Signal estimation in on/off measurements including event-by-event variables (BASIL)

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

Tim Unbehaun
Science club 23.03.2023

Fermi Gamma-ray Space Telescope
H.E.S.S.
ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Content

PHYSICAL REVIEW D **103**, 123001 (2021)

Signal estimation in on/off measurements including event-by-event variables

G. D'Amico,[1,*] T. Terzić,[2] J. Strišković,[3] M. Doro,[4] M. Strzys,[5] and J. van Scherpenberg[6]
[1]*Department for Physics and Technology, University of Bergen, Bergen NO-5020, Norway*
[2]*University of Rijeka, Department of Physics, 51000 Rijeka, Croatia*
[3]*Josip Juraj Strossmayer University of Osijek, Department of Physics, 31000 Osijek, Croatia*
[4]*Università di Padova and INFN, I-35131 Padova, Italy*
[5]*Institute for Cosmic Ray Research (ICRR), The University of Tokyo, Kashiwa, 277-8582 Chiba, Japan*
[6]*Max-Planck-Institut für Physik, D-80805 München, Germany*

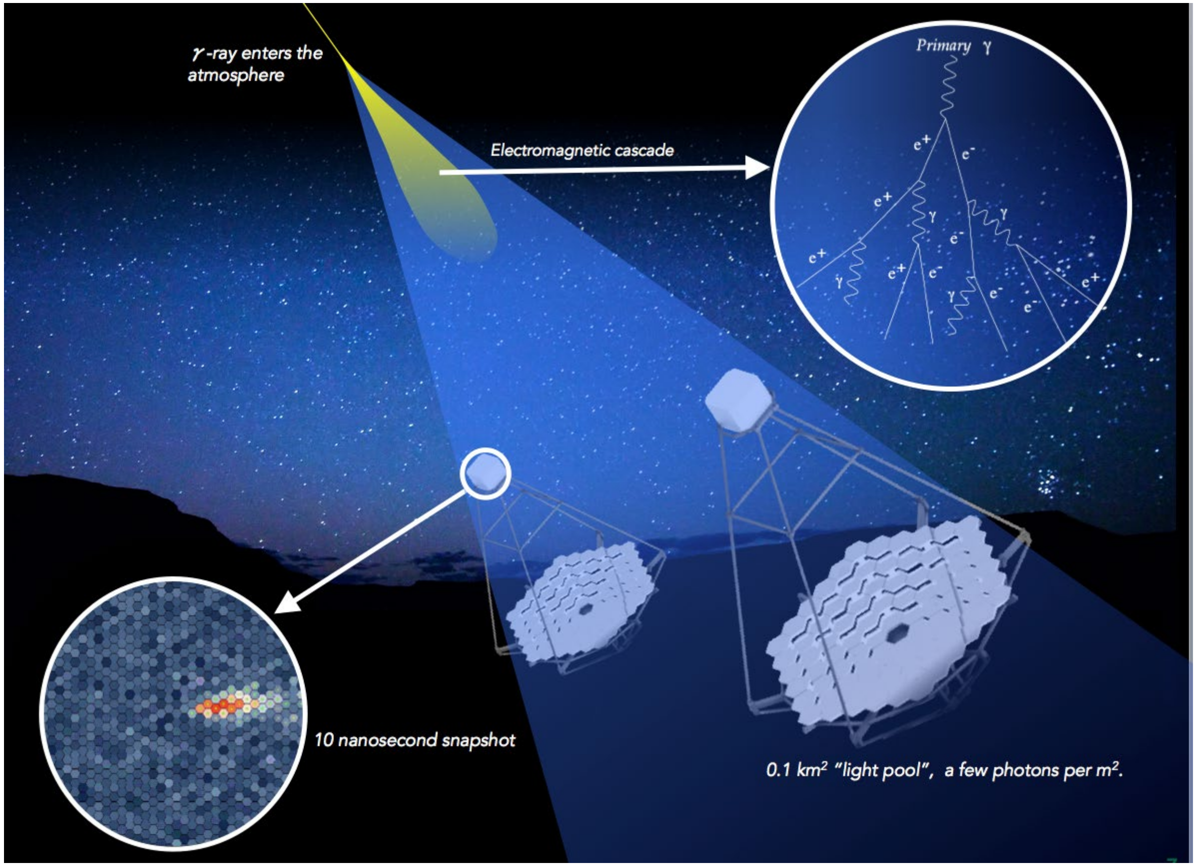
(Received 10 April 2021; accepted 30 April 2021; published 1 June 2021)

Signal estimation in the presence of background noise is a common problem in several scientific disciplines. An "on/off" measurement is performed when the background itself is not known, being estimated from a background control sample. The "frequentist" and Bayesian approaches for signal estimation in on/off measurements are reviewed and compared, focusing on the weakness of the former and on the advantages of the latter in correctly addressing the Poissonian nature of the problem. In this work, we devise a novel reconstruction method, Bayesian analysis including single-event likelihoods (dubbed BASiL), for estimating the signal rate based on the Bayesian formalism. It uses information on event-by-event individual parameters and their distribution for the signal and background population. Events are thereby weighted according to their likelihood of being a signal or a background event and background suppression can be achieved without performing fixed fiducial cuts. Throughout the work, we maintain a general notation that allows us to apply the method generically and provides a performance test using real data and simulations of observations with the MAGIC telescopes, as a demonstration of the performance for Cherenkov telescopes. BASiL allows one to estimate the signal more precisely, avoiding loss of exposure due to signal extraction cuts. We expect its applicability to be straightforward in similar cases.

DOI: 10.1103/PhysRevD.103.123001

IACTs

- Imaging air showers
- Reconstructing primary particle properties (energy, direction, time)
- Signal: γ-ray
- Background: protons
- ~ 1000 protons per γ
  → need discrimination

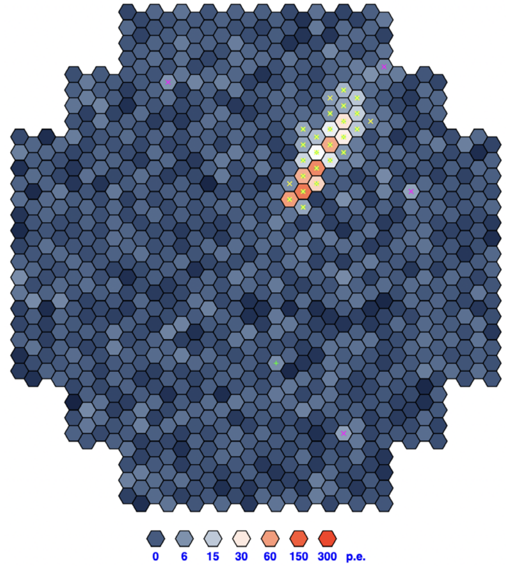


credit: CTAO

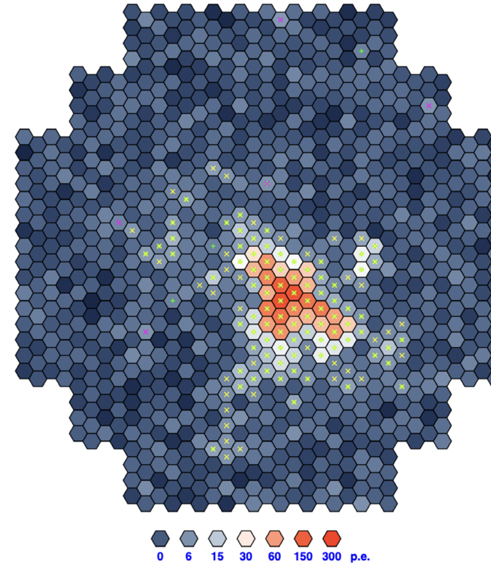
IACTs

credit: 2009 Völk et al.

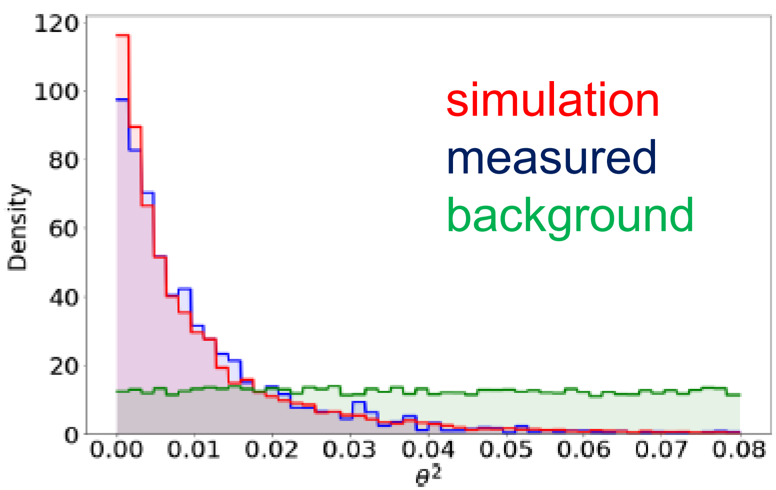
- Good discrimination between p and γ based on, i.e. different shape
- Train BDTs or NN
- Cut away most protons
- Some very “gamma-like” protons pass the cuts
- Need background estimation

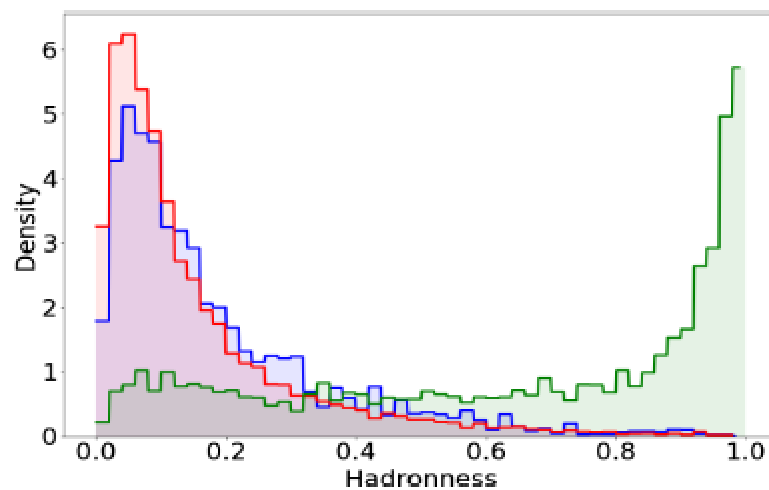


1.0 TeV gamma shower



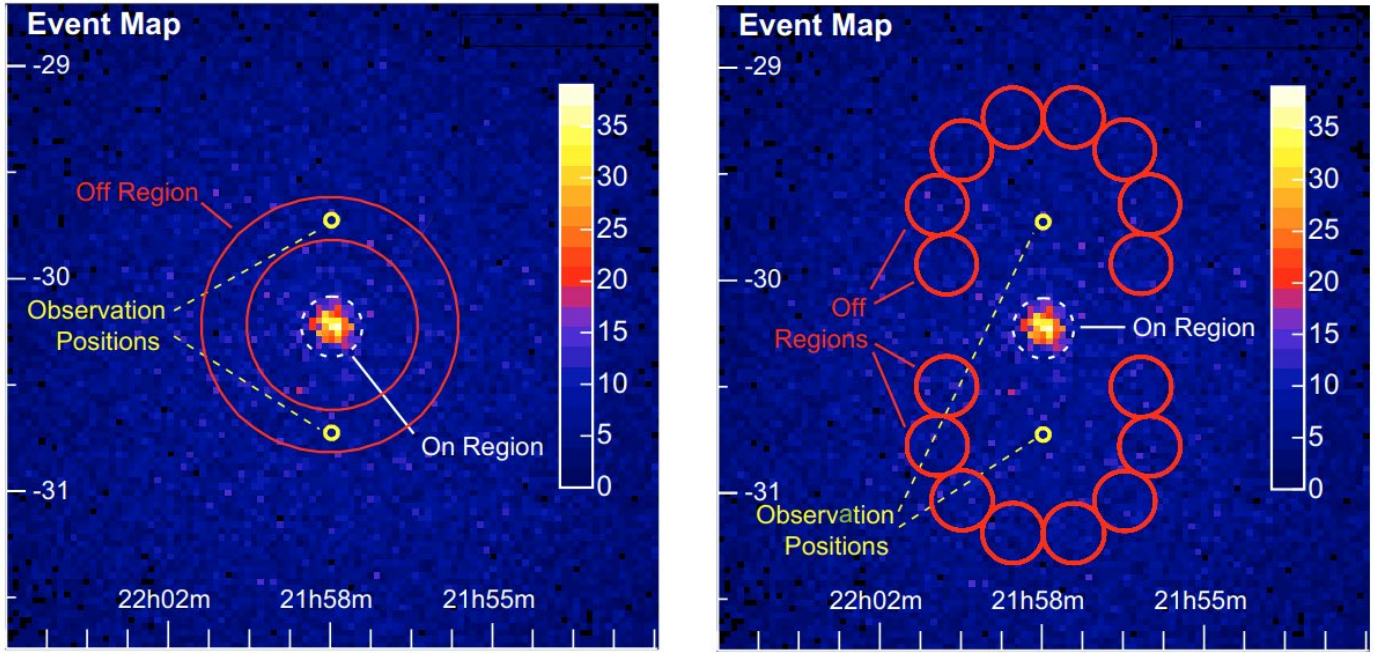
2.6 TeV proton shower





IACTs

A&A 466, 1219–1229 (2007)



Variable	Description	Property
N_{on}	Number of events in the on region	Measured
N_{off}	Number of events in the off region	Measured
α	Exposure in the on region over the one in the off regions	Measured
b	Expected rate of occurrences of background events in the off regions	Unknown
s	Expected rate of occurrences of signal events in the on region	Unknown
N_s	Number of signal events in the on region	Unknown

Analysis

Variable	Description	Property
N_{on}	Number of events in the on region	Measured
N_{off}	Number of events in the off region	Measured
α	Exposure in the on region over the one in the off regions	Measured
b	Expected rate of occurrences of background events in the off regions	Unknown
s	Expected rate of occurrences of signal events in the on region	Unknown
N_s	Number of signal events in the on region	Unknown

Poisson distribution[2]

$$p(N_s|s) = \frac{(\xi \cdot s)^{N_s} e^{-\xi \cdot s}}{N_s!}, \quad s \geq 0,$$

where

$$\xi = t_{\mathrm{eff}} \cdot A_{\mathrm{eff}}$$

Analysis

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

Variable	Description	Property
N_{on}	Number of events in the on region	Measured
N_{off}	Number of events in the off region	Measured
α	Exposure in the on region over the one in the off regions	Measured
b	Expected rate of occurrences of background events in the off regions	Unknown
s	Expected rate of occurrences of signal events in the on region	Unknown
N_s	Number of signal events in the on region	Unknown

Likelihood function:

probability of observing the on counts

probability of observing the off counts

$$p(N_{\text{on}}, N_{\text{off}} | s, b; \alpha) = p(N_{\text{on}} | s, \alpha b) \cdot p(N_{\text{off}} | b)$$

$$= \frac{(s + \alpha b)^{N_{\text{on}}}}{N_{\text{on}}!} e^{-(s+\alpha b)} \cdot \frac{b^{N_{\text{off}}}}{N_{\text{off}}!} e^{-b}, \qquad (3)$$

Analysis

$$p(N_{\mathrm{on}}, N_{\mathrm{off}} | s, b; \alpha) = p(N_{\mathrm{on}} | s, \alpha b) \cdot p(N_{\mathrm{off}} | b)$$
$$= \frac{(s + \alpha b)^{N_{\mathrm{on}}}}{N_{\mathrm{on}}!} e^{-(s+\alpha b)} \cdot \frac{b^{N_{\mathrm{off}}}}{N_{\mathrm{off}}!} e^{-b}, \qquad (3)$$

Frequentist solution: find s which maximizes the likelihood

$$s = (N_{\mathrm{on}} - \alpha N_{\mathrm{off}}) \pm \sqrt{N_{\mathrm{on}} + \alpha^2 N_{\mathrm{off}}}. \qquad (11)$$

Analysis

$$p(N_{\rm on}, N_{\rm off}|s, b; \alpha) = p(N_{\rm on}|s, \alpha b) \cdot p(N_{\rm off}|b)$$
$$= \frac{(s+\alpha b)^{N_{\rm on}}}{N_{\rm on}!} e^{-(s+\alpha b)} \cdot \frac{b^{N_{\rm off}}}{N_{\rm off}!} e^{-b}, \quad (3)$$

Bayesian solution: integrate over the nuisance parameter b

$$p(s|N_{\rm on}, N_{\rm off}; \alpha) \propto \sum_{N_s=0}^{N_{\rm on}} \frac{(N_{\rm on}+N_{\rm off}-N_s)!}{(1+1/\alpha)^{-N_s}(N_{\rm on}-N_s)!} \cdot \frac{s^{N_s}}{N_s!} e^{-s}. \quad (5)$$

$$p(N_s|N_{\rm on}, N_{\rm off}; \alpha) \propto \frac{(N_{\rm on}+N_{\rm off}-N_s)!}{(1+1/\alpha)^{-N_s}(N_{\rm on}-N_s)!}. \quad (8)$$

BASiL

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

- Advantage of Bayesian approach: Don't cut BUT
  Include single-event variables (likelihood of the event being a hadron)
- Previously

$$p(N_{\rm on}, N_{\rm off}|s, b; \alpha) = p(N_{\rm on}|s, \alpha b) \cdot p(N_{\rm off}|b)$$

- Now $p(\vec{\mathbf{x}}, N_{\rm on}, N_{\rm off}|s, b; \alpha)$

$$= p(\vec{\mathbf{x}}|N_{\rm on}, s, \alpha b) \cdot p(N_{\rm on}|s, \alpha b) \cdot p(N_{\rm off}|b). \quad (15)$$

$$p(\vec{\mathbf{x}}|N_{\rm on}, s, \alpha b) = \prod_{i=1}^{N_{\rm on}} [p(\mathbf{x}_i|\gamma) \cdot p(\gamma|s, \alpha b) + p(\mathbf{x}_i|\bar{\gamma}) \cdot p(\bar{\gamma}|s, \alpha b)], \quad (16)$$

prior probabilities of being gamma or bkg

event i being gamma

event i being background

previously

$$p(N_s|N_{\text{on}},N_{\text{off}};\alpha) \propto \frac{(N_{\text{on}}+N_{\text{off}}-N_s)!}{(1+1/\alpha)^{-N_s}(N_{\text{on}}-N_s)!}. \quad (8)$$

now:

$$p(N_s|\vec{\mathbf{x}},N_{\text{on}},N_{\text{off}};\alpha) \propto \frac{(N_{\text{on}}+N_{\text{off}}-N_s)!}{(N_{\text{on}}-N_s)!(1+1/\alpha)^{-N_s}} \frac{C(\vec{\mathbf{x}},N_s)}{\binom{N_{\text{on}}}{N_s}}. \quad (22)$$

$$C(\vec{\mathbf{x}},N_s) = \sum_{A\in F_{N_s}} \prod_{i\in A} p(\mathbf{x}_i|\gamma) \cdot \prod_{j\in A^c} p(\mathbf{x}_j|\bar{\gamma})$$

BASiL

$$C(\vec{\mathbf{x}}, N_s) = \sum_{A \in F_{N_s}} \prod_{i \in A} p(\mathbf{x}_i|\gamma) \cdot \prod_{j \in A^c} p(\mathbf{x}_j|\bar{\gamma})$$

Let us assume $N_{\mathrm{on}} = 3$ events in our on region and that we have also measured x_1, x_2, x_3 respectively for each event, with x a variable whose distribution is $p(x|\gamma)$ for a signal population and $p(x|\bar{\gamma})$ for a background population. Thus, when $N_s = 0, 1, 2, 3$ the combinatorial term will be respectively[12]

case of all bkg →

$$C(\vec{x}, 0) = p(x_1|\bar{\gamma}) \cdot p(x_2|\bar{\gamma}) \cdot p(x_3|\bar{\gamma}),$$

case of 1 signal →

$$\begin{aligned} C(\vec{x}, 1) = {} & p(x_1|\gamma) \cdot p(x_2|\bar{\gamma}) \cdot p(x_3|\bar{\gamma}) \\ & + p(x_1|\bar{\gamma}) \cdot p(x_2|\gamma) \cdot p(x_3|\bar{\gamma}) \\ & + p(x_1|\bar{\gamma}) \cdot p(x_2|\bar{\gamma}) \cdot p(x_3|\gamma), \end{aligned}$$

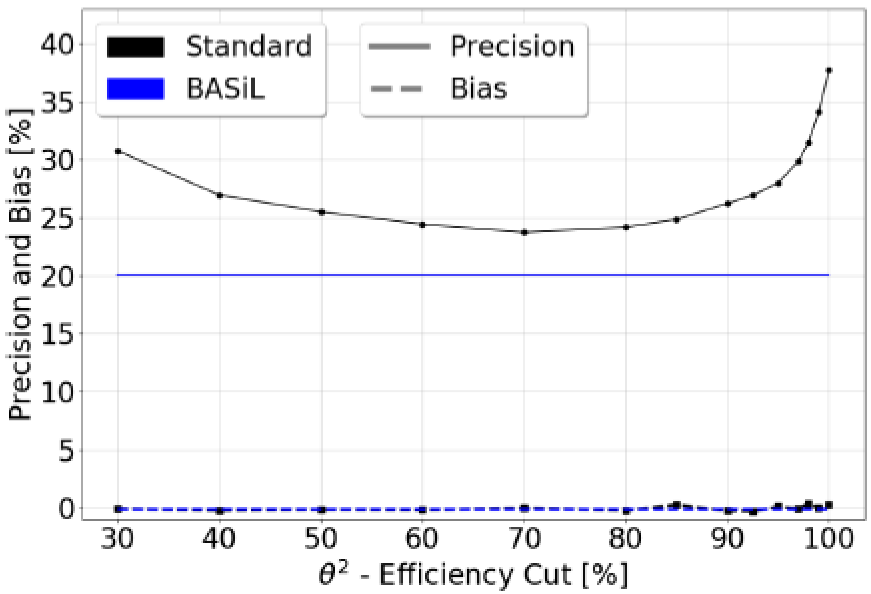
case of 2 signal →

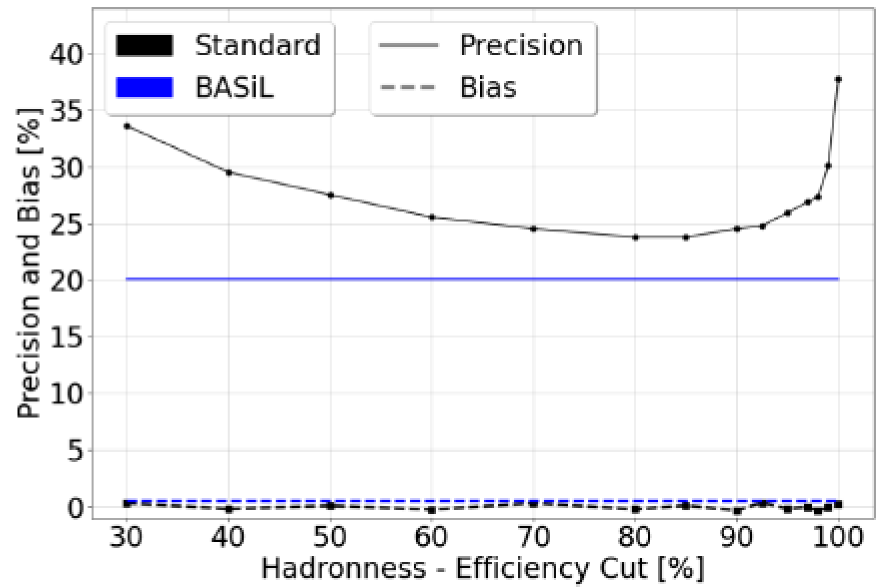
$$\begin{aligned} C(\vec{x}, 2) = {} & p(x_1|\gamma) \cdot p(x_2|\gamma) \cdot p(x_3|\bar{\gamma}) \\ & + p(x_1|\gamma) \cdot p(x_2|\bar{\gamma}) \cdot p(x_3|\gamma) \\ & + p(x_1|\bar{\gamma}) \cdot p(x_2|\gamma) \cdot p(x_3|\gamma), \end{aligned}$$

case of all signal →

$$C(\vec{x}, 3) = p(x_1|\gamma) \cdot p(x_2|\gamma) \cdot p(x_3|\gamma).$$

Results





- BASiL is performing better than the best cut

Results

- Improvement is dependent on the signal to noise ratio (SNR)
- Most benefits in the low SNR case where we are statistics limited