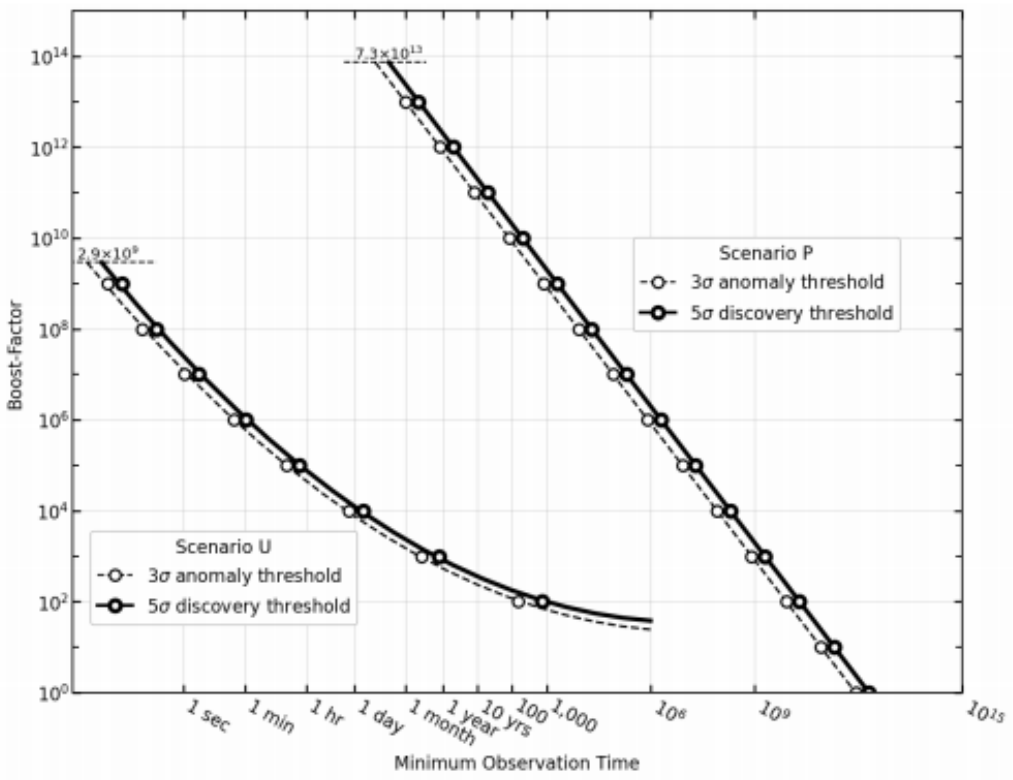

Figure 1.9: The CRAYFIS sensitivity to GZ-effect double showers under two scenarios: the upper bound of 100% worldwide adoption (Scenario "U") and a more pragmatic 10^6 active phones operating nightly. Minimum observation time for 3σ and 5σ significance are plotted against the "Boost factor", i.e. a hypothetical factor by which the true GZ-effect rate exceeds the predicted value. From Ref. [86].

Chapter 2

Image Sensors

The missing piece in CRAYFIS 's sensitivity, the detection cross sections $A_i\epsilon_i$ for muons, photons, and electrons, requires an understanding of both the hardware and software underlying modern smartphone cameras. The insight that began CRAYFIS—that these cameras, as CMOS active pixel sensors, resemble the well-studied detectors used in high-energy physics—is true to a certain extent. However, accounting for the different design in a sensor meant for photography, and how this affects its efficiency, is the key to determining CRAYFIS 's viability.

2.1 Photodiode theory

Each pixel is, at the simplest level, a silicon photodiode with several transistors managing readout. A photodiode is, as its name suggests, a p–n junction that can convert energy from an incoming particle—particularly a photon in the visible range—into a measurable current. Figure 2.1 demonstrates this process. At the junction, holes in the valence band of the p-layer recombine with electrons in the conduction band of the n-layer, creating a net charge on each side. This creates an electric field varying roughly linearly with distance to the junction and pointing towards the p-layer. This area with nonzero field is known as the "depletion zone", as it contains a greatly reduced majority carrier count.

When a particle deposits energy into the silicon, electrons are excited into the con-

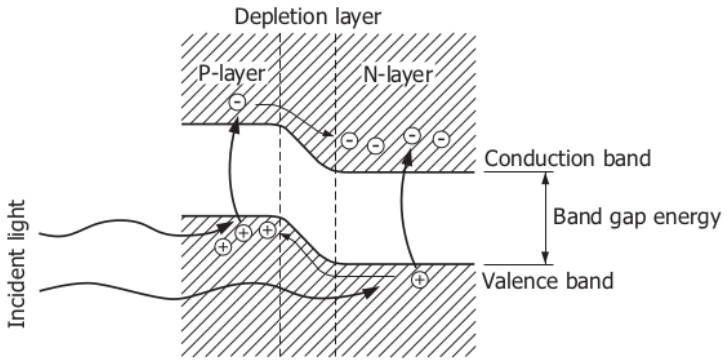

Figure 2.1: A diagram of the p–n junction in a photodiode. Incident light excites electrons into the conduction band, which are swept by the electric field of the depletion zone into the n-layer. The opposite will occur with holes, although these are not always collected. From Ref. [115].

duction band, creating mobile electron-hole pairs. The electrons diffuse throughout the silicon until they either recombine with a hole or are swept into the potential well made by the depletion zone field. At readout, the cathode and anode are electrically connected, generating a measurable photocurrent from these captured electrons. However, electron-hole pairs are also spontaneously generated—at a rate growing exponentially with temperature according to Arrhenius' Law—leading to an additional *dark current*. This and the roughly Gaussian noise from the readout response are the two major sources of noise in CMOS pixel detectors.

2.2 Pixel engineering

Figure 2.2 below shows a typical cross section of the pixel array in a modern smartphone camera. Each pixel is supplied with a micro-lens and a color filter on the opposite side of the potential well and readout architecture, known as a "backside-illuminated" (BSI) configuration. In early CMOS image sensors, the transistors were fabricated on the same side as the lens and color filter ("frontside-illuminated", or FSI) but this design interferes with the transmission of light to the active pixel region. Since 2009, when Sony patented a method to cheaply manufacture BSI sensors, BSI has been the industry standard [116].

Meanwhile, the dimensions of these pixels have followed an even more rapidly

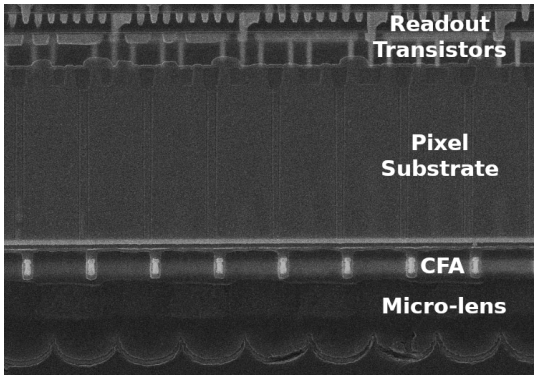

Figure 2.2: A cross section of the active sensor region in an ISOCELL S5K3P3 sensor manifesting the BSI architecture, taken by electron microscopy. The micro-lenses, color filter array (CFA), and readout architecture are also visible. For reference, the pixel pitch is 1.12 µm. From Ref. [117].

changing standard. The number of "megapixels" in a smartphone camera has been a profitable marketing tool over the last decade, and sensor manufacturers have heavily invested in fitting more pixels into a constant sensor area while maintaining image quality. Figure 2.4 shows some of the consequences of this trend: a decrease in pixel pitch will also decrease the full well capacity (FWC) of the photodiode, limiting the dynamic range, unless the pixel thickness is increased proportionately [118].

However, longer and narrower pixels do not necessarily produce sharper images. Such a geometrical shift would increase crosstalk between pixels without additional measures to limit diffusion; advances in deep trench isolation (DTI) have thus been instrumental in the shift to sub-micron pixels [117, 118]. DTI separates individual pixels with an oxide layer (e.g. SiO_2 [119]), and can be fabricated in one of two ways: extending from the transistor-side ("front", or F-DTI) or from the lens-side ("back", or B-DTI). These variants are both common in smartphone CMOS sensors today. However, as the electric field is strongest near the transistors, B-DTI provides more protection against crosstalk per unit-length [120]. It is therefore common for F-DTI structures to extend to the opposite end of the pixel for full isolation at the cost of reducing the well width and capacity.

For practical purposes, it is worth briefly discussing the main manufacturers of smartphone CMOS sensors: Samsung and Sony. These two corporations have dominated the industry since the advent of smartphones, and Omnivision is only recently

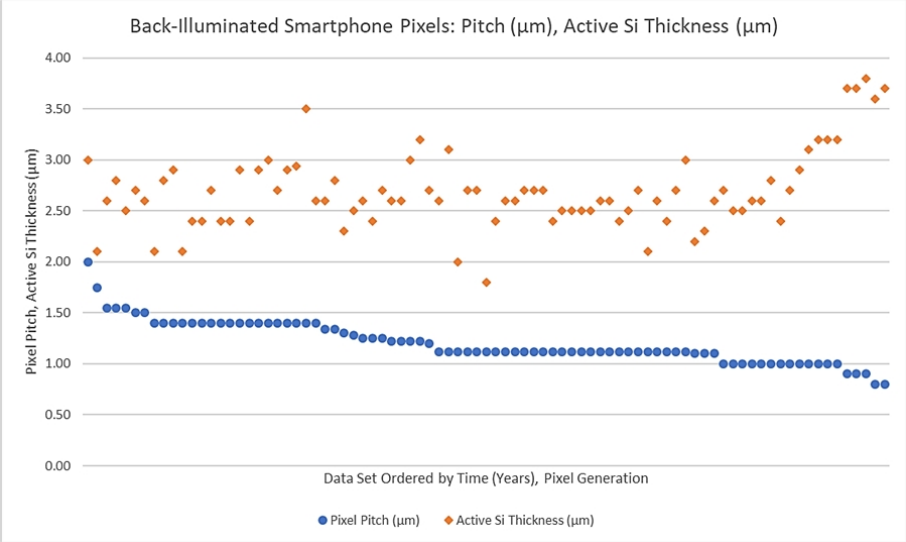

Figure 2.3: Active silicon thickness (orange) and pixel pitch (blue) for devices between 2009-2019. While pixel thickness remained roughly constant for devices up to the 1.12 µm generation, more recent sensors boasting sub-micron pixels have extended the pixel thickness to increase FWC. From Ref. [118].

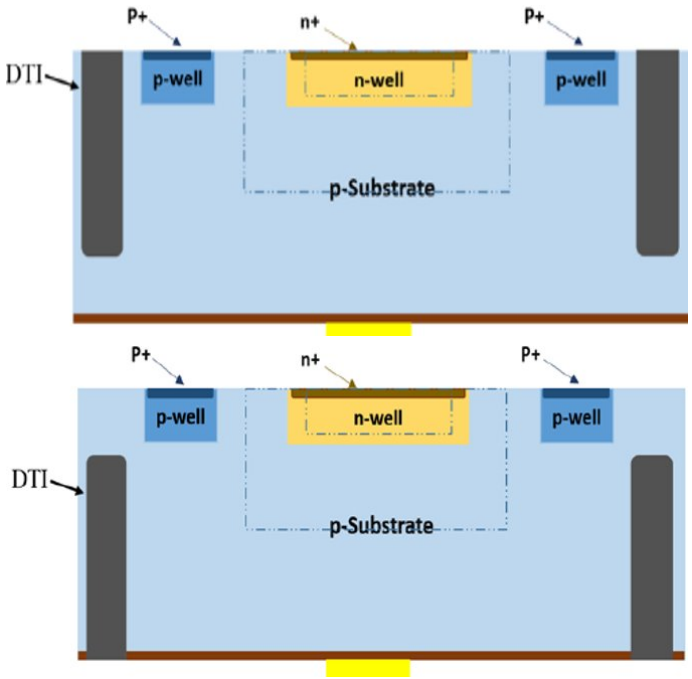

Figure 2.4: An illustration of F-DTI (top) and B-DTI (bottom). From Ref. [120].

breaking into the field as a minor third supplier. For many phone models—Samsung's included—similar sensors from Sony's "Exmor" line and Samsung's "ISOCELL" line are used interchangeably, often with identical pixel pitches and resolution. Yet although these sensor varieties are designed to be similar, they are certainly not identical. As a particularly visible example, Sony and Samsung have diverged in their approaches to reducing crosstalk for smaller pixels: while Sony has continued the use of B-DTI, Samsung has instead migrated to full-depth F-DTI beginning with the 1.0 µm generation [117]. This difference is shown in Figure 2.5.

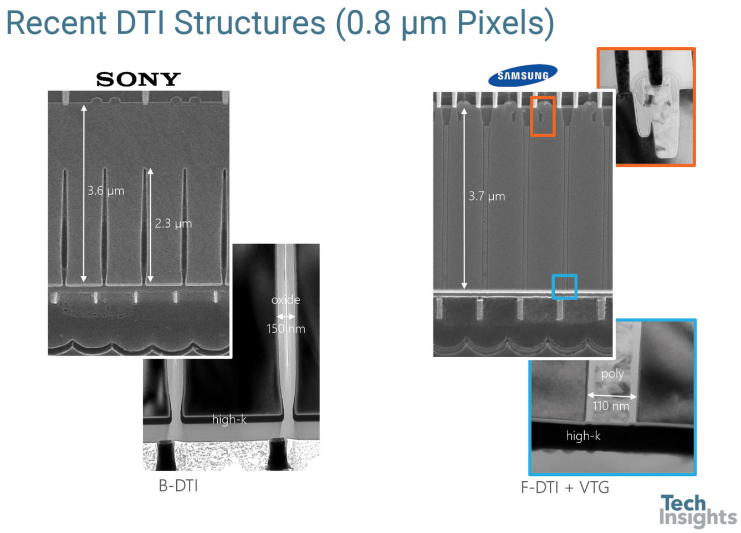

Figure 2.5: Diverging approaches in Samsung and Sony to reduce crosstalk in smaller, thicker pixels. Sony has continued to use B-DTI, while Samsung has gravitated towards full-depth F-DTI with a vertical transfer gate (VTG) to compensate for reduced well size. From Ref. [121].

Finally, it is worth mentioning how this design differs from the silicon pixels used in high-energy physics. The primary difference is size: pixels in the CMS detector, for example, measure 150 µm in pitch and 250 µm in thickness [122]. Since camera pixels are designed to absorb photons in the visible range, the extra stopping power provided 250 µm of silicon is wholly unnecessary. As Figure 2.6 shows, this limits the spectral information that smartphone pixels are able to convey. While a proposed CRAYFIS array might have a viable detection efficiency for different particle species, distinguishing between these species will likely be difficult, if not impossible.

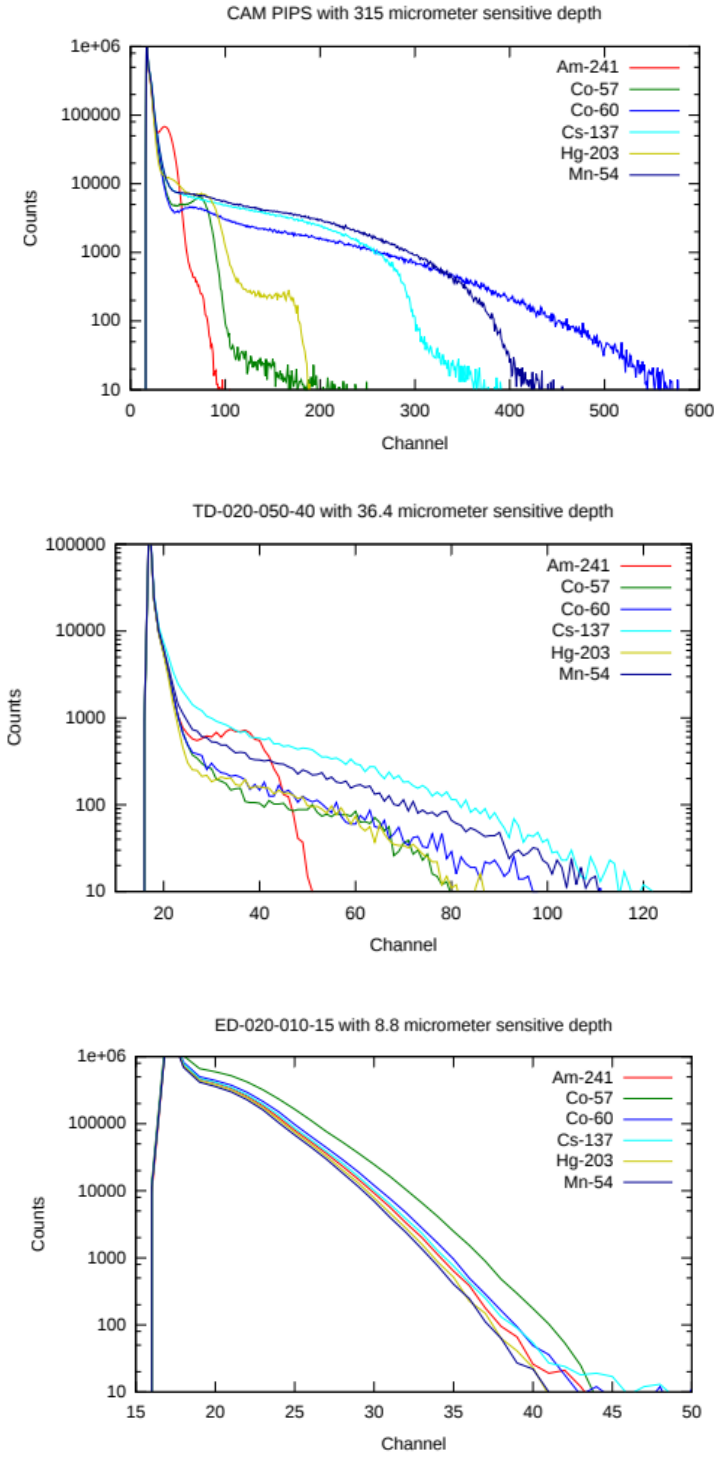

Figure 2.6: Energy deposited by gammas of various energies (roughly indicated by color) in silicon surface barrier detectors of thickness: (top) 315 µm, (middle) 36.4 µm, and (bottom) 8.8 µm. Distinct Compton edges can be seen in the thickest detector, which fade in the lower plots as fewer electrons are fully absorbed in the sensitive region. The last of these most closely approximates the pixels in a smartphone CMOS and loses almost all spectral information. From Ref. [123].

2.3 Digital processing and PDAF pixels

As mentioned in Section 2.1, each pixel will produce a photocurrent which is the basis of the values reported by the camera. This current is digitized and packaged into an image buffer by the following process:

1. A gain in digital numbers (DNs) per electron is applied, as determined by the programmable ISO setting.

2. Black-level subtraction is performed to compensate for dark current. The magnitude of this correction is determined by "dark pixels" which are masked from visible light, typically located along the edge of the physical sensor [125]. Occasionally, the black level is set at a nonzero value to show the full range of responses, both above and below the mean dark current.

3. The result is digitized. In smartphone cameras, DNs are usually given 10-bit precision.

In a CMOS sensor, this digitization process is executed in parallel through the per-pixel micro-transistors shown in Figure 2.2. In contrast to charge-coupled devices (CCDs), in which pixel readout occurs serially outside of the physical sensor, the CMOS design permits higher frame rates and eliminates charge-smearing in overexposed frames. As this approach requires greater numbers of amplifiers and analog-to-digital converters per device, CMOS sensors sacrifice uniformity in the pixel response and are more vulnerable to fixed-pattern noise. [124]

However, additional steps are often added to this processing pipeline. While the considerations thus far have been general to all silicon pixel detectors, building an effective camera can require a more intricate design. Autofocus is one such complication to this simple model which extends even to the level of hardware.

Figure 2.7 illustrates the operative principle behind this design. By definition, an object is in focus when light from a single point on the object will converge on a single point on the sensor; thus, if a given pixel is exposed to an in-focus point source,

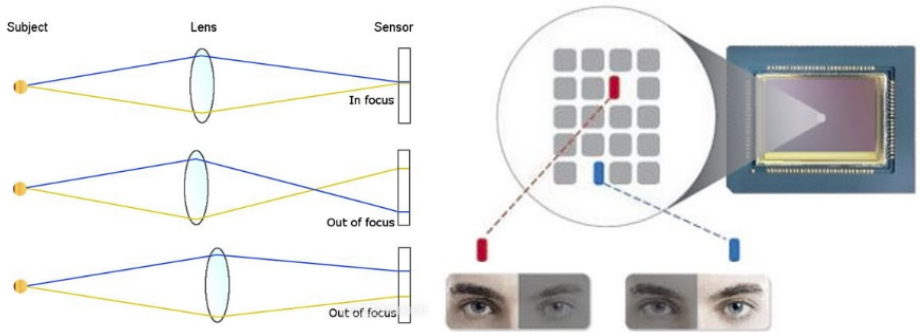

Figure 2.7: An illustration of PDAF function and design. Description in text. From Ref. [126].

the relative exposures from, e.g., left-entering and right-entering light should be a well-defined ratio depending on the lens geometry and pixel position. By measuring this ratio across the sensor, the camera can determine which parts of the image are in focus and how to adjust the lens accordingly. The directionality of incident light can be determined through clever engineering: since by design, the micro-lens maps the angles of incoming photons to specific regions of the pixel, the simplest method to create a "left-facing" or "right-facing" pixel is through a partial mask [117]. By comparing the responses of nearby left-masked and right-masked pixels (Figure 2.7b), the autofocus can be quickly and accurately adjusted. A masked pixel, however, will receive only half of the exposure of a regular pixel, creating a low signal-to-noise ratio in dark conditions; hence, such "phase-detection autofocus" (PDAF) pixels are sparsely distributed. Rather than report the true pixel value of PDAF pixels, a sensor may instead interpolate the value from nearby pixels.

This initial design was first observed in smartphones in 2014 [127], but in later generations, more creative PDAF approaches were developed without the use of masking. A common strategy is to instead use two or more photodiodes per pixel (2PD PDAF), often across the entire sensor, to measure both the total light exposure and its directional components [127]; see Figure 2.8b. However, this approach requires a larger pixel size and thus a lower image resolution—a difficult compromise in a market driven by megapixels. Nonetheless, the fundamental insight—that PDAF could be accomplished with multiple photodiodes per micro-lens—could be exploited without this

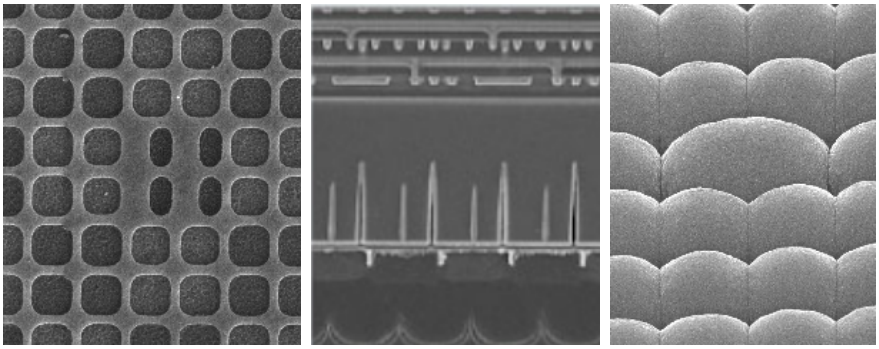

Figure 2.8: The three most common PDAF approaches as seen by scanning electron microscopy: (left) pixel masking, (center) dual-photodiode pixels, and (right) a 2×1 on-chip-lens. Images from Refs. [118, 127].

drawback by instead interspersing larger micro-lenses, covering several pixels simultaneously, across the sensor [127]. This third approach, on-chip-lens (OCL) PDAF, is shown in Figure 2.8c.

2.4 Color-space processing

This culmination of the digitization process is a RAW-format image. RAW is an uncommon format intended for professional photographers, as the color information, white balance, etc. are present only in the image metadata; while RAW is ideal for Crayfis in principle, it is therefore not always supported by the camera APIs. However, even when RAW is available, it is not necessarily optimal in practice. Due to their high resolution and bit-depth, RAW images are expensive to process for both memory and CPU. This cost is augmented by inefficiencies in the Android pipeline, which understandably was not built for computationally-intensive tasks on RAW-format video buffers.

Consequently, while RAW-format data in Crayfis is possible, the standard video buffers are required for higher frame rates or lower CPU loads. Counterintuitively, this is accomplished by greatly extending the camera pipeline; see Figure 2.9. After the RAW values have been computed, the camera converts RAW values to 8-bit RGB tuples. As mentioned in Section 2.2, each pixel is equipped with a color filter: typically, this follows a Bayer pattern shown in Figure 2.10. The camera first applies color-

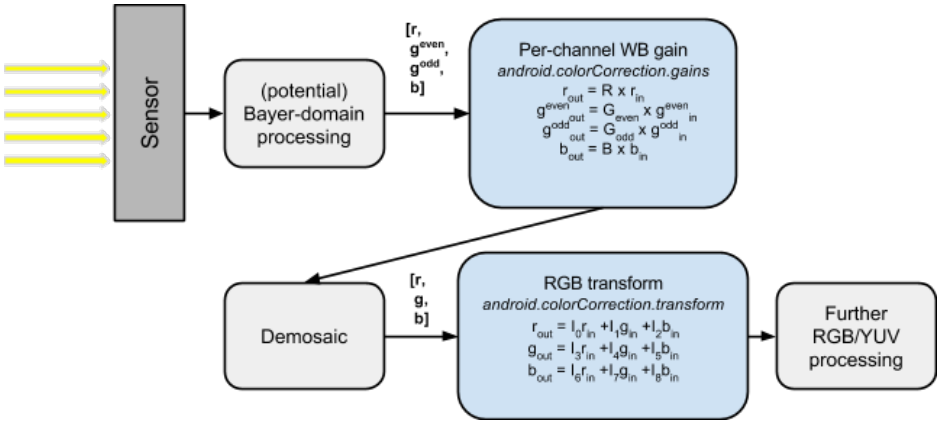

Figure 2.9: The color correction pipeline in Android phones. From Ref. [129].

specific gains to each pixel as part of a white-balancing algorithm. Next, demosaicing occurs: each pixel infers the other two colors from its neighbors. The simplest such algorithm is bilinear interpolation—averaging values from the 8 nearest neighbors—but more sophisticated and even nonlocal means may be used [125]. Finally, a linear transformation is applied to this tuple, mapping the native RGB space from the specific color filters to a standard RGB space (sRGB) used by displays [125].

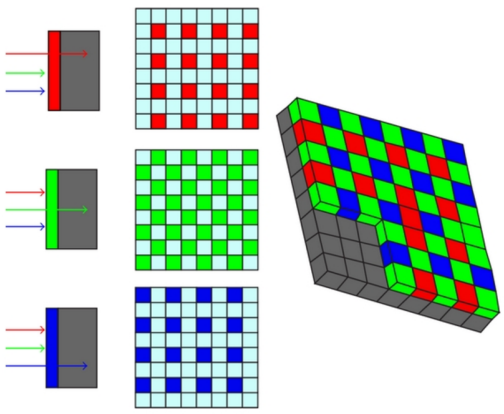

Figure 2.10: An illustration of a Bayer filter. From Ref. [128].

In order to reduce video size, an additional manipulation known as *chroma subsampling* is applied. As the human eye is more able to detect sharp differences in brightness rather than color, the color information can be presented at lower resolution with minimal reduction in perceived video quality [125]. Accordingly, the RGB signal is broken into luminance (Y') and chrominance (UV) pieces through the follow-

ing linear transformation:

$$\begin{bmatrix} Y' \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.14713 & -0.28886 & 0.436 \\ 0.615 & -0.51499 & -0.10001 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \tag{2.1}$$

and the U and V components are subsampled through one of several algorithms. These buffers can be easily accessed and manipulated in Android, and the Y' components can stand as a rough substitute for RAW values.

Lastly, downsampling takes place if necessary to further reduce the CPU load. The resolution is decreased through a combination of three processes: cropping, binning, and decimation. The frame is first cropped to reach the desired aspect ratio: either 4:3 or 16:9. The visible pixels at the lower resolution are then constructed from sets of physical pixels, typically no more than 2×2 in dimension, which are averaged together in each color coordinate. The remainder of pixels are ignored.

2.5 Image enhancement

Even beyond the color-space pipeline, many digital effects may be used to enhance an image, most often to the detriment of the signal. As of Android 10.0, there are about 11 different "modes" in the `camera2` API [129] which introduce distinct digital effects to the pipeline: "edge mode", "noise reduction mode", "distortion correction mode", etc. Fortunately, many of these can typically be disabled on newer phones. Here, several of the more important effects will be discussed.

Firstly *tonemapping* or *gamma correction* is a method of lending high dynamic range to an image by abandoning the linearity of the pixel response. This often takes the form $V_{out} = V_{in}$, where the input and output are normalized to the interval $[0, 1]$. While this can often be disabled, linearity is not a guarantee is not guaranteed in YUV buffers, even well-below saturation.

Figure 2.11: An image (a) before and (b) after lens-shading corrections, which digitally increase the brightness in the artificially dark corners. From Ref. [130].

Secondly, *lens-shading* is a method of digitally correcting for vignetting, that is, the natural decrease in brightness as distance from the aperture increases (Figure 2.11). Geometrically, the intensity of light passing through a small aperture coplanar to the sensor scales as $\cos^4\theta$, where θ is the angle between the pixel and the center of the sensor as measured from the aperture [125]. In practice, however, lens-shading values are typically calibrated by proprietary algorithms using image data. In some devices, these gains are coupled to the ISO gain, thereby appearing even at the RAW level.

Finally, *hot-pixel correction* is a mask for defective pixels that gravitate towards large values, often saturation. Certain types of hot-pixel correction may be helpful to Crayfis, but non-saturating hot pixels will frequently be overlooked in these algorithms. Both hot pixels and lens shading will be discussed in detail in Section 3.4.

The latter half of this chapter has shown that, although the response and noise of a typical pixel may be constrained by its physical construction, the detection efficiencies of cosmic ray secondaries are equally dependent on the software. Achieving a performance in line with the physical capabilities of the sensor requires both a precise handling of the camera APIs and a sophisticated and efficient trigger. Part II examines this element of the Crayfis experiment, from which the physical performance will be quantified in Part III.

Part II

Constructing the CRAYFIS Network

Chapter 3

The Android App

While CMOS sensors are the core of the Crayfis experiment, smartphone cameras are more than a slab of sensitive silicon. As the last chapter examined in detail, each frame is produced with multiple layers of digital processing. Intelligently querying the camera, efficiently accessing the buffers, and appropriately triggering the pixels are all necessary to harness the CMOS's full capabilities—and are no trivial matter. This chapter will explore the progression of the Crayfis app as it struggled through several efficiency tests and was subsequently enhanced on multiple fronts.

3.1 The initial release

By early 2015, the main features of the app were present, and a version was prepared for beta-testing. The app continually queried the sensor APIs for the timestamp, location, and orientation, and the back camera buffer for particle hits. When a hit was found, these components were bundled together and serialized using Google Protocol Buffers, and periodically, these events were uploaded to `crayfis.io` with the trigger settings that generated them.

In order to limit the data volume, the app saved single pixels rather than entire frames using nearly the simplest algorithm possible, outlined in Figure 3.1. The variables `thresh1` and `thresh2` characterize the level 1 and 2 triggers ("L1" and "L2"), determining the fractions of frames and pixels, respectively, that are saved. The former

```
int[] frameBuffer; // obtained from the camera API
int thresh1, thresh2; // calibrated separately

int max = 0;
for (int i=0; i < frameBuffer.length; i++) {
    if (frameBuffer[i] > max) {
        max = frameBuffer[i];
    }
}

if (max <= thresh1) return;

for (int i=0; i < frameBuffer.length; i++) {
    if (frameBuffer[i] > thresh2) {
        savePixel(i);
    }
}
```

Figure 3.1: Pseudocode for the DAQ trigger (2015 app)

was calibrated by the app to achieve the desired trigger rate—a hard-coded 1 frame per second—while the L2 threshold was set to `thresh1-1`.

Pixels passing this trigger were grouped with the corresponding sensor data and frame metadata and added to a queue of "`Event`s." Every two minutes, these in turn would be assembled into an `ExposureBlock`, a set of `Event`s with identical exposure and trigger settings, and would be uploaded to the server if Wi-Fi was available or cached otherwise; after this, the trigger thresholds would be recalibrated. This hierarchy is outlined below, with data structures in boldface:

ExposureBlock

- Run ID (associates with hardware, software information)
- Start/End time
- DAQ State (CALIBRATION/DATA)
- Sensor resolution
- Trigger configuration
- Dropped frames (e.g. due to CPU latency)
- **Event**(s)
  - Timestamp
  - GNSS latitude, longitude, altitude, precision
  - Mean, variance of all pixel values
  - Phone orientation vector
  - **Pixel**(s)

- Pixel coordinates (x, y)
- 8-bit brightness value
- Average value of adjacent pixels

This design for the DAQ was complemented with UI features and stabilized in late 2015 after a year of beta testing; yet, it remained unknown whether cosmic ray secondaries would trigger with a reasonable efficiency. Though an abundance of particle tracks were found in this beta testing period, it was unclear what fraction derived from cosmic rays rather than terrestrial radiation or pixel noise. Several experiments were devised to this end, and the results, while initially discouraging, set a clear course leading to a far more capable Crayfis app.

3.2 Early tests, early problems

To distinguish between cosmic ray secondaries and noise in the Crayfis app, three general approaches exist:

1. Tag particles with an independent, high-efficiency detector and measure the rate of coincidences.

2. Vary the flux and compare results. This can be done by taking data at different altitudes, either higher in the atmosphere (in an airplane or balloon) or below-ground.

3. Expose phones to a particle emitter (e.g. a beam or radioactive sample), from which the background can be subtracted, and the response to cosmic ray secondaries can be extrapolated.

The third method was the first to be attempted, using radioactive sources to estimate the photon efficiency [83]. The activities of three gamma-emitters— ^{226}Ra, ^{137}Cs, and ^{60}Co—were first measured with high-precision gamma spectroscopy; these sources were then kept at constant distance from a single phone in a wax assembly, producing a significant sensor response; see Figure 3.2. By comparing the number of

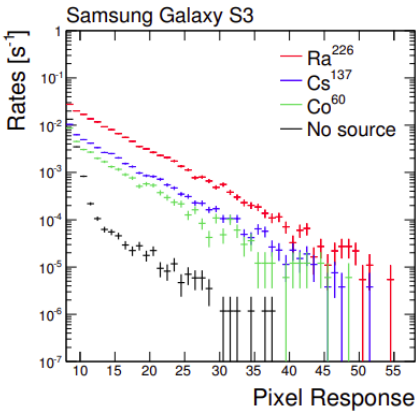

Figure 3.2: The response of a Samsung Galaxy S3 to three radioactive sources, each placed a fixed distance from the sensor, and with no source. From Ref. [83].

observed particle tracks to the computed gamma flux given an active sensor area of 2×10^{-5} m^2, a range of $(2.5 - 25) \times 10^{-9}$ m^2 was found for $A_\gamma \epsilon_\gamma$. These figures were likely an overestimate, as the ^{226}Ra activity was a factor of 3 higher than reported due to human error, and shielding from the lens alone may have been insufficient to block high-energy beta radiation, especially from the ^{226}Ra source. Within this $A_\gamma \epsilon_\gamma$ range, contributions from photons were essentially negligible for shower reconstruction under realistic levels of user adoption.

The reasons for this disappointing efficiency became more clear when the muon sensitivity was measured with the first of the three methods above. To identify cosmic muons, a hodoscope was constructed with three scintillator panels, and a Samsung Galaxy S4 running the CRAYFIS app was placed between the top two; see Figure 3.3. By comparing the times at which the scintillators and CRAYFIS app recorded muon candidates, and by computing the fraction of muons traveling through the scintillators which would also pass through the CMOS, the efficiency of the CRAYFIS trigger could in principle be reconstructed.

When the three scintillators registered a coincidence, an LED would blink into the camera 250 ms later, thereby consolidating the phone triggers and scintillator tags onto a single clock. These LED flashes could be distinguished from ordinary muon candidates registered by the app through the number of pixels in each event: while some longer particle tracks tracks were found in the sensor response, events with over

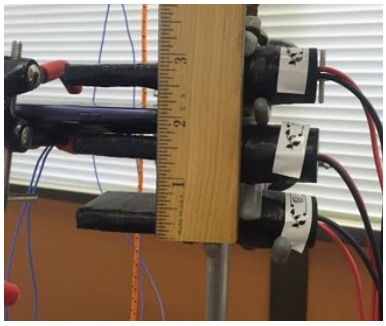

Figure 3.3: An image of the 2015 muon hodoscope construction, with a Galaxy S4 suspended between three scintillators.

35 excited pixels were derived exclusively from LED signals (Figure 3.4b). However, when the sensor area illuminated by the LED spanned across the position of the rolling shutter, the same LED flash would appear in consecutive frames, requiring consolidation of these LED timestamps.

The sample of muon candidates in the CRAYFIS app was substantially less clean, however. When LED flashes were removed, the vast majority of triggered events came from a small number of hot pixels (Figure 3.4a). Though these could be removed offline for data analysis, the damage to the CRAYFIS trigger was not entirely mitigated, as these these hot pixels significantly inflated the app's L1 thresholds, potentially excluding muon tracks. Attempts to override the device's calibration and manually reduce the thresholds were no more successful: the surplus of hot pixel events was a burden on the CPU, leading to skipped frames.

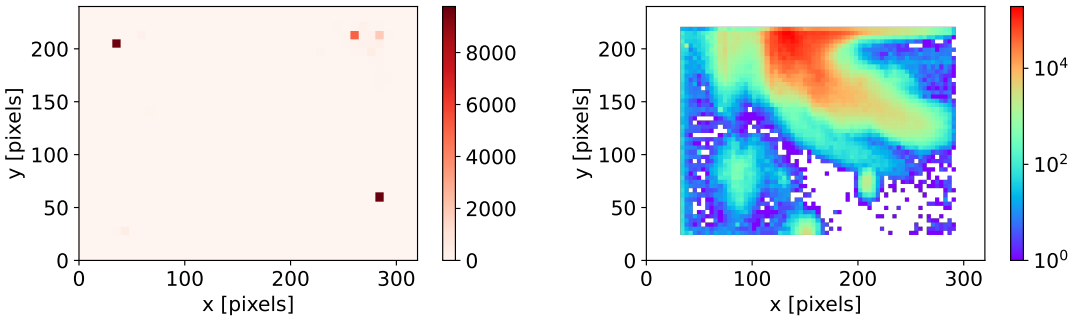

Figure 3.4: Two-dimensional histograms, in pixel coordinates, of events with (left) less than 10 pixels above threshold, and (right) greater than 35 pixels above threshold. The former is composed almost entirely of hot pixels, while the latter shows the various positions of the LED as adjustments were made.

Table 3.1: Efficiency loss from downsampling in a Galaxy S4, assuming a 2×2 binning scheme.

Downsampling losses			
Resolution	Dimensions	Physical pixels used	% ignored
Full	4224×3176	13.4M	0
1080p	1920×1080	8.3M	38
720p	1280×720	3.7M	73
Medium	640×480	1.2M	91
Low	320×240	300K	98

Even beyond hot pixels, however, the CRAYFIS trigger still fared poorly. The majority of triggered pixels were located along the edges of the sensor; due to lens-shading enhancement, the muon signal competed against digitally augmented noise. A mask was consequently applied to this border: the triggering region was restricted to the middle 80% of pixels in both x and y, reducing the efficiency by 36% outright.

However, the most significant deficiency in the CRAYFIS trigger was camera resolution. The Galaxy S4, running comfortably at low resolution (320×240) would map each visible pixel to a grid of physical pixels roughly 12×12 in size. This does not, however, imply that all 144 physical pixels would be used in computing each pixel value. While the Galaxy S4 sensor has some degree of binning [131], many of these pixels may also be skipped: as shown in Table 3.1, at least 98% percent of physical pixels will be ignored in a 2×2 binning scheme. However, larger bins would not necessarily be an improvement, as this would greatly dilute the signal.

After 44 days of running under this configuration, the results were consistent with zero efficiency. A histogram of all possible pairs of $\Delta t = t_{\text{LED}} - t_\mu$ was constructed, using LED timestamps both before and after each muon candidate and a bin size equal to the frame duration. The bin counts for negative, unphysical Δt were then subtracted from those of positive Δt, thereby removing the combinatorial background from spurious phone-scintillator coincidences; see Figure 3.5. No statistically significant excess was found in bins 3-4 as anticipated (at 15 FPS, the LED delay was 3.75 frames); instead, only 20 phone-scintillator pairs were observed, compared to a possible 3500 at perfect efficiency.

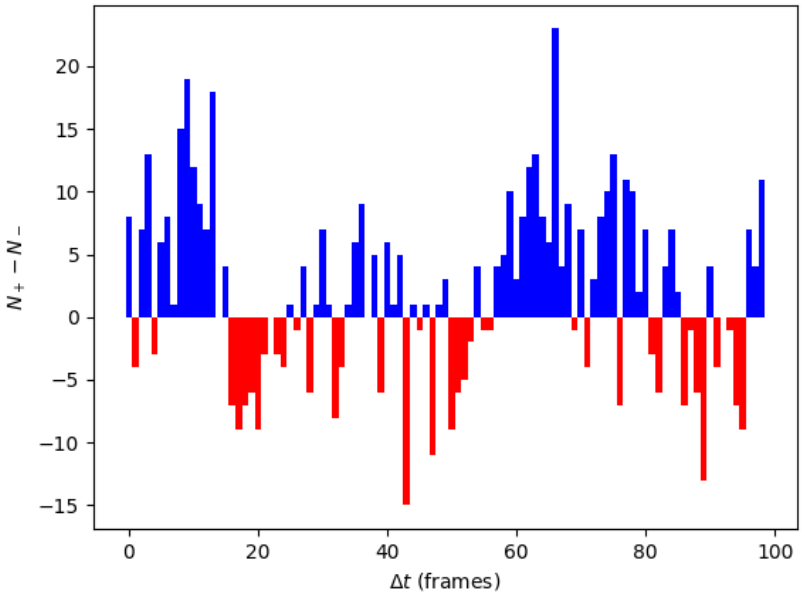

Figure 3.5: Difference between bin counts $N_+ - N_-$ for $\Delta t_\pm = \pm(t_{\text{LED}} - t_\mu)$.

These measurements were consequently attempted with a higher resolution, as increasing to 720p or 1080p would reduce the fraction of ignored pixels (again assuming a 2×2 binning scheme), to 72% or 38%, respectively. However, at these higher resolutions, the phone would overheat, cease charging, and the battery would quickly die.

3.3 Streamlining the app

Clearly, the app was in need of three major modifications:

- A drastically reduced burden on the CPU, allowing for high-resolution streaming and more sophisticated trigger algorithms, such as:
- Online hot pixel masking, and
- Spatially dependent thresholds to correct for lens shading.

Decreasing the CPU consumed by the app was the first obstacle to be overcome if Crayfis was to be salvaged. Although a few poorly designed segments of code led to inefficiencies, over 90% of the CPU load resulted from two design choices: disabling

sleep mode to keep the DAQ running, and triggering frames through a simple loop over frame buffers.

The app was first refactored into separate background (DAQ) and foreground (UI) components, allowing the DAQ to continue after the phone entered sleep mode. This eliminated the burden of both the Crayfis UI and system foreground processes during nightly runs. To assure that, as a background process, the DAQ would not run indefinitely and consume the user's battery, frames were additionally screened for a flat camera orientation: multiple consecutive fails would stop the DAQ if the UI was inactive.

Frame processing was likewise made far more efficient by employing the GPU through Android's Renderscript APIs. By design, the GPU is optimized for asynchronous operations across large arrays, and was able to perform the same trigger operations far more efficiently than the Java layer. In particular, Renderscript featured a highly-optimized method for constructing histograms of frame buffers, which reduced processing time for the L1 trigger by two orders of magnitude while simultaneously providing more detailed data. However, Renderscript is specifically designed for asynchronous, element-wise computation such as array-to-array mappings. Several operations during calibration fell outside this domain; here, native OpenCV [133] was substituted to improve efficiency.

To evaluate these improvements, battery temperature was queried with the other sensor fields. With this vastly improved pipeline, a number of phones in the 2012-2014 generation—which previously would exceed 45℃ at 1080p and 30 FPS—instead reached a plateau at 35-37℃. As a safeguard, a shutoff/cooldown mechanism was implemented for temperatures above 41℃. With this reduced CPU load, the app could devote more processing power to masking hot pixels, correcting for lens-shading, and ultimately decreasing thresholds over the majority of the sensor.

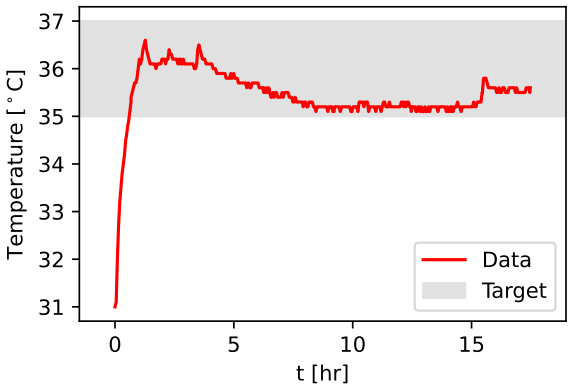

Figure 3.6: Temperature fluctuations in a Samsung Galaxy S6 using a grid of scale-factors for triggering. A plateau is observed in the same temperature band as was observed with a flat trigger.

3.4 A more sophisticated trigger

In both hot pixel and lens-shading corrections, three distinct steps were required:

- A short online calibration
- Implementation of the calibration in the trigger,
- Server-side adjustments to these calibrations with larger sets of uploaded data.

For the second of these, a common solution was found for both calibrations. By multiplying each frame buffer by a calibrated grid of scale-factors, trigger rates in the high-gain edges could be dampened, while individual hot pixels could be suppressed entirely. As Renderscript is optimized for such operations, this advanced triggering produced very little impact on the run temperature; see Figure 3.6.

3.4.1 Lens-shading corrections

To calibrate this grid for lens-shading, a simple model is employed. All pixels are assumed to have identical analog noise distributions X, which are modified by spatially dependent digital gains λ_i, normalized such that $\min_i \lambda_i = 1$. These gains, and the corresponding corrections $w_i = \lambda_i^{-1}$, can be found by measuring the relative responses of each pixel in dark conditions. To first order, the gain factors are proportional to

the responses:

$$E(\lambda_i X) = \lambda_i E(X) \propto w_i^{-1} \,. \tag{3.1}$$

However, as the pixel responses are digitized; the computed mean of each pixel is in fact $E(\lfloor \lambda_i X \rfloor)$, for which the equality above does not hold. Particularly when the computed means were low, this simple algorithm did not adequately correct for lens-shading.

To enhance this approach, a functional form of $\lambda_i X$ is assumed, from which the relationship between $\lambda_i E(X)$ and $E(\lfloor \lambda_i X \rfloor)$ can be inferred. Denoting the probability density function of the undigitized noise $\lambda_i X$ as $f_i(x)$, the mean of the truncated responses becomes:

$$E(\lfloor \lambda_i X \rfloor) = \sum_{n=0}^{\infty} \int_n^{n+1} dx \, n f_i(x) = \sum_{n=1}^{\infty} \int_n^{\infty} dx \, f_i(x) \,. \tag{3.2}$$

This is in general difficult to evaluate; however, with a convenient assumption that $X \sim \mathrm{Exp}(1/\mu)$, (3.2) can be simplified, yielding:

$$E(\lfloor \lambda_i X \rfloor) = \frac{1}{e^{\frac{1}{\lambda_i \mu}} - 1} \,, \tag{3.3}$$

which, inverted, provides a closed-form correction for lens-shading:

$$(\lambda_i \mu)^{-1} = \log(1 + E(\lfloor \lambda_i X \rfloor)^{-1}) \propto w_i \,. \tag{3.4}$$

To justify that this is not overly sensitive to the arbitrary choice of noise model, this result is numerically evaluated for other models as well. As mentioned in Chapter 2, pixel noise is the sum of two independent contributions—Poisson-distributed dark current and Gaussian readout noise—both amplified by a gain factor. Gaussian noise models with $\sigma^2 = g\mu$ are therefore used to approximate a continuous analog of the Poisson distribution, with an applied gain g in pixel values per electron. The results are plotted in Figure 3.7 for $g^{-1} = \{1, 4, 16\}$ alongside the exponential result. As

expected, the curves converge when the measured means are large, but display quite different behavior near the origin. In this limit, the exponential model falls between the Gaussian models with gains of 1 and $\frac{1}{4}$, and hence should perform reasonably well for gains on the order of a few electrons per 8-bit pixel value. In the sample of phones used for testing, the camera APIs typically reported gains within this range for RAW buffers near the maximum analog ISO.

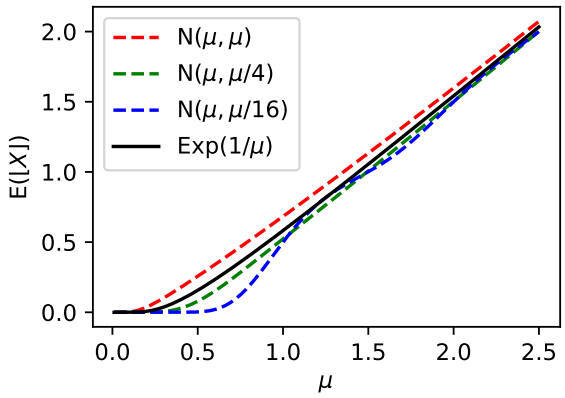

Figure 3.7: $E(\lfloor X \rfloor)$ vs. $E(X)$ for exponential noise and Gaussian noise with three different variances.

Scale-factors were computed by applying (3.4) to downsampled blocks and applying bicubic interpolation to the results. The resulting pixel triggers were more well-divided across the sensors of most test phones used, with the exception of the yet unquenched hot pixels. Whereas flat L1 thresholds on a typical Galaxy S4 oscillated between 9–10, the scaling algorithm reduced the effective thresholds to a range of 3–6 (see Figure 3.8) across the majority of the sensor while maintaining the same trigger rate.

3.4.2 Hot-pixel masking

While rigorous statistical methods such as a Kolmogorov-Smirnov test were first considered for the masking algorithm, many hot pixels fired too infrequently for these metrics to attain significance without a prohibitively large sample size. Rather, a more practical approach was needed: pixels firing at relatively high values above a

[1]As the highest pixel values could be caused by cosmic rays or terrestrial radiation, the second-highest value was more robust.

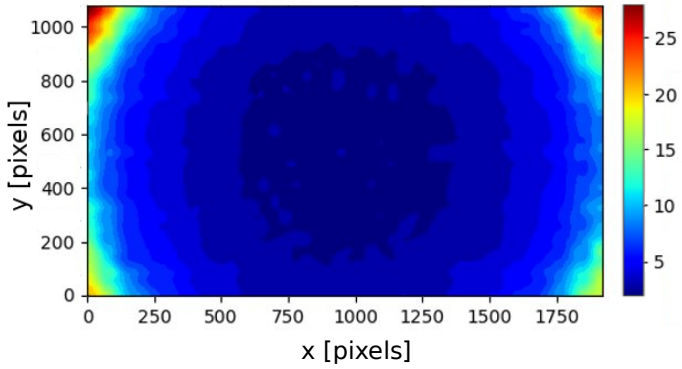

Figure 3.8: Effective L1 thresholds on unscaled digital pixel values (shown on color axis) in a Samsung Galaxy S4 using the calibration scheme above.

certain frequency were the intended targets. The second highest value[1] each pixel attains over a set of calibration frames proved to be a simple but effective metric for hot pixel classification; outliers in this distribution are accordingly masked. An initial calibration attempt, sampling 5000 frames and masking 0.01% of pixels, identified the most prominent hot pixels and decreased the thresholds as expected; however, the trigger rate was then determined by 20–30 less-frequent hot pixels in place of the most active 2 or 3. A more intensive calibration with 50,000 frames and 0.2% of pixels masked decreased the thresholds further, but likewise produced a sensor response dominated by a small subset of pixels, many of which had not once exceeded the trigger thresholds in the calibration data.

For a deeper cleansing, a far larger sample size was required, one more suited to the server. The on-device calibration would only serve as a first-pass. Consequently, these cuts were not unduly aggressive: the hot pixel tail was bound by a threshold of $< 0.2\%$ of pixels per bin, correlating the number of pixels masked to the kurtosis of the distribution. To prevent excessive cuts in irregular data, this threshold is restricted such that no more than 1% of total pixels are masked. Figure 3.9 illustrates these cuts in a Samsung Galaxy S7.

3.4.3 Calibration data

These calibrations were vital metadata that needed to be preserved, but the full grid of scale-factors could occupy several megabytes—with a million phones in active use, each

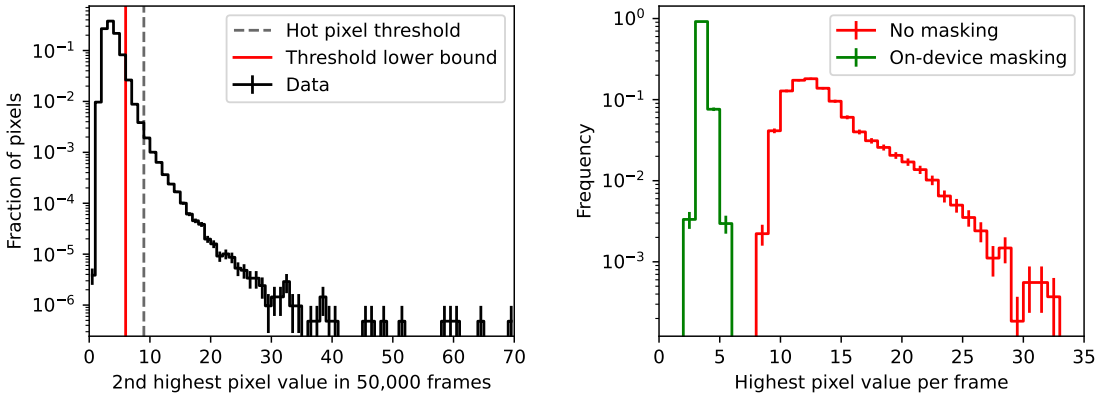

Figure 3.9: Left: a histogram of second highest pixel values in a Samsung Galaxy S7, used in hot-pixel calibration, with differential and integral thresholds. Right: resulting distribution of highest pixel values per frame, used to calibrate the L1 trigger. No lens-shading calibrations are applied in either plot.

submitting one grid per day, the experiment would reach the petabyte scale far more quickly than desired. As the calibration for hot pixels was particularly time-intensive—spanning 50,000 frames—repeated on-device calibrations would also unnecessarily limit Crayfis's duty cycle. Several modifications were therefore necessary.

First, the frequency of on-device calibrations was reduced. As changes to the underlying lens-shading pattern are unlikely, and hot pixels are mostly stable on the order of weeks, calibrations could be cached in device storage and reused with minimal reduction in data quality. These cached results are considered valid for two weeks, after which they are discarded.

Given the smoothness of the lens-shading pattern, downsampling and compression also were effective in minimizing data volume. By downsampling the lens-shading corrections, re-normalizing the values as 8-bit integers, and compressing the resulting grid as a JPEG, the total size of a typical lens-shading calibration was reduced to less than a kilobyte with no more than a 0.5% loss in precision. To keep the metadata consistent between the phones and the server, the app draws its lens-shading corrections from the same JPEG buffer which is sent to the server. Likewise, the locations of masked pixels are stored in a sparse format, from which the full grid of scale-factors can be reconstructed.

All of these models and algorithms, however, only amounted to a successful imple-

mentation of the first item at the beginning of the section: a reasonable initial guess from a relatively small sample of frames. To further enhance the lens-shading corrections and to check the emergence of new hot pixels, a feedback loop on the server was necessary. This next step is presented in the following chapter.

3.5 Other additions

The modifications described above addressed the main problems presented in Section 3.2, but as the app was being modified, a number of other improvements were made. These included not just enhancements to the DAQ, but to the user experience as well.

First, the camera pipeline was migrated to the upgraded `camera2` API, granting manual control over latency between exposures (which previously, may have been substantial) and over more of the digital processing discussed in Section 2.5. This API also provided direct access to firmware-level frame timestamps, which are significantly more accurate than those reported by application threads when the CPU is under stress.

Two new datatypes characterizing the sensor response were also added to the Protocol Buffers. The descriptively named `ZeroBiasSquare` was the output of a new trigger, which periodically saved 10×10 blocks of pixels at random locations; from this data, the spatial noise dependence could be quantified on both large and small scales. A new L2 trigger algorithm (`ByteBlock`) was also implemented which, rather than saving the mean of the surrounding pixel values, instead saves the individual values themselves. Because only the 8-bit values of surrounding pixels are saved and not their coordinates (which are determined by the order in which these values are saved), this algorithm yields more useful data than the default trigger without a significant increase in data volume.

Several other features were added to make the app as convenient to the users (and consequently, to the experimenters) as possible. To assure regular, high-quality data

with minimal user interaction, the user is prompted upon loading the app for the first time to schedule daily data-taking intervals, such that when the phone is charging in between certain hours of the night, the app will automatically start. Then, rather than assume the rear camera is covered, the app uses the gravity sensor to select whichever sensor is facedown. If it is determined that the phone is not lying flat, the user will be prompted to reposition the phone. If this fails after several attempts (such as when the phone is being actively used apart from the Crayfis UI), the app will reattempt several minutes later if the phone is still charging; otherwise, the app will quit. This assures that the Crayfis app does not interfere with the user's phone use and does not unnecessarily discharge the battery, both of which are essential to a positive user experience and acquiring a large user base.

Lastly, a cryptographic hash was instantiated to verify the data received by the server, computed with a salt appended to the uploaded request body. This prevents malicious data from contaminating the experiment: without the salt, an attacker is unable to create passable data that the server will accept. With this safeguard, the app can be safely open-sourced, allowing for the public to participate as developers as well as users in accord with the democratizing spirit of Crayfis.

Chapter 4

The CRAYFIS backend

For Crayfis, the processing pipeline does not stop once the data is uploaded—a vast architecture exists beneath `crayfis.io` that at least equals that of the Android app. This was not, however, strictly necessary—the data could have simply been deposited in the cloud and retrieved by a local cluster for analysis—but some initial benefits steered the experiment in this direction. At first, it was realized that a Crayfis website was useful for user engagement and could be employed to manage beta-testers and distribute the app. Once the website was established, it was natural to use it as an endpoint for data submission: this would allow blacklisting of spurious data and tracking of user statistics. Then, if the users could be associated with their data, a UI could be built around user scores, data tracking, and leaderboards, supplying users with abundant incentives to run the app.

All of these features were available in 2015 as beta testing was underway. The insight missing at that time, though, was that the megabytes of data collected by the server from each phone could be used to enhance the quality of the DAQ. A mechanism had already been developed for passing commands to the app—adjusting the resolution or frame rate, for example—for remote debugging; this feature needed only to be automated and grown to a much larger scale.

4.1 Web app components

Before these modifications can be discussed, a sketch of the web app will prove useful—an illustration is also provided in Figure 4.1 below.

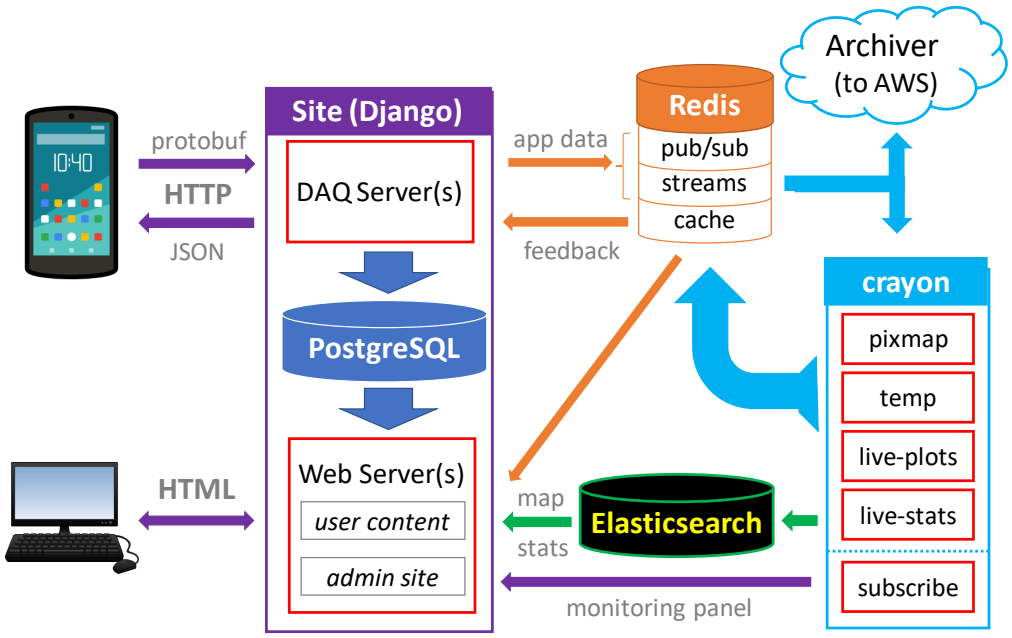

Figure 4.1: A schematic of the current web app infrastructure.

The web app is composed of the CRAYFIS website, three databases (Redis, PostgreSQL, and Elasticsearch), five types of "`crayon`" (CRAYfis ONline) worker nodes, and the "Archiver," a node routing data beyond the web app into offline data storage. These components are containerized through Docker and run in a self-scaling and self-healing Kubernetes ecosystem on the Amazon Web Services (AWS) cloud, a design discussed in Section 4.4.

At the simplest level, the web app receives data from the Android app, sends this data to long-term storage (discussed in Section 4.5), and processes the data to generate user content on the website and adjustments to each phone's DAQ settings. The three databases assist with this flow in different ways. PostgreSQL stores structured experimental metadata, linking each run to a particular user. By design, it can efficiently process simple queries, such as the statistics populating each user's personal dashboard. Elasticsearch handles much of the same data but is capable of more sophisticated queries, which will be discussed in Section 4.3. Redis, by contrast, stores

unstructured data and occupies several diverse roles in the web app ecosystem. Most prominently, uploaded application data is channeled through the web app via two Redis Streams: queue-like objects that support load balancing between copies of the same worker node. One Stream leads to offline storage through the Archiver, while the other routes data into the five `crayon` workers for online processing.

The `crayon` workers can be divided into three groups. The first two in Figure 4.1, `crayon-pixmap` and `crayon-temp`, are feedback loops discussed in the next section, which compute adjustments to particular devices and store the results in Redis. The next two, `crayon-live-plots` and `crayon-live-stats`, are responsible for generating figures on the website—the former by caching serialized plot data in Redis and the latter by populating Elasticsearch. Finally, `crayon-subscribe` is not tied into the Redis Stream, but instead uses Redis's publish/subscribe feature to create a detailed live-monitoring interface on the admin site for devices of interest.

4.2 Server-side trigger corrections

As discussed, the main goal of the web app development was a feedback loop to modify the existing calibrations in Redis based on incoming app data. This required two elements: an algorithm for computing corrections and an interface for the phone to retrieve the results.

4.2.1 Adjustments with crayon-pixmap

A new worker node, `crayon-pixmap`, accomplished the first of these by tracking the spatial distribution of triggered pixels. Trigger rates for each device are recorded in sparse matrices with the necessary single-pixel precision for hot-pixel masking; however, as more pixels are triggered over months of data-taking, a sparse format becomes increasingly memory-intensive. To avoid wasting memory on pixels with normal trigger rates, a rolling average is maintained when the number of `ExposureBlock`s exceeds 250. Explicitly, if a set of k new `ExposureBlock`s with trigger occupancies $n(x, y)$ is

added to the existing sparse matrix $N_i(x, y)$, the updated matrix $N_{i+k}(x, y)$ is given by:

$$N_{i+k}(x, y) = \frac{250}{250 + k}(N_i(x, y) + n(x, y)) \,. \tag{4.1}$$

Values below 0.01 are then purged to limit the number of nonzero entries. Pixels with trigger rates above a certain threshold are then masked; currently this limit stands at 15 per 250 `ExposureBlock`s, or 1 trigger in 60,000 frames.

Adjustments to lens-shading calibrations required a slightly different approach. These are evaluated based on the uniformity of the spatial distribution of triggers,[1] though the digitization of the pixel response makes a perfectly uniform distribution impossible for non-trivial lens-shading. As the underlying lens-shading calibrations are downsampled from the total sensor resolution, a histogram with the same downsampled dimensions was sufficient to compute the needed correction—a sparse matrix and the rolling-average approach were therefore unnecessary. However, as a downsampled histogram was vulnerable to bias from a cluster of pixels triggered by a single particle, bin counts are incremented by no more than 1 per frame.

Evaluating the uniformity of the trigger distribution also required a nuanced approach. For a sizeable histogram, a standard χ^2 test would require a large sample of triggers to achieve significance—at least a week of runtime between iterations. While this is tolerable for fine-tuning the result, more significant faults with the initial calibration require faster interventions. Instead, this statistical distance can be computed more rapidly by noting that each trigger count in the histogram follows a Poisson distribution with rate $\lambda(x, y)$; hence, if the response were spatially uniform, these rates would be identical:

$$\lambda(x, y) = \lambda = \frac{N_{\text{triggers}}}{N_{\text{bins}}} \,. \tag{4.2}$$

Therefore, the distribution of triggers per bin—specifically, its statistical distance from

[1] In fact, perfectly correcting for lens-shading may not produce a uniform response due to spatially-

a Poisson distribution—can also be used as a measure of spatial uniformity.

The simplest statistic to measure this distance is the variance-mean ratio of the distribution of trigger occupancies. As is well-known, a Poisson distribution has a variance equal to its mean; however, when there is variability in the means $\lambda(x, y)$ being sampled, this will produce a wider distribution than a standard Poisson. As demonstrated by Figure 4.2, when the $\lambda(x, y)$ are considered as rates and $\mu = \lambda\Delta t$, the variance-mean ratio will increase over time; hence, this ratio can be used to quantify the statistical significance of the non-uniformity.

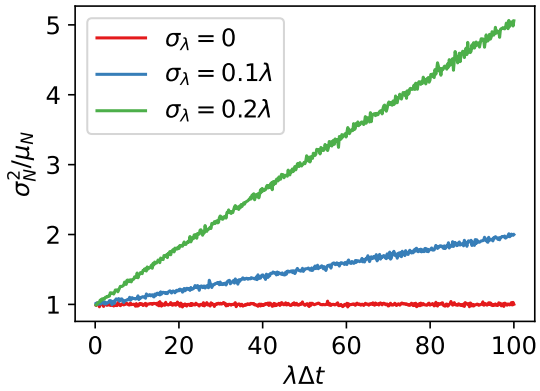

Figure 4.2: 10,000 Poisson rates λ_i are drawn drawn from a distribution $\Lambda \sim N(\lambda, \sigma_\lambda^2)$. For various values of Δt, a single sample N_i is drawn from each of 10,000 Poisson distribution with means $\lambda_i \Delta t$. The resulting variance-mean ratio σ_N^2/μ_N is plotted against runtime for three values of σ_λ.

This insight is implemented as follows: when the variance-mean ratio for the distribution of trigger occupancies exceeds a threshold (currently $\sigma^2/\mu > 1.8$), the trigger non-uniformity is significant enough to apply a correction. Then (4.2) is used to find λ, the mean trigger count per bin, and the corresponding Poisson CDF is used to calculate p-values for each individual bin count. These p-values in turn determine the adjustments to the lens-shading calibrations $w(x, y)$ via a logit (inverse logistic) response function:

$$\Delta w(x, y) = \frac{\alpha}{\sqrt{N}} \log(p(x, y)^{-1} - 1) \tag{4.3}$$

dependent thermal noise. In this case, the pixels with higher signal-to-noise ratios would be under-represented, which is most likely undesirable for Crayfis.

where α sets the rate of convergence and N is the total number of triggers in the histogram. This function has the desirable property that corrections approach positive and negative infinity as the p-value approaches 0 and 1, respectively; furthermore, as the trigger response becomes more uniform, larger amounts of data will be required to trigger a correction, and thus the $N^{-1/2}$ dependence will lead to smaller adjustments. To avoid overcorrection, the magnitude of $|\Delta w|$ is limited to 32 when scaled to a $[0, 255]$ range; conversely, as a check against undercorrection for smaller sample sizes, Δw is set to 1 for empty bins. Then, in order to renormalize the results, an offset is applied such that at least 5% of the resulting calibrations $w(x, y)$ are saturated at 255.

However, one more critical feature was needed. Because the L1 thresholds were calibrated to reduce the trigger rate below a fixed target, a single hot pixel, firing at a constant value above this target rate, could set the threshold without once being triggered.[2] Consequently, such a hot pixel would never appear to `crayon-pixmap`, and so could never be masked. To make such below-threshold hot pixels visible to the server, a modified L1 trigger was introduced in the Android app. This trigger oscillates between two adjacent threshold values so as to achieve the target pass rate, thereby overcoming such quantization barriers.

Figure 4.3 shows the stages of the calibration process for two physical devices streaming at 1080p. Each plot shows the positions of roughly 20,000 triggers with no masking or scaling (top), using the first-pass results from the devices (middle), and after server-side refinements from 7 days of continuous running (bottom), while not accepting any further refinements. The first-pass calibrations performed relatively poorly in these particular phones, slightly over-correcting for lens-shading and, as a consequence of the order of calibrations, leaving many less-active hot pixels in the center of the sensor. The final round of plots shows a nearly uniform sensor response—remarkably so, given the significant quantization effects inherent to low pixel values.

This does not guarantee, however, that any of these methods are optimal, or will necessarily converge on all device models, present and future. This would require

[2]While this scenario may seem contrived, it can be particularly troublesome in RAW buffers where a single hot pixel saturates in every frame, pushing the thresholds up to saturation as well.

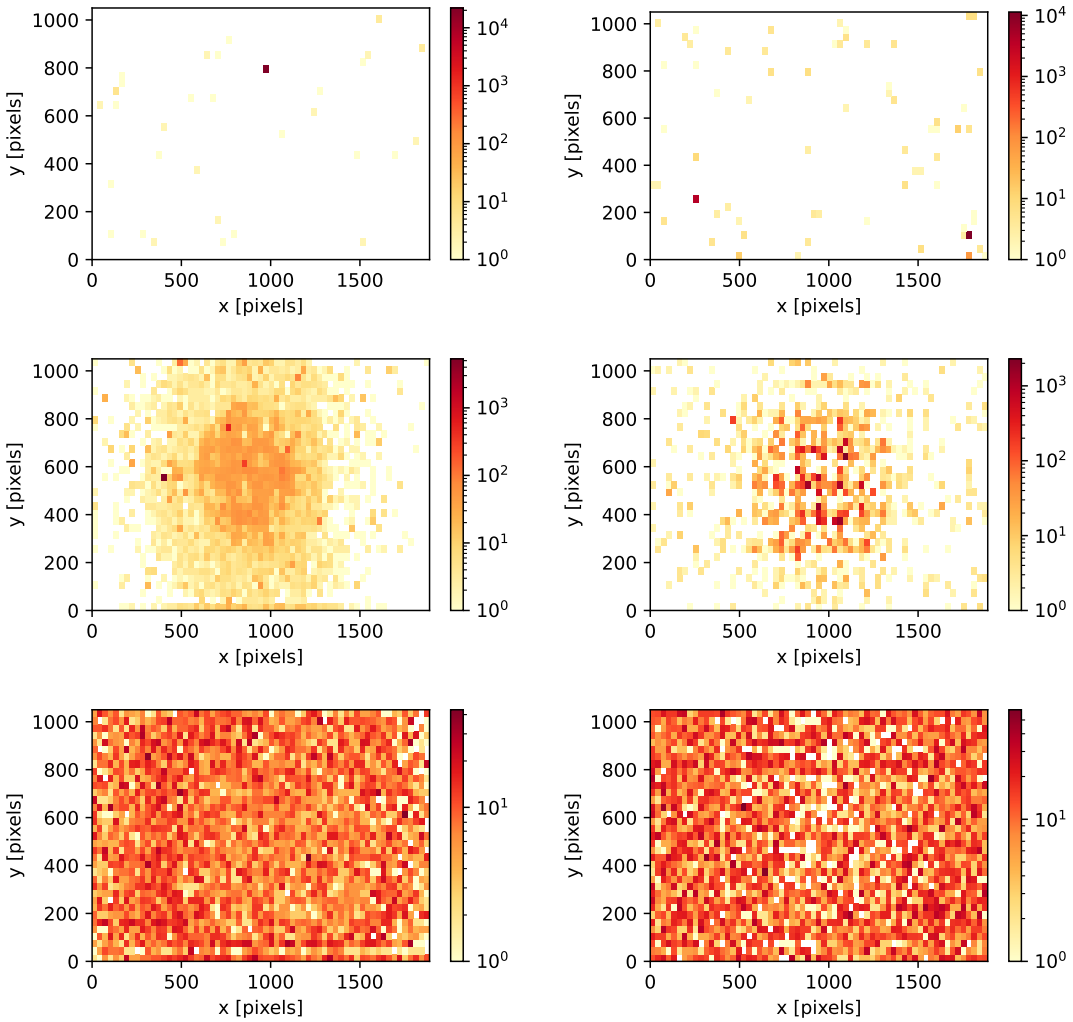

Figure 4.3: The results of the crayon-pixmap algorithm on two phones, shown in the left and right columns. Each image is a histogram of triggered pixel positions, with axes in 1080p pixel coordinates. Description in text.

both a larger set of beta testers and a more appropriate architecture for analyzing larger offline datasets. Nonetheless, an emulator was designed to test the algorithm's robustness on non-trivial lens-shading patterns, such as in Figure 4.4. A realistic design was used to generate sample data: the assigned lens-shading pattern and hot-pixel locations determined the emulator's spatial PDF as a function of threshold; then, as in the production app, a threshold was selected to achieve a target trigger rate. In all cases tested, `crayon-pixmap` converged on a near-uniform solution within several days of simulated runtime.

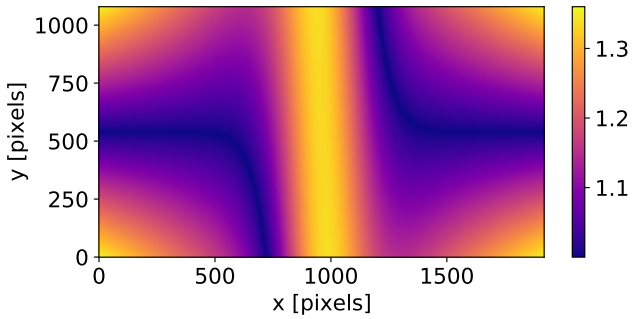

Figure 4.4: An example lens-shading pattern tested through an emulator. This pattern is given explicitly by: $\lambda(x, y) = 1 + \frac{1}{3}\left|\frac{xy}{100} - e^{-x^2/3}\right|$, where x and y are scaled to the ranges $[-16, 16]$ and $[-9, 9]$, respectively.

4.2.2 Persistence of pixel calibrations

As mentioned in Section 3.4, each device generates initial estimates of the hot-pixel and lens-shading corrections, and the results are saved to the phone's memory for future access. For the server to edit these calibrations, however, it was appropriate for the primary copy to be held by the server (specifically Redis) instead. The Android app was modified to use the server as a first recourse for loading lens-shading corrections and hot pixels; only if this failed (e.g. from connectivity issues), would the phone need to use its own copy of these calibrations. Then, if neither of these attempts succeeded (e.g. because no data exists for a particular camera or resolution) a device would compute these calibrations manually.

To implement this strategy, a new server endpoint was constructed to which a phone would post its unique identifier, the camera in use, and the sensor resolution,

and the server would respond with a serialized set of calibrations. This simplified the process of updating calibrations: rather than attempting to post new values to a phone directly, the server would update its values internally and notify the phone that new calibrations are available. If the identifiers for the phone's calibrations and for those of the new update differed, the phone would simply pull the new version from the server endpoint. This design ensured that missed updates would be handled gracefully, and that devices would never need to update and recalibrate unnecessarily.

4.2.3 Temperature management

A second feedback loop was instantiated on a new `crayon-temp` node to regulate each phone's pixel processing rate, i.e., the frame rate multiplied by the sensor resolution. Though the Crayfis app would ideally sample frames at the maximum resolution and frame rate, this inevitably overheats the phone as discussed in Chapter 3. High battery temperatures should trigger a response which lowers the resolution, frame rate, or both. As the relative priority of the frame rate and resolution, as well as the range of acceptable temperatures, is primarily determined by the phone model, the server was found more suitable for this task.

The `crayon-temp` process saves battery temperature statistics in the Redis database for the most recent 250 `ExposureBlocks` of each device. If the temperature exceeds 41℃ on a particular phone or the time-average exceeds 37℃, the phone is instructed to lower the pixel processing rate. Conversely, if the average temperature decreases below 35℃, the pixel processing rate is increased. Because the relative importance of time resolution and pixel resolution is yet unknown, currently the resolution alone is adjusted and the frame rate is kept constant; by implementing this algorithm on the server, however, these priorities can easily be tailored to each phone model, both manually and automatically.

To test this algorithm, a temperature feature was added to the emulator. For a given emulated device, the relationship between the pixel-processing rate r and plateau temperature $T_0(r)$ was first constructed using a power model with randomly

selected parameters in a realistic parameter space. Over the course of an emulated `ExposureBlock`, the change in device temperature is given by:

$$\Delta T = k(T - T_0(r)) + n \tag{4.4}$$

where k sets the rate of the temperature convergence and n is a normally-distributed noise component with zero mean. On all emulated devices, the generated temperatures quickly reached a resolution with inside the requested range—a result replicated on several physical devices.

4.3 Managing the experiment

To develop these feedback loops, a number of useful tools for the experiment were generated as by-products: most notably, a monitoring application built into admin interface of `crayfis.io`. Though a monitoring site had previously been constructed for the experiments of Section 3.2, it existed outside the scope of the web app. Integrating these two applications was not only instrumental for developing feedback loops, but made a number of new features possible as well.

The new monitoring interface was powered by a new "`crayon-subscribe`" worker. With this tool, experimenters could subscribe to specific devices—whether real phones or emulated devices on a test cluster—by entering their unique device identifiers into a form or by searching for devices on the main website with an admin account. Data submitted by these devices would thereafter be cached in Redis, from which detailed graphs of performance metrics could be created on the admin site; see Figure 4.5.

Moreover, integration with the web infrastructure opened up a wealth of new potential for the monitoring app. In particular, linking the monitoring interface with Elasticsearch allowed both mass subscriptions with targeted queries and mass feedback to certain device classes. This was implemented through an "ElasticCommand" application, which allowed, for example, all phones of a specific model or region, or a set of phones exhibiting unusual behavior, to be collectively monitored and manip-

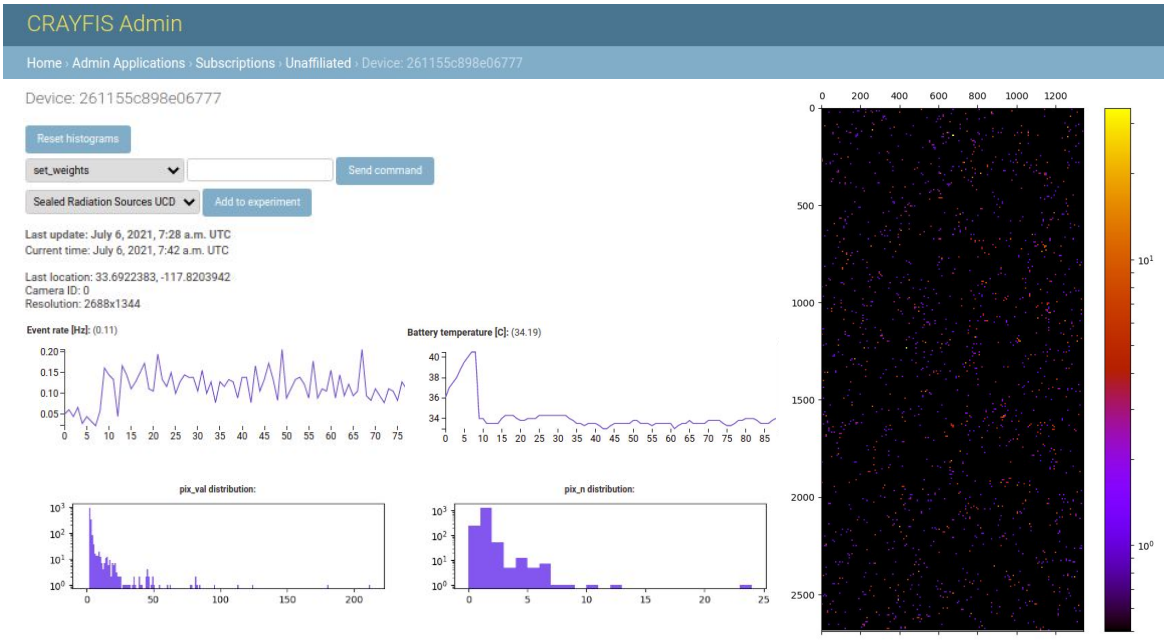

Figure 4.5: A partial view of the monitoring panel for a particular device. Recorded metrics include average pixel values, battery temperature, and thresholds vs. time, as well as histograms of pixel values and hit coordinates.

ulated remotely. Moreover, this interface can be used to optimize and standardize performance: for instance, when the lens-shading pattern of a particular phone model is well-known through controlled laboratory testing, the computed corrections can be directly applied to all phones of that model, bypassing the standard online calibration entirely. This principle can also be extended to optimizing frame rates, resolutions, trigger rates, ISO gain, and even the camera buffer settings particular to a specific model.

PostgreSQL also proved useful in conjunction with this new tool. As the primary store of experiment metadata—records of user- and device-level interaction with the web app—it was a natural choice for an ElasticCommand logbook, in which every device modification made through the admin website would automatically be stored.

4.4 Kubernetes and cluster management

With this new infrastructure made production-ready, the cloud architecture was also revisited during its implementation. The web app was migrated to Kubernetes, a

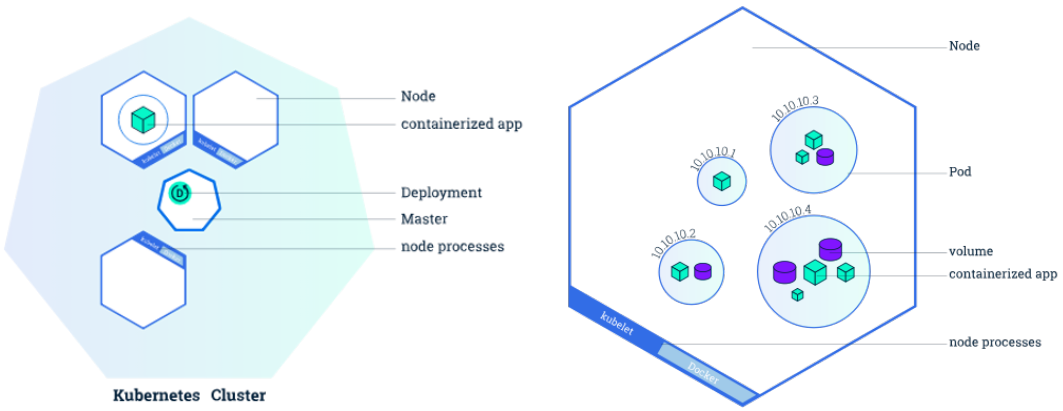

Figure 4.6: An illustration of a Kubernetes cluster. Each node is able to run several pods, each of which is managed by a master deployment controller. Images from [132].

cluster management platform designed by Google with a host of advantages. In a Kubernetes cluster, each node is able to run several *pods*, units with their own private IP addresses containing one or more containerized applications or volumes (see Figure 4.6). This allowed the entire web app—the three databases, five crayon workers, and two load-balancing website containers—to run on just three physical nodes for beta-testing. Furthermore, one of these (the *master* node) runs as a deployment controller, intelligently scaling the application with surges in traffic and restarting pods if they are terminated. Such a cluster is thus able to operate in a cost-efficient way with very minimal oversight.

Most of the cluster design was trivial: the individual containers were exposed to the correct ports, connected to the proper AWS credentials, and backed with underlying Elastic Block Store (EBS) volumes on AWS. However, several other additions were prudent. A high-level cluster management interface was implemented: operations such as deploying new additions, updating credentials, and restoring databases from AWS volumes in the event of a crash were made as simple as possible for future experimenters. As a further measure, redundancy in the data was achieved by scheduling periodic snapshots of the EBS volumes on the AWS cloud. These snapshots were designed to not be tagged as part of the cluster, thereby remaining safe from accidental cluster deletion.

4.5 Long-term data storage

While Kubernetes had many advantages, it unfortunately prevented the data from being offloaded in the same way as the initial cluster. Before this migration, the data was temporarily stored on the website servers' file systems and moved onto a local machine, nicknamed "`craydata`". Kubernetes, however, was not compatible with this simple design. The cluster publicly exposed the website containers as a single *Service* containing a public IP address differing from the pods' private addresses; retrieving the data directly from the pods was thus no longer possible.

As it was difficult to pull the data from the web servers' containers, the alternative was to offload the data to a more accessible location. An additional worker node, the "Archiver", was built for this task. The long-term storage branch of the web app was routed through a Redis Stream into the Archiver, which would hourly pull new files from this Stream, package them into a compressed archive, and upload them to AWS's Simple Storage Service (S3). From here, `craydata` could then synchronize these files from the cloud for offline analysis.

The data record was thus made accessible and redundant, but also expensive to maintain on a petabyte scale. Instead of accruing costs of $0.021 per GB-month on S3, each month the data are migrated into S3 Glacier, a cheaper storage solution ($0.004 per GB-month) designed for infrequent access. However, to assure that the data records on S3 Glacier and in local storage are identical, this migration is performed with care; regularly migrating data across AWS could leave gaps in `craydata`'s data store, especially if `craydata` should for any reason go offline. Rather than copy the data directly over AWS, `craydata` instead queries S3 for files matching those in its filesystem, and only these are subsequently uploaded to S3 Glacier. With this assurance, the corresponding files on S3 can then safely be deleted.

While a number of needed improvements have been discussed in this chapter and the last, the DAQ enhancements are the focal point to which these all converge. With-

out a successful trigger—one which achieves a reasonably high detection efficiency—there is no experiment to manage, no devices to correct, and no data to properly store. The viability of Crayfis, greatly improved though it may be, depends upon a best-case measurement of the sensor response, undertaken in Part III.

Part III

CMOS Efficiency Measurements

Chapter 5

Hodoscope Revisited

In Part I, this work presented the motivation for a smartphone array to detect UHECR showers and tied its hypothetical performance to two key metrics: the user density and the sensitivity to muons, electrons, and photons, parametrized by the detector cross sections $A_i\epsilon_i$. Part II then explored the design of such a detector to maximize the efficiency. Now, the fundamental question for Crayfis can be revisited: with reasonable user adoption, will this detector be able to contribute to our knowledge of UHECR physics?

As mentioned in Chapter 3, there are three main strategies for distinguishing cosmic rays from noise on a smartphone, and this chapter presents a refined attempt at the first: tagging muon coincidences with an independent second detector. The advantage of this approach is that it is the most direct: rather than extrapolate the cosmic ray muon efficiency from a pixel model fit to other data (e.g. radioactive sources or beam spills), the muon efficiency is measured with the cosmic ray flux itself. Though some systematic adjustment may nonetheless be warranted—for instance, to correct for the hodoscope's finite angular acceptance—such corrections are comparatively minor. However, the rate of data-collection is rather slow—a single beam spill might generate the same particle count as a full year of cosmic ray muons on a smartphone camera—hence, this method has a more limited potential for modelling the sensor. To supplement this deficiency, a beam test was performed, albeit with disappointing

results, as will be discussed in Chapter 7.

This chapter explores the design and results of this more refined version of the hodoscope experiment in Section 3.2. The calibrations and DAQ for both the phones and scintillators will be detailed, and the efficiency as a function of threshold value will be presented for several different phone models. These results are supplemented with a fit to a GEANT4 pixel model in Chapter 8.

5.1 Experimental Design

5.1.1 Hardware

A construction of two phones between three scintillators was employed for this experiment, illustrated in Figure 5.1 below.

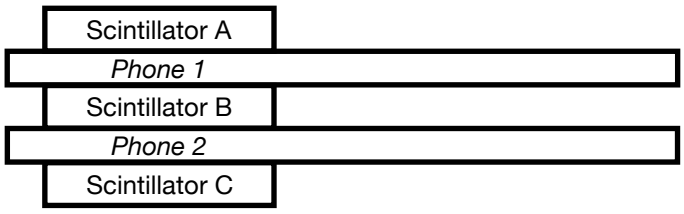

Figure 5.1: A diagram of the hodoscope construction.

A coincidence in two scintillators serves as a *tag*, indicating the passage of a muon, which can be used to probe the response of the smartphone CMOS sensor between them. The strategy of the experiment is to assemble a set of cosmic muons tagged via coincidences in adjacent scintillators; from this nearly-pure sample, the muon efficiency of the CMOS sensors can be measured as a function of a threshold on the per-pixel response. Though only a single phone between two scintillators was necessary, this configuration allowed for more efficient running and for a valuable geometrical cross-check: the outermost scintillators (A and C) can also be used to tag muons, though with a significantly lower acceptance.

The hodoscope detectors are composed of LYSO scintillators measuring, in millimeters, 16×14×8, attached to PMTs via rectangular waveguides. To shield against

light contamination, each scintillator, waveguide, and PMT were wrapped in multiple layers of aluminum foil and secured with foil tape; shown below in Figure 5.2.

Figure 5.2: A photograph of the foil-wrapped PMT, waveguide, and scintillator.

To support these detectors and to aid in their alignment, 3D-printed modules were designed for both the phones and the PMT assemblies, illustrated in Figure 5.3. Square pegs and holes anchor these modules in a way that minimizes the gaps between sensors, thus maximizing the solid angle sampled; 1.2 mm-thick walls above and below each component were the only spacing required. To more accurately align both phones and scintillators within the modules, clear plastic was used for the 3D prints and holes were added at the intended sensor positions; the foil on the scintillators and electrical tape on the cameras provided sufficient shielding from light.

Three phone models were selected to run in the hodoscope. Two models, the ISOCELL-variant Galaxy S6 and the Exmor-variant Galaxy S7, were chosen because their pixel geometries were made publicly available through electron microscopy images (Figure 5.4), allowing for more precise pixel modelling in GEANT4. The advertised pixel pitches of each sensor set the scale for these images, from which the depth of the pixels and dimensions of the DTI structures could be easily ascertained. As shown, the S6 has a pixel pitch of 1.12 µm and B-DTI extending roughly halfway into a 2.6 µm-deep substrate, while the S7 has a pixel pitch of 1.4 µm, a depth of 2.9 µm, and 1.9 µm

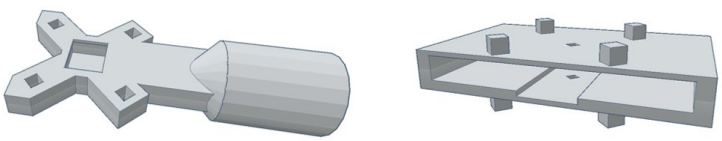

Figure 5.3: Illustrations of 3D-printed modules for (left) the PMT-scintillator assembly and (right) an example phone: the Samsung Galaxy S6.

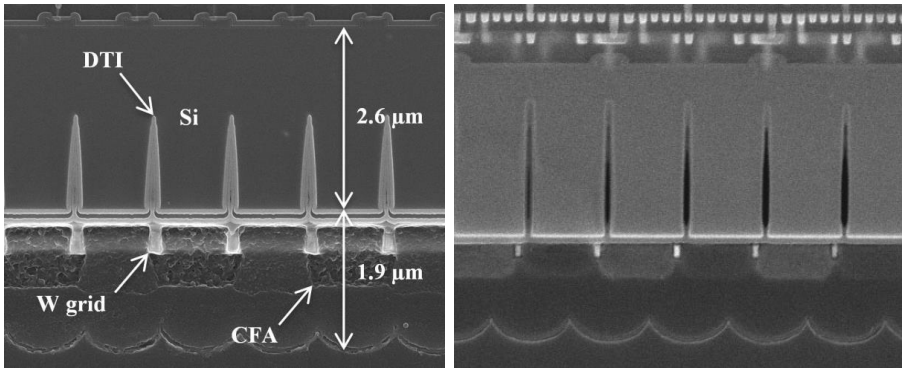

Figure 5.4: Images of active pixels for (left) the ISOCELL S5K2P2 [134], found in the Galaxy S6, and (right) the Sony IMX260 in the Galaxy S7 [117].

B-DTI. According to Figure 2.3, these pixel depths are fairly typical of the 2015–2016 generation; however, the S7's larger, dual-photodiode pixels were a novelty at the time, attaining more widespread adoption over the following years.

The LG V20 was selected as the remaining device model. Although their exact pixel geometry is unknown, these phones were used in the beam test in Chapter 7 due to their manual control over lens-shading and powerful CPUs. One of these phones—containing the Sony IMX298 sensor—was used in both the hodoscope and the beam test, while the other used the ISOCELL-equivalent S5K2P7.

5.1.2 The FishStand app

Several compounding issues made Crayfis app suboptimal for the phone DAQ despite its many improvements. Most importantly, the app was unnecessarily bulky: the polished UI, gravity sensors, location services, and data quality checks were no longer needed in this context and could be dispensed with, reserving more of the CPU for frame processing. The production app was also inflexible: the frame pipeline and server interface were designed for a simple online calibration under dark conditions, followed by pixel triggering, with no user input. Here, a more precise calibration by hand was preferable, though incorporating this functionality would require radically reinventing the state machine. Instead, a streamlined version of the app was preferable—one which would utilize the advances of the production app while keeping the structure as flexible as possible.

This "FishStand" app ("Crayfis test-stand") contained a camera architecture similar to that of the Crayfis app, though RAW-format images at low frame rates were preferred in this context. The frame processing, however, was kept perfectly flexible, allowing vastly different types of measurements to be consolidated under a single app. These "`Analyses`"—experiments in FishStand's repertoire—include:

- *PixelStats*: A dedicated calibration run which keeps running statistics for each pixel, and from which offline calibrations can be performed.
- *Cosmics*: A run similar to the Crayfis production app, which saves bright individual pixels and their neighbors.
- *TriggeredImage*: A design specifically for the beam test in Chapter 7, which saves full RAW frames to file when large numbers of bright pixels are present.

These three were linked by a common method of configuring runs, common camera architecture, and a common backend.

Unlike the production app, FishStand is configured and calibrated entirely by hand, giving the experimenters full manual control of each experiment. Instead of importing the `ExposureBlock` structure from the production app, premised on periodic and automatic recalibrations, the fundamental unit of FishStand data is a *run*: an uninterrupted camera stream with a fixed set of parameters. Each run is associated with a corresponding configuration file, containing the number of frames to sample (or to run indefinitely), the camera exposure time, ISO gain setting, compression level for the output, and a host of other settings proper to each `Analysis`. An simple and secure backend for FishStand was also developed: rather than routing traffic through a separate endpoint at `crayfis.io`, the data is directly uploaded to S3, where it is copied to `craydata` for offline analysis and live-monitoring.

5.1.3 Hodoscope readout

The PMT output was directed to an Arduino micro-controller, which functioned as a remotely adjustable discriminator. Its two primary roles, to set the three PMT

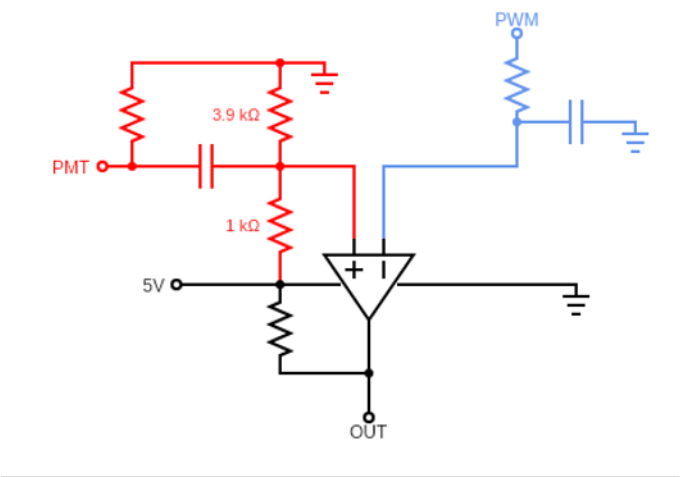

Figure 5.5: A circuit diagram of the Arduino discriminator for a single PMT, divided into input (red), thresholding (blue), and output (black).

thresholds and to register and record triggers, were both driven oven a serial connection by a Raspberry Pi computer.

The focal point of the discriminator circuit (Figure 5.5) is the comparator IC chip, which outputs either HIGH (5V) or LOW (GND) depending on the relative magnitudes of the PMT signal (red) and threshold voltage (blue). The former is a combination of the ∼4V DC component set by the labelled resistors and the high-frequency AC component of the raw PMT pulses, while the latter is supplied by a PWM signal passing through an integrator circuit. With a large value of $\tau = RC$, the threshold voltage is essentially constant, with a magnitude determined by the PWM duty cycle.

As the raw PMT pulses are negative, a trigger occurs when the PMT signal drops below the threshold voltage, changing the discriminator output from HIGH to LOW. A more restrictive threshold thus corresponds to a lower voltage and lower duty cycle. A more typical threshold scale—with higher values being more restrictive—was thus defined as $T = 255 - T_{DC}$, where T_{DC} is the 8-bit duty cycle value set by the Arduino. The 4V baseline shifts the zero-point of the voltage scale to approximately 52.

The Arduino communicates the resulting PMT triggers to the Raspberry Pi, which saves the Arduino's microsecond timestamps to file. To relate the Arduino microsecond clock (which cycles back to zero after roughly an hour) to the Raspberry Pi wall clock, a dummy "heartbeat" channel on the Arduino is triggered by the Raspberry Pi at

fixed intervals. By matching the corresponding heartbeat timestamps on the Arduino and Raspberry Pi, the conversion between the two clocks was in principle established. Offline, the hodoscope trigger times are converted to Raspberry Pi wall clock time via linear interpolation.

5.2 Calibration

5.2.1 Lens-shading and system gain

As in the production app, corrections for both lens-shading gains and hot pixels were necessary to optimize the FishStand trigger; here, a more rigorous calibration was made possible by the PixelStats `Analysis`. Whereas the CRAYFIS app extracts these calibrations from dark frames without user input, the FishStand app could employ more exact techniques using controlled light exposures.

The pixel gains are measured using a standard approach for image sensors: through a well-understood relationship between the mean and variance of the pixel response known as the *photon transfer curve* [135–137]. As mentioned in Section 2.1, photons incident upon a pixel excite electrons into the conduction band, which then drift into the depletion region and are stored on a capacitor. The overall efficiency of this process—the absorption of photons and diffusion of electrons—is known as the quantum efficiency $\eta = n_e/n_p$, and can reach 90% for some visible wavelengths in modern image sensors [138]. The resulting charge, including a contribution from the dark current, is modified by an adjustable readout amplifier and then digitized, creating an overall gain K in DNs per electron. This process is outlined in Figure 5.6.

In the simple case where the range of digitized values is unbounded (either by saturation above or the zero point below), this process follows straightforward Poisson statistics. For a single pixel under a particular exposure, the mean digitized count μ_y is given by:

$$\mu_y = K(\mu_e + \mu_d) = K\mu_e + \mu_{y,\text{black}} \tag{5.1}$$

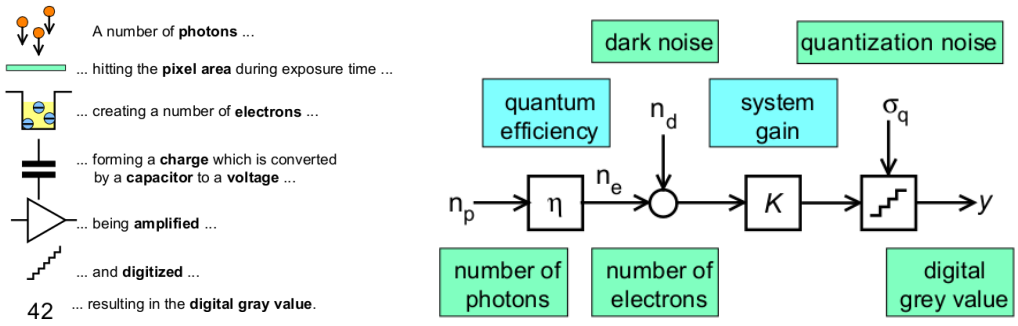

Figure 5.6: A schematic of the process of photon capture and conversion to digital values, taken from Ref. [137].

where μ_e and μ_d are the mean number of photoelectrons and dark current electrons per image, respectively, and $\mu_{y,\text{black}}$ is the black level, which is subtracted by the device either before or after quantization. Likewise, the variance is given by:

$$\begin{aligned}\sigma_y^2 &= \sigma_q^2 + K^2(\sigma_e^2 + \sigma_d^2) \\ &= \sigma_q^2 + K^2(\mu_e + \mu_d)\end{aligned} \tag{5.2}$$

where σ_q^2 is the variance induced by digitization and readout. Combining these two equations yields a linear relationship between the mean and variance of the digitized response:

$$\sigma_y^2 = (\sigma_q^2 + K^2\mu_d) + K(\mu_y - \mu_{y,\text{black}})\,. \tag{5.3}$$

However, near either the zero point or saturation, the variance sharply drops due to the restricted range of counts; see Figure 5.7. By collecting pixel data under light exposures well within these limits, the overall gain K can be found by linear regression.

Due to the applied lens-shading gains, however, K is not constant, but instead varies spatially—typically with a radial symmetry. This dependence is parametrized as:

$$K(r) = K_0\lambda(r) \tag{5.4}$$

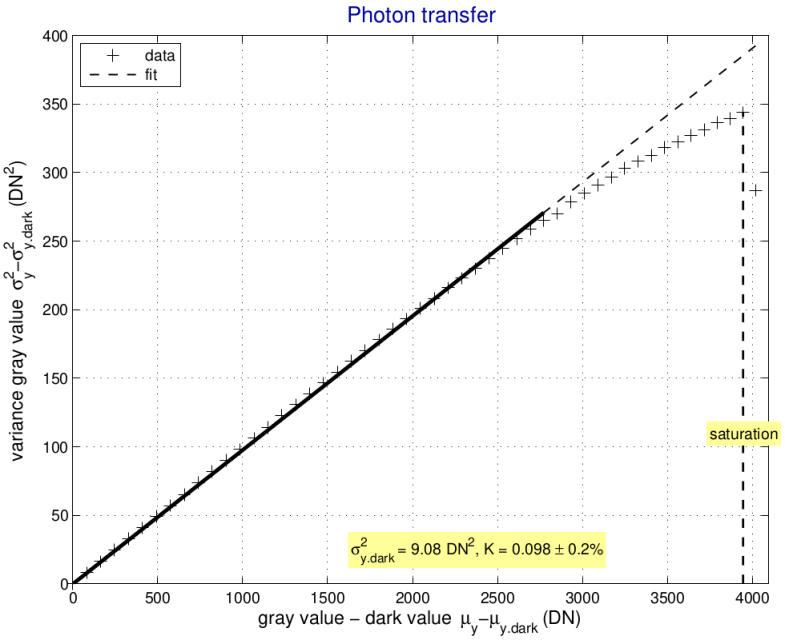

Figure 5.7: A fit of the gain K from the photon transfer curve of a typical image sensor. The origin has been shifted to the mean and variance under dark conditions. From Ref. [135].

where $\lambda(r)$ has been normalized such that $\min_r \lambda(r) = 1$. Scale-factors $w(r)$ in the range $[0, 1]$ can then be found such that, as in Section 3.4:

$$w(r) = \lambda(r)^{-1} \,. \tag{5.5}$$

Pixels calibrated by these scale-factors will then share a common system gain of K_0, from which digitized pixel values are converted into numbers of readout electrons. This both establishes a common scale between the various sensors and allows for meaningful modelling of the physical response in GEANT4.

Lens-shading corrections could be disabled on the V20s, and hence $w_{\mathrm{V20}}(r) = 1$. To find $w(r)$ for the S6s and S7s, one phone of each model received varying degrees of exposure in runs of 5000 frames each. Their rear cameras were secured several inches away from a vertically-oriented white sheet of paper, while the lab was illuminated by constant artificial lighting (slightly biased towards longer visible wavelengths) and kept isolated from any additional light sources, natural or artificial. To modulate the exposure, varying thicknesses of white and colored paper were taped over the lenses,

constructed in such a way that the means of nearby red, green, and blue channels were comparable. The PixelStats `Analysis` recorded the sample means and variances of each pixel, from which the per-pixel gains were calculated by linear regression as in (5.3); see Figure 5.8. When points were sampled from a strict range of pixel means, sufficiently far from both the zero point and saturation, the linear fit described the data well, with $R^2 > 0.985$ for over 99% of pixels.

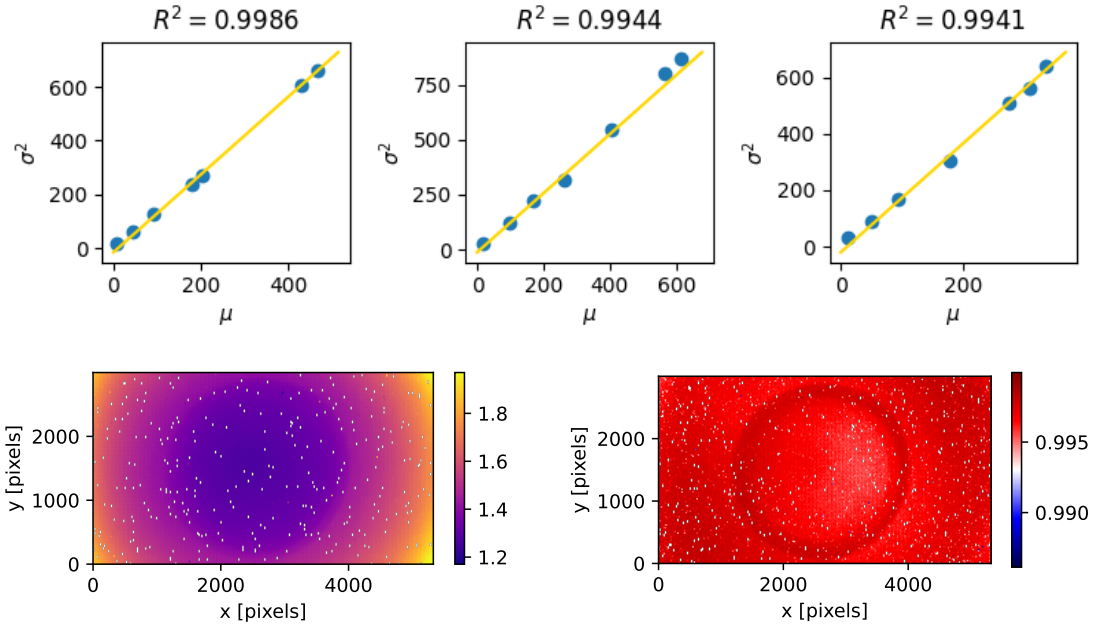

Figure 5.8: Linear regressions of the photon transfer curve for one of the Galaxy S6s. Top: the linear regression is shown for three randomly chosen pixels. Bottom: the spatial dependence of the fitted gain (left) and the R^2 value (right) are plotted against pixel coordinates.

The smoothness and clear radial symmetry of $\lambda(r)$ were then used to further constrain the error. First, the fitted gains were grouped into 4×4 blocks and downsampled according to the median of each block, excluding pixels with low R^2 values and PDAF pixels. These downsampled gains were then plotted against radius, and a piecewise linear fit of $\lambda(r)$ was performed; see Figure 5.9a. Identical $\lambda(r)$ curves were found between different phones of the same model, and this identity held up to a normalization factor when the ISO gain was adjusted. Figure 5.9b shows the resampled system gain $K(r)$.

However, while smoothing the fitted gains was appropriate for estimating the lens-

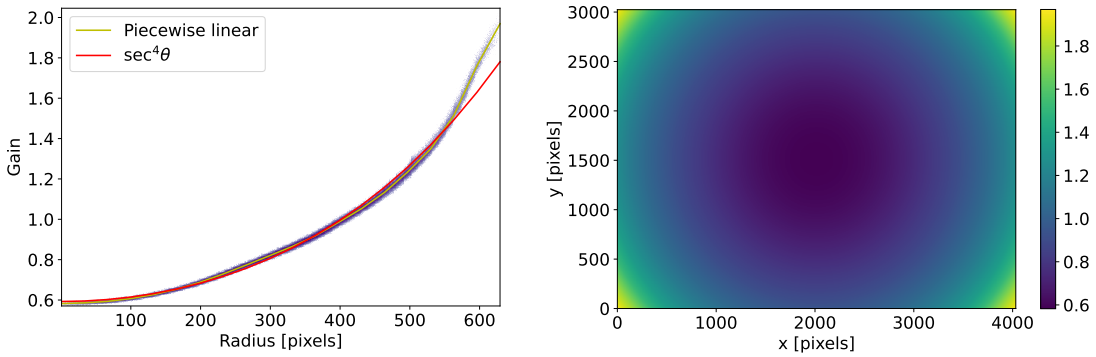

Figure 5.9: A fit of $K(r)$ on the Galaxy S7. Left: individual pixel gains are fit to a piecewise linear function; a naive $\sec^4\theta$ function, the appropriate correction for a lens-less aperture, is also shown. Right: this fit is then re-presented as a function of x and y across the sensor.

shading pattern, this erased any small-scale variation due to, e.g., PDAF pixels or additional gains specific to each color channel. To properly measure the system gain K_0, the photon transfer curves for each phone were linearized by dividing the pixel means by $\lambda(r)$ and the variances by $\lambda(r)^2$; see Figure 5.10. Linear fits were found for each Bayer channel with a Hough transformation,[1] from which K_0 could be read as the slope according to (5.3). The per-channel gains were indistinguishable in the V20 and S7, whereas a slight discrepancy was found in the S6, attributable to small ($< 10\%$) non-linearities in the photon transfer curve. These were presumably a reflection of slightly non-linear readout.

5.2.2 Hot-pixel masking

With the pixel gains calculated, cuts for the online hot-pixel mask could ensue. It is worthwhile to first explore the goal of these cuts: while some pixels are grossly defective and attain values near saturation on nearly every frame, a large number of pixels are only marginally more noisy than the rest. More aggressive cuts can enable running at lower trigger thresholds (i.e. while keeping the trigger rate constant), but at the cost of removing viable pixels.

Finding an appropriate degree of hot-pixel masking is thus inextricably related to

[1]This approach excluded outliers such as the PDAF curves and the saturation tail

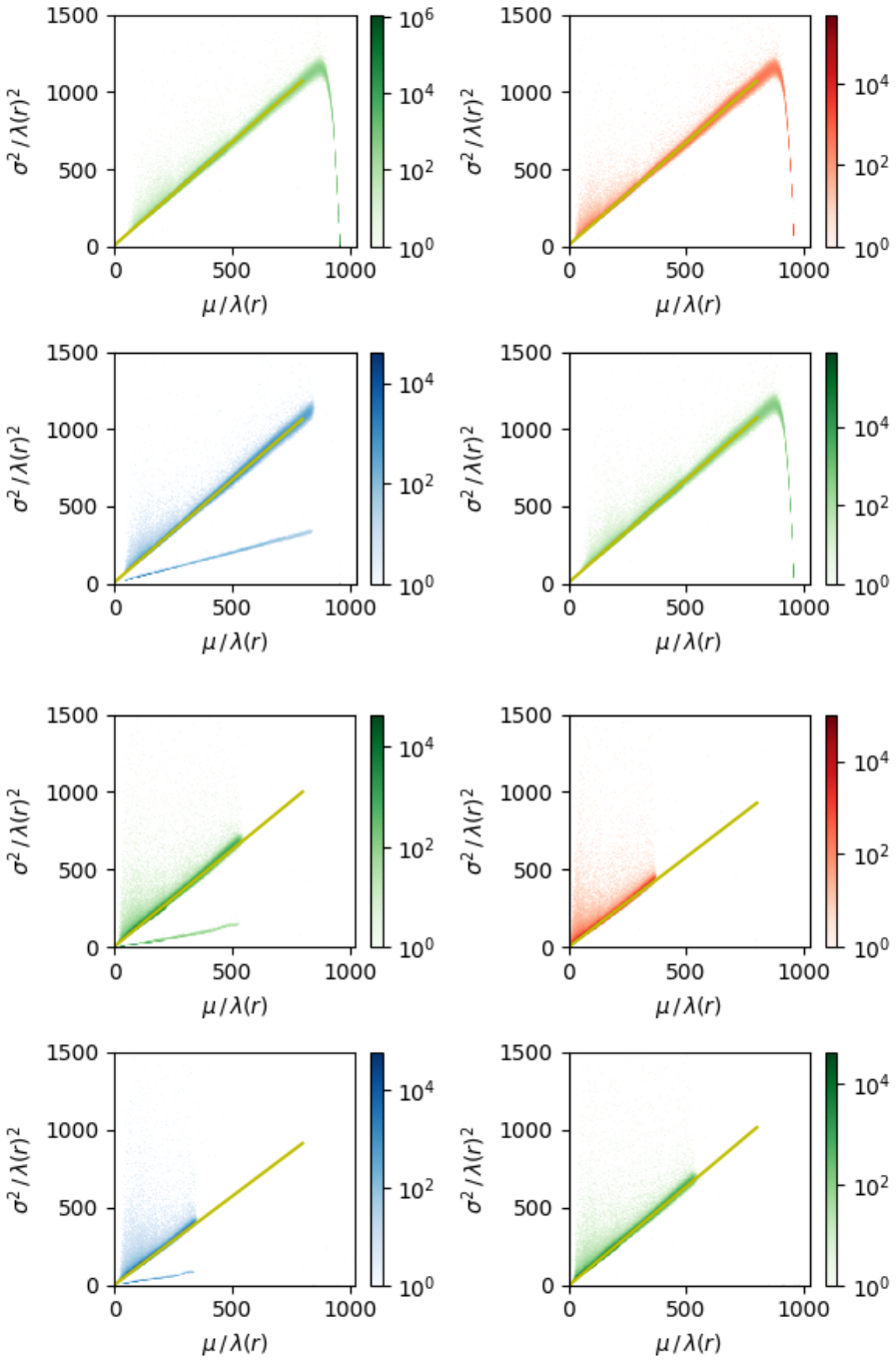

Figure 5.10: Photon transfer curves for (top four) the ISOCELL-sensor LG V20 and (bottom four) a Galaxy S6, illustrated with two-dimensional histograms. The former has no lens-shading gains, while lens-shading corrections have been applied to both the mean and the variance of the latter. Each data point is a sample mean and variance for a single pixel with the appropriate location in the Bayer filter pattern. Two runs with different light exposures are superimposed to show a larger range.

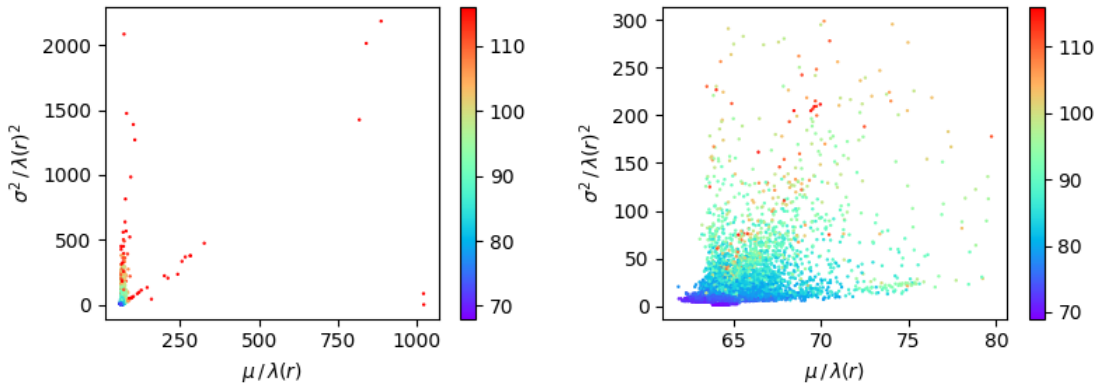

Figure 5.11: Sample variance vs. mean for a random subset of pixels, with second-highest value on the color axis, before (left) and after (right) preliminary cuts on the ISOCELL-sensor V20. The scale of the x and y axes are different in the two plots while the color axis remains constant.

the desired trigger rate and trigger threshold. Although the S6, for example, could save 10–15 pixels per frame without overheating, such a trigger rate is unusable without significant offline cuts: with a 100% frame pass rate, comparing frame and hodoscope timestamps becomes meaningless. Rather, a corresponding *pixel* trigger rate below 1 per frame is necessary to measure frame-hodoscope coincidences; still, this does not preclude using higher trigger rates for the online trigger, provided more aggressive cuts are then applied offline.

The objective of online hot-pixel cuts, then, is to mask enough pixels to run without overheating at a target threshold, at which the trigger rate can be further reduced by a factor of 10 offline. The PixelStats `Analysis` was first used under dark conditions to cut extreme outliers in mean, variance, and second-highest value, shown in Figure 5.11. The following iterative procedure was then undertaken for more exhaustive cuts:

1. FishStand Cosmics was run in dark conditions for 50,000 frames.

2. Any pixels that triggered more than N times were masked, where $N > 10$ is chosen by hand from the shape of the pixel frequency distribution (see Figure 5.12a).

3. If less than 10 triggers per frame remained (i.e. 500,000 total), the threshold was lowered, and another iteration ensued with Step 1.

4. If a further cut of $N = 10$ would lower the rate below 1 trigger per frame, another

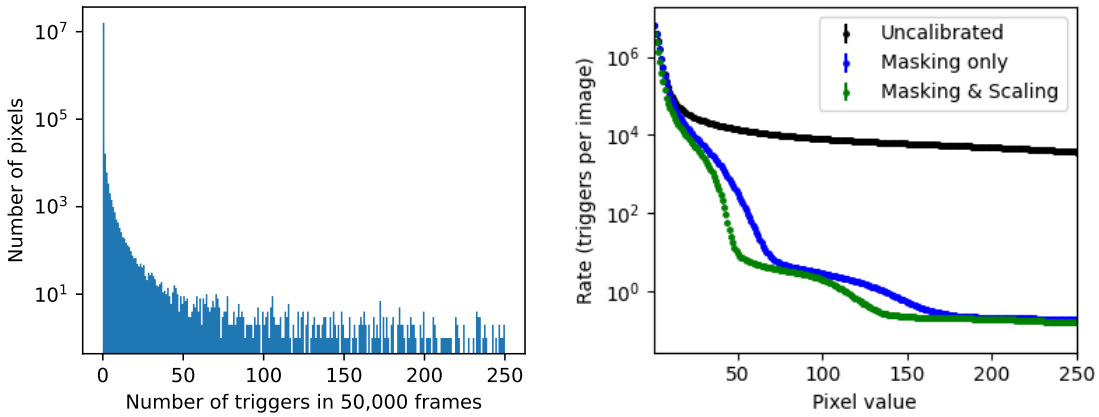

Figure 5.12: An iteration of hot-pixel masking in a Galaxy S6. Left: the occupancy of each pixel from a 50,000 frame run is compiled into a histogram, from which a hand-picked cutoff is imposed to mask hot pixels. Right: the trigger rate is plotted against threshold in green (for comparison, rates of triggers on unscaled pixel values are plotted in black, and in blue with hot pixel cuts). If either a lower threshold would produce a rate under 10 triggers per frame or an occupancy cut of $N = 10$ would reduce the rate below 1 trigger per frame, another iteration is required.

iteration would begin with step 1. Otherwise, the threshold was considered sufficiently low for the online trigger.

While the particular value of N in step 2 for a given iteration is somewhat arbitrary, less aggressive cuts would simply require more iterations to reach the same goal threshold and rate; for 5–10 iterations. Larger values of N were preferred as the threshold approached its target value to minimize unnecessarily aggressive online cuts. Moreover, for most phones, additional iterations were performed beyond the nominal goal threshold to provide a buffer for more aggressive offline cuts. This procedure is graphically detailed in Figure 5.12.

Offline cuts then aimed to mask 1% of the total pixels for each sensor, which greatly reduced thresholds without substantially reducing the efficiency. With this degree of masking, the online thresholds attained trigger rates well above 1 trigger per frame, necessitating more restrictive offline thresholds.

5.2.3 Hodoscope thresholds

The PMTs likewise required an appropriate choice of trigger thresholds. In this calibration, the three scintillators were stacked vertically with minimal separation between them, aligned precisely by their cases. The efficiency of the middle PMT—the fraction of top-bottom coincidences in which the middle PMT also triggered—was then computed, shown in Figure 5.13. Importantly, this measurement quantified the scintillator efficiency of only muons passing through both the top and bottom surfaces of the middle scintillator; the omnidirectional efficiency—also including muons entering or exiting through the sides—is treated in Section 5.3.3.

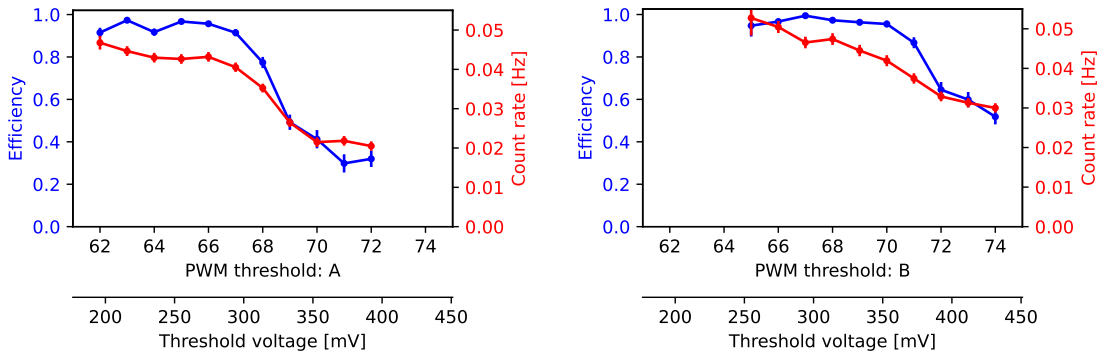

Figure 5.13: Vertical efficiency estimates vs. middle PMT threshold, for PMT A (left) and B (right). The middle-PMT trigger rate in plotted in red, neglecting the small combinatorial noise.

The permutations ABC, BAC, and ACB were used to calibrate all three PMTs, with only the threshold of the middle PMT varying. As expected, a plateau of $\epsilon \approx 1$ was found in each ordering, with a sharp decline beginning around thresholds of 67 in A and 69-70 in B and B and C. Thresholds of 62 for A and 64 for B and C were selected for hodoscope operation, below which the electronic noise in the Arduino became significant. As suggested by the higher trigger rates in Figure 5.13, these lower thresholds attained greater omnidirectional efficiencies than those at the edge of the plateau, and hence expanded the solid angle of the hodoscope.

The difference in these curves was primarily caused by a slowly shifting supply voltage, to which the PMTs are incredibly sensitive, rather than inherent differences in the PMTs or scintillators. It was observed that all three curves shifted progres-

sively leftward over the months of the hodoscope's operation, and the order of plateau thresholds—B and C, followed further below by A—also corresponded to the chronological order in which these samples were taken. The consequences of this shift will be discussed in Section 5.4.

5.2.4 Clock synchronization

To compare the timestamps generated by the phone and hodoscope, it was first necessary to synchronize their respective clocks. The FishStand app draws from the nanosecond timestamps associated with each frame buffer, which precisely measure intervals, but do not accurately report the wall clock time. This generates a nearly constant drift, which can be calibrated with a linear adjustment.

To compute the drift correction, all possible pairs of phone triggers and hodoscope coincidences are used to populate a plot of $t_{\text{phone}} - t_{\text{hodo}}$ vs. t_{hodo}. Under a more restrictive offline threshold, a band of phone-hodoscope coincidences is clearly visible; see Figure 5.14a.

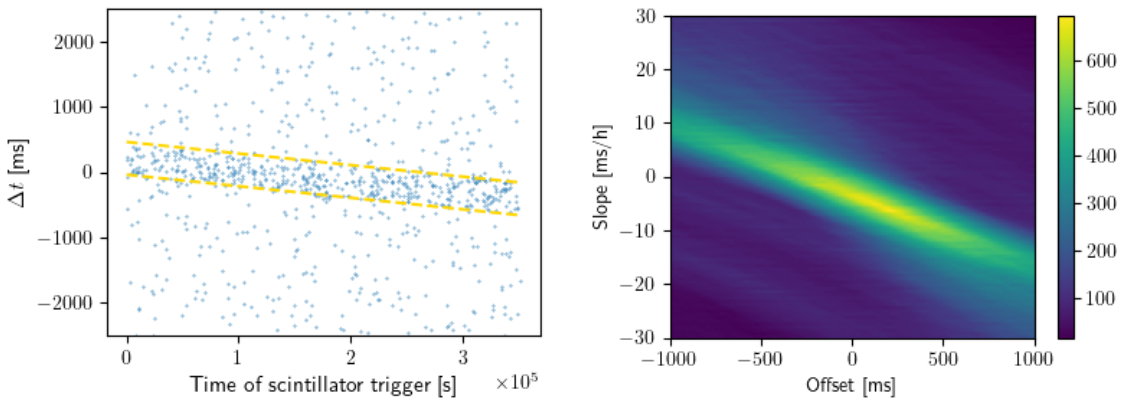

Figure 5.14: The drift correction procedure for the hodoscope, from the ISOCELL-equipped V20. Left: nearby phone triggers are found for each hodoscope coincidence, and the time differences between such pairs are plotted against the hodoscope timestamps. Right: the linear fit for the clearly visible band is computed through a Hough transform, from which the gold lines in the left plot are drawn.

A Hough transformation is then applied to find the corresponding linear fit. Writing

the fit as:

$$y = a + bx\,,$$

each data point (x_0, y_0) can be associated with a line in parameter space:

$$a = -x_0 b + y_0\,.$$

In a true linear relationship without uncertainty, all such lines would intersect at the true value of (a, b).

However, the uncertainty induced by the finite frame duration—appearing as the band width in Figure 5.14a—required a non-standard treatment. As the actual muon hits are equally likely to occur at any time within a frame, the values of Δt should be uniformly distributed within this band: a simple best-fit line was therefore unsuitable for this situation. Thus, each data point was instead mapped to a *band* in parameter space rather than a line, with a width along the a-axis equal to the frame duration. An accumulator grid was used to find the point of intersection of these parameter-space bands; see Figure 5.14b. The inverse transformation maps this point to the region bounded by gold lines in Figure 5.14a.

The resulting trigger timestamps could then be associated with hodoscope tags. A histogram of the corrected Δt was first constructed to visualize the quality of the calibration: the result, shown in Figure 5.15, reasonably approximates the expected uniform distribution. Frames were then tagged if the nearest hodoscope timestamp t_{hodo} satisfied:

$$|t_{\text{phone}} - t_{\text{hodo}}| < \tau/2$$

for a coincidence window τ. To avoid selection bias, the value of τ was fixed at 120% of the frame duration; this histogram confirmed that such a window would be sufficiently wide.

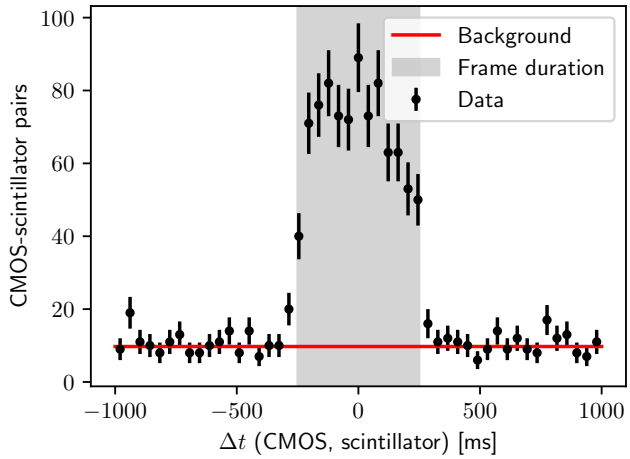

Figure 5.15: A histogram of $t_{\text{phone}} - t_{\text{hodo}}$ after drift corrections have been applied, with the expected combinatorial background in red.

5.3 Statistical analysis

In order to extract the efficiency from the hodoscope data, this efficiency must first be defined in rigorous statistical language proper to the experiment. A combination of muons, background radiation, and electronic noise may trigger the phone; some of these muons trigger the hodoscope while others do not; some will furthermore create secondaries, which will trigger distant regions of the sensor. To navigate through the myriad possibilities, a basic unit of analysis will first be defined: an *event* will hence refer to an independent cause of excited pixel values. In this framework, secondaries created by a particle scattering on the phone glass are thus still part of a single event. As such, events, regardless of their source, constitute a Poisson process. With the following notation:

μ = the event is a muon

H = the event (a particle) creates a coincidence in the hodoscope

C = the event (a particle) is incident upon the CMOS

T = the event triggers the phone (i.e. excites a pixel above the chosen threshold)

the efficiency takes the form:

$$\epsilon = P(T \mid \mu \cap C)\,. \tag{5.6}$$

To characterize this efficiency using phone-hodoscope coincidences, two assumptions are employed:

1. Hodoscope coincidences are exclusively caused by muons passing through the hodoscope.

2. Differences in the angles, energies, etc. of this population of muons do not bias the efficiency estimate.

The former is handled as a systematic in Section 5.3.3, while the impact of the latter—at least, the angular cutoff imposed by the hodoscope—is visible in the results of Section 5.4, particularly Figure 5.25. Under these conditions, the efficiency can instead be written as:

$$\epsilon = P(T \mid \mu \cap C \cap H) = P(T \mid C \cap H)\,. \tag{5.7}$$

Using the definition of conditional probabilities and the fact that $T \subset C$, this becomes:

$$\epsilon = \frac{P(T \mid H)}{\alpha} \tag{5.8}$$

where $\alpha = P(C \mid H)$ is the geometrical acceptance of the hodoscope. This acceptance can be calculated by Monte Carlo given the construction of the hodoscope, while the numerator is determined entirely by the observed coincidence rates.

5.3.1 Geometry-independent efficiency

The numerator $\epsilon\alpha$ can be recast in terms of counts, that is:

$$P(T \mid H) = \epsilon\alpha = \frac{N_{T\cap H}}{N_H}\,. \tag{5.9}$$

Although the total number of hodoscope coincidences N_H is easy to quantify, finding $N_{T\cap H}$—that is, the number of muons triggering both the hodoscope and the sensor—requires a careful treatment of the combinatorial background.

As mentioned in Section 5.2.4, frames are tagged when $|t_{\text{phone}} - t_{\text{hodo}}| < \tau/2$ for a given coincidence window τ. When τ is larger than the frame duration, two (or more) adjacent frames may be tagged by the same hodoscope hit, introducing a statistical dependence in their outcomes; this is most evident in the zero-noise, perfect efficiency case, where exactly one of the two frames will be triggered. Rather than treat these statistical dependencies, each hodoscope coincidence will instead define a "block" of tagged frames; by treating these independent blocks rather than the individual, dependent frames comprising them, simple binomial statistics can be recovered. This, however, requires that all blocks are disjoint—that is, no frame is tagged by more than one hodoscope coincidence. To avoid this additional complexity, blocks sharing a frame with any other block are discarded. This procedure is illustrated in Figure 5.16.

Because each block is by construction independent, blocks with the same number of frames n should have identical probabilities p_n of at least one frame being triggered. Considering both signal and noise, the probability of a block containing at least one triggered frame is given by:

$$p_n = 1 - (1 - \epsilon\alpha)(1 - \eta)^n\,. \tag{5.10}$$

Here $\epsilon\alpha$ and η are, respectively, the probabilities that a muon triggering the hodoscope also triggers the phone, and that a single frame is triggered by any other kind of event. If B_n is the total number of blocks with n frames, then the number of those blocks

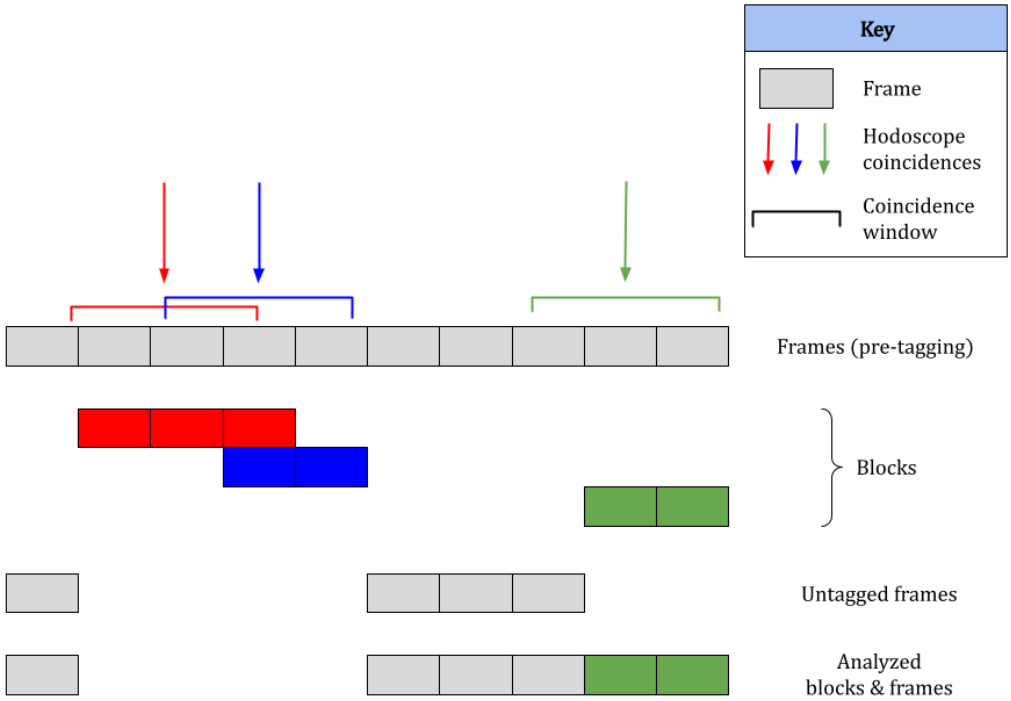

Figure 5.16: Assembly of frames into blocks from three hodoscope coincidences, shown in red, blue, and green. Coincidence windows (here, roughly 2.6 frame lengths) are applied to the timestamps at the start of each frame. As the red and blue blocks share a frame, neither are used in the analysis.

which are triggered follows the binomial distribution:

$$B_{n,T} \sim \text{Binom}(B_n, p_n)\,, \tag{5.11}$$

and so p_n can be estimated as:

$$\begin{aligned} \widehat{p}_{\text{n}} &= \frac{B_{n,T}}{B_n} \\ \widehat{\sigma}^2_{p_n} &= \frac{\widehat{p}_n(1-\widehat{p}_n)}{B_n}\,. \end{aligned} \tag{5.12}$$

The untagged frames (explicitly, frames further than some threshold $\tau/2$ from the nearest hodoscope timestamp) can then be used to quantify η. As the block formalism was introduced to handle the statistical dependence introduced by tagging, no such treatment is necessary here: under the assumption that all events are Poisson and thus independent, individual untagged frames will follow simple binomial statistics.

Denoting the numbers of triggered and total untagged as U and U_T, respectively, the requisite probability is, similarly to (5.12):

$$\begin{aligned} \widehat{\eta} &= \frac{U_T}{U} \\ \widehat{\sigma}_\eta^2 &= \frac{\widehat{\eta}(1-\widehat{\eta})}{U}\,. \end{aligned} \tag{5.13}$$

With this measurement of η, (5.10) can be solved for the desired quantity $\epsilon\alpha$. As each value of n will yield a different estimate, a weighted average for all values of n is computed with the block counts B_n as weights. This yields the main result:

$$\widehat{\epsilon\alpha} = 1 - \frac{1}{B_{\text{tot}}} \sum_n \frac{B_n(1-\widehat{p}_n)}{(1-\widehat{\eta})^n} \tag{5.14}$$

$$\begin{aligned} \widehat{\sigma}_{\epsilon\alpha}^2 = \frac{1}{B_{\text{tot}}^2} \Bigg[&\sum_n \frac{\widehat{p}_n}{B_n(1-\widehat{p}_n)} \left(\frac{B_n(1-\widehat{p}_n)}{(1-\widehat{\eta})^n} \right)^2 \\ &+ \frac{\widehat{\eta}}{U(1-\widehat{\eta})} \left(\sum_n \frac{nB_n(1-\widehat{p}_n)}{(1-\widehat{\eta})^n} \right)^2 \Bigg] \end{aligned} \tag{5.15}$$

where $B_{\text{tot}} = \sum_n B_n$.

A toy Monte Carlo was constructed to verify eqs. (5.14) and (5.15). Trigger and hodoscope timestamps were first generated according to Poisson statistics with a predefined $\epsilon\alpha$, hodoscope rate, and noise rate. Signal and noise timestamps were then associated with frames, and information about their multiplicities was destroyed. The analysis was applied to these hodoscope timestamps and frame triggers—first by tagging frames according to a chosen value of τ, and then by computing a test statistic for $\epsilon\alpha$. The distributions of these test statistics are shown in Figure 5.17 with 100,000 samples of 10^6 frames for various values of $\epsilon\alpha$, λ_{hodo}, λ_{noise}, and τ.

Yet on larger time scales, triggers are not truly Poisson-distributed. The phones exhibit day-night temperature shifts, which in turn creates cyclical variations in the thermal noise. Nonetheless, this algorithm is relatively insensitive to a time-varying

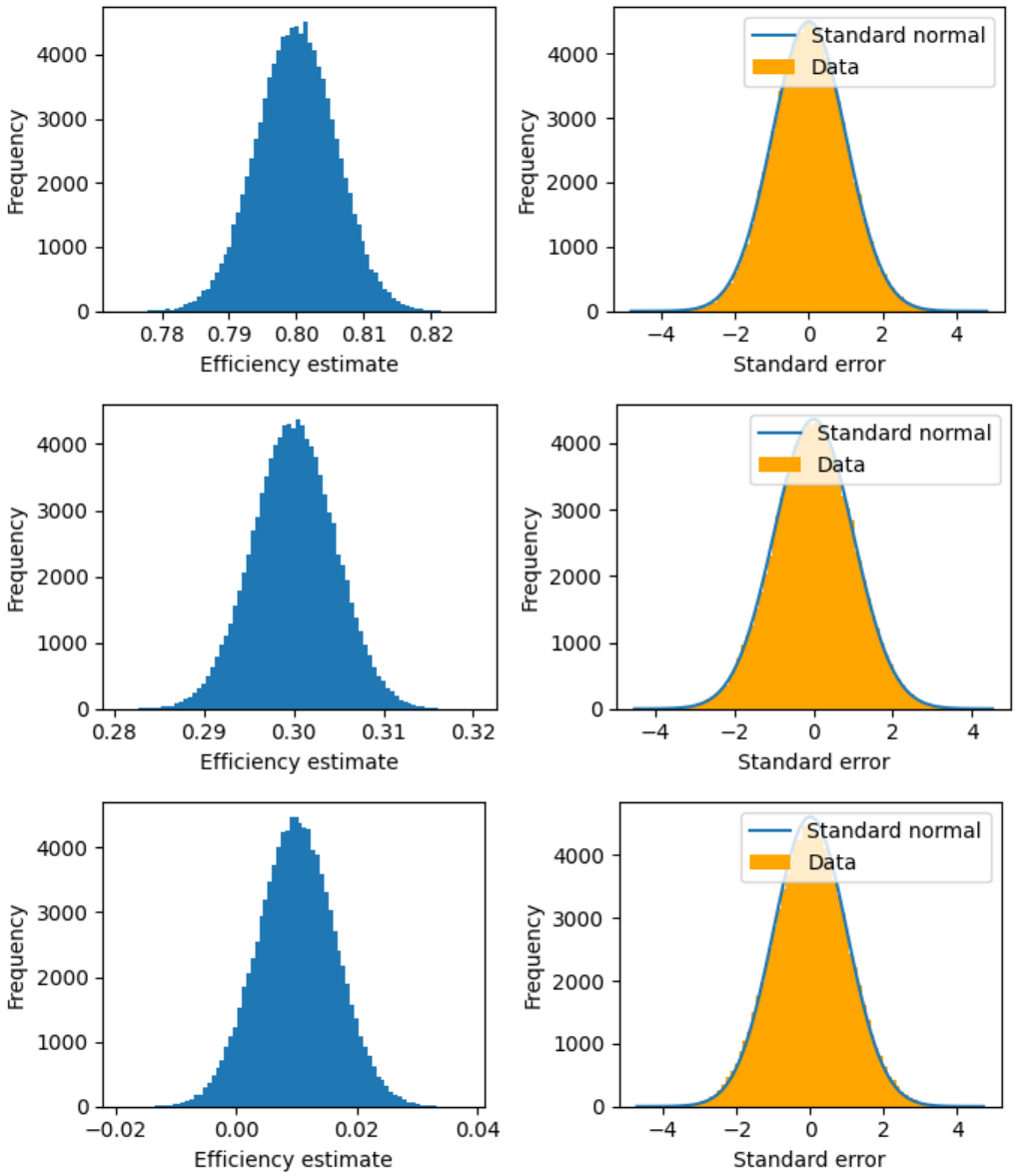

Figure 5.17: Efficiency estimates (time units in multiples of t_F) for:

(a) $\epsilon\alpha = 0.8$, $\lambda_{\text{hodo}} = 0.005$, $\lambda_{\text{noise}} = 0.05$, $\tau = 1.2$

(b) $\epsilon\alpha = 0.3$, $\lambda_{\text{hodo}} = 0.02$, $\lambda_{\text{noise}} = 0.2$, $\tau = 1.2$

(c) $\epsilon\alpha = 0.01$, $\lambda_{\text{hodo}} = 0.02$, $\lambda_{\text{noise}} = 0.1$, $\tau = 5$

The left plots are histograms of efficiency values reconstructed from Monte Carlo simulations, and the right plots show the pull distribution of these efficiency measurements overlaid with a standard normal curve.

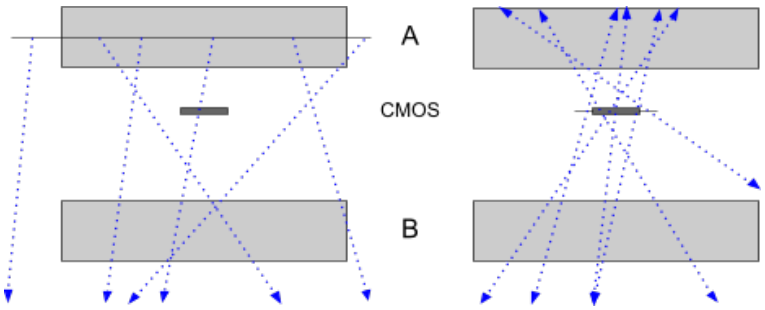

Figure 5.18: Illustrations of the geometric Monte Carlo, with particles generated from scintillator A (left), and the CMOS (right).

Poisson rate: repeating the simulations in Figure 5.17, but removing all noise in the first half of the runtime, biases the efficiency downward by only a few percent. The impact of far more smaller shifts in the rate ($< 20\%$) during the experiment will therefore be left as a systematic uncertainty.

5.3.2 Geometrical acceptance

In order to calculate the geometric acceptance α in (5.8), muons were simulated passing through two rectangular scintillators with an infinitesimally thin CMOS between them. These virtual particles were generated uniformly on planes bisecting each of the three surfaces with a directional intensity:

$$I(\theta) = I_0 \cos^n \theta \tag{5.16}$$

where a range of $n = 2.0 \pm 0.1$ is considered, as described below. A threshold on the muon path length through the scintillator volumes was used for counting hits: a value of $L = 3.7 \pm 0.3\,\text{mm}$ was fit to the observed coincidence rates assuming $n = 2$. Both the values of n and L are described in the following section.

Ray-tracing was then used to compute the relevant acceptances. Through Bayes' theorem, the value of $\alpha = P(\text{CMOS} \mid A \cap B)$ was obtained from the probabilities $P(A \cap B \mid \text{CMOS})$ and $P(B \mid A)$, requiring the left and right geometries in Figure 5.18, respectively.

To set the CMOS position, a representative device of each model was disassembled

to find the location of the CMOS chip within the phone. By comparing the dimensions of the individual modules to those of the total assembly, the average gap size between the modules was found to be 0.4 mm. These gaps were included in the geometry and a 0.4 mm uncertainty was also assigned to all distances within the simulation.

To propagate these uncertainties to the computed geometric acceptance, each simulation was repeated 100 times with independent Gaussian displacements in the measured quantities: the phone and scintillator positions and the path length threshold L. The value and uncertainty of the acceptance were estimated from the mean and standard deviation of these trials. To estimate the additional uncertainty from the zenith angle distribution, this analysis was then repeated for $n = 1.9$ and $n = 2.1$, requiring separate fits for L.

As a final consideration, lens-shading gains in the S6s and S7s produced lower saturation levels in photo-electrons at the outer edges of the sensor; consequently, the highest thresholds in these devices sampled only a small fraction of the sensor area. This dependence was absorbed into α, which was scaled by the fraction of the pixels with saturation levels below each threshold.

5.3.3 Hodoscope systematics

Several systematic uncertainties are inherent to the hodoscope response. These include the rates of non-muonic coincidences, either from electronic noise or other cosmic ray secondaries; the omnidirectional efficiency, determined by muons entering or exiting the sides of the scintillators; and the zenith angle distribution, characterized by n in (5.16). Non-muonic coincidences bias the calculation of $\epsilon\alpha$, premised on a pure population of tagged muons, while the others influence the geometrical acceptance α.

By isolating the three scintillators and counting two-PMT coincidences, the rate of electronic coincidences was found to be negligible: in 6 hours of running with the lowest viable thresholds, only a single coincidence was registered among the three possible PMT pairs.

The hodoscope efficiency to muons only partially traversing the scintillators was

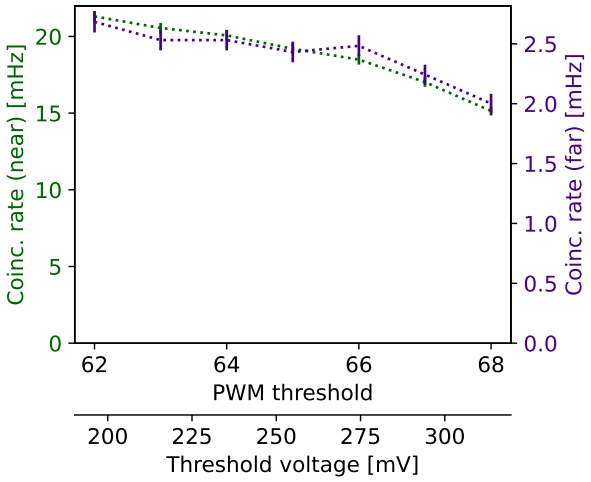

Figure 5.19: Coincidence rates (corrected for combinatorial background) for PMTs B and C vs. bias voltage. Centers of the scintillators are separated by 9.6 mm and 43.4 mm, plotted on the left and right axes, respectively. PMT thresholds are restricted to the plateau in Figure 5.13.

measured by varying the PMT thresholds and separations between the scintillators, both of which modify the relative abundance of these muon tags. Figure 5.19 shows the dependence of the observed coincidence rates on these two variables. Notably, coincidence rates fall more sharply in the near geometry: a greater fraction of these tags are due to partially-traversing muons, which are disproportionately excluded by higher thresholds.

This systematic was parametrized in the toy Monte Carlo by a threshold L on the muon path length through each scintillator. To relate L to the coincidence rates in Figure 5.19, the ratio of far-geometry and near-geometry rates was determined by the Monte Carlo for $0 \leq L \leq 6\,\text{mm}$ (the LYSO crystal height), negating the dependence on the unknown flux; see Figure 5.20a. The inverse mapping was then applied to the ratios of the curves in Figure 5.19, producing estimates of L, shown in Figure 5.20b. A linear fit was found between the bias voltage and L (shown on a non-linear axis), from which $L = 3.7 \pm 0.3\,\text{mm}$ was obtained for a threshold of 64.

This analysis, however, is dependent upon the value of n in (5.16). As the simulated ratios for $n = 2$ encompass the range of experimental values, the standard zenith angle dependence of cosmic muons is supported in the hodoscope response. For comparison, the same simulation was performed with an isotropic source ($n = 0$), which was

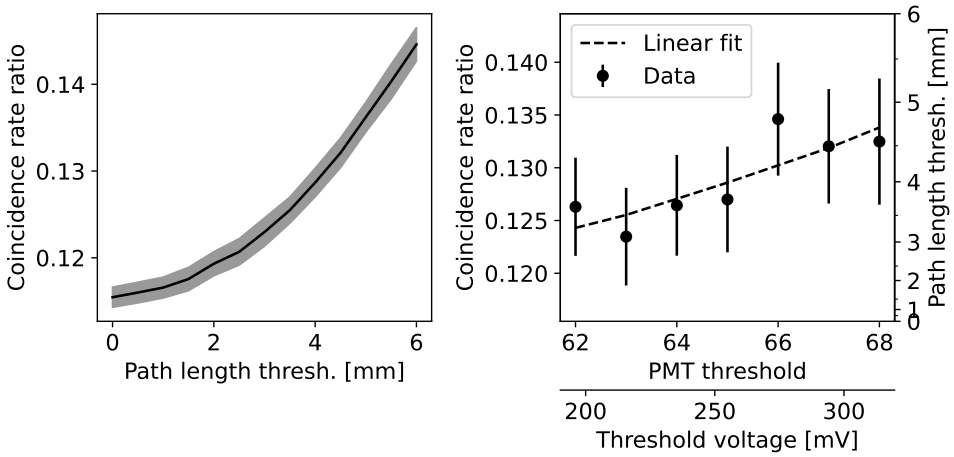

Figure 5.20: The conversion of the rates in Figure 5.19 to effective path length thresholds. Description in text.

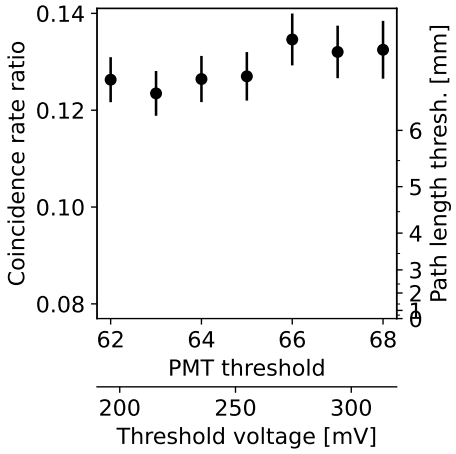

Figure 5.21: A revision of Figure 5.20b, with the second vertical axis instead generated from a population of isotropic particles. Because no overlap exists between the experimental and MC-generated ratios, no fit between the two could be constructed.

clearly excluded by the data; see Figure 5.21. The anticipated proportional relationship between the bias voltage and effective path length threshold was also evidenced by the data.[2] The L-intercept was roughly 0.7 mm, a reasonable fit when geometric and electronic systematics are considered.

With an established relationship between L and the threshold voltage, a third systematic could be estimated: the rate of coincidences caused by non-muonic cosmic ray secondaries. As muons have greater penetrating power than the "soft" component

[2]While the errorbars admittedly appear to be quite large for the fit, the one-sided p-value for the χ^2 statistic was a moderate 0.05. When the same analysis was applied on several slices of the data, no apparent overfitting was present.

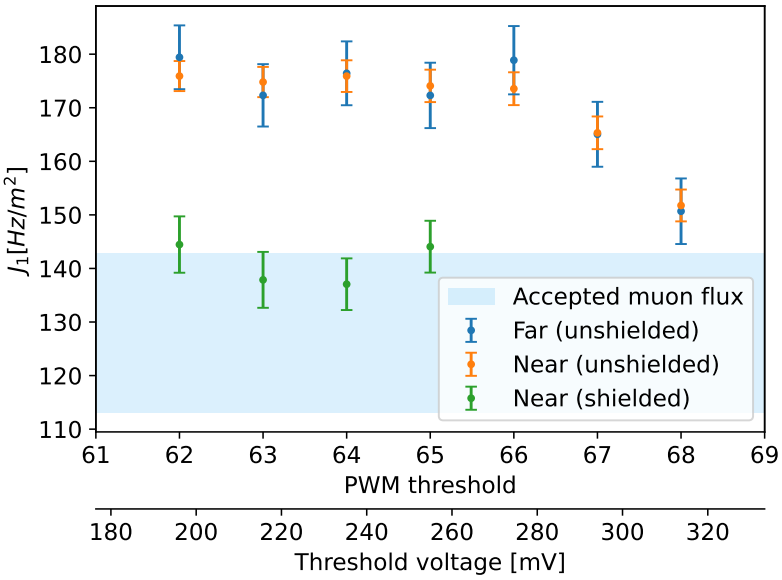

Figure 5.22: Estimated flux vs. PMT threshold voltage. By construction, the flux values for the near and far geometries are very close relative to the random error involved. The blue band represents the range of experimental muon flux values at sea level in Table 3.12 of Ref. [139], excluding those with energy cutoffs above 0.4 GeV.

of the cosmic ray flux—high-energy photons, electrons, and positrons typically blocked by lead shielding—this would be observed as an excess flux above well-established values. To measure the muon flux from the observed hodoscope rates, the Monte Carlo was used with the fitted values of L to find the acceptances of both the the near and far geometries as a function of threshold voltage. The resulting flux values are presented in the blue and orange datasets in Figure 5.22.

As the experiment was performed at an altitude of only 119 m, the discrepancy with other experiments at sea level suggests a significant non-muonic component. The soft component of cosmic ray secondaries is characterized by a similar angular dependence to muons at sea level—hence the excellent agreement with $n = 2$ in Figure 5.20—and can constitute up to 30–40% to the total flux [139], in agreement with the observed excess. For a purer set of muon tags, a lead brick measuring 8”×4”×2” was suspended several millimeters above the scintillators, with the 4” side extending vertically, producing the green dataset in Figure 5.22 within the accepted range.

Lastly, the value of n, quantifying the zenith angle distribution, was determined from Figure 5.23. As the observed flux with shielding was still at the upper edge of

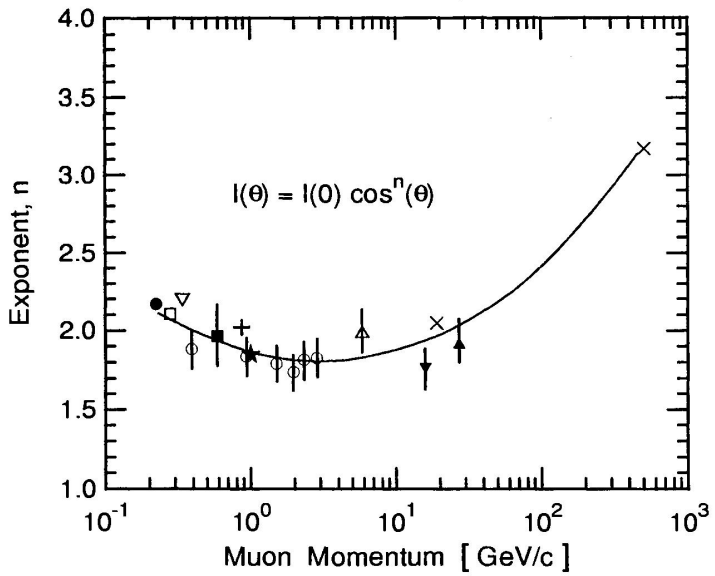

Figure 5.23: The exponent n in the zenith angle distribution vs. minimum muon momentum sampled, taken from Figure 3.60 in Ref. [139].

the range of accepted values, the muon momentum cutoff is expected to be among the lowest of the values cited in Ref. [139]. Consequently, the value of n was estimated at 2.0 ± 0.1.

5.4 Results

With the lead brick above PMT A as before, the three pairs of phones were run in the hodoscope for roughly two weeks each, at which point the geometric uncertainty became dominant; see Table 5.1. Following the analysis in Section 5.3, the efficiency was calculated for each threshold at which a significant population of frames ($> 20\%$) were not triggered; below this level, the higher random error and systematic error from time-dependent thermal noise made the resulting estimates unreliable.

The resulting efficiencies are plotted in Figures 5.24 to 5.26, and several values are listed in Table 5.2. Importantly, the errors across each curve are highly correlated: as the threshold increases, smaller subsets of the same dataset are used. Systematic errors in α, moreover, scale each curve as a whole and dominate the total uncertainty.

Table 5.1: Estimates of the uncertainty in muon efficiency for an S7 smartphone for various thresholds on N_e, the number of photoelectrons collected. Statistical uncertainty (Eq. (5.15)) and systematic uncertainty due to experimental alignment, partially-traversing muons, and the zenith-angle distribution (Eq. (5.16)) are provided.

Uncertainty in efficiency			
Contribution	$N_e = 20$	$N_e = 50$	$N_e = 100$
Statistical	.039	.023	.019
Exper. alignment	.041	.034	.020
Partial trav. muons	.022	.018	.011
Zenith angle distr.	.025	.021	.012
Total	.067	.051	.033

Table 5.2: Measured efficiencies for several phone models, as a function of a threshold on the number of photo-electrons. The first column is the minimum viable threshold, also given, below which the sensor noise and corresponding uncertainties rapidly increase. Uncertainties include all statistical and systematic uncertainties, dominated by the uncertainty on the geometric acceptance. For the two S6s and S7s, the average values of the (nearly identical) means and uncertainties are provided.

Selected efficiencies			
Phone	Min threshold	$N_e = 50$	$N_e = 100$
S6	$.74 \pm .07$ @ $43e$	$.72 \pm .06$	$.47 \pm .04$
S7	$.87 \pm .07$ @ $20e$	$.70 \pm .05$	$.43 \pm .03$
V20 ISOCELL	$.73 \pm .06$ @ $33e$	$.69 \pm .05$	$.44 \pm .03$
V20 Exmor	$.27 \pm .02$ @ $21e$	$.19 \pm .02$	$.14 \pm .01$

As a check on the systematic uncertainty, the efficiencies of the two S6 and S7 devices are plotted together in Figure 5.24. An agreement within 1σ is observed[3] across nearly the entire range of sampled thresholds, confirming the degree of precision with which the phones were aligned and the uniformity within sensors of the same model. Furthermore, this demonstrates that the differences between the scintillator-PMT modules, even under slightly different thresholds, were not ultimately large enough to substantially influence the results: the magnitude of this effect is manifestly insignificant compared with other sources of error.

Since the V20s have two different sensors, however, the systematics are instead checked by computing efficiency curves for all three pairs of scintillators using the appropriate acceptances; Figure 5.25 shows these curves for the ISOCELL and Exmor sensor, respectively. Unlike in Figure 5.24, these three curves should not necessarily be

[3]Again, this level of agreement is caused by the the lack of independence within each curve.

identical, as including the far scintillator in coincidences greatly restricts the solid angle sampled by the hodoscope. Instead, the two far-geometry curves are closely aligned for both phones, while diverging from the curve derived from the adjacent scintillators: a pattern also seen in analogous graphs for the S6s and S7s. This divergence is small for the Samsung sensor (and the Galaxy S6 and S7, not pictured), while the Sony sensor shows a more substantial ($> 2\sigma$) gap near the lowest thresholds.

Even beyond this gap, the efficiencies from this last sensor appear to be an anomaly, with the far-geometry efficiency a factor of 5 lower than the other three phones. In Chapter 7, a possible explanation for this poor performance will be explored. These other three, however, show excellent agreement despite their slightly differing pixel designs; see Figure 5.26). The relationship of this curve to the pixel geometry will be further explored in Chapter 8.

5.5 Discussion

For all but the LG V20 with the Sony sensor, the values of $A_\mu \epsilon_\mu$ at the lowest feasible thresholds are within the range considered by Ref. [83], falling between (1.4 – 2.0) $\times$ $10^{-5}\,\mathrm{m}^{-2}$, when the minimal acceptance $A_\mu = A_{\mathrm{CMOS}}$ is considered. Given that the optimistic bound in Ref. [83] was twice the area of the largest sensor, such an efficiency would not be observable in this experiment. However, this physical limitation does not necessarily exclude such higher values: secondaries generated when muons pass through nearby shielding—roofs, walls, etc.—may also contribute to the acceptance. Extrapolating the true $A_\mu \epsilon_\mu$ from the "strict" $A_{\mathrm{CMOS}} \epsilon_\mu$ is attempted in Chapter 8.

Furthermore, these efficiencies depend both on RAW buffers being sampled, and on unrealistically low thresholds being used. While the lowest thresholds here—triggering just under a rate of 1 per frame—can be used to measure the sensor response, such rates would produce unreasonably high combinatorial noise for reconstructing showers. Fortunately, the background noise drops rapidly past these nominal thresholds, suggesting that a more realistic trigger might perform nearly as well.

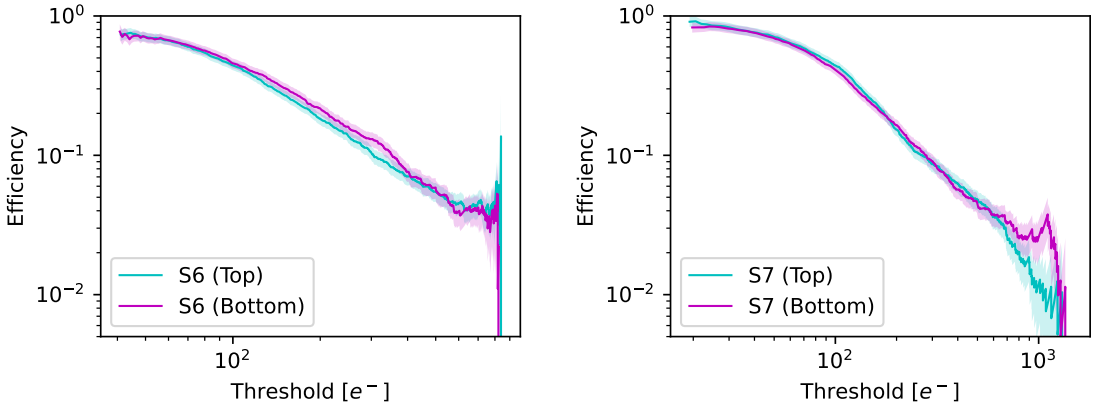

Figure 5.24: Efficiency vs. threshold as for the Samsung Galaxy S6 (left) and Samsung Galaxy S7 (right). The vertical scales are fixed, the horizontal scales in electrons are different in the two plots due to the different sensor gains.

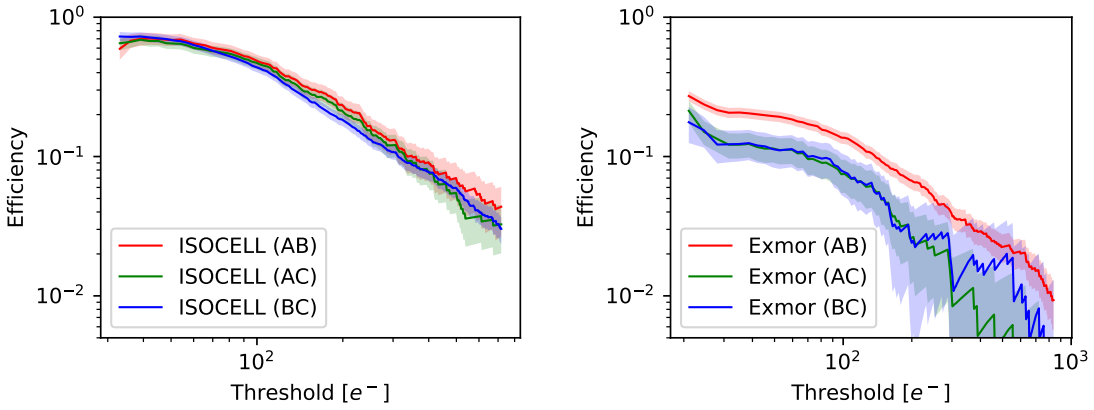

Figure 5.25: Efficiency vs. threshold for the LG V20, equipped with a Samsung sensor (left) and Sony sensor (right). Curves are shown using coincidences and corresponding acceptances from all three pairs of scintillators.

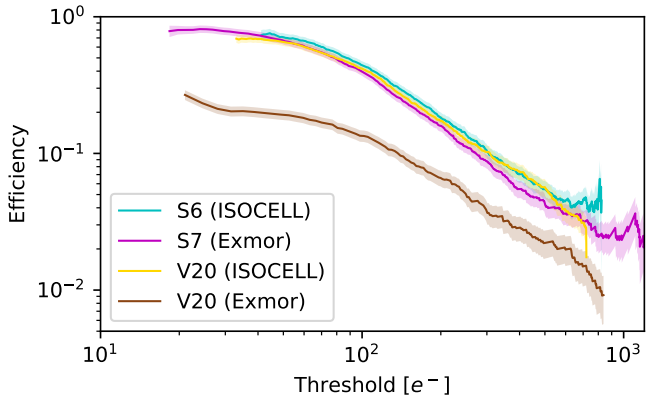

Figure 5.26: Agreement between efficiencies in Figures 5.24 and 5.25. The top S6 and bottom S7 are plotted with the two V20s.

The combinatorial noise is also a function of the timing resolution, which was kept exceptionally low to comfortably stream RAW frames rather than downsampled YUV. A standard video speed of 30 FPS would be preferable for shower reconstruction, but would potentially cost a large fraction of the efficiency; further studies are needed to discern the relationship between efficiency and downsampling. However, the results suggest that higher frame rates, higher efficiencies, and a less noisy trigger are not mutually exclusive: the S7 performed the best of all three models, potentially due to its shorter frames incurring smaller dark currents. Beyond the scope of this dissertation, finding the optimal balance between the efficiency and shower noise will be the next logical step in quantifying Crayfis's performance.

Chapter 6

Radioactive Source Tests

The successes of the previous chapter could be extended to other equally valuable measurements of the sensor response with the completed FishStand app; modifying its structure for other experiments while using the same calibrations was a straightforward endeavor. The third method for distinguishing signal from noise in Chapter 3—exposing the phones to a particle emitter of known composition—was consequently attempted under two different approaches. In this chapter, the photon efficiency measurements from Ref. [83] are repeated with the many advances of the FishStand app, and in the next, a set of phones are exposed to 120 GeV protons in the MTest beamline at the Fermi National Accelerator Laboratory (FNAL).

6.1 Statistical analysis

The measurement of $A_\mu \epsilon_\mu$ of the previous chapter assumed straight muon paths through the scintillators and CMOS, with energy being deposited primarily through soft collisions [151]. Under this model, the acceptance A_μ was taken to be equal to the active sensor area A_{CMOS}, and the efficiency ϵ_μ could be understood as the fraction of muons passing through the sensor which produced a detectable pixel response. Though some hard-scattered electrons from the scintillators or plastic also may have contributed to the CMOS response, this minor systematic was absorbed into the efficiency. Such a treatment was not a defect in the analysis, however: electrons scattered from walls or

ceilings contribute to CRAYFIS's shower sensitivity, and can rightfully be included in the device efficiency.

For high-energy photons, which deposit energy in the CMOS indirectly through Coulomb-scattered or pair-produced electrons, this simple picture is no longer adequate. As this scattering can occur within the sensor itself, the lens glass, or other intervening materials, the acceptance A_γ may be substantially larger than A_{CMOS}, and ϵ_γ quantifies the CMOS efficiency to scattered electrons rather than to incident photons. As many such electrons may be scattered by a single photon, some ambiguity arises in the definitions of A_γ and ϵ_γ. Instead, the full detector cross section $A_\gamma \epsilon_\gamma$ is taken as fundamental, defined as the rate of photon detections divided by the photon flux. This quantity accordingly depends on the material between the radioactive sources and the CMOS.

For a radioactive point source, the detector cross section is given by:

$$A_\gamma \epsilon_\gamma = \frac{4\pi d^2}{R_\gamma}(\lambda_{\text{src}} - \lambda_{\text{bg}}) \tag{6.1}$$

where d is the separation between the source and the sensor, λ_{src} and λ_{bg} are the rate of detections with and without a source present, respectively, and R_γ is the source activity. As a single photon may produce multiple secondaries, the multiplicity of tracks in a frame cannot be used as a proxy for the number of photon detections; instead, the approximately[1] Poisson rate of triggered frames is used to reconstruct λ_{src} and λ_{bg}.

The rates of triggered frames and detected events (photons or background) are related by:

$$\frac{T_i}{N_i} = 1 - e^{-\lambda_i t_F} \tag{6.2}$$

where t_F is the frame duration, $\lambda_i = \{\lambda_{\text{src}}, \lambda_{\text{bg}}\}$ for runs with and without a source

[1] As two photons are emitted per ^{60}Co decay, this rate does not strictly follow Poisson statistics. However, the probability of detecting both photons in a single frame is negligible given the small sensor area.

present, respectively, and T_i and N_i are the number of triggered and total frames in the corresponding run. By inverting this relationship and inserting λ_{src} and λ_{bg} into (6.1), the detector cross section can be measured directly in terms of frame statistics:

$$A\epsilon_\gamma = \frac{4\pi d^2}{R_\gamma t_F} \log\left(\frac{1 - T_{\mathrm{bg}}/N_{\mathrm{bg}}}{1 - T_{\mathrm{src}}/N_{\mathrm{src}}}\right) . \tag{6.3}$$

To employ (6.3), the photon activity R_γ and distance d must be precisely known. The activities of the three sources from Ref. [83], ^{226}Ra, ^{137}Cs, and ^{60}Co, were obtained from previous gamma spectroscopy measurements adjusted for 6.6 years of decay. R_γ was then computed from each activity using both X-ray and gamma-ray lines. Due to the varying beta spectra of the three sources, however, differing degrees of shielding were included in the geometry, which influenced the value and precision of d.

6.2 Shielding and geometry

A proper measurement of the photon efficiency requires ample shielding for the beta radiation produced by these sources. As the photon efficiency is expected to be much lower than that of MeV-scale electrons, even a small amount of penetrating beta radiation could substantially bias the result.

For ^{60}Co, this effect was essentially negligible: the primary beta line is at 318 keV—which the lens glass alone would be sufficient to stop—while all other lines compose approximately 0.1% of decays. ^{137}Cs and ^{226}Ra, however, both contain prominent beta lines with energies above 1 MeV; for instance, ^{214}Bi, a daughter isotope of ^{226}Ra, emits a 3.3 MeV beta in 18% of decays. Especially for ^{226}Ra, screening out beta radiation entirely would inevitably block a considerable number of X-rays and gammas as well.

Consequently, two measurements of the detector cross section were performed: a moderately-shielded measurement analogous to the experimental configuration in Ref. [83] and a heavily-shielded measurement for comparison with the Monte Carlo pixel model in Chapter 8. In the first approach, holes were drilled through 2" wax bricks to secure the sources on one face and to align the phones on the opposite

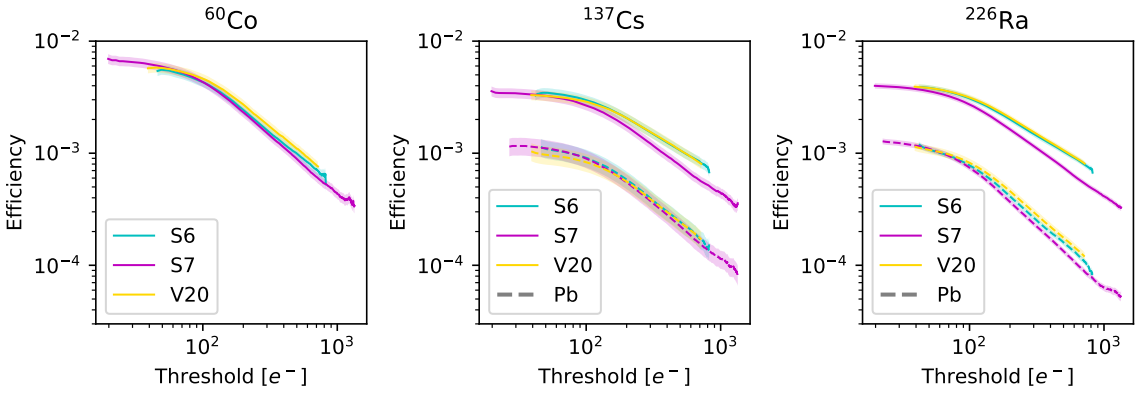

Figure 6.1: Measurements of the detector cross section, normalized by the sensor area, versus trigger thresholds for ^{60}Co (left), ^{137}Cs (center), and ^{226}Ra (right). Thresholds are provided in terms of collected electrons using the gains calculated in Section 5.2.1. Measurements with wax/epoxy shielding are indicated by solid lines, and lead shielding by dashed lines.

face; empty space within the brick was then filled with an epoxy to shield against beta radiation. In the second approach, 1/2" of lead shielding was instead used to fully screen out betas for ^{137}Cs and ^{226}Ra, though with a less precise alignment. The activity of the ^{226}Ra was sufficiently high that the wax and lead shielding could be stacked, generating a smaller relative uncertainty in d while maintaining trigger rates well above background.

6.3 Detector cross section measurements

One model of each phone from the hodoscope experiment was exposed to each of the three sources, the results of which are presented in Figure 6.1. As in Section 5.4, $A_\gamma \epsilon_\gamma$ is computed at each threshold value, producing curves with many of the same features and caveats as previously discussed. To better compare the responses of different phones, the detector cross section is normalized by the sensor areas, yielding the unitless naive efficiency plotted above. Varying saturation levels (in electrons) from lens-shading in the S6 and S7 were handled using the same first-order scaling as in Section 5.4. The total error is dominated by the uncertainty in d, though systematic errors in R_γ are the least precisely quantified and may be larger than assumed here.

The values of $A_\gamma \epsilon_\gamma$ with wax shielding fall between $(0.6 - 1.6) \times 10^{-7}\,\mathrm{m}^{-2}$, an order of magnitude above the previous measurements. Noticeably in all three plots, the S7 curve falls more steeply than the other two, which are in near-perfect agreement. This pattern is more pronounced where some beta contamination is suspected, and can also be seen, albeit to a lesser degree, in the muon efficiencies in Figure 5.26. Such a discrepancy is not entirely surprising given the differing sensor designs: both the S6 and V20 feature 16 MP sensors (albeit with different aspect ratios) with a 1.12 µm pixel pitch, while the S7 increases both the individual pixel size and overall sensor area.

6.4 Extrapolation to cosmic-ray gammas

These photon efficiencies—though valuable in their own right—would not directly translate to an $A_\gamma \epsilon_\gamma$ value for CRAYFIS even without the complications from beta-radiation. Though the wax-shielded detector cross sections for the three sources in Figure 6.1 are within an order of magnitude, significant discrepancies arise from their different photon spectra, which are in turn far different from the photon spectrum of UHECR showers. The spectral lines for the simple decay chains of ^{60}Co and ^{137}Cs are presented in Table 6.1, while the more complex spectrum of ^{226}Ra is shown graphically in Figure 6.2; to broadly simplify, the ^{60}Co spectrum is exclusively composed of MeV-scale photons, ^{137}Cs features a single prominent gamma line at 662 keV with substantial X-ray activity, and ^{226}Ra encompasses a broad photon spectrum from tens of keVs to several MeVs. Conversely, the photon spectrum in UHECR showers (in this case, generated through bremsstrahlung) extends many orders of magnitude, with a broad peak between 1–50 MeV as shown in Figure 6.3. As MeV-scale photons, both in ^{60}Co and the Pb-filtered ^{226}Ra spectrum, produce steeper efficiency curves, the response is undoubtedly sensitive to the photon energy. Further increasing the energy scale by a factor of 10 or 100 may yield a significantly different $A_\gamma \epsilon_\gamma$, for better or for worse.

However, this is not the only complication in extending these values of $A_\gamma \epsilon_\gamma$ to

Table 6.1: Significant lines in the photon spectra (gamma and X-ray) of ^{60}Co and ^{137}Cs, taken from Ref. [140].

Prominent spectral lines		
Isotope	E [keV]	Prob. per decay [%]
^{60}Co	1173.2	>99.9
^{60}Co	1332.5	>99.9
^{137}Cs	661.7	85.1
^{137}Cs	32.2	3.76
^{137}Cs	31.8	2.04

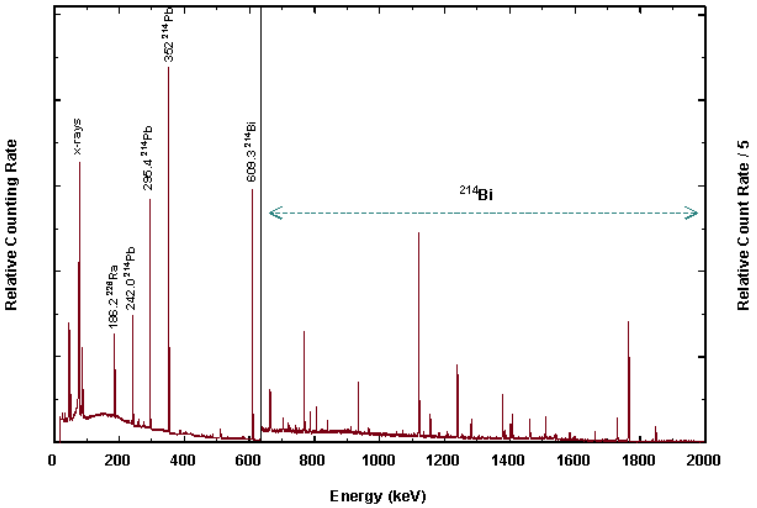

Figure 6.2: Photon spectrum for ^{226}Ra as seen in a high-purity germanium (HPGe) detector. Prominent lines are marked according to the daughter isotope which generated them: primarily ^{214}Pb and ^{214}Bi. From Ref. [141]

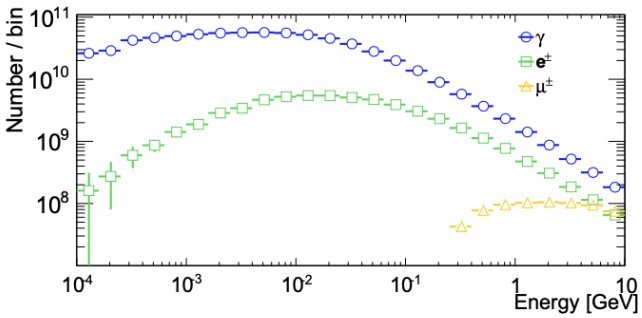

Figure 6.3: Simulated UHECR spectra for the three main secondary components at sea level, from Ref. [83].

CRAYFIS. As the photons produced by the radioactive sources can only generate sub-MeV secondaries, the lens glass alone will shield out almost all such electrons generated outside the glass. This ensures that only photons incident on the lens glass are capable of triggering a sensor response, and furthermore, that a slightly thicker lens would not significantly affect the detector cross section. Cosmic-ray photons, however, can scatter electrons at much higher energies and with much greater penetrating depths, allowing electrons generated in the overburden to trigger the sensor. As the distance between this overburden and the sensor will typically be several meters for a CRAYFIS detector, secondaries created by a single photon (e.g. through multiple Compton scatters) may spread an area much larger than the sensor, increasing the acceptance relative to the sensor size. Such effects will be considered quantitatively in Chapter 8 through simulation.

Chapter 7

MTest beam exposure

As a final characterization of the sensor, phones were exposed to the MTest beam at the Fermilab National Accelerator Laboratory (FNAL). Although the primary beam was composed of 120 GeV protons rather than a prevalent constituent of cosmic ray showers at sea level, the incredibly high flux—200,000 per 4.2 s spill—generated far higher statistics than either of the previous experiments with which models of the sensor response could be fine-tuned. Unfortunately, the resulting efficiency characterized an unusual digital filter rather than the physical response of the sensor.

7.1 Experimental design

In the tests discussed in Section 3.2, the initial Crayfis app exhibited two primary flaws: a low resolution and an excessively restrictive online pixel trigger. The FishStand app had already resolved the former: with efficient triggering in RenderScript, FishStand could comfortably process full-resolution RAW buffers, albeit at a low frame rate. To eliminate the latter concern entirely, it was decided to save entire RAW buffers captured during a spill rather than individual pixels, preserving the full sensor response for offline analysis. The remainder of the experiment was designed around this choice.

First, a suitable phone model was chosen. In the 1 min period between spills, a phone could typically process and upload 3–4 RAW buffers to the AWS backend; to keep pace with the beam without dropping frames, a frame rate below 1 Hz was

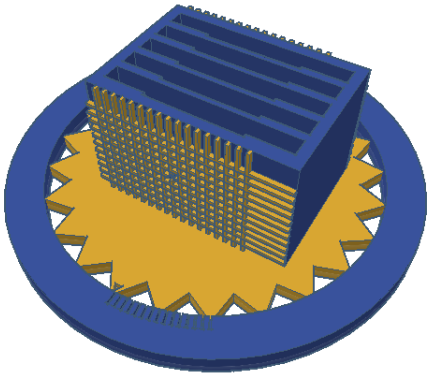

Figure 7.1: A rendering of the phone case used in the beam test at FNAL. The grids along the faces and the ticks on the turntable were designed for manual repositioning guided by the laser alignment system, but a motion table was secured for this purpose instead.

therefore required. Of the three models calibrated for the hodoscope, only the LG V20 had this capability, and the lack of lens-shading on this model was an additional advantage. However, without sufficient knowledge at the time, the Sony IMX298 variant was chosen.

In order to extract an efficiency from the sensor response, a tracker was also needed. Although a silicon telescope was held at the Fermilab Test Beam Facility (FTBF) for this purpose, its capabilities were suboptimal for this particular design. With such long frame exposures, fluxes on the order of 50,000 particles per frame were expected, leaving only 10–20 microns between incident particles on each buffer; in this regime, the telescope's 5 µm spatial resolution was significant, while its nanosecond precision was gratuitous for such a low frame rate. Instead, several phones aligned normally to the beamline were used as a tracker—by construction, possessing the desired spatial resolution—from which the efficiency of each sensor plane could be measured. A case, shown in Figure 7.1, aligned five LG V20s to roughly 0.2 mm precision, with the phones themselves inverted to allow access to their charging ports. A motion table was then used to align this case to within 1° of the beamline.

7.2 Trigger algorithms and calibrations

Two separate triggers were required for this experiment: an online trigger to select frames exposed to beam spills and an offline trigger to identify hits in the uploaded RAW frames. The first of these took the form of an additional `Analysis` in the FishStand app, designed to identify, compress, and upload frames during a beam spill. Because the beam would presumably excite a small fraction ($< 1\%$) of pixels above background, frames were triggered when the mean of the brightest n pixels (denoted as q_n) exceeded some threshold T, where both n and T were be fine-tuned during the beam runs. Such fine-tuning was made possible by two features: remotely adjustable configuration files stored on S3 and a log of q_n values for each frame, both above and below the trigger threshold. The latter proved useful both as a monitoring tool and a source of metadata.

No hot-pixel masking was used in the frame trigger, as it was expected that the beam signal would overwhelm any contribution from hot pixels; however, this assumption was undermined by the lackluster pixel response on the IMX298-equipped V20s. The spills at peak intensity were indeed easy to distinguish, but capturing frames in the rising and falling tails required a precisely calibrated threshold. Hence, while $n = 487$ (i.e. .003% of the total pixels) was used for the entire experiment, T was continually adjusted with the background levels, which were very sensitive to hot pixels and temperature fluctuations.

In order to locate particle hits offline, a simple threshold at 80 DN (roughly 15 e^-) was first applied. At this threshold, a tolerable $\mathcal{O}(10)$ triggers were expected per frame due to background noise, a rate confirmed through additional frames saved between spills. This low rate was made possible by aggressive hot-pixel cuts: PixelStats was run in dark conditions both before and after the beam test, and cuts on the second brightest values for each pixel excluded slightly over 1% of each sensor. An occupancy cut was then performed on the offline pixel triggers, removing any remaining over-represented pixels in the beam data; see Figure 7.2.

These excited pixels were then grouped into clusters, with which a proper analysis

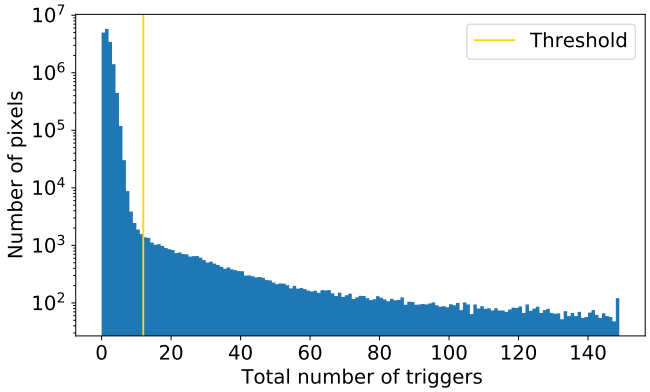

Figure 7.2: Pixel occupancies after cuts on second brightest values for dark frames. An additional cut removed the tail past 12 counts out of roughly 17,000 beam-exposed frames.

of particle hits could be performed. The DBSCAN algorithm [142] was used to connect pixels within a certain Euclidean distance, forming extended chains; a number of these stretched hundreds or even thousands of pixels in straight tracks across the sensor. To appropriately set the distance threshold for this clustering algorithm, the spectra (maximum pixel values per cluster) resulting from integer thresholds from 0 (no clustering) to 9 were compared in Figure 7.3. A significant drop is seen between black (no clustering) and purple (clustering adjacent pixels), indicating that large numbers of excited pixels were accompanied by others nearby. However, the spectral differences fade quickly as the clustering threshold increases past 2 (blue), demonstrating particles incident on the sensor very rarely create disjoint regions 3 or more pixels apart.[1] Accordingly, this clustering algorithm was applied with a threshold of 3. For a more straightforward analysis, these clusters were later considered as points, using the location of the brightest pixel in each cluster.

However, the tenfold drop in the spectrum above a per-cluster maximum pixel response of ∼100 DN required an explanation. When the same clustering algorithm was implemented with pixel thresholds above this cutoff, no analogous feature was present; hence, this was not an artifact of the thresholding or clustering by themselves. Furthermore, the spatial profile of the clusters above and below this value were identical,

[1]This excludes secondaries created in the lens which strike a separate region of the sensor, which could conceivably be very common. Rather, this analysis is restricted to pixels excited by a single particle leaving tracks in the silicon.

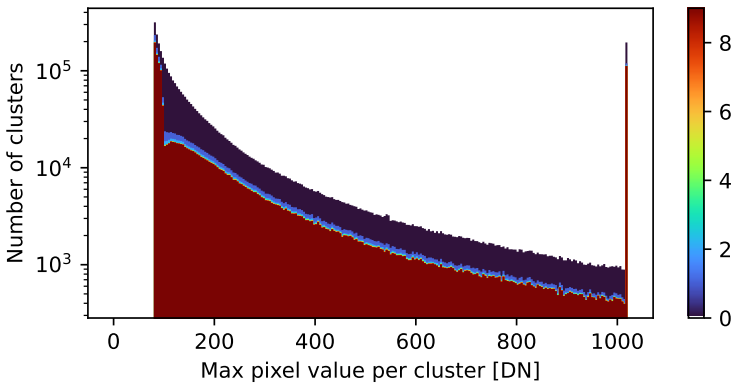

Figure 7.3: The maximum pixel response per cluster for Euclidean distance thresholds between 0-15 in a particular run. Spectra for all other runs are virtually identical.

suggesting that the excess events were not caused by noise or hot pixels. This feature was also observed on all frames throughout the various beam runs. This irregularity appears to be a true feature of the pixel response, a possible explanation of which will be offered in Section 7.5.

7.3 Alignment

To locate tracks with the five sensor planes, it was first necessary to reconstruct the alignment of the frames, both in time and in space. While the efficiency analysis (to be discussed in Section 7.4) only requires that the relative ordering of frames is discovered, the spatial alignment must be precise to the scale of a single pixel.

7.3.1 Timestamp synchronization

The Android API contains two different types of timestamps: nanosecond timestamps, which precisely measure time differences but do not have a defined zero point, and millisecond timestamps, which measure the wall clock time but with less precision. These two were blended into a hybrid measurement in the FishStand app: the first frame of a run was assigned the millisecond timestamp at capture, and subsequent frames used this initial measurement as a fixed point to which nanosecond time differences are added. For calibrating drift, this preserved the precision of the nanosecond timestamps

while providing a reasonable reference point to align the various clocks. Due to the continual threshold adjustments, however, this design functioned poorly; because the reference timestamp would be recalculated with each new run, and each new threshold adjustment *created* a new run, hundreds of new runs—and hundreds of independent offsets—were generated per phone throughout the experiment.

To calibrate the many separate run clocks for each phone, the beam intensity as a function of time was used as a common reference. The trigger statistic q_n was by design a suitable indicator of the beam flux: sharp rises in pixel activity could be associated with the beginning of a spill, while sharp declines corresponded to its end. Due the slowly fluctuating background levels, Δq_n between adjacent frames proved to be a more effective discriminator for beam activity than the magnitude of q_n alone: in Figure 7.4, Δq_n is plotted against frame timestamps in an interval encompassing several spills, from which the spill times can easily be identified. Spills were then identified by pairs of Δq_n values, the first positive and the second negative, in nearby frames satisfying $|\Delta q_n| > 30$. The full duration of each spill was then established by incrementally applying a lower Δq_n threshold of 6 to frames at the beginning and end of each discovered interval.

However, these timestamps were only as precise as the (considerable) frame duration. For a more refined estimate of the spill times, a "midpoint" was derived for each spill by weighting the timestamps by the corresponding background-subtracted q_n values. With these more precise spill timestamps, spills could be more finely matched across phone clocks.

However, using the spill times as a reference requires that a spill on one phone can be matched with the *same* spill on the others. To assure that the frame timestamps had not drifted further than the spill interval of 1 min, the timestamps were compared to the corresponding upload times on the AWS backend—the one common reference clock shared between all phones. Delays between capture (phone clock) and upload (AWS clock) were found to be roughly 20 ± 10 s, where the error encompasses both the timestamp drift and varying upload rates. As the drift was much smaller than

than the beam interval, spills found on one clock could rightly be associated with the spills on the others with the nearest reported times.

In the absence of a true standard clock, these timestamps were compared to one of the phone clocks (hereafter denoted as the "reference phone"), which experienced the least downtime during the experiment. A roughly linear drift, shown in Figure 7.5, was found between the estimated spill times, with some variation across runs. In the final night, two of the phones were rebooted, and thus separate fits were performed.

As discussed, residuals from this drift fit were primarily caused by the hundreds of different offsets calculated with each new run. Conceptually, this can be written as:

$$\begin{aligned} t_i &= t_{\text{spill}} + r_i + \delta t_i \\ t_{\text{ref}} &= t_{\text{spill}} + r_{\text{ref}} + \delta t_{\text{ref}} \,, \end{aligned} \tag{7.1}$$

and so:

$$\Delta t_{i,\text{ref}} = (r_i - r_{\text{ref}}) + (\delta t_i - \delta t_{\text{ref}}) \tag{7.2}$$

where t_{spill} is the true spill time, t_i is the drift-corrected spill time measured by phone i (or the reference phone), and r_i and δt_i are the run offset and random error associated with t_i.

To correct for the run offsets, the residuals $\Delta t_{i,\text{ref}}$ of the fits from Figure 7.5 were computed and grouped by the run number on the reference phone; see Figure 7.6a. For each run, the mean of all $\Delta t_{i,\text{ref}}$, including all values of i, is denoted by gold horizontal lines. As $E(\delta t_i - \delta t_{\text{ref}}) = 0$ and $E(r_i - r_{\text{ref}}) = 0$, averaging $\Delta t_{i,\text{ref}}$ across multiple spills and phones should reasonably estimate the run offsets specific to the reference phone. Any remaining residuals were attributed to the run clocks of the other four phones, one of which is shown in Figure 7.6b.

Though still inexact, this degree of precision was sufficient for reconstructing the spatial alignment. However, more exact timestamp corrections were required for the efficiency measurements, which at least preserved the ordering of frames in different

phones. For each spill, the number of two-sensor particle tracks were computed for all possible pairs of frames (discussed below) and compared to combinatorial noise. The run offsets were then further adjusted so that excess numbers of tracks were found only in frames overlapping in time.

7.3.2 Spatial alignment

With the clocks reasonably well synchronized, the spatial alignments of the five phones could be reconstructed. Fundamentally, this entailed transforming the pixel coordinates from one sensor to another such that a significant number of corresponding hits could be found between the two. This process requires two essential steps: finding an appropriate parametrization for the transformation, and the computing the best fit.

Parametrization

As the divergence of the beam is small, and thus small phone displacements in the longitudinal direction are insignificant, the most general transformation relating sensor planes is a transverse translation coupled with a three-dimensional rotation. However, as only the phones' *relative* positions are relevant for reconstructing tracks, several of these degrees of freedom are unnecessary: it is therefore beneficial to reparametrize these transformations in a more useful form.

This general transformation from the sensor to the lab frame can be written in terms of Eulerian angles as:

$$\begin{aligned} x_l &= (x_s + x_{s,0})(\cos\psi\cos\theta\cos\phi - \sin\psi\sin\phi) \\ &\quad -(y_s + y_{s,0})(\sin\psi\cos\theta\cos\phi - \cos\psi\sin\phi) \\ y_l &= (x_s + x_{s,0})(\cos\psi\cos\theta\sin\phi + \sin\psi\cos\phi) \\ &\quad -(y_s + y_{s,0})(\sin\psi\cos\theta\sin\phi + \cos\psi\cos\phi) \end{aligned} \tag{7.3}$$

where (x_0, y_0) is the spatial translation in the sensor frame, and the lab frame is centered and perfectly normal to the (nominal) beamline. Since the angular alignment

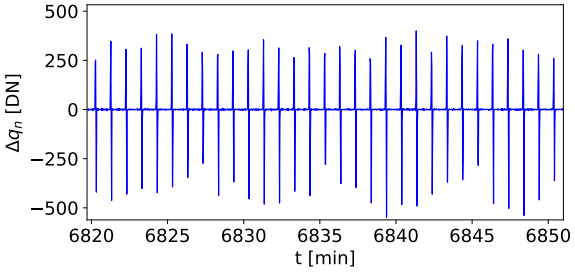

Figure 7.4: Values of Δq_n for adjacent frames. An interference pattern can be observed from the beam frequency and the frame rate.

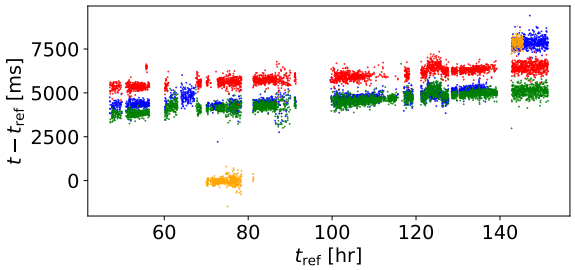

Figure 7.5: Differences in spill times recorded by the reference phone and the other four, plotted against time. The phone plotted in orange had a faulty charging port, and was consequently inactive for most of the experiment.

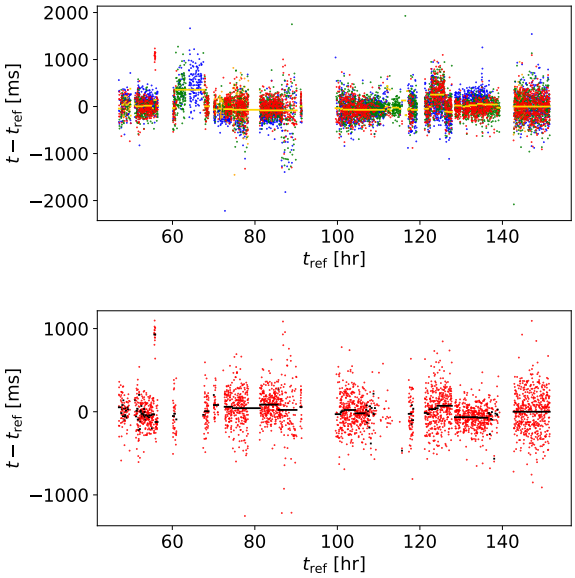

Figure 7.6: Corrections for run offsets, for the reference phone (top) using the residuals for for the other four combined, and the other four individually (one of which is shown at bottom), based on any residuals remaining from this first step. Horizontal lines denote a single run, and are positioned vertically at the computed offset. Colors shown are the same as in Figure 7.5.

of the phones is fairly precise in the transverse direction, $\psi \approx -\phi$. Writing this as $\psi = -\phi + \delta$ and redefining the translations in terms of the lab frame, this becomes, in matrix notation:

$$\begin{pmatrix} x_l \\ y_l \end{pmatrix} = \begin{pmatrix} x_{l,0} \\ y_{l,0} \end{pmatrix} + \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \cos\theta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \cos\delta & -\sin\delta \\ \sin\delta & \cos\delta \end{pmatrix} \begin{pmatrix} x_s \\ y_s \end{pmatrix}. \tag{7.4}$$

The rightmost of these four matrix transformations is a rotation in the lab space, whereas the others compress the sensor coordinates about some axis. The transformations in ϕ and θ can be reorganized as:

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \cos\theta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} = \mathbb{1} - \mathbf{v} \cdot \mathbf{v}^T \tag{7.5}$$

for $\mathbf{v} = (\cos\phi, \sin\phi)$. While the spatial offset $\mathbf{x}_{l,0}$ may be quite large relative to the pixel scale, the precision of the alignment assures that $\delta \approx 0$ and $\mathbf{v} \approx \mathbf{0}$. Consequently, a grid search in this parameter space is much simpler than with Eulerian angles, in which θ and ϕ are unconstrained.

Moreover, in the approximation that the beam divergence is negligible, the two-dimensional (i.e. transverse) lab coordinates for each particle hit will be identical on all five sensors. Therefore, these hits can be evaluated in the sensor frame of the reference phone, rather than in the lab frame, with no loss of information. Practically, this entails a forward transformation to the lab frame, followed by an inverse transformation to the desired sensor frame. To first order, this can be reparametrized as:

$$\mathbf{x}' = \mathbf{x_0} + R(\delta)(\mathbb{1} - U)\mathbf{x} \tag{7.6}$$

where R is the 2-dimensional rotation matrix, and

$$U = \begin{pmatrix} u_x & u_{xy} \\ u_{xy} & u_y \end{pmatrix}$$

absorbs the two off-plane transformations previously represented by $\mathbf{v}$. Hence, only 6

parameters are needed to characterize the alignment.

Computation

To find these transformation parameters between pairs of phones, an objective function was devised to score the goodness of fit across the parameter space. Several considerations were included in this choice. First, for any pair of frames on different sensors, their corresponding time intervals will only overlap partially; thus, even with perfect sensor efficiency, only a fraction of hits will have a partner on the opposite sensor. An ideal objective function would therefore gain from closely aligned pairs of hits, while not suffering substantially from unpaired hits. Second, the standard for a "good" or "bad" fit for a pair of hits will vary according to the degree of precision of the alignment. An initial coarse alignment would benefit from a more lenient standard, while more precise refinements would gain from more stringent requirements. Hence, an adjustable length scale would also be advantageous.

To incorporate these elements, the degree of closeness between a pair of hits is first quantified. For uncorrelated and uniformly distributed sets of points A and B on the same coordinate system, the distance between any point in A to its nearest neighbor in B will follow the cumulative distribution function:

$$F(r) = 1 - e^{-\pi \rho_B r^2} \tag{7.7}$$

where ρ_B is the local density of points in B. From this distribution, an appropriate objective function can be written as follows:

$$\mathcal{L}_\alpha(\mathbf{x}_A, \mathbf{x}_B) = \alpha^{-1} \sum_A \max\left\{ e^{\pi \rho_B r_{\min}^2} - (1-\alpha), 0 \right\} . \tag{7.8}$$

For each cluster in sensor A, the CDF is first computed given the nearest cluster in B. The resulting p-value is then compared to a threshold $1-\alpha$: clusters with p-values above α are ignored, while those with exceptionally close neighbors in B contribute towards the total. Naturally then, the value of α can be fine-tuned to the specific

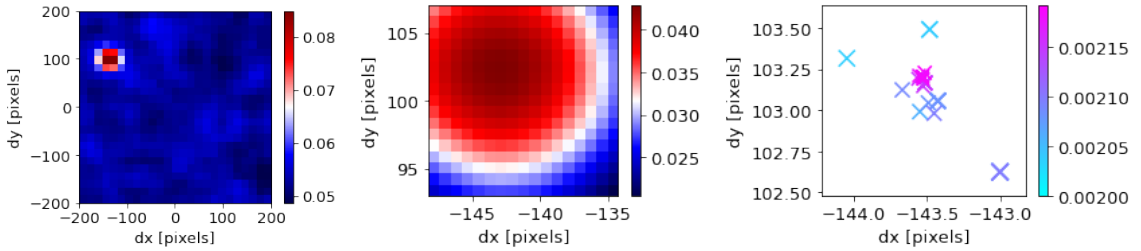

Figure 7.7: The alignment process for pairs of phones. Left: a grid search is first performed in x and y using a larger value of $\alpha = 0.1$, from which the region of interest can be located. Middle: successive grid searches are performed with higher resolutions, and including δ. Right: when the scale in parameter space is small enough so that the objective function is well-behaved, Nelder-Mead optimization is used to precisely locate the best fit in all six variables, using a larger sample of frames and α values around 0.01. (x, y, δ) tuples of successive iterations are indicated by markers decreasing in size.

stage of the calibration.

Using a combination of grid searches and Nelder-Mead optimization (Figure 7.7), best fits were found between the reference phone and the other four. As Figure 7.8 shows, the reconstructed alignment was exceptionally precise. The observed variation was characteristic of an angular alignment of the case to the beamline within 0.5° and an alignment of phones within the case to 0.2 mm. The former was roughly the precision of the angular motion table on which the case was situated, and the latter was the approximate layer size of the 3D-printer used to create the case; hence, such an alignment was not unreasonable. This approach was further verified through robust Monte Carlo simulations, in which the alignment was typically reconstructed to single-pixel precision.

7.4 Statistical Analysis

7.4.1 Formalism

In the simplest approach, efficiency can be measured as probability that a particle tagged on two sensors $T = \{T_1, T_2\}$ will likewise be found on a sensor S between them.[2]

[2]Though more sophisticated methods are possible involving more than three sensor planes, systematic uncertainty increases significantly with the number of sensor planes in the approach used here, and the minimal case will therefore be treated exclusively.

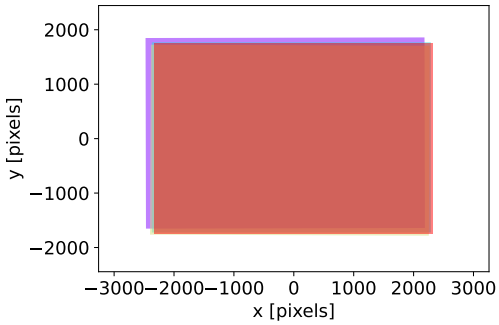

Figure 7.8: The positioning of the sensors as seen from the beamline in a typical run, in units of the pixel pitch (1.12 µm). Due to phones' positioning within the case, the x axis was oriented vertically in the lab; thus, this display is rotated 90°from the actual lab view.

Assuming a particle's probability of triggering a particular sensor is independent of all other sensors, the efficiency can be written as:

$$\epsilon = P(S \mid T) = \frac{N_{S \cap T}}{N_T} \,. \tag{7.9}$$

When a track is found on T_1 and T_2, however, the exact frame on S corresponding to this particle is not typically known; due to the unsynchronized phone clocks, each hit may potentially be found in one of two frames. Both the numerator and denominator in (7.9) must therefore be construed as sums over all possible sets of overlapping frames.

Tracks between the three sensor planes are constructed through spatial coincidence windows, i.e., two hits are considered a track when $|\Delta x| \leq c_x$ and $|\Delta y| \leq c_y$, for values of (c_x, c_y) calibrated to each pair of sensors.[3] Three-sensor tracks are then assembled from pairs of two-sensor tracks sharing a common hit on the middle sensor. This preference given to the middle sensor is not arbitrary: due to the finite beam divergence, larger coincidence windows are needed as the physical distance between phones increases. Computing from the middle sensor therefore yields less combinatorial noise and smaller uncertainties.

To compute the combinatorial background—the expected number of false tracks from the distribution of hits—the simplest case is first considered: that of the two-

[3]As all tracks satisfying these proximity requirements are counted, a single hit may be included in several tracks. Yet, rather than being a defect, this substantially simplifies noise calculations.

sensor coincidence count Q_2. If there are N_1 and N_2 uncorrelated, uniformly distributed hits on two frames, F_1 and F_2, N_1N_2 possible "tracks" can be drawn between them; however, only a small fraction will be within the coincidence window, selected with area $C = (2c_x + 1)(2c_y + 1)$. Since each the N_1N_2 tracks has a probability C/A of being counted, where A is the area of intersection between the two sensors relative to the beamline, the mean number of coincidences found will therefore be:

$$E(Q_2) = N_1 N_2 C A^{-1}\,. \qquad \text{(No true tracks)}$$

For the case of a nonuniform distribution of hits $P(\mathbf{x})$ on both sensors, it can be shown that the factor of A^{-1} is replaced by $\int_A P(\mathbf{x})^2\,dA$.

If correlated hits—true tracks—are then introduced, where N_{12} hits out of both N_1 and N_2 will by construction be counted as coincidences, the total number of coincidences becomes:

$$E(Q_2) = N_{12} + (N_1 N_2 - N_{12})\, C \int_A P(\mathbf{x})^2\,dA\,. \tag{7.10}$$

As $C\int_A P(\mathbf{x})^2 dA \ll 1$, the number of true tracks can be estimated from the observed coincidence count q_2 as:

$$N_{12} = q_2 - N_1 N_2 C \int_A P(\mathbf{x})^2\,dA\,. \tag{7.11}$$

Extending this to three-sensor coincidences, combinatorial noise may be caused by three unrelated hits, or more likely, by real two-sensor tracks with a spurious hit on the third. The corresponding background correction becomes:

$$\begin{aligned} N_{123} = q_3 &- (N_1 N_{23} C_1 + N_2 N_{13} C_{13} + N_3 N_{12} C_3) \int_A P(\mathbf{x})^2 dA \\ &- N_1 N_2 N_3 C_1 C_3 \int_A P(\mathbf{x})^3\,dA \end{aligned} \tag{7.12}$$

where N_{ij} is calculated from (7.11), and C_1 and C_3 are the areas of the coincidence windows between the middle and outer sensors. C_{13}, meanwhile, quantifies the prob-

ability of a real track through F_1 and F_3 being matched with an unrelated hit on F_2. Using $\boldsymbol{\imath}_i = \mathbf{x}_i - \mathbf{x}_2$, this is calculated as:

$$C_{13} = \int_{C_1} d\,\boldsymbol{\imath}_1 \int_{C_3} d\,\boldsymbol{\imath}_3\, P(\boldsymbol{\imath}_1 \mid \boldsymbol{\imath}_3) \tag{7.13}$$

where $P(\boldsymbol{\imath}_i \mid \boldsymbol{\imath}_3)$ is the probability that if a hit is found at $\boldsymbol{\imath}_3$ on F_3, a hit will likewise be found at $\boldsymbol{\imath}_1$ on F_1, averaged over all values of $\mathbf{x}_2$.

7.4.2 Intermediates and systematics

Several key elements in the previous section are non-trivial to extract from the data. In particular, $\int P(\mathbf{x})^2\, dA$ and $\int P(\mathbf{x})^3\, dA$ must be calculated from the observed beam profile, and the coincidence windows (c_{xi}, c_{yi}) require a careful calibration—from which the rather opaque Eq. (7.13) must be applied. Additionally, the analysis neglects that two or more particle hits may be clustered into one, for which a correction to first-order is proposed.

Beam profile

In order to calculate $\int_A P(\mathbf{x})^k\, dA$, two-dimensional histograms of the hit occupancy—that is, a non-normalized $P(\mathbf{x})$—were first constructed for each phone, using only the overlapping area of each sensor; see Figure 7.9a. Since the hits per bin will be approximately Poisson-distributed,[4] an unbiased estimator for the mean number of hits squared in a particular bin is:

$$\begin{aligned} E(N^2 - N) &= E(N^2) - E(N) \\ &= E((N - \bar{N})^2) + E(N)^2 - E(N) \\ &= \sigma_N^2 + E(N)^2 - \bar{N} \\ &= E(N)^2\,. \end{aligned} \tag{7.14}$$

[4]Since only a single hit per pixel can be recorded per frame, the distribution is actually binomial, but these distributions converge as the number of frames approaches infinity.

For higher order terms, it can likewise be shown that $E\left(\prod_{i=0}^{n-1}(N-i)\right) = E(N)^n$. The integral can then be calculated by normalizing this statistic and computed the sum of the entire grid.

However, the accuracy of this measurement is dependent upon the chosen binning size. Small bins become overly sensitive to hot pixels—both cuts on normal pixels with high trigger counts and mild hot pixels which escaped these cuts will distort this statistic—while larger bin sizes will more poorly capture the curvature of the profile. For each phone, then, the calculated $\int_A P(\mathbf{x})^k \, dA$ is evaluated as a function of chosen bin size in Figure 7.9b. To present this in more relatable units, the reciprocal of this integral (i.e. the effective overlap area) is instead plotted. A bin size of 8 on the edge of the plateau was selected for subsequent analysis.

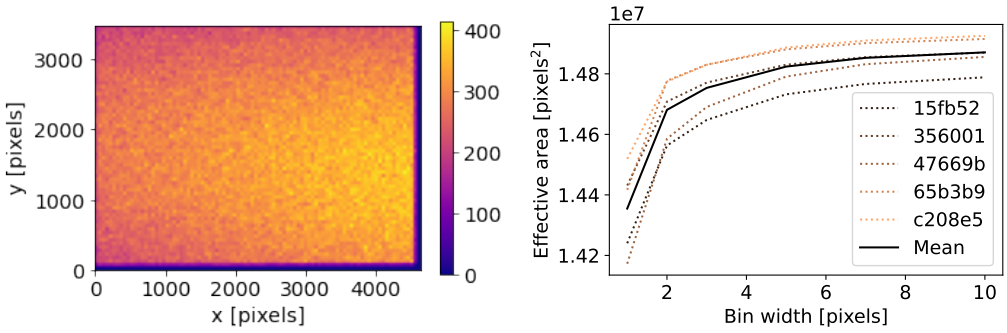

Figure 7.9: Measurement of $\int_A P(\mathbf{x})^k \, dA$. Spatial histograms of hits were first constructed for each phone (at left, the phone furthest upstream is shown). The bottom and right edges are excluded, as this region in physical space is not sampled by one of the other sensors. This profile is used to estimate the $\int_A P(\mathbf{x})^k \, dA$ (right) as a function of the bin width; specifically, each mark n on the horizontal axis corresponds to bins of dimension $n \times n$.

Several features are noteworthy in Figure 7.9. A larger variation is present in the per-device curves than was found in simulated data, and the curves, from darkest to lightest, fall in order the phones were positioned from upstream to downstream. Likewise, the total hits per spill also monotonically increases with distance downstream. Both of these trends are presumably due to scattered secondaries, which decrease the sharpness of the beam profile.

Coincidence windows

To estimate the spatial spread between corresponding hits on different sensors, the sensors were evaluated in pairs: for each event on the first sensor, all nearby hits on the second (using both co-temporal frames, if they were recorded) were added to a two-dimensional histogram, and the combinatorial background from (7.11) was subtracted from the resulting counts. This is displayed for a typical run in Figure 7.10. The shape of these excesses mirror the observed beam profile, and their length scales correspond to the physical distance between the each pair of phones.

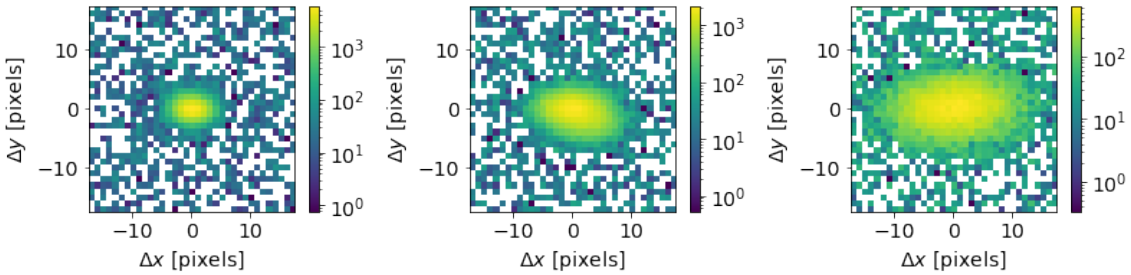

Figure 7.10: Histograms of relative spacing between hits on different sensors, corrected for background. The spread increases from adjacent phones (left) to phones two (middle) and three (right) spaces apart in the case. The pixel pitch was used as the bin size, although the transformed coordinates were no longer integer values.

From these plots, coincidence windows were chosen encompassing the entire excess for each pair of sensors; for the plots in Figure 7.10, for example, windows of dimension 9×7, 19×11, and 23×13 were selected. Moreover, these histograms, when normalized, provide the probability $P(\boldsymbol{r}_i \mid \boldsymbol{r}_j)$ needed in (7.13). By integrating over the appropriate limits set by the choice of coincidence windows, this last remaining term in (7.12) could be computed.

Clustering corrections

Finally, a correction was devised for the effects of clustering, which may combine two nearby particle hits into one. As the analysis above assumes that all hits correspond to only one particle, both the number of two-sensor and three-sensor tracks in (7.11) and (7.12) will be underestimated. Consequently, a relationship between the number

of clusters n and the true number of triggered particles N is sought, from which $N(n)$ can be applied in these formulae.

The inverse $n(N)$ is first approached by considering N particles sequentially triggering a sensor. According to (7.7), the ith hit has a probability

$$p_i = e^{-\pi\rho_{i-1}L^2} \tag{7.15}$$

of not being clustered with any of the previous hits, where L is the clustering threshold and, assuming a uniform spatial distribution, the previous density of hits ρ_{i-1} is simply $(i-1)/A$. This additional hit will either increment the current number of clusters by 1 (with probability p_i) or leave it unchanged;[5] after N hits, the number of clusters can therefore be written as a sum of Bernoulli random variables with probabilities p_i:

$$n = \sum_{i=1}^{N} X_i \tag{7.16}$$

for which the expected value (i.e. $n(N)$) is:

$$\begin{aligned} E(n) &= \sum_{i=1}^{N} p_i \\ &= \sum_{i=1}^{N} e^{-(i-1)\pi L^2/A} \\ &= \frac{1-e^{-N\pi L^2/A}}{1-e^{-\pi L^2/A}}\,. \end{aligned} \tag{7.17}$$

Inverting this yields the desired quantity:

$$N = -\frac{A}{\pi L^2}\log\left(1-n\left(1-e^{-\pi L^2/A}\right)\right)\,. \tag{7.18}$$

With a toy Monte Carlo, this formula was shown to be accurate to within a few percent for the range of hit densities observed in this experiment. To correct for a nonuniform beam profile, A^{-1} is again replaced by $\int_A P(\mathbf{x})^k\,dA$.

[5]While there is a small probability that a hit will "bridge" two other clusters and decrease the total number of clusters, this outcome is not considered here.

7.4.3 Verification with Monte Carlo

To verify that these methods accurately reproduce the particle efficiency, the analysis was applied to simulated data. In this Monte Carlo, particles were ray-traced through four sensor planes, and pixels intersecting with these particle trajectories were saved with a fixed efficiency and assigned digital response values from a uniform distribution. The particles were assigned a Gaussian beam profile with a flux, spot size, and divergence configured to the observed beam parameters. Moreover, the sensor planes were given fully general displacements and rotations with a slightly poorer alignment than was measured experimentally. Hits were grouped into frames by a realistic buffering system, and clustered with the same threshold used in Section 7.2.

However, several noteworthy features of the experiment were not replicated in the simulation. While a small drift was applied to the timestamps, the process of correcting run offsets was bypassed. No hot pixels were included in the simulation, nor were scattered secondaries. As one last point of departure, hits were created as single excited pixels rather than extended clusters.

As the full analysis was time-intensive, trials were performed individually rather than in aggregate through a pull distribution; an example trial is shown in Figure 7.11. In this particular simulation, the spatial alignment algorithms in Section 7.3.2 were first applied to a set of 300 spills with an efficiency of $\epsilon = 0.2$ at the lowest threshold. Using the analysis of this section, curves of efficiency vs. threshold were constructed for four permutations of (T_1, S, T_2), shown in Figure 7.11a. As a uniform distribution of triggered pixel values was used, the true efficiency decreases linearly with the applied threshold, clearly reflected in the efficiency reconstruction. Averaging over the four curves, the accuracy was roughly within 1% of the true value, as shown in Figure 7.11b. Systematic error derives principally from the clustering of particle hits; when clustering was excluded from the simulation, the accuracy improved by an order of magnitude. Although the clustering correction proposed in the previous section shifts the result in the correct direction, the magnitude of this correction is far too small.

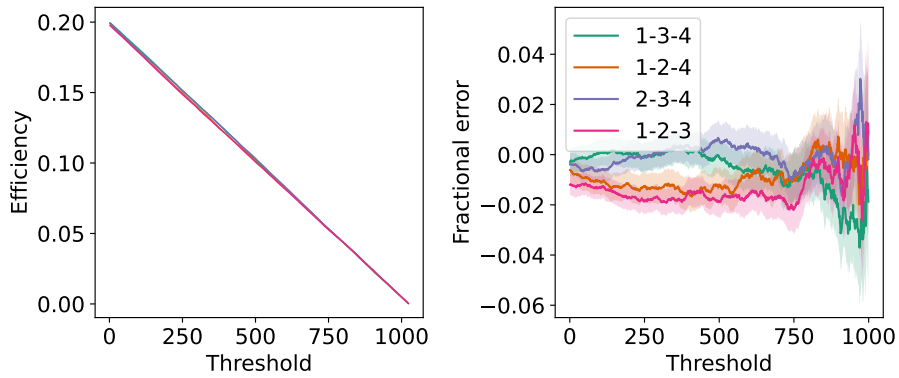

Figure 7.11: Computed efficiency vs. threshold for simulated data (left), and fractional error from the true efficiency curve (right). The labels denote which of the four sensors were used in each curve, in the format T_1-S-T_2, with sensor 1 being the furthest upstream.

7.5 Results and discussion

Four runs were analyzed, lasting several hours each with different beam profiles and intensities. Since the computed values of $\int_A P(\mathbf{x})^k \, dA$ were a significant source of systematic error, efficiency measurements were restricted to sets of 3 adjacent phones, under which the coincidence windows and the combinatorial background were minimized. As one of the phones was offline for most of the experiment, only seven such datasets out of a possible 12 were available over these four runs, presented in Figure 7.12 below. The computed efficiencies were comparable to the far-geometry configurations in Figure 5.25b, with a similar ankle in the power law spectrum at roughly the same threshold.

This confirms the results of Chapter 5 for this particular sensor, the Sony IMX298. As the hodoscope results showed, this is not a general feature of smartphone sensors or of the physical device, but a particular feature of this sensor model. As the IMX298 and its twin, the Samsung S5K2P7, are identical in pixel pitch and likely pixel depth while displaying vastly different responses (Figure 5.25), the physical pixels are unlikely to be at fault for this unusually low efficiency. Rather, the camera software is most likely responsible.

An unusual form of digital processing can can be inferred from the IMX298 pixel statistics plotted in Figure 7.13. Here the per-pixel variance and mean are plotted a

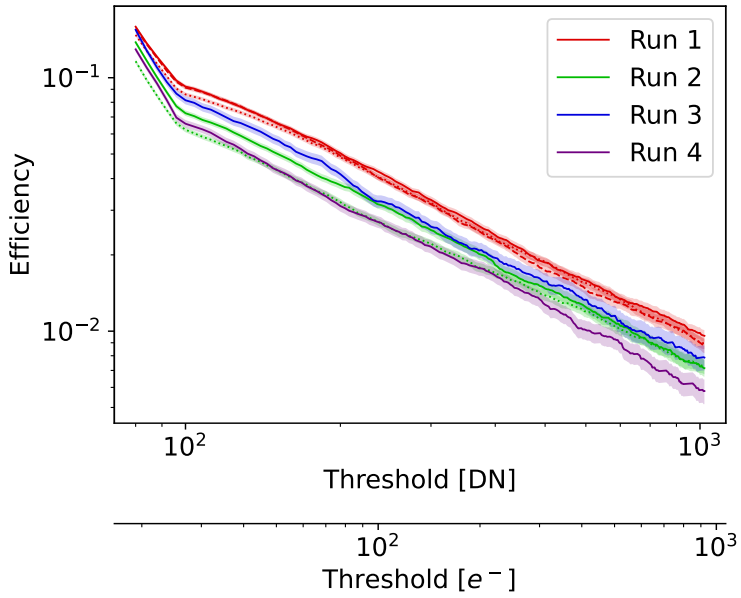

Figure 7.12: Efficiency vs. threshold for four beam runs, arranged in (nonlinear) order of intensity by color, with red corresponding to the highest flux. In the terminology of Section 7.4.3, the phone configurations 1-2-3 (dotted), 2-3-4 (solid), and 3-4-5 (dashed) are presented where available. The displayed uncertainties reflect only the random error observed among spills, ignoring the systematic error from $\int_A P(\mathbf{x})^k \, dA$.

run in dark conditions, with the second brightest value in the run on the color axis. While a number of notable features are present in this plot, the clearly visible loop on the right is the most unexpected. Typically in such plots, a long tail of hot pixels extends to the right, ending at saturation; here the tail is bent into a counterclockwise loop instead.

This is likely caused by the dynamic masking of hot pixels in the IMX298 [143]. If a hot pixel reads out a high value, but is instead masked by the camera and replaced with a low value, this single measurement would increase the pixel's variance rather than its mean (i). However, when many of these replacements occur, the mean is eventually pulled lower while the variance stays high (ii). For the most active pixels, the majority of measurements are suppressed, decreasing the variance as well as the mean, until at the far end of the spectrum, every value is masked (iii). This interpretation is supported by the second maximum statistic, which increases along the loop until the cluster marked at (iii). If such an algorithm also targeted solitary bright pixels created by incident protons or muons, the IMX298 would naturally perform poorly in the beam

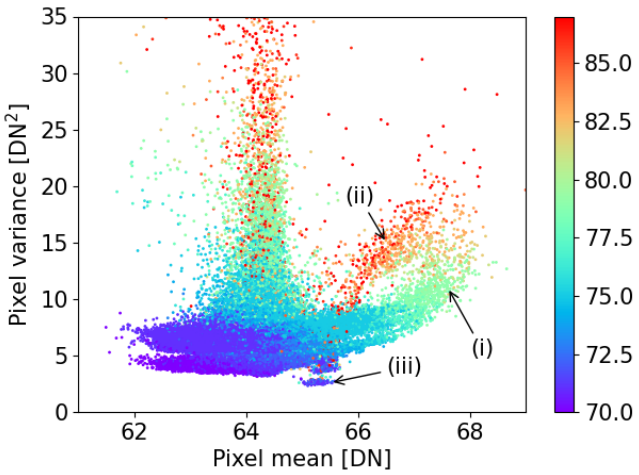

Figure 7.13: Sample variance vs. mean for a dark run of the Sony IMX298. Description of labels in text.

and hodoscope experiments.

Such an explanation is supported by several features of the experimental data. In Figure 5.25b, muons tagged by scintillators AC and BC—entering at narrower zenith angles, and thus exciting fewer pixels—exhibited significantly lower efficiencies than those tagged by AB. Likewise, slicing the spectrum in Figure 7.3 (i.e. with the selected distance threshold of 3) by the number of pixels per cluster reveals that this sharp decline is present specifically among solitary pixels, although clusters of 2 experience a smaller dip at the same cutoff;[6] see Figure 7.14. While a larger decline would naturally would naturally be expected in solitary pixels—events with large energy depositions would be more likely to excite neighboring pixels through electron drift or by delta rays—the sharpness in the 1-pixel spectrum and discontinuity in the 2-pixel spectrum clearly suggest a digital effect. As this decomposition shows, the ankle features in Figures 5.25 and 7.12 are formed by the population of 2-pixel clusters overtaking the suppressed tail of solitary pixels.

The variation among these efficiency curves—much larger than the systematic errors observed in the simulation—likewise merits discussion. This variation is caused, at least in part, by systematic error in $\int_A P(\mathbf{x})\, dA$ discussed in Section 7.4.3, which

[6]Under the interpretation provided here, this feature may be caused by, e.g., would-be clusters of three in which the brightest pixel was masked. The left shoulder of the 2- and 3-pixel distributions is simply due to combinatorics, as the maximum pixel value per cluster is plotted.

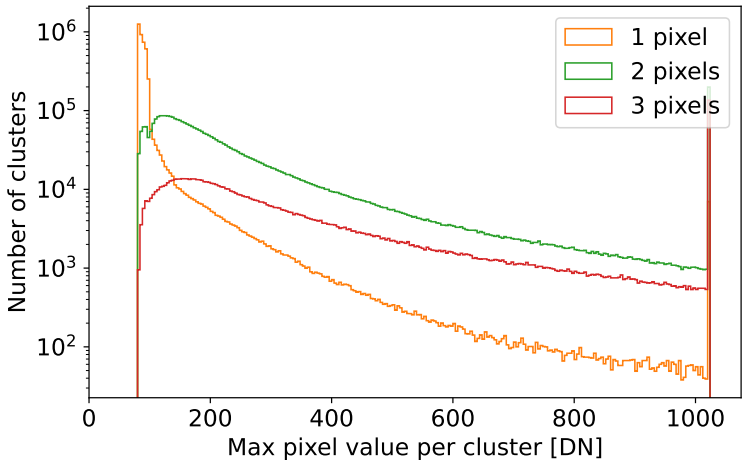

Figure 7.14: Histograms of the highest pixel values in clusters of 1, 2, and 3 pixels, excluding hot pixels, in the spills labelled "Run 1" in Figure 7.12.

the simulation fails to fully consider. Beam runs with lower flux—and hence lower combinatorial noise—are less exposed to this error, and would therefore be more accurate. However, in light of the dynamic masking discussed above, another explanation is possible: if the masking algorithm relies on nearby pixel values, or on the average pixel response of the sensor as a whole, a higher beam flux might lead to fewer tracks being masked. Such a dependence on flux is indeed observed in Figure 7.12. In either scenario, the lower curves would more accurately reflect the phones' sensitivity to more common doses of radiation.

While the case of the IMX298 demonstrates that innovative camera firmware can drastically reduce the efficiency, the trend (for Android, at least) has always been towards more manual camera control. During the development of the Crayfis app, features such as RAW buffer access, customizable lens-shading patterns, hot-pixel masking, and even some control over hot-pixel and black-level algorithms have been added to the developer's toolkit. While not all devices support each of the camera options afforded by the Android platform, in general, newer devices tend to be more flexible. Thus, it can be expected that such sensors limited by digital effects will be less prevalent in the future.

Chapter 8

Modelling the Response

With precise efficiency measurements for muons and photons, the shower efficiency of Crayfis could be estimated to a limited degree. However, such an estimate would suffer from several systematic biases, some of which have been previously discussed. First, the photons from the radioactive sources were on the order of 1 MeV, far below typical photon energies produced by UHECR showers. Second, the muon efficiency found in Chapter 5 does not incorporate the full acceptance of secondary electrons being scattered from the overburden in a typical home; while electrons below several hundred keV will be stopped by the lens, GeV-scale muons can certainly scatter electrons above this cutoff. Third, the $e^{\pm}$ component of the showers has thus far been entirely neglected due to its limited range, though GeV-scale electrons are numerous and capable of penetrating through a typical roof. Finally, these results were obtained for only four sensor models, all of which were designed in the years 2015-2016; according to Figure 2.3, more recent smartphone sensors typically contain longer and narrower pixels. All of these effects can be quantified through careful modelling. Beyond such extrapolation, however, simply tethering the calculated efficiencies to the known pixel structure would be a strong confirmation of the results.

Though a large body of literature exists on Monte Carlo simulations of pixel sensors, these typically treat either visible photons and dark noise an image sensor [144–146], or ionizing radiation on significantly larger pixels in a dedicated particle

detector [147–149]. In both cases, the exact geometry and doping profiles of the p–n junction are often known, allowing for a precise simulation of the electric field and of the resultant photocurrent. By contrast, the hybrid case of ionizing radiation on an image sensor is considered here, where the n-well size is considerable relative to the p-substrate. In addition, the fine details of the geometry are not known, and so a general model is instead sought to replicate a broad class of image sensors. To this end, a simulation in GEANT4 was developed, incorporating the basic elements of pixel design from Chapter 2 and the experimental designs of Chapters 5 and 6. The details of the simulation and the fit to experimental data are first presented, followed by the consequences of recent trends in sensor design.

8.1 Building the simulation

Five fundamental processes connect the emission of a particle to a digitized sensor response. These are, in order:

1. Propagation through shielding
2. Energy deposition in the sensitive silicon region
3. Conversion of this energy to numbers of electron-hole pairs
4. Diffusion of electrons to collection regions, with some recombination
5. Digitization of the resulting photocurrents.

This last step has already been characterized to first-order by the gains found in Section 5.2.1; the remaining four are the key elements of this simulation.

8.1.1 Particle propagation

In brief, the simulation consists of particles traversing through layers of shielding toward an array of identical pixels with a surface area comparable to the physical sensors. The two experiments required different approaches to particle generation; see

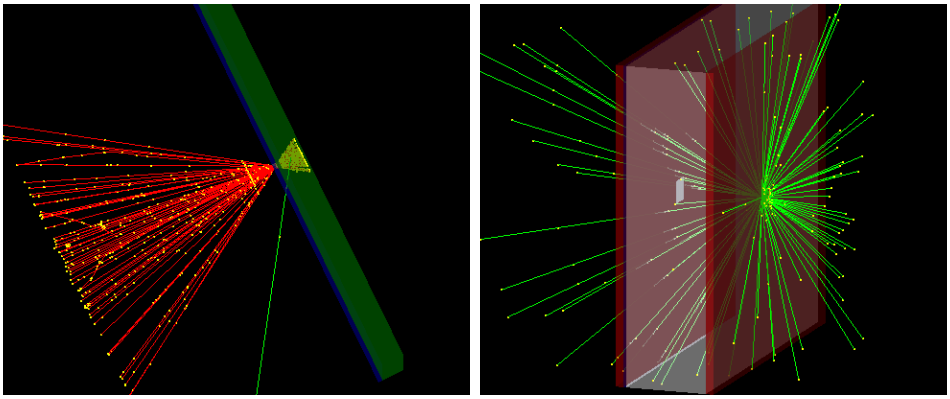

Figure 8.1: An illustration of the geometries of the muon hodoscope simulation (left) and the radioactive source simulation (right). The silicon sensor has been enlarged for illustration at right, while its position can be deduced from the convergence of muon tracks at left. Glass is shaded in light blue, plastic in green, wax in white, and aluminum in red.

Figure 8.1.

To model the hodoscope experiment, $\mu^{\pm}$ primaries are constructed on a plane beyond layers of glass and plastic shielding with trajectories targeting a randomized point on the central pixel of the array. These muons are supplied with the same $\cos^3\theta \sin\theta$ zenith angle distribution discussed in Section 5.3.2 and energies from the PARMA/EXPACS model [150] with the proper latitude and altitude. An energy cutoff at 100 MeV—the approximate low edge of the minimum ionizing region shown in Figure 8.2—is imposed to account for the lead shielding, and the angular distribution is likewise cut at 40°, approximating the solid angle of the hodoscope. To incorporate electrons scattered from the LYSO scintillators and 3D-printed modules, 1 cm of PLA plastic was added as a zeroth-order approximation; reassuringly, this produced no discernible effect on the resulting efficiency.

By contrast, the radioactive source tests generate particles from an isotropic point source behind layers of wax and lead shielding, mirroring each of the lower-bound measurements in Chapter 6. To simulate backscattering from the casing of the sources and from the body of the phone, additional thin layers of aluminum are positioned on both ends of the simulation geometry, and the full 4π solid angle is simulated. Discrete photon spectra for each source are taken from Ref. [141], using both X-ray and gamma lines.

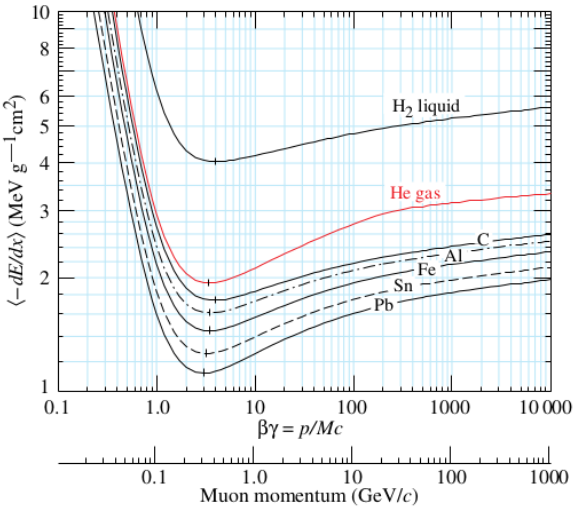

Figure 8.2: $-\frac{dE}{dx}$ for muons in various substances in the minimum ionizing region, taken from Ref. [151].

To propagate these secondaries, the standard Geant4 physics packages were used except for the electromagnetic interactions, which required a more careful approach. To accurately model the propagation of low-energy secondaries, the Livermore physics models [152] are used below 100 MeV; likewise, GEANT4's silicon-specific electromagnetic models [152, 153] are employed within the sensor volume in their verified energy domains—down to 16.7 eV for electrons. To adequately track particles on the scale of individual pixels, secondaries with ranges above 100 nm were propagated by the simulation.

8.1.2 Pixel construction and diffusion

Each individual pixel is modelled as a rectangular block of silicon with two regions: a *depletion region* and the *diffusion region* comprising the remainder of the pixel. In the diffusion region, excited electrons are free to traverse into neighboring pixels, whereas it is assumed that the fields in the depletion region keep conduction-band electrons contained. Reflecting the near-universal BSI design (Section 2.2), the diffusion region is adjacent to the camera lens: downstream for the hodoscope and upstream for the radioactive sources.

When a particle deposits energy in a pixel, this is first converted into a number of

electron-hole pairs. As part of this energy excites vibrational modes of the silicon, the effective ionization energy of $3.7\,\mathrm{eV}/\mathrm{e}^-$ [154, 155] is used to compute electron counts rather than the band gap of 1.12 eV. The resulting electron count is then rounded to an integer value, and the remainder contributes to the energy deposited in the ionizing particle's next step. If the energy was deposited in the depletion region, the electrons are automatically associated with that pixel; in the diffusion region, however, the electrons are randomly assigned positions along the particle step, and are propagated individually to the depletion regions of nearby pixels.

This diffusion algorithm is illustrated in Figure 8.3. Each electron follows a three-dimensional random walk until reaching the depletion region of one of the pixels. Those passing either the simulated DTI structures or the far edge of the diffusion region are instead reflected across the boundary. In principle, recombination occurs with a fixed probability after each step; however, as the diffusion length of epitaxial silicon is on the order of centimeters [156], an essentially infinite 100 µm is used. In a real sensor, most recombination is caused by Shockley-Read-Hall traps in the silicon lattice [156] which cannot be modelled with a simple recombination probability per step. A more sophisticated treatment of recombination is left for future refinements.

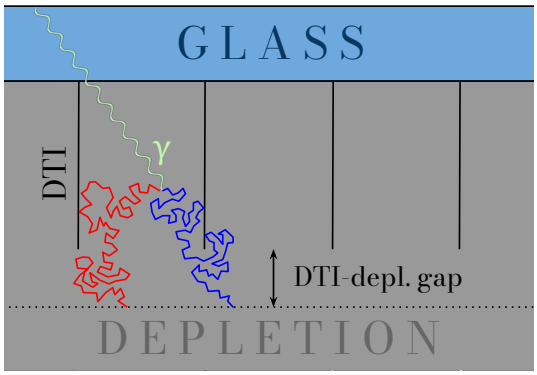

Figure 8.3: An illustration of the random walk and pixel geometry used for electron diffusion. Two electrons created in the same simulation step are shown arriving at the depletion regions of two different pixels. Inter-pixel diffusion is limited to the gap between the DTI and depletion zone, also labelled.

8.2 Experimental agreement

This model was compared to the sensor muon response (Chapter 5) and photon response (Chapter 6) through two features: the detection efficiency and the diffusion of electrons in the conduction band to nearby pixels. In both cases, simulated data was generated with the S7 pixel dimensions found in Figure 5.4: a 2.9 µm pixel depth, 1.4 µm pixel pitch, and DTI extending 1.9 µm into the pixel from the lens-side (i.e., B-DTI). The relative thicknesses of the diffusion and depletion regions, determining the size of the DTI-depletion gap where electrons can diffuse across pixel boundaries (see Figure 8.3), is fit to the data.

8.2.1 Inter-pixel diffusion

To quantify inter-pixel diffusion and to fit the DTI-depletion gap length, the maximum pixel response in each frame is compared to number of electrons collected in the surrounding 4-, 8-, and 20-pixel regions; see Figure 8.4 below.

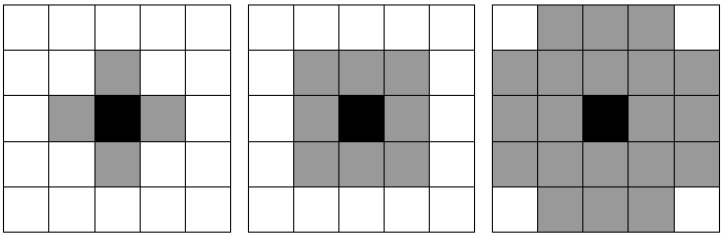

Figure 8.4: The three regions (gray) around the maximum pixel of each frame (black) used for quantifying pixel diffusion.

Figure 8.5 plots the ratio of electrons collected in each of these three regions $N_{\{4,8,20\}}$ to those collected by the central (brightest) pixel $N_{\max}$. Histograms of these ratios are drawn from the S7 ^{137}Cs data and the corresponding best-fit simulation data, though the agreement is similar for ^{60}Co and ^{226}Ra, and for the other phones. Muons are not considered here due to the low signal-to-noise ratio in the hodoscope data.

The data in Figure 8.5 include adjustments for noise and saturation inherent to physical devices. To minimize these effects, a cut of $150 \leq N_{\max} < 900$ was applied to both experimental and simulated datasets: the low limit increased the signal-to-noise

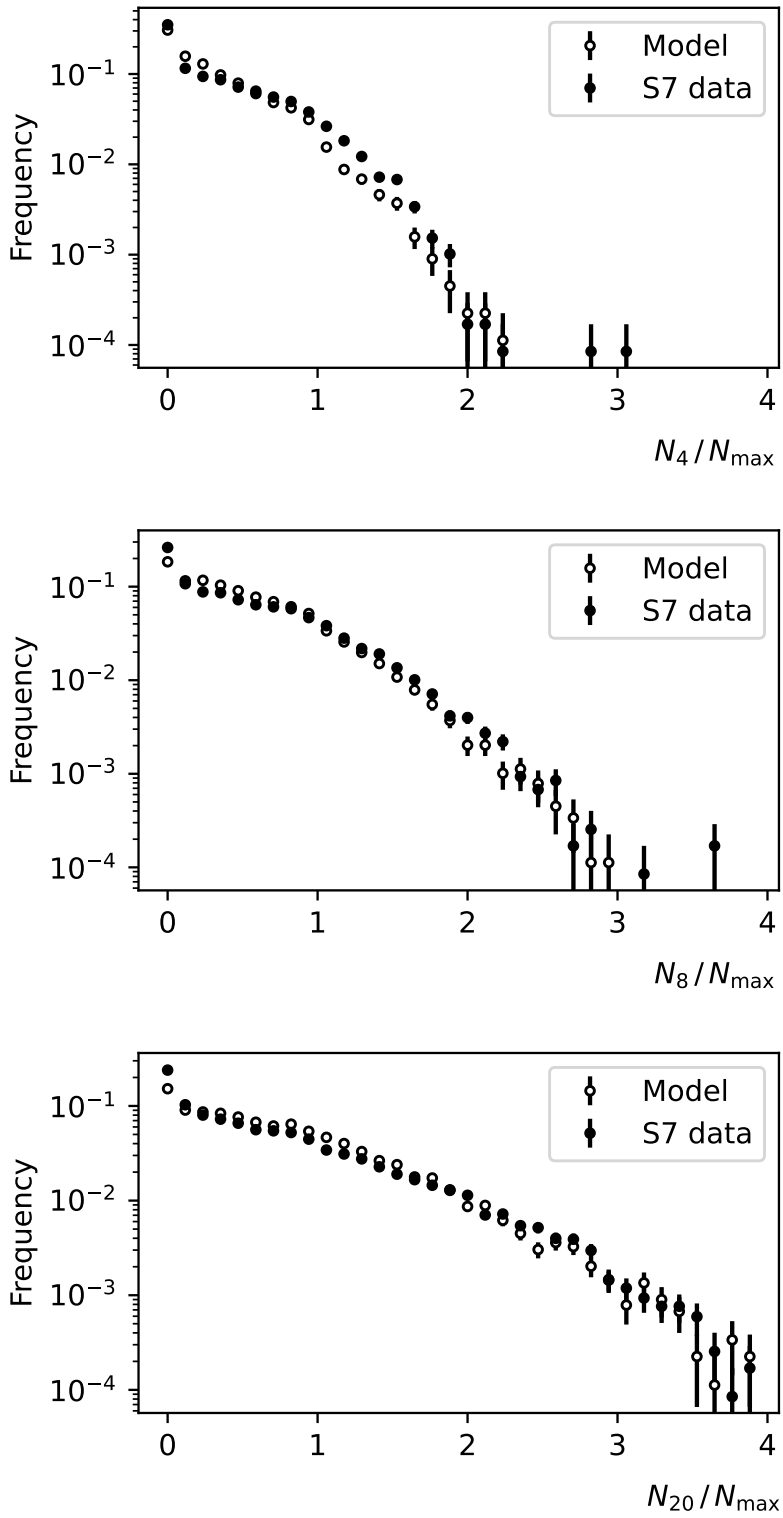

Figure 8.5: Normalized histograms of the ratio $N_i/N_{\max}$ for $i = 4$ (top), $i = 8$ (center), and $i = 20$ (bottom) for the simulated and experimental ^{137}Cs data.

ratio in the experimental data, while the high limit removed events near saturation.[1] Any remnants of noise in the experimental data were then quantified through a sample of dark frames with no radioactive exposure. Dark currents were added to simulated pixels from this sample of dark frames; likewise, rates of background events such as cosmic muons have been removed from the experimental distributions in Figure 8.5.

The simulated data in Figure 8.5 also incorporate a best fit of the DTI-depletion gap, which almost exclusively determines the extent of inter-pixel diffusion. With a fixed 2.9 µm pixel thickness and 1.9 µm-long DTI, this gap is constrained below 1.0 µm, the limiting case where the depletion region is infinitesimally thin. Several DTI-depletion gap lengths were simulated by varying the relative size of the depletion and diffusion regions, and as Figures 8.6 and 8.7 show, the data strongly prefer that this gap be close to nonexistent (< 20 nm). Though the true depletion depth is proprietary, publicly available sensor specifications such as Ref. [157] suggest that such a result—that the depletion region extends past the DTI—is to be expected. Figure 8.5 was accordingly generated with a nonexistent gap, eliminating all inter-pixel diffusion for conduction-band electrons. The spread in these plots is due exclusively to *free* electrons traversing pixel boundaries, such as those Compton-scattered by the photon primaries.

As Figure 8.5 shows, the model captures the trends of the experimental data, with the notable exception of the leftmost bins in each plot. Instead, excesses of solitary bright pixels are found, increasing with the number of adjacent pixels included in the ratio. As this part of the distribution is sensitive to dark currents on the order of single electrons, small black-level offsets, and hot-pixel masks, the observed divergence is unsurprising.

8.2.2 Sensor efficiency

Having fit the depletion region thickness in the previous section, the detection efficiency is likewise modelled. As in Section 6.3, naive efficiencies are computed by normalizing $A_\gamma \epsilon_\gamma$ by the sensor area. As shown in Figure 8.8, the ^{60}Co and ^{226}Ra efficiencies were

[1] Due to lens-shading effects, small numbers of pixels on the edges of the sensor nonetheless reached saturation within this range and were removed from the analysis.

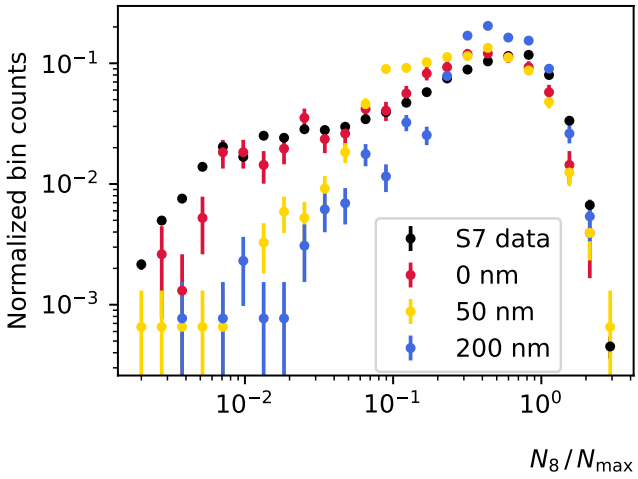

Figure 8.6: Agreement with experimental ^{226}Ra data in simulations with three different depletion-DTI gap lengths; to better resolve the leftmost edge, logarithmic bins are used unlike in Figure 8.3. The middle 8-pixel region from Figure 8.4 is used for illustration.

accurately reproduced by the model, while the model ^{137}Cs curve was lower than the experimental data by roughly 30%. Significant disagreement is likewise found in the muon efficiency, though just as in the photon curves, the simulation becomes more accurate at higher thresholds. As the agreement found in ^{60}Co and ^{226}Ra was only achieved through the silicon-specific models in GEANT4—currently applicable only to electrons, protons, and heavy ions—a more precise model of muon propagation in silicon may be necessary to properly simulate the muon response.

As the source efficiencies—especially for ^{226}Ra—are pulled downward by X-rays and low-energy gamma rays which fail to penetrate the wax or lead shielding, Figure 8.8 can be a misleading characterization of the absolute sensor efficiency. As an alternate vantage point, Figure 8.9 plots the simulated efficiency as a function of photon energy, from which the spectral components of these sources can be untangled. This may yield some insight into the poorer ^{137}Cs agreement. In both the ^{60}Co and lead-shielded ^{226}Ra, the sensor response is dominated by MeV-range photons, which are near an efficiency plateau; meanwhile the 662 keV line from ^{137}Cs is located at a steeper part of the curve, and therefore may be more sensitive to inaccuracies in the scattering models used. Alternately, a systematic error in the ^{137}Cs activity may explain the observed

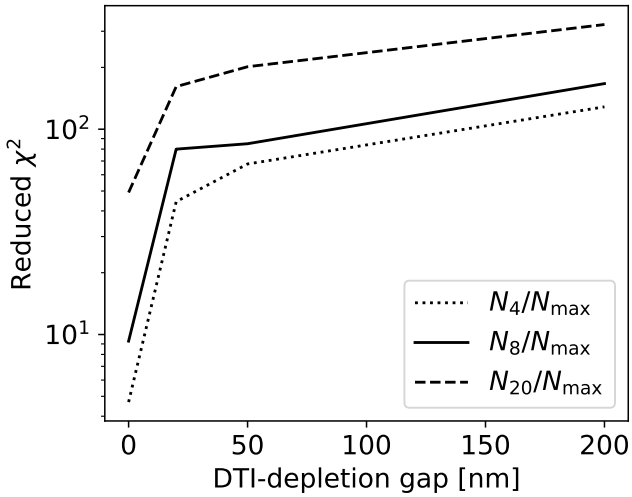

Figure 8.7: Reduced χ^2 statistic of fits in Figure 8.6, including all three regions in Figure 8.4 and an additional gap length of 20 nm.

discrepancy.

8.3 Simulated CRAYFIS sensitivity

While this agreement is worthwhile in its own right, a reasonable pixel model can also provide a benchmark for CRAYFIS's shower reconstruction efficiency. This requires two forms of extrapolation: modified particle spectra for UHECR showers (including the addition of electrons, which have so far been indirectly measured through Compton scattering), and modified sensors, incorporating recent design trends.

8.3.1 Extrapolated particle efficiencies

The three phone models used in Chapters 5 and 6 were typical of their generation (2015-2016), and their similar design naturally yielded similar results. In recent years, however, three diverging trends in sensor design have emerged. In many new models, the pixel pitch has continued to decrease (most recently, down to 0.7 µm [158]) as noted by Figure 2.3, while in others, the greater optical noise of smaller pixels [159] has prompted a return the same 12 MP, 1.4 µm dual-photodiode design of the Galaxy S7 [160, 161]. Meanwhile, a branch of both Sony and Samsung sensors have featured

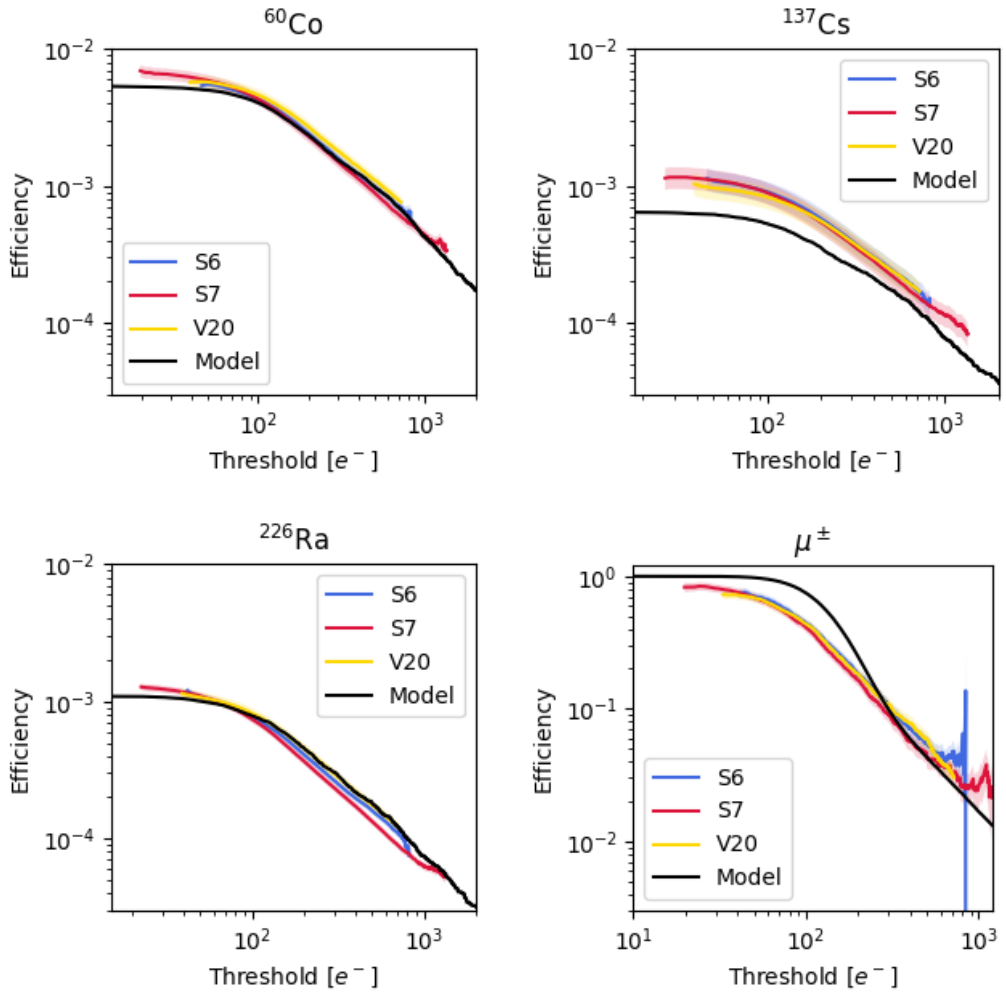

Figure 8.8: Comparison of experimental data and simulated efficiencies for ^{60}Co (top left), ^{137}Cs (top right), ^{226}Ra, (bottom left), and muons (bottom right, with a modified scale).

drastically larger surface areas: as one very recent example, the IMX700 on the Huawei P40 Series [162] measures 74 mm^2 with a 1.22 µm pixel pitch, roughly triple the size of those studied in the last few chapters. Samsung has even attempted to consolidate the first and third trends, with the ISOCELL HM3 offering 108 MP on a 69 mm^2 sensor with a 0.8 µm pixel pitch [163]. It is particularly the sub-micron pixels in which Figure 2.3 had noted an increasing pixel depth; the pixel depths of the other two classes of sensors are not yet publicly available. Yet given the long history of pixel depths between 2.5–3 µm, it is reasonable to expect this trend will continue.

While the efficiency of the 12 MP, dual-photodiode sensors is likely similar to that of the Galaxy S7, the implications of these other two trends—denoted here as shrinking-

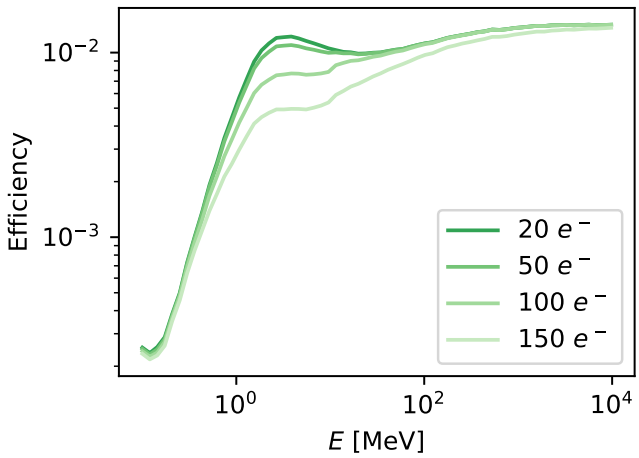

Figure 8.9: Simulated efficiency as a function of photon energy for four thresholds. Photons were generated at normal incidence to the sensor with shielding from only the 2 mm lens.

pixel (SP) and growing-sensor (GS)—can be studied with this simulation. Figure 8.10 shows the predicted effective areas for muons, gammas, and electrons for example SP, GS, and combination SP/GS sensors, juxtaposed to the S7-like "classic design" (CD). The energy distributions are drawn from a CORSIKA simulation of 10^{19} eV UHECR showers, and 5 cm of dense wood is included as shielding: roughly the equivalent overburden of a roof with asphalt shingles and a second-story hardwood floor. Particles are generated uniformly along a large plane behind this shielding layer. The same zenith angle distribution is used as in Figure 8.1a; however, the azimuthal angle and vertex position are both independently sampled from uniform distributions, and particle trajectories are no longer constrained to target the sensor. The trigger levels in electrons—a function of the thermal and readout noise—has been assumed to remain the same.[2]

The SP curve diverges from the CD by no more than 15%, with the greatest losses observed in thresholds slightly above 100 e^-; however, as the model diverges from experimental data at the lowest thresholds, it is uncertain whether the SP design performs even more poorly as thresholds continue to decrease. Importantly though, these efficiencies are predicated on triggering bright individual pixel values. In this metric,

[2]A greater number of pixels requires stricter thresholds to keep a constant trigger rate, but also allows hot-pixel masking in finer regions of the sensor. It is not obvious which effect will dominate.

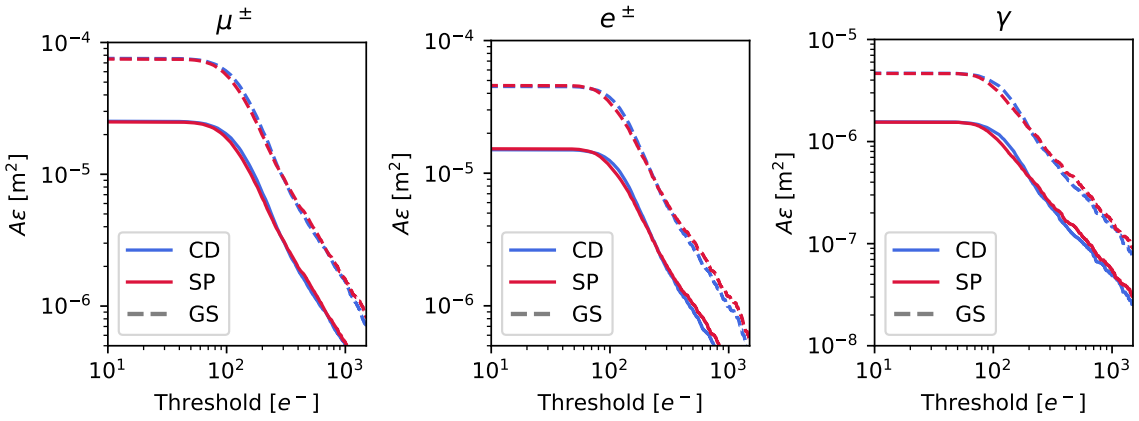

Figure 8.10: Plots of $A\epsilon$ for muons (left), electrons and positrons (middle), and gammas (right, with modified scale) for various sensor trends. The CD sensor is identical to the IMX260 on the Galaxy S7, while the SP sensor has a modified pixel pitch of 0.8 µm and a pixel depth of 3.8 µm. The GS trend simply scales the areas by a factor of 3.

the thicker substrate in an SP sensor does not compensate for the more distributed tracks produced by a thinner pixel pitch. With an appropriate convolutional filter incorporated into the trigger, however, the SP trend may yet yield higher efficiencies than the CD standard.

One last trend is worthy of mention here: the inclusion of multiple rear cameras per device. By this design, a user can choose from a wide selection of sensors to optimize each photograph given the subject (e.g. landscape vs. portrait) and light levels. The effect is similar for Crayfis: this provides flexibility to choose between several specialized sensors in order to optimize performance. In particular, if the GS trend is particularly advantageous for low-light photography and becomes common in phones with two or three rear cameras, this may vastly improve Crayfis's shower reconstruction capabilities.

8.3.2 Shower reconstruction efficiency

Finally, these efficiencies are used to quantify Crayfis's power to resolve UHECR showers. The lateral densities of muons, photons, and electrons at sea level are fit to CORSIKA simulations using the NKG model [164], as in Ref. [86]. With "optimistic" and "pessimistic" per-phone detection efficiencies drawn from Figure 8.11 (see

Table 8.1: Benchmark $A\epsilon$ values used to evaluate CRAYFIS's sensitivity. "Optimistic" and "pessimistic" values are drawn from Figure 8.11, incorporating the simulation's overestimate of the muon efficiency and underestimate of the photon efficiency. These are compared to the original estimates from Ref. [83].

Selected $A\epsilon$ benchmarks				
Benchmark	$A_\mu\epsilon_\mu$ [mm^2]	$A_\gamma\epsilon_\gamma$ [mm^2]	$A_e\epsilon_e$ [mm^2]	Duty cycle
Optimistic	30	2	24	0.33
Pessimistic	12	1	8	0.33
Upper, original	50	.01	0	1
Lower, original	10	.001	0	1

Table 8.1), these particle densities are converted to trigger probabilities as a function of distance from the shower core.

Figure 8.11 shows the resulting distribution of triggers per shower. The optimistic efficiencies are plotted in purple, and the pessimistic in yellow, for primary energies of 10^{18}, 10^{19}, 10^{20}, and 10^{21} eV, shaded from dark to light. For each of these energies, showers from ^{4}He, ^{16}O, and ^{56}Fe nuclei were simulated, and the width of the bands correspond to the span of this mass range. For reference, the gray band marks 1% of the population density of Los Angeles, with the right edge corresponding to the urban core, and the left end to the outlying areas. In Figure 8.11b, a threshold of 5 triggers per shower is used for direct comparison with Ref. [83]. Though a combinatorial background (and requisite threshold) can be calculated from the noise and frame rates from Chapters 5 and 6, the low timing resolution of RAW streaming is likely suboptimal for a real CRAYFIS observatory. Balancing frame rate and resolution is left for future work.

These results show more optimistic detection efficiencies than even the upper bound in Ref. [83]. This is largely due to the much higher photon efficiency, attained partly by improvements to the app, and partly by extrapolating to higher photon energies and including scattering from the overburden. When combined with the addition of electrons, the resulting detection efficiencies are no longer primarily determined by the muonic component—instead, photon and electron detections outnumber those of muons by an order of magnitude, as shown in Figure 8.12.

Figure 8.11 assumes, however, that data-taking phones are uniformly spread through-

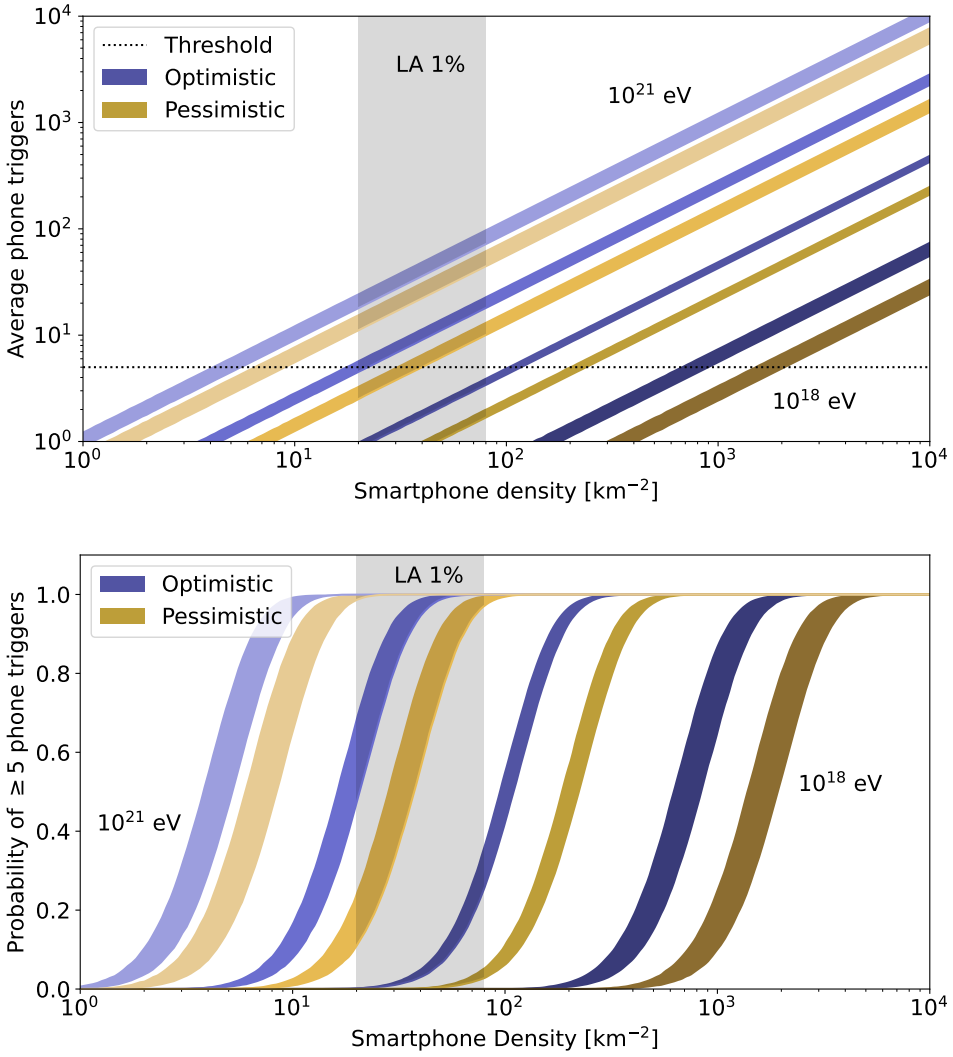

Figure 8.11: The average number of phones triggered per shower (top) and the probability of at least five phones being triggered (bottom) as a function of user density. Description in text.

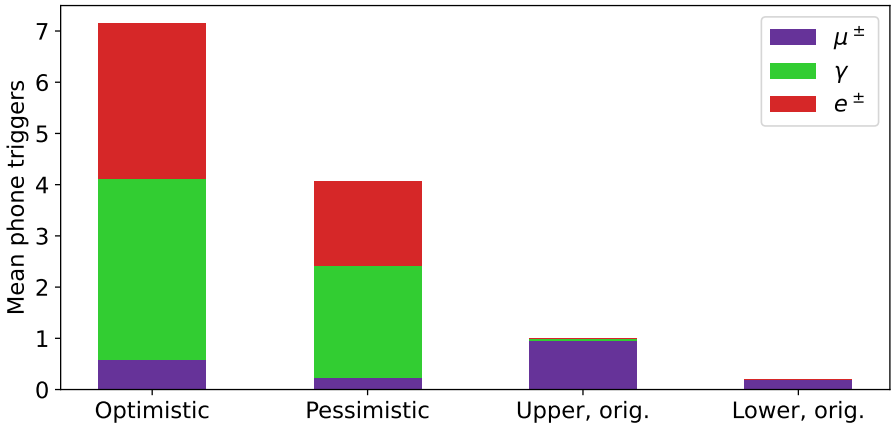

Figure 8.12: Cumulative (stacked) abundances of muon, photon, and electron triggers per shower under the four $A\epsilon$ benchmarks from Table 8.1, excluding differences in duty cycle. Values are generated from 10^{20} eV He primaries with a user density of 30 km^{-2}.

out the area of each shower. In a real Crayfis observatory, phones would instead be clustered; in an urban core, for example, public and commercial spaces would be devoid of detectors at night, while multi-story residential buildings would contain a much higher density of phones than the city as a whole. Even in lower-density single-family homes, it is likely that if one device is running the Crayfis app, other members of the household would install it as well. Though the average number of triggered phones per shower would remain unchanged by this fine structure in the detector distribution, the curves in Figure 8.11b would exhibit gentler slopes, and 10^{18} eV showers would have a more reasonable probability of being detected even with at the benchmark Los Angeles density shown. Ref. [83] demonstrates these general trends of device clustering, though a more precise modelling of these effects will be left for future refinements.

Lastly, the total effective area of the Crayfis observatory was computed from these device efficiency estimates through a Monte Carlo simulation. For each primary energy and mass, 10^9 showers were generated across the earth's surface with a uniform spatial distribution. A map of global population from 2020 [165] with a resolution of 1 arc-minute was combined with the lateral shower densities to compute each shower's detection efficiency, assuming a 100% global participation rate. These detection efficiencies were adjusted for lower levels of user adoption using binomial statistics, and the resulting effective areas were multiplied by the duty cycles listed in Table 8.1.

Even assuming such a reduced duty cycle, the efficiency estimates presented in this chapter exceed those of Ref. [83] by a substantial margin, as Figures 8.13 and 8.14 show. With the efficiencies in the pessimistic scenario, the benchmark area of Auger can be matched with roughly a tenth of the global number of users quoted in Ref. [83] for 10^{20} eV UHECRs.[3] This adjusted goal—5.8M nightly users—can likely be further reduced as suggested in Ref. [83] with a targeted campaign in large metropolitan areas. With even modest levels of adoption, this novel approach to UHECR detection may vastly improve the global sensitivity to super-GZK events.

[3]Part of this difference is due to discrepancies in the lateral density models, Monte Carlo approach, and current global population.

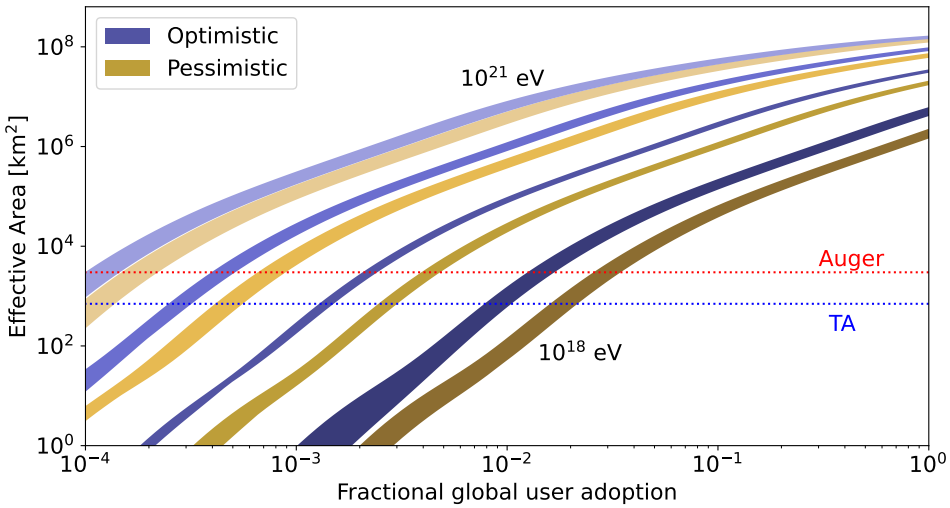

Figure 8.13: The effective area of the CRAYFIS experiment as a function of global user adoption and UHECR energy. The colored bands represent the same parameters as their counterparts in Figure 8.11.

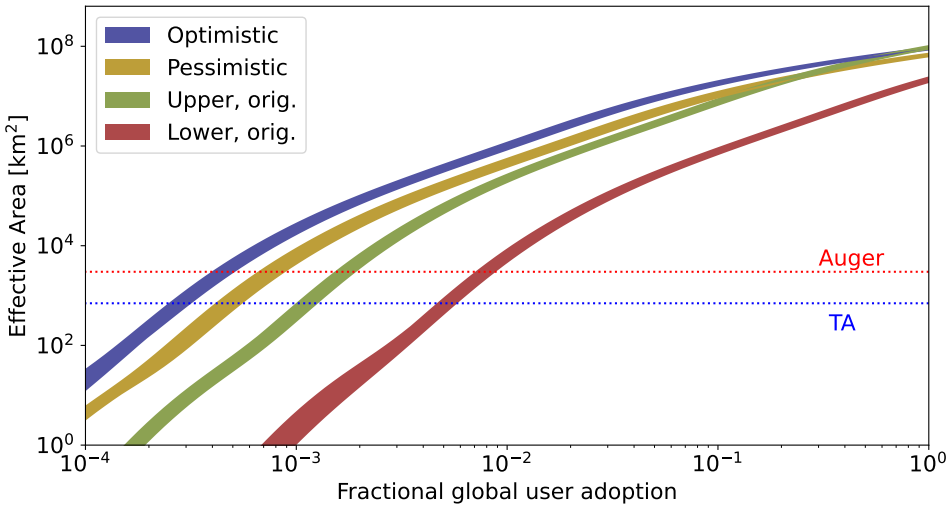

Figure 8.14: A comparison of effective areas for all four benchmark efficiencies in Table 8.1 with 10^{20} eV primaries, incorporating differences in duty cycles.

Epilogue

The Future of CRAYFIS

In the prologue, it was stated that the gap between theory and practice for Crayfis was particularly wide, and this thesis has made significant strides to bridge this gap. Through the improvements to the app discussed in Part II, the Crayfis trigger attained substantially-improved trigger efficiencies for photons and a high muon efficiency as well. Then, through careful modelling, the sensitivity of a Crayfis network to cosmic ray secondaries was found to greatly exceed the original estimates from Ref. [83]. However, several obstacles remain before these results can be considered representative of a real Crayfis observatory.

The efficiencies measured in Part II were attained by lowering the thresholds to allow trigger rates well above initial projections. This in turn leads to higher combinatorial noise than what Ref. [83] considered, upon which assumption the detection threshold of five phones per shower in Figure 8.11 was predicated. Though Part III has shown that moderately higher thresholds are not fatal to the efficiency, balancing the combinatorial noise with detection efficiency is one of the primary remaining challenges for the Crayfis experiment.

There are three primary paths to achieve such a balance. First, the online trigger noise can be reduced, e.g. through lower trigger rates or higher fractions of masked pixels. Both of these measures would marginally decrease the trigger efficiency; hence, further studies are needed to optimize these trigger parameters. Second, the timing

resolution of triggers could be improved by increasing the frame rate while keeping the trigger rate constant. The frame rates used in the FishStand app in Part III were held at 2 and 7 FPS, though the production Crayfis app can attain 30 FPS at modest temperatures with YUV buffers downsampled to 1080p. It is yet unknown to what degree that more modest downsampling would degrade the efficiency, and similar efficiency studies with YUV data are essential to finding an optimal frame rate and resolution. Lastly, noise can be reduced through an offline classifier, such as was explored in Ref. [13].

With these last steps extending the efficiencies of Part III to the context of shower-detection, Crayfis will at last be ready to reach a global scale. As suggested by Figure 8.11, the density of active users is of vital importance in making the Crayfis experiment viable. A targeted marketing campaign to densely-populated communities—urban cores and large college campuses, for example—would grow the experiment sufficiently to begin cataloguing UHECR showers. At that stage, Crayfis's future will depend on enthusiastic base of users, scientists, and even volunteer developers even more so than optimized triggers, but for now, this work constitutes a large step forward towards a fully-realized Crayfis observatory.

Bibliography

[1] C. Moskowitz, Track Cosmic Rays with Smartphone App. Scientific American Podcast: 60-Second Space. (2014). https://www.scientificamerican.com/podcast/episode/track-cosmic-rays-with-smartphone-app/

[2] J. Timmer, Cosmic ray particle shower? There's an app for that. Ars Technica. (2014). https://arstechnica.com/science/2014/10/cosmic-ray-particle-shower-theres-an-app-for-that/

[3] F. Diep, This App Turns Your Phone Into A Cosmic Ray Detector. Popular Science. (2014). https://www.popsci.com/article/gadgets/app-turns-your-phone-cosmic-ray-detector/

[4] K. Jepsen, Reading the heavens with your phone. Symmetry. (2014) https://www.symmetrymagazine.org/article/reading-the-heavens-with-your-phone

[5] J. Fang, Turn Your Phone Into A Cosmic Ray Detector. IFLScience. (2014). https://www.iflscience.com/space/your-phone-already-cosmic-ray-detector/

[6] M. Lemonick. Seeking the Source of Cosmic Rays. The New Yorker. (2014). http://www.newyorker.com/tech/elements/seeking-source-cosmic-rays

[7] B. Sullivan et al., The eBird enterprise: An integrated approach to development and application of citizen science. Biological Conservation 169:31-40 (2014)

[8] M. Fienen and C. Lowry, Social.Water—A crowdsourcing tool for environmental data acquisition. Computers and Geosciences 49:164-169 (2012)

[9] https://www.zooniverse.org/

[10] https://pressurenet.io

[11] J. Vandenbroucke et al., Measurement of cosmic-ray muons with the Distributed Electronic Cosmic-ray Observatory, a network of smartphones. JINST 11, P04019 (2016). arXiv:1511.00660

[12] M. Meehan et al., The particle detector in your pocket: The Distributed Electronic Cosmic-ray Observatory. PoS (ICRC 2017), 375 (2017), arXiv:1708.01281 [astro-ph.IM]

[13] M. Winter et al., Particle identification in camera image sensors using computer vision. Astropart. Phys., 104, 42-53 (2019). arXiv:1803.04493 [astro-ph.IM]

[14] https://github.com/dkirkby/cosmic

[15] P. Homola et al., (CREDO), Cosmic Ray Extremely Distributed Observatory. Symmetry, 12(11), 1835 (2020). arXiv:2010.08351 [astro-ph.HE]

[16] B. R. Dawson, M. Fukushima, and P. Sokolsky, Past, present, and future of UHECR observations. Prog. Theor. Exp. Phys. 12 A101 (2017)

[17] K. H. Kampert and A. A. Watson, Cosmic rays, gamma rays and neutrinos: a survey of 100 years of research. Eur. Phys. J. H, 37, 359–412 (2012), arXiv:1207.4827.

[18] Thoudam et al., Cosmic-ray energy spectrum and composition up to the ankle: the case for a second Galactic component. Astron & Astrophys. 595, A33 (2016). arXiv:1605.03111 [astro-ph.HE]

[19] Kawamura, et al., Observation of Heavy Cosmic Ray Primaries by Means of Balloon - Borne New Type of Emulsion Chamber. Proc. 23rd ICRC 3, 85 (1990)

[20] O. Adriani et al. (PAMELA), PAMELA Measurements of Cosmic-Ray Proton and Helium Spectra. Science 332, 69 (2011)

[21] M. Aguilar et al. (AMS), Precision Measurement of the Proton Flux in Primary Cosmic Rays from Rigidity 1 GV to 1.8 TV with the Alpha Magnetic Spectrometer on the International Space Station. Phys. Rev. Lett. 114, 171103 (2015)

[22] M. Aguilar et al. (AMS), Precision Measurement of the Helium Flux in Primary Cosmic Rays of Rigidities 1.9 GV to 3 TV with the Alpha Magnetic Spectrometer on the International Space Station. Phys. Rev. Lett. 115, 21, 211101 (2015)

[23] K. Abe et al., Measurements of Cosmic-Ray Proton and Helium Spectra from the BESS-Polar Long-duration Balloon Flights over Antarctica. Astrophys. J. 822, 2, 65 (2016)

[24] M. J. Christ et al., Cosmic-ray proton and helium spectra: results from the JACEE experiment. Astrophys. J. 502, 278 (1998)

[25] A.D. Panov et al. (ATIC Collab.), Cosmic-ray proton and helium spectra: results from the JACEE experiment. Bull. Russian Acad. of Science, Physics, 73, 564 (2009).

[26] V. A. Derbina et al. (RUNJOB), Cosmic-ray spectra and composition in the energy range of 10-1000 TeV per particle obtained by the RUNJOB experiment. Astrophys. J. 628, L41 (2005).

[27] H. S. Ahn et al., Energy spectra of cosmic-ray nuclei at high energies. Astrophys. J. 707, 593 (2009).

[28] J.J. Engelmann et al. (HEAO3-C2 Collab.), Charge composition and energy spectra of cosmic-ray nuclei for elements from Be to NI-Results from HEAO-3-C2. Astron. & Astrophys. 233, 96 (1990).

[29] D. Müller et al. (CRN Collab.), Energy spectra and composition of primary

cosmic rays. Astrophys. J. 374, 356 (1991).

[30] M. Ave et al., Composition of primary cosmic-ray nuclei at high energies. Astrophys. J. 678, 262 (2008).

[31] F. Aharonian et al. (H.E.S.S.), First ground-based measurement of atmospheric Cherenkov light from cosmic rays. Phys. Rev. D75, 042004 (2007).

[32] A. Archer et al. (VERITAS), Measurement of the iron spectrum in cosmic rays by VERITAS. Phys. Rev. D98, 2, 022009 (2018).

[33] D. Heck et al., CORSIKA: a Monte Carlo code to simulate extensive air showers., (Forschungszentrum Karlsruhe Report FZKA 6019 (1998)

[34] T. Bergmann, R. Engel, D. Heck, et al., One-dimensional Hybrid Approach to Extensive Air Shower Simulation. Astropart. Phys., 26, 420–432 (2007), arXiv:0606564.

[35] F. Cohen and K. Kasahara, A Fast and Accurate Monte Carlo EAS Simulation Scheme in the GZK Energy Region and Some Results for the TA experiment. Proc. 30th ICRC 4, 585-588 (2008).

[36] R. Engel, D. Heck, and T. Pierog. Extensive Air Showers and Hadronic Interactions at High Energy. Annu. Rev. Nucl. Part. Science, 61, 467–489 (2011).

[37] D. d'Enterria, R. Engel, T. Pierog, S. Ostapchenko, and K. Werner, Constraints from the first LHC data on hadronic event generators for ultra-high energy cosmic-ray physics. Astropart. Phys., 35, 98–113 (2011). arXiv:1101.5596.

[38] V. B. Petkov, The knee of EAS size spectrum, missing component and very high energy muons. Int. J. Mod. Phys. A20, 6846–6848 (2005).

[39] A. A. Petrukhin, The knee and unusual events at PeV energies. Nucl. Phys. Proc. Suppl. 151, 57–60 (2006).

[40] R. Barcel´o, M. Masip, and I. Mastromatteo, Cosmic ray knee and new physics at the TeV scale. JCAP 0906, 027 (2009). arXiv:0903.5247 [hep-ph].

[41] A. Dixit, P. Jain, D. W. McKay, and P. Mukherjee, New physics, the cosmic ray spectrum knee, and *pp* cross section measurements. Eur. Phys. J. C 68, 573-580 (2010), arXiv:0912.0980 [astro-ph.CO].

[42] A. Ariga (Faser Collab.), Studying neutrinos at the LHC: FASER and its impact to the cosmic-ray physics. PoS (ICRC2021), 1025 (2021).

[43] P. Abreu, M. Aglietta, E. J. Ahn, et al., Measurement of the proton-air cross-section at $\sqrt{s} = 57$ TeV with the Pierre Auger Observatory. Phys. Rev. Lett., 109(6), 062002 (2012), arXiv:1208.1520.

[44] R. U. Abbasi, M. Abe, T. Abu-Zayyad, et al., Measurement of the Proton-Air Cross Section with Telescope Array's Middle Drum Detector and Surface Array in Hybrid Mode. Phys. Rev. D92, 3, 032007 (2015), arXiv:1505.01860

[45] A. Aab et al., The Pierre Auger Cosmic Ray Observatory. Nucl. Instrum. Methods Phys. Res., Sect. A, 798, 172–213 (2015).

[46] C. Bonifazi (Auger Collab.), The angular resolution of the Pierre Auger Observatory. Nucl. Phys. B Proc. Suppl., 190, 20–25 (2009), arXiv:0901.3138.

[47] T. Abu-Zayyad, R. Aida, M. Allen, et al., The surface detector array of the Telescope Array experiment. Nucl. Instrum. Methods Phys. Res., Sect. A, 689, 87–97 (2012), arXiv:1201.4964.

[48] H. Tokuno, Y. Tameda, M. Takeda, et al., New air fluorescence detectors employed in the Telescope Array experiment. Nucl. Instrum. Methods Phys. Res., Sect. A, 676, 54–65 (2012), arXiv:1201.0002.

[49] M. Abou Bakr Othman et al., Radar Detection of UHECR Air Showers at the Telescope Array. PoS (ICRC 2011), 1314 (2011)

[50] J. Bellido et al., Fluorescence and Hybrid Detection Aperture of the Pierre Auger Observatory. Proc. 29th ICRC 8, 113 (2005), arXiv:astro-ph/0507103

[51] Y. Tsunesada et al., Telescope Array Aperture: Mono, Stereo, and Hybrid. Proc. 30th ICRC 4, 409-412 (2008).

[52] E. Kido et al. (Telescope Array Collab.), Status and prospects of the TAx4 experiment. EPJ Web Conf. (proceedings of UHECR 2018), 210, 06001 (2019)

[53] M. Tanabashi et al. (Particle Data Group), Review of Particle Physics. Phys. Rev. D98, 030001 (2018)

[54] R. Alfaro et al. (HAWC), All-particle cosmic ray energy spectrum measured by the HAWC experiment from 10 to 500 TeV. Phys. Rev. D96, 12, 122001 (2017).

[55] Y. A. Fomin et al., The Energy Spectrum and the Mass Composition of the Primary Cosmic Rays Around the Knee. Proc. 22nd ICRC 2, 85 (1991).

[56] M. Amenomori et al., The all-particle spectrum of primary cosmic rays in the wide energy range from 1014 to 1017 eV observed with the Tibet-III air-shower array. Astrophys. J. 268, 1165 (2008).

[57] M. Nagano et al., Energy spectrum of primary cosmic rays between $10^{14.5}$ and 10^{18} eV. J. Phys. G10, 1295 (1984).

[58] M. A. K. Glasmacher et al., The cosmic ray energy spectrum between 10^{14} and 10^{16} eV. Astropart. Phys. 10, 291 (1999).

[59] T. Antoni et al. (KASCADE), KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems. Astropart. Phys. 24, 1 (2005).

[60] W. D. Apel et al. (KASCADE Grande), Kneelike structure in the spectrum of the heavy component of cosmic rays observed with KASCADE-Grande. Phys. Rev. Lett. 107, 171104 (2011).

[61] K. Andeen, M. Plum et al., (IceCube Collab.), Cosmic Ray Spectrum and Composition from PeV to EeV from the IceCube Neutrino Observatory. PoS (ICRC2019), 172 (2019).

[62] R. U. Abbasi et al. (HiRes), First Observation of the Greisen-Zatsepin-Kuzmin Suppression. Phys. Rev. Lett. 100, 101101 (2008).

[63] V. Verzi et al., (Auger Collab.), Measurement of the energy spectrum of ultra-high energy cosmic rays using the Pierre Auger Observatory. PoS (ICRC2019), 450 (2019).

[64] D. Ivanov et al., (Telescope Array Collab.), Energy Spectrum Measured by the Telescope Array. PoS (ICRC2019), 298 (2019).

[65] D. Kazanas and A. Nicolaidis. Cosmic Ray "Knee": A Herald of New Physics?. PoS (ICRC 2001), 1760 (2001). arXiv:astro-ph/0103147

[66] K. Greisen, End to the cosmic-ray spectrum?. Phys. Rev. Lett., 16, 748 (1966).

[67] G.T. Zatsepin and V.A. Kuz'min, Upper Limit of the Spectrum of Cosmic Rays. JETP Lett., 4, 78–80 (1966).

[68] A. Yushkov, (Auger Collab.), Mass Composition of Cosmic Rays with Energies above $10^{17.2}$ eV from the Hybrid Data of the Pierre Auger Observatory. PoS (ICRC 2019), 482 (2019).

[69] D. Allard, Extragalactic propagation of ultrahigh energy cosmic-rays. Astropart. Phys., 39, 33–43 (2012), arXiv:1111.3290.

[70] D. Ivanov, (Auger Collab.), Report of the Telescope Array - Pierre Auger Observatory Working Group on Energy Spectrum. PoS (ICRC2017), 498 (2017).

[71] D. Ivanov et al., (Telescope Array and Auger Collabs.), EPJ Web Conf. 210 (2019) 01002.

[72] O. Deligny, (Telescope Array and Auger Collabs.), The energy spectrum of ultra-high energy cosmic rays measured at the Pierre Auger Observatory and at the Telescope Array. Proceedngs of Science (ICRC2019), 234 (2019)

[73] D. J. Bird et al. (HiRes), The Cosmic-Ray Energy Spectrum Observed by the Fly's Eye. Astrophys. J. 424, 491 (1994).

[74] V. Berezinsky, A. Z. Gazizov and S. I. Grigorieva, On astrophysical solution to ultrahigh energy cosmic rays. Phys. Rev. D74, 043005 (2006).

[75] S. Wykes et al., UHECR propagation from Centaurus A. (2017), arXiv:1706.08229

[76] J. Abraham et al. (Auger Collab.), Correlation of the highest-energy cosmic rays with the positions of nearby active galactic nuclei. Astropart.Phys. 29, 188 (2008), arXiv:0712.2843. [astro-ph]

[77] A. Aab, P. Abreu, M. Aglietta, et al., Searches for Anisotropies in the Arrival

Directions of the Highest Energy Cosmic Rays Detected by the Pierre Auger Observatory. Astrophys. J., 804, 15 (2015), arXiv:1411.6111.

[78] M. Wirtz, (Auger Collab.), Search for magnetically-induced signatures in the arrival directions of ultra-high energy cosmic rays measured by the Pierre Auger Observatory. PoS (ICRC2019), 469 (2019)

[79] A. Aab et al. (Auger Collab.), Indication of anisotropy in arrival directions of ultra-high-energy cosmic rays through comparison to the flux pattern of extragalactic gamma-ray sources. The Astrophysical Journal Letters 853 (2018), arXiv:1801.06160 [astro-ph.HE]

[80] L. Caccianiga, (Auger Collab.), Anisotropies of the highest energy cosmic-ray events recorded by the Pierre Auger Observatory in 15 years of operation. PoS (ICRC2019), 206 (2019).

[81] R. U. Abbasi et al. (Telescope Array Collab.), Indications of intermediate-scale anisotropy of cosmic rays with energy greater than 57 EeV in the northern sky measured with the surface detector of the Telescope Array experiment. ApJL, 790, L21 (2014)

[82] K. Kawata et al., (Telescope Array Collab.), Updated Results on the UHECR Hotspot Observed by the Telescope Array Experiment. PoS (ICRC2019), 310 (2019).

[83] D. Whiteson, M. Mulhearn, C. Shimmin, K. Cranmer, K. Brodie, and D. Burns, Searching for ultra-high energy cosmic rays with smartphones, Astropart. Phys.79, 1(2016), arXiv:1410.2895 [astro-ph.IM].

[84] M.J. Berger, J.S. Coursey, M.A. Zucker, and J. Chang, ESTAR, PSTAR, and ASTAR: Computer Programs for Calculating Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions (version 1.2.3). National Institute of Standards and Technology, (2005). `http://physics.nist.gov/Star`

[85] N. Gerasimova, and G. Zatsepin, Disintegration of cosmic ray nuclei by solar photons. Soviet Phys.,JETP, 11, 899 (1960)

[86] E. Albin and D. Whiteson, Feasibility of Correlated Extensive Air Shower Detection with a Distributed Cosmic Ray Network, (2021), arXiv:2102.03466 [astro-ph.HE].

[87] R. I. Epstein, The acceleration of interstellar grains and the composition of the cosmic rays. Monthly Notices of the Royal Astronomical Society, 193, 723 (1980), doi: 10.1093/mnras/193.4.723

[88] L. Spitzer, On the origin of heavy cosmic-ray particles. Phys. Rev., 76, 583 (1949), doi: 10.1103/PhysRev.76.583

[89] N. C. Wickramasinghe, On the Injection of Grains into Interstellar Clouds. Monthly Notices of the Royal Astronomical Society, 159, 269 (1972), doi: 10.1093/mnras/159.3.269

[90] N. C. Wickramasinghe, Electric Charge and Accleration of Suprathermal Grains. Astrophysics and Space Science, 28, L25 (1974), doi: 10.1007/BF00641936

[91] P. Goldreich and P. Morrison, On the absorption of gamma rays in intergalactic space. JETP, 18, 239 (1964)

[92] R. J. Gould and G. P. Schréder, Opacity of the universe to high-energy photons. Phys. Rev., 155, 1408 (1967), doi: 10.1103/PhysRev.155.1408

[93] J. V. Jelley, High-Energy γ-Ray Absorption in Space by a 3.5°K Microwave Field, Phys. Rev. Lett., 16, 479 (1966), doi: 10.1103/PhysRevLett.16.479

[94] A. I. Nikishov, Absorption of High-Energy Photons in the Universe. JETP, 14, 393 (1962). http://www.jetp.ac.ru/cgi-bin/e/index/e/14/2/p393?a=list

[95] F. W. Stecker, The cosmic gamma-ray spectrum from secondary-particle production in the metagalaxy. APJ, 157, 507 (1969), doi: 10.1086/150091==z

[96] V. S. Berezinsky and G. T. Zatsepin, Cosmic rays at ultra high energies (neutrino?). Phys. Lett., 28B, 423 (1969), doi: 10.1016/0370-2693(69)90341-4

[97] D. Fargion, B. Mele, and A. Salis, Ultra-high-energy neutrino scattering onto relic light neutrinos in the galactic halo as a possible source of the highest energy extragalactic cosmic rays. Astrophys. J., 517, 725 (1999), doi: 10.1086/307203

[98] E. Roulet, Ultrahigh energy neutrino absorption by neutrino dark matter. Phys. Rev. D, 47, 5247 (1993), doi: 10.1103/PhysRevD.47.5247

[99] T. Weiler, Resonant Absorption of Cosmic-Ray Neutrinos by the Relic-Neutrino Background. Phys. Rev. Lett., 49, 234 (1982), doi: 10.1103/PhysRevLett.49.234

[100] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, Gravi-burst: super-GZK cosmic rays from localized gravity. Phys. Lett., B549, 267 (2002),

[101] V. Berezinsky and A. Vilenkin, Cosmic necklaces and ultrahigh energy cosmic rays. Phys. Rev. Lett., 79, 5202 (1997), doi: 10.1103/PhysRevLett.79.5202

[102] P. Bhattacharjee, Cosmic strings and ultrahigh-energy cosmic rays. Phys. Rev. D, 40, 3968 (1989), doi: 10.1103/PhysRevD.40.3968

[103] P. Bhattacharjee and N. Rana, Ultrahigh-energy particle flux from cosmic strings. Phys. Lett. B, 246, 365 (1990), doi: https://doi.org/10.1016/0370-2693(90)90615-D

[104] P. Bhattacharjee. and G. Sigl, Monopole annihilation and highest energy cosmic rays. Phys. Rev., D51, 4079 (1995), doi: 10.1103/PhysRevD.51.4079

[105] R. H. Brandenberger, Topological defects and structure formation. Int. J. Mod. Phys., A9, 2117 (1994), doi: 10.1142/S0217751X9400090X

[106] D. J. H. Chung, E. W. Kolb, and A. Riotto, Superheavy dark matter. Phys. Rev., D59, 023501 (1998), doi: 10.1103/PhysRevD.59.023501

[107] P. Dave, and I. Taboada, (IceCube). Neutrinos from Primordial Black Hole Evaporation. PoS (ICRC2019), 863 (2019), arXiv:1908.05403 [astro-ph.HE]

[108] C. T. Hill, Monopolonium. Nucl. Phys. B, 224, 469, (1983). doi: https://doi.org/10.1016/0550-3213(83)90386-3

[109] C. T. Hill, D. N. Schramm, & T. P. Walker, Ultra-high-energy cosmic rays from superconducting cosmic strings. Phys. Rev. D, 36, 1007 (1987), doi: 10.1103/PhysRevD.36.1007

[110] M. B. Hindmarsh, & T. W. B. Kibble, Cosmic Strings. Rept. Prog. Phys., 58, 477 (1995), doi: 10.1088/0034-4885/58/5/001

[111] T. W. B. Kibble, Topology of cosmic domains and strings. Journal of Physics A: Mathematical and General, 9, 1387, (1976)

[112] L. Masperi and M. Orsaria, Hard component of ultra-high energy cosmic rays and vortons. Astropart. Phys., 16, 411, (2002)

[113] L. Masperi and G.A. Silva, Cosmic rays from decaying vortons. Astropart. Phys., 8, 173 (1998)

[114] A. Vilenkin, Cosmic strings and domain walls. Physics Reports, 121, 263, (1985)

[115] Opto-semiconductor Handbook, Hamamatsu, Chapter 2: Si Photodiodes (2014) [Online] `https://www.hamamatsu.com/resources/pdf/ssd/e02_handbook_si_photodiode.pdf`

[116] H. Yamanaka, Method and apparatus for producing ultra-thin semiconductor chip and method and apparatus for producing ultra-thin back-illuminated solid-state image pickup device. S patent 7521335, Sony Corporation (2009)

[117] R. Fontaine (TechInsights), A Survey of Enabling Technologies in Successful Consumer Digital Imaging Products. Proc. Int. Image Sensor Workshop (IISW 2017). R06 (2017)

[118] R. Fontaine (TechInsights), The State-of-the-Art Smartphone Imagers. Proc. Int. Image Sensor Workshop (IISW 2019). R01 (2019)

[119] Tournier, A. et al., Pixel-to-Pixel isolation by Deep Trench technology: Application to CMOS Image Sensor, Proc. Int. Image Sensor Workshop (IISW 2011). R5 (2011)

[120] M. Ghormishi, A. Pil-ali, & M.A. Karami, Crosstalk Characterization in Backside Illuminated Photodiodes. Iranian Conference on Engineering Electromagnetic (ICEEM2014). (2014). doi:10.13140/2.1.2412.8325.

[121] R. Fontaine, The state of the art of smartphone imagers. Part 3: Back-Illuminated Active Si Thickness, Deep Trench Isolation (DTI). TechInsights Blog. 23 July 2019. `https://www.techinsights.com/blog`

[122] The CMS Collaboration. Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons. JINST 5:T03007 (2010). doi: 10.1088/1748-0221/5/03/T03007

[123] J. Cogliati, K. Derr, & J. Wharton. Using CMOS Sensors in a Cellphone for Gamma Detection and Classification. (2014). arXiv:1401.0766

[124] CCD or CMOS: can CMOS sensors replace CCDs in all cases? Allied Vision (2016). https://www.stemmer-imaging.com/media/uploads/cameras/avt/12/120483-Allied_Vision_White_Paper_CCD_vs_CMOS.pdf

[125] K. Pulli, Camera Processing Pipeline. CS231M: Mobile Computer Vision. Stanford University. (2015). https://web.stanford.edu/class/cs231m/lectures/lecture-11-camera-isp.pdf

[126] R. Triggs, What is PDAF and how does it work? Phase Detection Autofocus explained. Android Authority. (2021). https://www.androidauthority.com/how-pdaf-works-1102272/

[127] R. Fontaine, The state of the art of smartphone imagers. Part 4: Non-Bayer CFA, Phase Detection Autofocus (PDAF). (2019). https://www.techinsights.com/blog

[128] Y. Fan & Y. Chiang. Discrete Wavelet Transform on Color Picture Interpolation of Digital Still Camera. VLSI Design. (2013). doi:10.1155/2013/738057.

[129] Android API Reference. (2021). https://developer.android.com/reference/android/hardware/camera2/CaptureRequest

[130] http://www.fujifilmusa.com/products/digital_cameras/x/fujifilm_x_s1/features/page_02.html

[131] IMX135-xxW Product Brief. Sony Corporation. https://www.uctronics.com/download/Image_Sensor/IMX135-xxW_DataSheet(E)_Partial.pdf

[132] The Kubernetes Authors, Kubernetes Documentation. (2021) https://kubernetes.io/docs/tutorials/kubernetes-basics/

[133] G. Bradski, The OpenCV Library, Dr. Dobb's Journal of751Software Tools (2000).

[134] R. Fontaine, The State-of-the-Art of Mainstream CMOS Image Sensors, Proc. Int. Image Sensor Workshop (IISW 2015). Introductory Paper (2015)

[135] W. S. Boyle and G. E. Smith, Charge coupled semiconductor devices. The Bell System Technical Journal, 49, 587 (1970).

[136] J. C. Mullikin, L. J. van Vliet, H. Netten, F. R. Bod-deke, G. van der Feltz, and I. T. Young, Methods for CCD camera characterization. in Image Acquisition and Scientific Imaging Systems, Vol. 2173, edited by H. C. Titus and A. Waks, International Society for Optics and Photonics (SPIE, 1994) pp. 73 – 84.

[137] European Machine Vision Association. EMVA Standard 1288: Standard for Characterization of Image Sensors and Cameras. (2010). https://www.emva.org/wp-content/uploads/EMVA1288-3.0.pdf

[138] J. Park et al., Pixel Technology for Improving IR Quantum Efficiency of

Backside-illuminated CMOS Image Sensor. Proc. Int. Image Sensor Workshop (IISW 2019). R14 (2019)

[139] P.K.F. Grieder, Cosmic Rays at Earth – Researcher's reference manual and data book. (2001)

[140] S.Y.F. Chu, L.P. Ekström, & R.B. Firestone, The Lund/LBNL Nuclear Data Search. Version 2.0. (1999). http://nucleardata.nuclear.lu.se/

[141] J.T. Cessna, Radium-226 Gamma Ray Spectrum from High Purity Germanium (HPGe) Detector. NIST, (2009). https://www.nist.gov/pml/marie-curie-and-nbs-radium-standards/1927-nbs-gold-leaf-electroscope

[142] M. Ester, H. P. Kriegel, J. Sander, & X. Xu, A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, Portland, OR, AAAI Press, 226–231. (1996)

[143] IMX 298 Product Brief. Sony Corporation. (2016). https://www.sony-semicon.co.jp/products/common/pdf/ProductBrief_IMX298_20160210.pdf

[144] R. D. Gow et al., A Comprehensive Tool for Modeling CMOS Image-Sensor-Noise Performance. IEEE Transactions on Electronic Devices, 54(6), 1321-1329 (2007)

[145] C. Leyris, F. Roy, & M. Marin, Modeling of the Temporal Pixel to Pixel Noise of CMOS Image Sensors. Proc. Int. Image Sensor Workshop (IISW 2007), 12 (2007)

[146] Yuichiro Yamashita et al., Quantum Efficiency Simulation with Boltzmann Transport Equation for Motion Modeling of Individual Particles in Photodiodes. Proc. Int. Image Sensor Workshop (IISW 2017), R15 (2017)

[147] S. Li, H. S. Matis, N. Xuong, and S. Kleinfelder, Modeling and Analysis of Charged-Particle CMOS Image Sensor Arrays. IEEE Transactions on Nuclear Science. 56(3), 1062-1068 (2009)

[148] D.P. Magalhaes & A. Tomal, Monte Carlo simulation of hybrid pixel detectors. Radiation Physics and Chemistry, 167, 108296 (2020)

[149] D. Dannheim, et. al., Combining TCAD and Monte Carlo methods to simulate CMOS pixel sensors with a small collection electrode using the Allpix2 framework. Nucl. Instrum. Methods Phys. Res., Sect. A, 964, (2020). arXiv:2002.12602 [physics.ins-det]

[150] T. Sato, Analytical Model for Estimating Terrestrial Cosmic Ray Fluxes Nearly Anytime and Anywhere in the World: Extension of PARMA/EXPACS. PLOS ONE 10(12): e0144679. (2015) https://doi.org/10.1371/journal.pone.0144679

[151] P.A. Zyla et al. (Particle Data Group), Prog. Theor. Exp. Phys. 2020, 083C01 (2020)

[152] V.N. Ivanchenko, et al., Geant4 electromagnetic physics: improving simulation performance and accuracy, SNAMC 2013 Joint International Conference on Supercomputing in Nuclear Applications Monte Carlo, Paris, EDP Sciences, 03101 (2014),

[153] M. Raine, M. Gaillardin, & P. Paillet, Geant4 physics processes for silicon microdosimetry simulation: Improvements and extension of the energy-range validity up to 10 GeV/nucleon, Nucl. Instrum. Methods Phys. Res., Sect. B, 325, 97-100, (2014)

[154] Emery, F. E. & Rabson, T. A., Average Energy Expended Per Ionized Electron-Hole Pair in Silicon and Germanium as a Function of Temperature, Phys. Rev. A140, 2089, (1965)

[155] B.G. Lowe & R.A. Sareen, A measurement of the electron–hole pair creation energy and the Fano factor in silicon for 5.9keV X-rays and their temperature dependence in the range 80–270K. Nucl. Instrum. Methods Phys. Res., Sect. A, 576, 367-370, (2007) https://doi.org/10.1016/j.nima.2007.03.020.

[156] E.C. Auden et al., Modeling charge collection efficiency degradation in partially depleted GaAs photodiodes using the 1- and 2-carrier Hecht equations. Nucl. Instrum. Methods Phys. Res., Sect. B, 399 (2017)

[157] J. Yun et al., A Small-size Dual Pixel CMOS Image Sensor with Vertically Broad Photodiode of 0.61μm pitch. Proc. Int. Image Sensor Workshop (IISW 2019), R03 (2019)

[158] ISOCELL GH1, Samsung Corporation. [Accessed: 28 May 2021] https://www.samsung.com/semiconductor/minisite/isocell/mobile-image-sensors/isocell-slim-gh1/

[159] A. Marino, Marc Levoy on the balance of camera hardware, software, and artistic expression. The Vergecast. (2020) https://www.theverge.com/21427188/marc-levoy-interview-adobe-smartphone-computational-photography-vergecast

[160] R. Fontaine, Samsung Galaxy S9 Camera Teardown. TechInsights Blog. 9 April 2018. https://www.techinsights.com/blog/samsung-galaxy-s9-camera-teardown

[161] C. Gartenberg, Google announces the Pixel 5 for $699. The Verge. (2020) https://www.theverge.com/2020/9/30/21456181/google-pixel-5-features-price-release-date-announcement

[162] A. Jhaveri, Huawei P40 series could sport a custom Sony IMX700 image sensor. TechRadar. (2020) https://www.techradar.com/news/huawei-p40-series-could-sport-a-custom-sony-imx700-image-sensor

[163] ISOCELL HM3, Samsung Corporation. [Accessed: 28 May 2021] https://www.samsung.com/semiconductor/minisite/isocell/mobile-image-sensors/isocell-hm3/

[164] K. Griesen, Progress in Cosmic Ray Physics 3 1 (1956).

[165] Center for International Earth Science Information Network - CIESIN - Columbia University. 2018. Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 11. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H4JW8BX5. Accessed 6 June 2021.