



THE UNIVERSITY
of ADELAIDE

Systematic and Statistical Uncertainties in Cosmic Ray Arrival Direction Reconstruction

Philip Wahrlich

A thesis submitted to the University of Adelaide in fulfilment of the
requirements of the degree of Doctor of Philosophy

January 2013

School of Chemistry and Physics
The University of Adelaide

Contents

1	Introduction	15
2	Ultra-High Energy Cosmic Rays	18
2.1	Introduction	18
2.1.1	Cosmic Ray Spectrum	19
2.1.2	Chemical Composition	19
2.2	Possible UHECR sources	21
2.2.1	Bottom-up Acceleration	21
2.2.2	Top-down Scenarios	28
2.3	Propagation of UHECR	31
2.3.1	Energy Losses	31
2.3.2	Magnetic Deflection	35
3	Extensive Air Showers	38
3.1	Heitler Models	38
3.2	Cascade Theory	43
3.3	Lateral shower development	45
3.4	Air Shower Detectors	46
3.4.1	Cherenkov Telescopes	46
3.4.2	Fluorescence Detectors	46
3.4.3	Ground Arrays	48
3.5	Air Fluorescence Detection	50
3.5.1	Air Fluorescence	50
3.5.2	Cherenkov Radiation	52
4	The Pierre Auger Observatory	55
4.1	Surface Detector	55
4.1.1	Calibration	56
4.1.2	Trigger System	58
4.2	Fluorescence Detector	60
4.2.1	FD Electronics and Hardware Trigger System	64
4.2.2	Calibration	66
4.3	HEAT	69
4.4	AMIGA	70
4.5	Atmospheric Monitoring	70

5	Hybrid Reconstruction	76
5.1	Pulse Finding	76
5.2	Shower-Detector Plane	78
5.3	FD Time Fit	82
5.4	Hybrid Time Fit	87
5.5	Angular Resolution	89
5.6	Longitudinal Profile Reconstruction	91
5.7	Hybrid Event Simulation	97
5.7.1	Air Shower Simulation	97
5.7.2	Detector Simulation	99
 6	 A Summary of Recent Results from the Pierre Auger Observatory	 106
6.1	Arrival Directions	106
6.2	Primary Composition	108
6.3	Neutrino Flux Limits	108
6.4	Photon Flux Limits	110
6.5	Energy Spectrum	111
 7	 Systematic and Statistical Uncertainties in Hybrid Reconstruction	 115
7.1	SDP Uncertainty	115
7.1.1	Simulating Shower Images	120
7.1.2	Toy Model	124
7.1.3	Conclusions and Discussion	131
7.2	Alternative SDP Reconstruction Algorithm	140
7.2.1	Amplitude Fitting	140
7.2.2	Toy MC	141
7.2.3	Implementation	143
7.2.3.1	Flattening the Light Flux	143
7.2.3.2	Non-pulsed Pixels	147
7.2.3.3	Outlier Rejection	148
7.2.3.4	Resolution from Simulated Showers	148
7.2.4	Problems with Wide Tracks	149
7.2.5	Discussion	149
7.3	Global Hybrid Geometry Regression	155
7.3.1	Weighted χ^2 Components	156
7.3.2	Discussion	157
7.4	Pulse Centroid Time Uncertainty	160
7.4.1	FD Shower Geometry Reconstruction	161
7.4.2	Determining the Pulse Centroid	162

7.4.3	Previous Studies	162
7.4.4	Pulse Centroid Time - Statistical Uncertainty	163
7.4.4.1	Signal Uncertainty	163
7.4.4.2	Signal Autocorrelation	165
7.4.4.3	Timing Uncertainty	167
7.4.5	Validation with FD Monocular Data	168
7.4.6	Data Set	170
7.4.7	Reduced χ^2 Distribution	170
7.4.8	χ^2 Probability Distribution	171
7.4.9	Correlation with Reconstructed Shower Parameters	172
7.4.10	Pulse Length Dependence	175
7.4.10.1	Pulse Finding Systematics	175
7.4.10.2	Cleaning Cut Systematics	175
7.4.11	Results	176
7.4.12	Discussion	177
7.5	Atmospheric Refraction	181
7.5.1	Ray-tracing	181
7.5.2	Atmospheric Refractive Index Profile	181
7.5.3	Curved Ray Trajectories and SD-FD Time Synchronicity Measurements	182
7.5.4	Correction for SD-FD time offset measurements with Laser Shots	183
7.5.5	Correction for SD-FD time offset measurements with Showers	184
7.5.6	Conclusions	185
7.6	Pixel Chord Asymmetry	186
7.6.1	Chord Asymmetry Estimator	187
7.6.2	Simulation	188
7.6.3	Discussion	189
8	Synchronising the Hybrid Detector	191
8.1	FD-SD time offset	192
8.1.1	Calculation of τ_{FD}	194
8.1.2	Calculation of τ_{SD}	194
8.1.3	Results	194
8.2	End-to-End Time Synchronisation	197
8.2.1	Weighted Least Trimmed Squares Curvature Fit	198
8.2.2	Measuring Time Synchronisation	198
8.2.3	Station 203: ‘Celeste’	200
8.2.4	Discussion	203
8.3	Telescope-Telescope Time Synchronisation	204

8.3.1	Calculating Telescope-Telescope Time Offsets	204
8.3.2	Methodology	205
8.3.2.1	Reconstruction and Analysis Cuts	205
8.3.3	Discussion	206
9	Galactic Centre Point Source Search	209
9.1	Galactic Centre Neutron Search	211
9.1.1	Data Selection	211
9.1.2	Sky Map	211
9.1.3	Significance	213
9.1.4	Results	213
10	Conclusion	216
A	Appendix A	220
A.1	Statistical Tools	220
A.1.1	Linear Least Squares Regression	220
A.1.2	Weighted Least Squares	222
A.1.3	Non-Linear Least Squares	222
A.1.3.1	Error Estimation	223
A.1.4	Diagnostics: Properties of the χ^2 minimisation	224
A.2	Least Trimmed Squares	225
A.3	The Bootstrap Method	225
A.3.1	Estimating the Variance of the ADS for Laser Shots	227
A.4	Angular Uncertainty under Coordinate Transformations	228
A.5	Standard Geometry Cuts	229
A.5.1	Basic Cuts	229
A.5.2	Geometry Cuts	229

List of Figures

2.1	The all-particle differential cosmic ray spectrum.	20
2.2	The flux of low energy cosmic ray nuclei, in particles per energy-per-nucleus, measured by various (mostly balloon-borne and satellite based) experiments. Figure from [1].	22
2.3	The average depth of shower maximum as measured by HiRes, the Pierre Auger Observatory, and the Telescope Array. Overlaid are lines representing the expectations from three hadronic interaction models for proton primaries (upper lines) and iron primaries (lower lines). Plot from [2].	23
2.4	Diagram showing a cosmic ray entering and exiting a molecular cloud.	24
2.5	Diagram showing a particle moving across the edge of a region of shocked gas.	26
2.6	The Hillas plot.	29
2.7	$P\gamma$ total cross section.	33
2.8	Attenuation of UHECR protons.	34
2.9	Galactic magnetic field models from [3]. See text for details.	37
3.1	Simulation of a 10^{12} eV proton-initiated EAS using CORSIKA [4].	39
3.2	Heitler's qualitative model for an electromagnetic cascade.	41
3.3	Matthews [5] describes a Heitler-type qualitative model for a hadronic cascade.	42
3.4	The Whipple Observatory IACT.	47
3.5	The Fly's Eye.	48
3.6	Nitrogen fluorescence spectrum from [6].	51
3.7	Energy levels for the 2P and 1N systems. Figure from [7].	51
3.8	Cherenkov radiation at the Advanced Test Reactor at Idaho National Laboratory.	53
3.9	Cherenkov radiation is emitted along the surface of a cone.	54
4.1	Schematic view of an SD station.	56
4.2	The layout of Auger South.	57
4.3	3ToT station configurations.	60
4.4	4C1 station configurations.	60
4.5	T5 trigger probability from [8]	61
4.6	FD building schematic.	62
4.7	An illustration of a Mercedes Star light collector.	62
4.8	The light collection efficiency of a Mercedes Star light collector from [9]	63
4.9	SLT fundamental patterns.	64

4.10	An illustration of the TLT cursory SDP fit.	65
4.11	T3 efficiency relative to TLT efficiency	67
4.12	T3 trigger efficiency.	67
4.13	The three HEAT telescopes are housed in separate enclosures which pivot, allowing for operation in either high or low elevation modes.	70
4.14	AMIGA is located near to, and in the field of view of, the Coihueco FD and HEAT telescopes.	71
4.15	The CLF, with the Celeste SD station in the background.	72
4.16	FRAM and its enclosure, located about 30 m from the Los Leones site.	73
5.1	An example shower track as detected by an FD camera. Colours represent relative pulse timing information, hotter indicating later pulses.	78
5.2	SDP normal vector convention.	82
5.3	An illustration of the shower geometry, defined in terms of the FD ob- servables.	83
5.4	Diagram of a downgoing event, landing in front of the FD site.	85
5.5	Diagram of a downgoing event, landing behind the FD site.	86
5.6	Diagram of an upgoing event, emerging from in front the FD site.	86
5.7	Diagram of an upgoing event, emerging from behind the FD site.	86
5.8	The resolution of hybrid reconstruction and FD monocular reconstruction.	89
5.9	Hybrid angular resolution for simulated events.	91
5.10	Cherenkov light contributions.	93
6.1	The arrival directions of 27 events which correlate with AGN from the VCV catalogue [10].	107
6.2	The mean X_{\max} and $\text{RMS}(X_{\max})$ as a function of energy from [11].	109
6.3	Single flavour UHE neutrino diffuse flux limits at 90% confidence from the Pierre Auger Observatory [12].	110
6.4	The Fisher response distributions for a simulated proton and photon events with energies between 10^{18} eV and $10^{18.5}$ eV.	112
6.5	Photon flux upper limits at 95% confidence [13].	113
6.6	The UHECR energy spectrum from [14].	114
7.1	The empirical distribution of $\frac{\sqrt{q_i}\Omega_i}{\text{RMS}(\sqrt{q_i}\Omega_i)}$ for golden hybrid events.	117
7.2	The dependence of L^2/N with the width of the track image for recorded events.	118
7.3	The dependence of L^2/N with the width of the track image for simulated events.	119
7.4	The image of a simulated shower.	121

7.5	The distribution of $\text{pull}_{\text{SDP}_\theta}$ for an ensemble of simulated shower events. The error bars represent the RMS of $\text{pull}_{\text{SDP}_\theta}$ in bins of $\text{SDP}_\theta^{\text{sim}}$	122
7.6	The distribution of the residuals ΔSDP_θ for an ensemble of simulated shower events.	123
7.7	The RMS of the pull distribution of SDP_θ as a function of various reconstructed shower parameters.	125
7.8	The RMS of the pull distribution of SDP_θ as a function of other various reconstructed shower parameters.	126
7.9	An example of a simulated shower image. Pulsed pixels (red) and non-pulsed pixels (black) are shown.	127
7.10	The distribution of $\text{pull}_{\text{SDP}_\theta}$ for an ensemble of toy Monte-Carlo simulated shower events. The error bars represent the RMS of $\text{pull}_{\text{SDP}_\theta}$ in bins of $\text{SDP}_\theta^{\text{sim}}$	129
7.11	The distribution of the residuals ΔSDP_θ for an ensemble of toy Monte-Carlo simulated shower events. The error bars represent the RMS of ΔSDP_θ in bins of $\text{SDP}_\theta^{\text{sim}}$	130
7.12	An illustration showing an example of an SDP reconstruction systematic error which occurs due to insufficient pixel resolution.	131
7.13	The distribution of $\text{pull}_{\text{SDP}_\theta}$ for an ensemble of toy Monte-Carlo shower events.	132
7.14	The distribution of the residuals ΔSDP_θ for an ensemble of toy Monte-Carlo simulated shower events.	133
7.15	An illustration showing the pathology of a geometry reconstruction error.	134
7.16	An example of spatial aliasing. There is a loss of information because of the discrete resolution of the camera.	136
7.17	An illustration of the scheme in which the upper and lower halves of the camera are horizontally offset by a half pixel width to reduce the effect of aliasing.	137
7.18	An example of a problematic SDP reconstruction for a simulated event.	139
7.19	The distributions of the differences between the reconstructed SDP parameters ($\text{SDP}_\theta^{\text{rec}}$ and $\text{SDP}_\phi^{\text{rec}}$) and the actual simulated SDP parameters ($\text{SDP}_\theta^{\text{sim}}$ and $\text{SDP}_\phi^{\text{sim}}$) resulting from the amplitude weighting reconstruction	144
7.20	The distributions of the differences between the reconstructed SDP parameters ($\text{SDP}_\theta^{\text{rec}}$ and $\text{SDP}_\phi^{\text{rec}}$) and the actual simulated SDP parameters ($\text{SDP}_\theta^{\text{sim}}$ and $\text{SDP}_\phi^{\text{sim}}$) resulting from the amplitude fitting reconstruction	145

7.21	The distributions of the differences between the reconstructed SDP parameters ($\text{SDP}_\theta^{\text{rec}}$ and $\text{SDP}_\phi^{\text{rec}}$) and the actual simulated SDP parameters ($\text{SDP}_\theta^{\text{sim}}$ and $\text{SDP}_\phi^{\text{sim}}$) resulting from the amplitude fitting reconstruction with the actual simulated ADS known a priori.	146
7.22	The light spot falls off the edge of the camera and some light fails to be detected, so the luminosity of the shower is not well known while the light spot falls on the outermost pixels (red dots).	147
7.23	The resolution of the fitted SDP using amplitude fitting in the absence of statistical fluctuations in the pixel signals.	150
7.24	The resolution of the fitted SDP using amplitude weighting in the absence of statistical fluctuations in the pixel signals.	151
7.25	The resolution of the fitted SDP using amplitude fitting in the presence of statistical fluctuations in the pixel signals.	152
7.26	The resolution of the fitted SDP using amplitude weighting in the presence of statistical fluctuations in the pixel signals.	153
7.27	An example of a problematic SDP reconstruction.	154
7.28	The difference between the size of the errors in the reconstructed shower axis parameters for global shower axis regression and sequential shower axis regression.	158
7.29	The difference between the size of the errors in the reconstructed SDP vector components for global shower axis regression and sequential shower axis regression.	159
7.30	Illustration of the geometrical shower reconstruction from the observables of the fluorescence detector [15].	161
7.31	A typical ADC trace showing the pulse centroid as indicated with the red dot and dashed line (its uncertainty is contained within the dot). The shaded region indicates the portion of the ADC trace flagged as containing the pulse.	162
7.32	A typical ADC trace from an air shower event.	166
7.33	Noise correlation coefficient distributions from data taken with Coiheuco (Fig. 7.33a) and HEAT (Fig. 7.33b) for consecutive samples ($\rho_{i,i+1}$), and samples which are separated by one sample ($\rho_{i,i+2}$).	167
7.34	The distribution of $\Delta q_i/\sigma_{q_i}$ has mean ~ 0 . That the RMS of the distribution is larger than unity is the result of statistical bias in the estimator σ_{q_i}	169
7.35	The distribution of $\Delta \bar{t}/\sigma_{\bar{t}}$ is well fit by a Gaussian with mean ~ 0 and standard deviation ~ 1 . Therefore the error model is appropriate for simulation.	169

7.36	The distribution of reduced χ^2 from timing fits without pulse length correction.	171
7.37	The pulse length dependence of the timing fit reduced χ^2 is well fit by a quadratic. The x-axis is the mean pulse length per event.	172
7.38	The distribution of reduced χ^2 from timing fits with pulse length correction of the centroid variance.	173
7.39	The χ^2 probability distribution for all events (blue) and events with cores falling further than 2000 m of an FD site (orange).	173
7.40	The reduced χ^2 of the timing fit as a function of reconstructed χ_0 , energy, and core-eye distance.	174
7.41	The distribution of reduced χ^2 for Coihueco events has unit expectation.	177
7.42	The χ^2 probability distribution for Coihueco events is approximately uniform.	177
7.43	For Coihueco events, the mean reduced χ^2 is relatively independent of reconstructed shower geometry parameters χ_0 and R_p	178
7.44	The distribution of reduced χ^2 for HEAT events has an almost unit expectation. A slight deviation from unity is not unexpected, since only nominal rather than measured calibration values were used.	178
7.45	The χ^2 probability distribution for 2464 HEAT events is approximately flat. Note the lack of significant peaks at $P(\chi^2)=0$ and $P(\chi^2)=1$ which would otherwise indicate inadequacies in how the model deals with particular classes of events.	179
7.46	For HEAT events the mean χ^2 is relatively independent of χ_0 and R_p	179
7.47	Angular difference between the straight-line path and the real path for light emitted in the field of view of the telescope.	183
7.48	Deviation from the expected light travel time if straight-line light propagation is assumed.	184
7.49	Expected SD-FD time offset correction for hybrid showers with a vertical SDP measured by Los Leones with refracted light paths taken into account.	184
7.50	Chords across circular and hexagonal shapes.	186
7.51	Hexagonal pixel dimensions for a pixel of width W . Also shown is a chord vector Ω traversing the hexagon.	188
7.52	The time residual as a function of pixel impact parameter expected from theory 7.52a and seen in simulated events 7.52b. A 6th order polynomial is fitted to the simulated data for visualisation purposes only. Note the similiary in the shape of the impact parameter dependence in both plots.	190

8.1	SD-FD time offset measured with vertical laser shots observed from Los Leones (LL), Loma Morados (LM), Loma Amarilla (LA), and Coiheuco (CO).	195
8.2	SD-FD time offset measured with inclined laser shots observed from Los Leones (LL), Loma Morados (LM), Loma Amarilla (LA), and Coiheuco (CO).	196
8.3	An illustration showing how station timing residuals can be expressed as a weighted, directed graph.	199
8.4	Time offset (a) and associated uncertainty (b) of each SD station with respect to Celeste.	201
8.5	For stations within 15 km of each FD site, the distribution of time offsets with respect to Celeste are shown, along with the distribution of deviations from zero.	202
8.6	An example of a single-telescope reconstruction of a multi-telescope event.	206
8.7	Relative time offsets between telescopes at the Los Leones (LL), Los Morados (LM), Loma Amarilla (LA), and Coihueco (CO) sites. The time offsets measured from reconstruction A (red), reconstruction B (blue), and their weighted average (black) are shown.	207
9.1	The density of selected hybrid events with arrival directions pointing back to a small region of the sky around the Galactic centre region. The location of Sag A* is shown (cross) in the centre of the plot.	212
9.2	The distribution of number of events within 1.8° of the Galactic centre, generated from 300 resampled data sets. 39 recorded events are found to lie within 1.8° of the Galactic centre, shown as the red line.	214
9.3	The p-values associated with various radii of circular search windows about the Galactic centre.	215
A.1	Millions of phone calls in Belgium, 1950-1973. Figure from [16]. Shown are three fitted lines. It can be seen that the LTS regression is robust against outliers.	226
A.2	Semi-hexagon distribution.	227

List of Tables

5.1	Systematic energy uncertainties.	96
7.1	The centroid error algorithm currently implemented in the Auger Offline Framework.	163
7.2	Noise correlation coefficients for consecutive samples $(\rho_{i,i+1})$, and samples which are separated by one sample $(\rho_{i,i+2})$	168
8.1	Time offsets between those stations within 15 km of each FD site and Celeste.	202
8.2	Time offsets between telescopes at Los Leones.	206
8.3	Time offsets between telescopes at Los Morados.	206
8.4	Time offsets between telescopes at Loma Amarilla.	207
8.5	Time offsets between telescopes at Coiheuco.	208
A.1	Standard Geometry Cuts	229

Abstract

The Pierre Auger Observatory (PAO) was constructed to study the highest energy cosmic rays (UHECR). A hybrid of ground array and fluorescence detector, it is the largest ultra-high energy cosmic ray detector to date. As such, the PAO detects UHECR in unprecedented amounts, offering unique insights into the nature and origin of these most extraordinarily energetic particles in the universe.

In this thesis we improve the accuracy of arrival direction uncertainty estimates for reconstructed events. We validate these improved uncertainty estimators through a number of statistical techniques, involving both recorded and simulated data. Furthermore, we identify and correct a number of systematic errors which arise in algorithmic corner cases. We propose novel techniques for measuring the time synchronisation of each detector in the using recorded air shower data. We use these techniques to measure the synchronisation of each detector across the PAO surface detector and fluorescence detector. Finally, we perform a cursory search for a point source of UHECR at the Galactic centre. A slight over-density of events is measured from the direction of the Galactic centre, however this over-density is not substantial enough to indicate a departure from isotropy.

Statement of Originality

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed: Date:

Acknowledgements

I would like to express my gratitude to the people who have helped me through the journey which has culminated in this thesis.

Firstly, to my supervisor, Prof. Bruce Dawson, who has been for me both mentor and role model throughout my stay. To my co-supervisor, Prof. Roger Clay, whose honest critique has always been timely and valuable.

I would also like to acknowledge the wider Adelaide Astrophysics Group, past and present - thank you for being such a great bunch of people to share a drink with. I recall my time spent with you with much fondness.

It has been a tremendous opportunity and a satisfying learning experience for me to work on the Pierre Auger Observatory. I am indebted to all those researchers and engineers whose efforts have made this possible.

Lastly, I would like to thank my parents, Michele and Gary, for their unwavering support and encouragement.