Statistics in X-ray Polarimetry

Herman L. Marshall (MIT)

Background (Marshall 2020, in press)

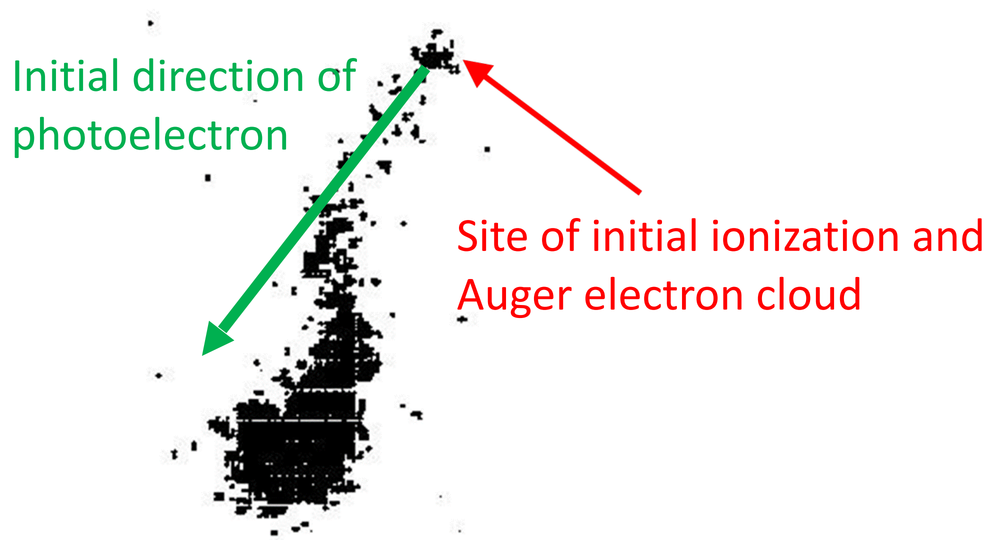
- Expected counts in $dEd\psi$ in time *T* for $Q = qI, U = uI$:
  - $\lambda(E, \psi; n_0, q, u)dEd\psi = [1 + \mu_E(q\cos 2\psi + u\sin 2\psi)]n_E A_E T dEd\psi$, where
    - $A_E = A\alpha(E)$ is the instrument effective area (independent of q or u by definition)
    - $n_E = n_0\phi(E)$ has units of ph/cm^2/s/keV per unit (measured) phase angle, ψ
  - Require $\Pi^2 \equiv q^2 + u^2 \leq 1$ physically (Π is fractional linear polarization)
  - Define $\phi_0 = \tan^{-1}(u/q) = 2\varphi$
- Modulated: $\mu_E(q\cos 2\psi + u\sin 2\psi)]n_E A_E T dEd\psi = \mu_E \Pi\cos(2\psi - 2\varphi)n_E A_E T dEd\psi = C(\psi)$
- Extrema of counts are $\lambda_{\max} = (1 + \mu_E\Pi)n_E A_E T dEd\psi, \quad \lambda_{\min} = (1 - \mu_E\Pi)n_E A_E T dEd\psi$
- Thus $\mu_E \equiv \dfrac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}$ for Π=1
- Likelihood:

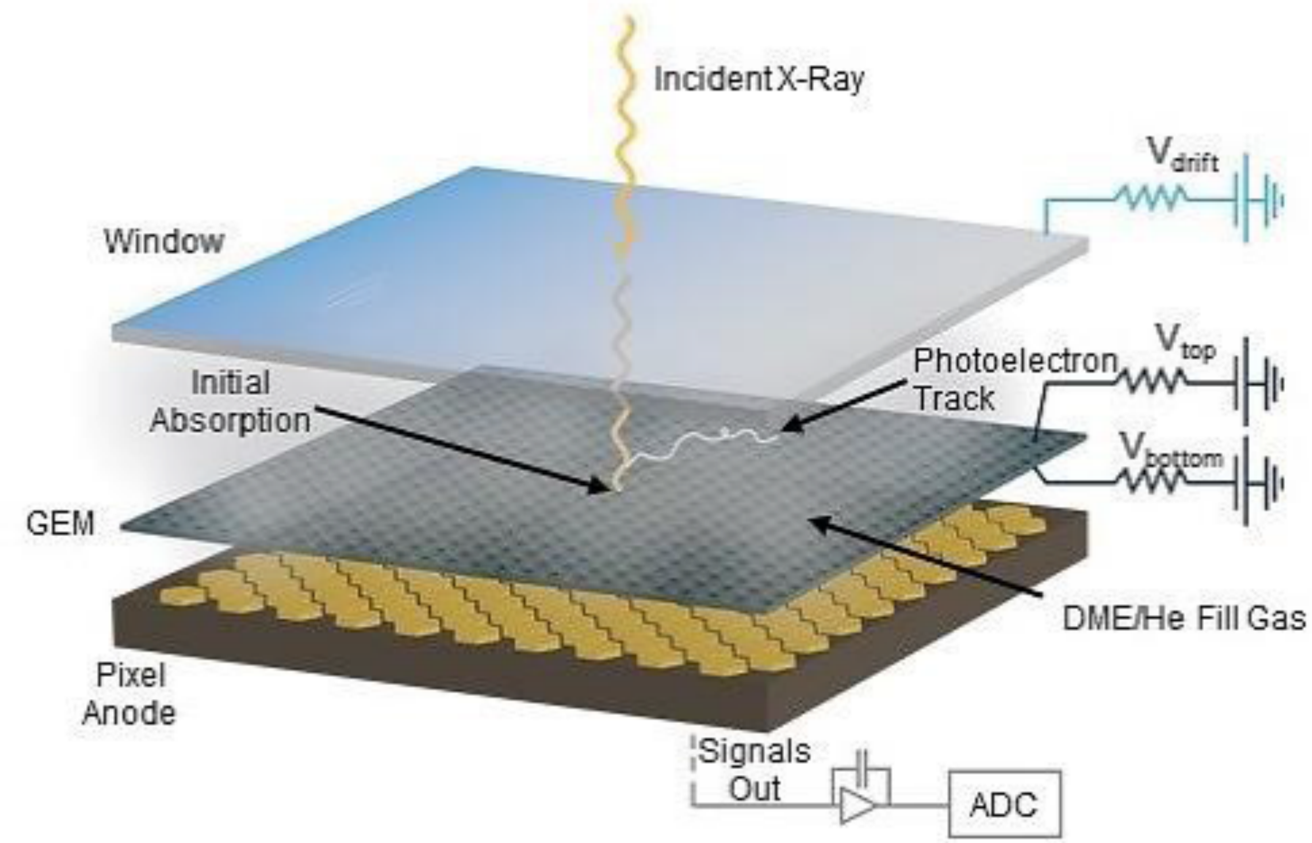
  $S(n_0, q, u) = -2\ln\mathscr{L} = -2\sum_i \ln\lambda(E_i, \psi_i) + 2T\int f_E A_E dE \int_0^{2\pi}[1 + \mu(E)(q\cos 2\psi + u\sin 2\psi)]d\psi$

  or $\tilde{S}(q, u) = -2\sum_i \ln(1 + q\mu_i\cos 2\psi_i + u\mu_i\sin 2\psi_i)$
- MDP$_{99}$ $= 4.29/\sqrt{\sum \mu_E^2 C(E)}$ for small $\Pi\mu$

Imaging Polarimetry Detector

- Photons ionize atoms in the detector gas
- Direction of the photoelectron is related to the photon's polarization angle
- The electron loses energy through the gas; charge is proportional to E
- Charge drifts to anode





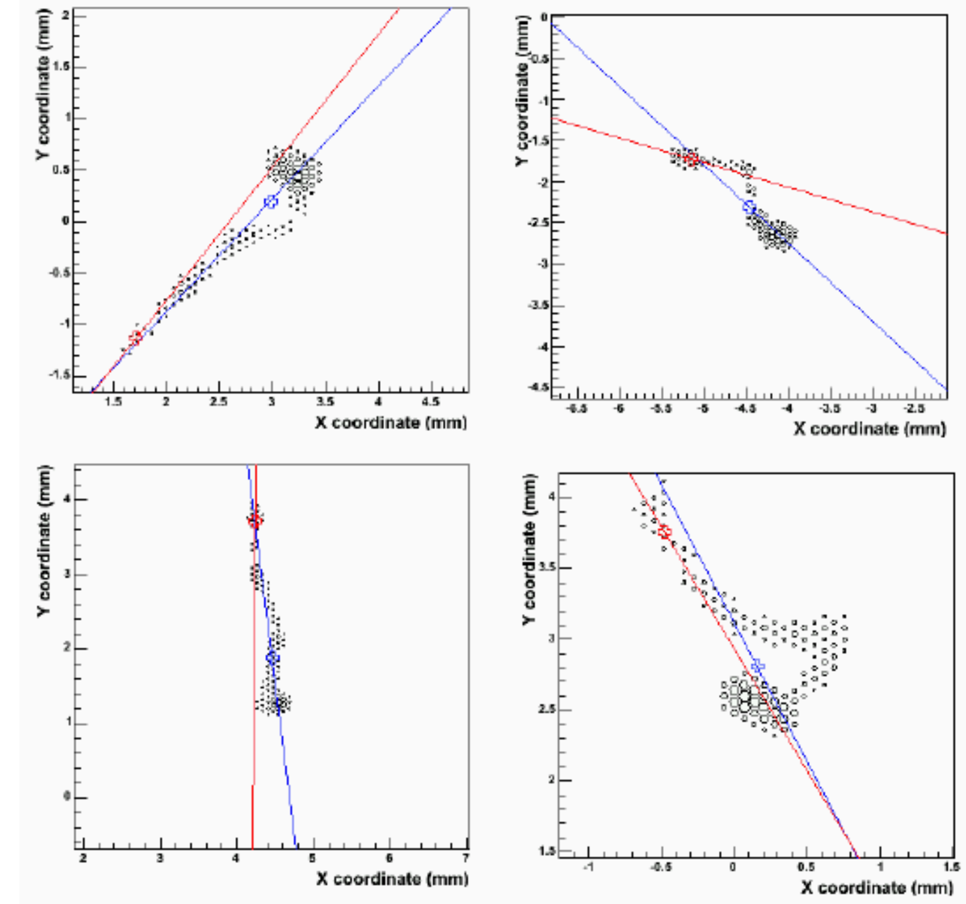
$$\frac{d\sigma}{d\Omega} = f(\zeta)\, r_0^2 Z^5 \alpha_0^4 \left(\frac{1}{\beta}\right)^{7/2} 4\sqrt{2}\,\sin^2\theta\,\cos^2\varphi \;, \text{ where } \beta \equiv \frac{E}{mc^2} = \frac{h\nu}{mc^2}$$

Measuring IXPE Event Tracks

- Current photoelectron track measurement: moments based

- How to do better — full track modeling?

Bellazzini+ '06



Bellazzini+ '06

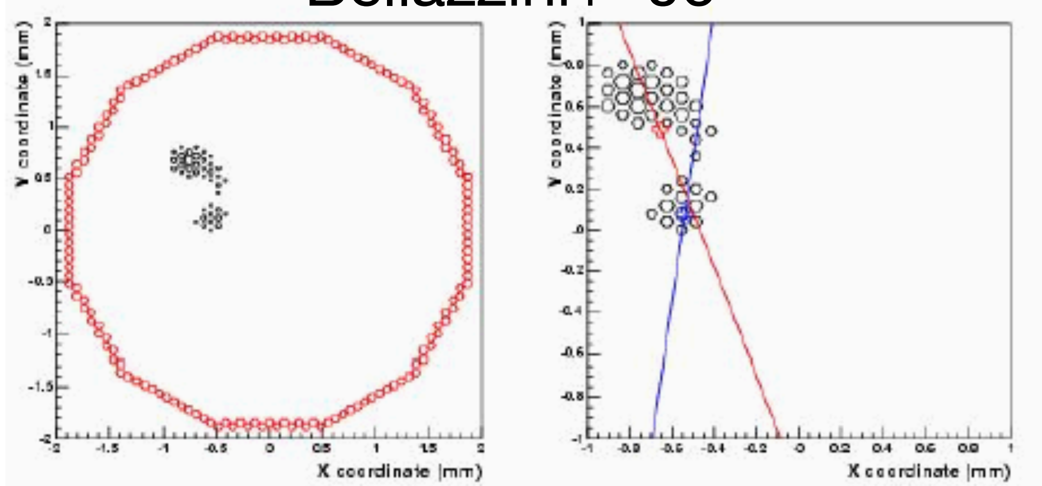
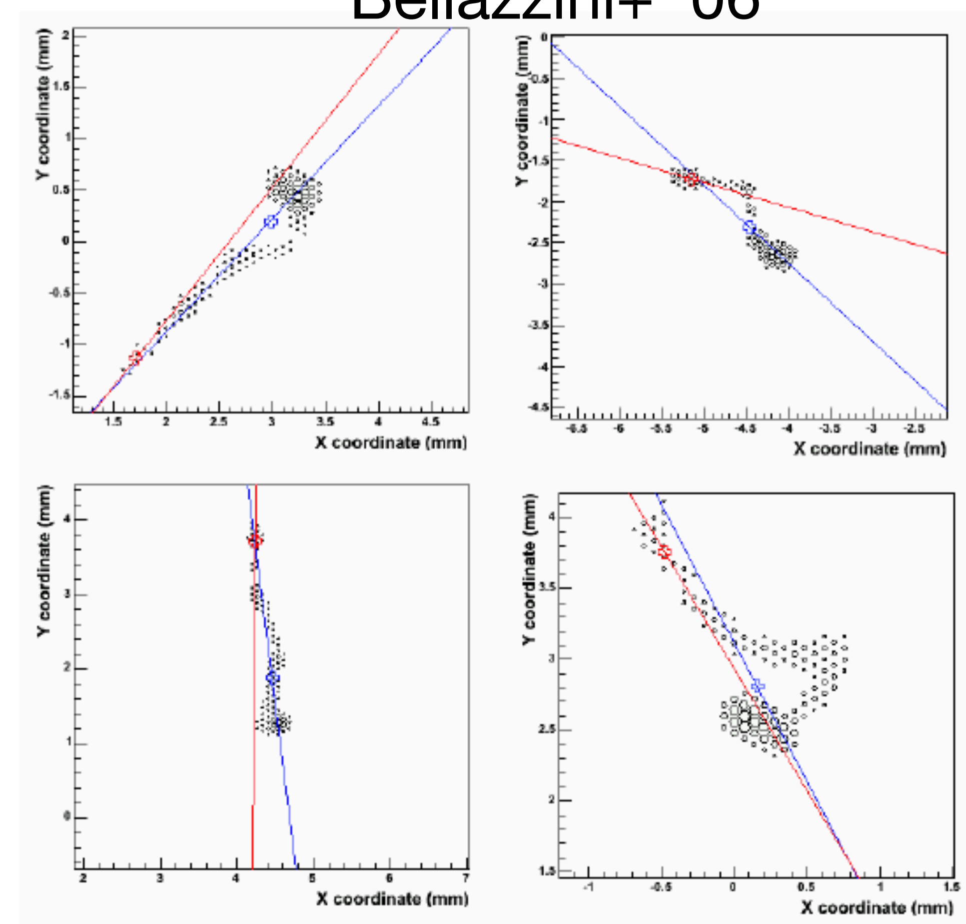


Fig.2. Real track produced in the gas by a 5.9 keV photon. The reconstruction algorithm develops in the following steps: 1) barycenter evaluation of the charge distribution (red cross), 2) reconstruction of the principal axis direction (red line), 3) conversion point evaluation (blue cross), 4) emission direction reconstruction (blue line). The polarization is derived from the photoelectrons angular distribution.

Measuring IXPE Event Tracks

- Current photoelectron track measurement: moments based
- How to do better — full track modeling?

Bellazzini+ '06



Bellazzini+ '06

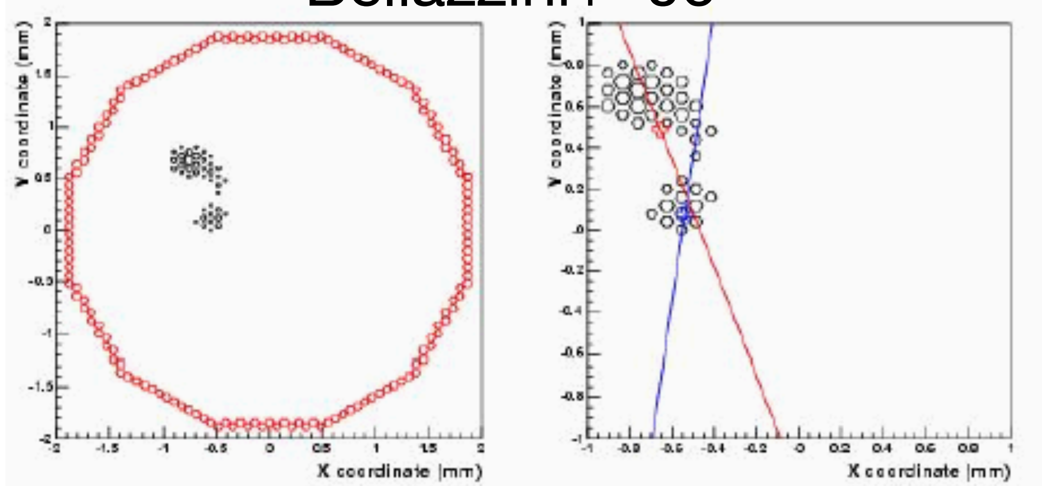
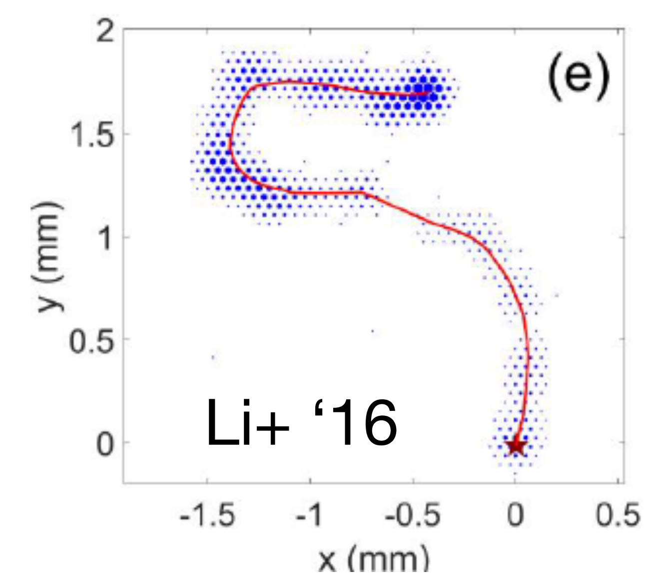


Fig.2. Real track produced in the gas by a 5.9 keV photon. The reconstruction algorithm develops in the following steps: 1) barycenter evaluation of the charge distribution (red cross), 2) reconstruction of the principal axis direction (red line), 3) conversion point evaluation (blue cross), 4) emission direction reconstruction (blue line). The polarization is derived from the photoelectrons angular distribution.

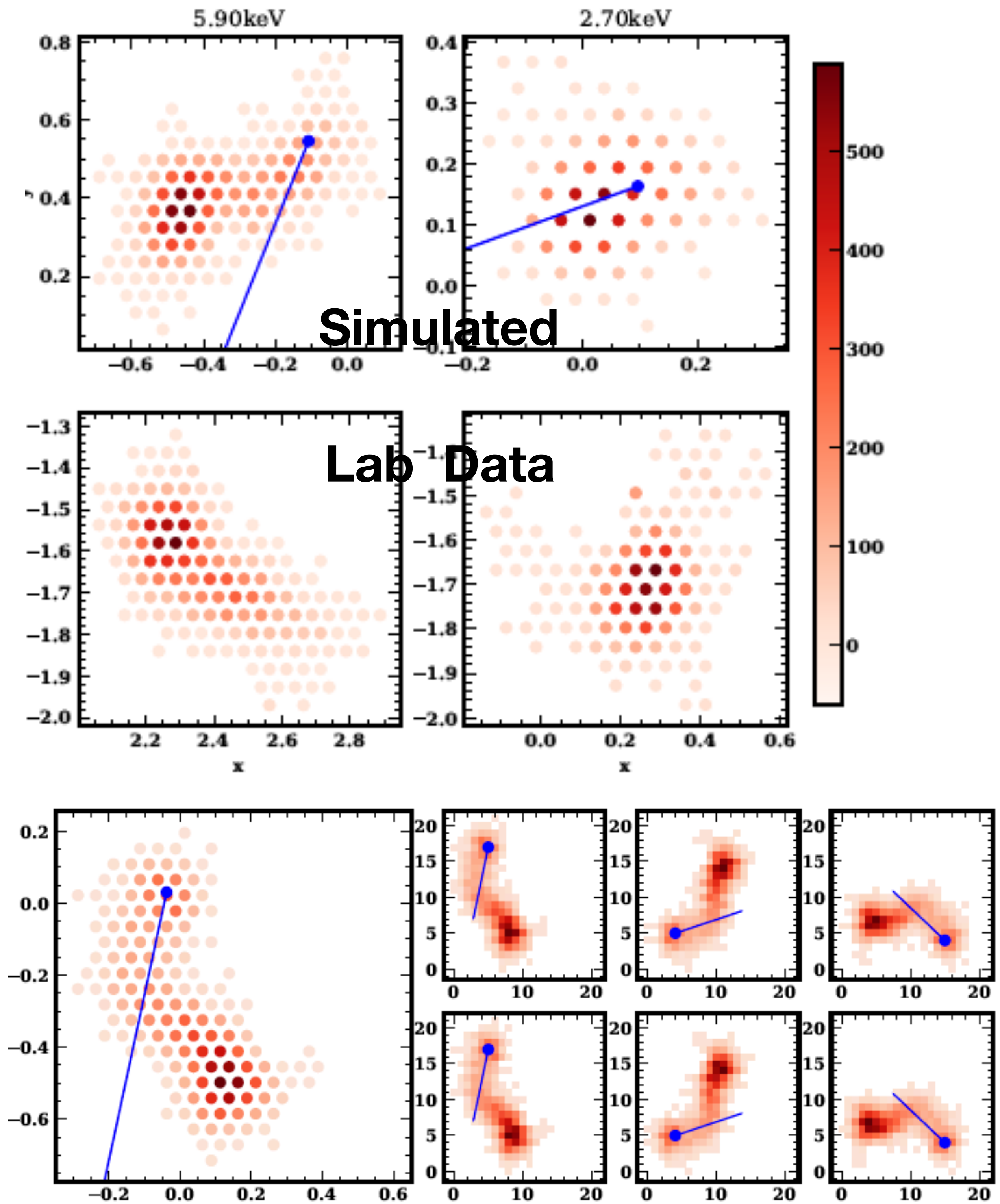


A Neural Network Approach

(Peirson+ 2020, in press)

- Convolutional Neural Net (CNN): N events, M networks
- Train to minimize angle errors on simulated data
- Estimate uncertainties in track angles
- Validate on additional simulated data
- Optimize using lab data
- Optimize for best nets using “importance-weighted” likelihood

$$\begin{aligned} \underset{\text{over}\,\mu,\phi}{\text{minimize}} \quad & -\sum_{j=1}^{M}\sum_{i=1}^{N}\sigma_{ij}^{-\lambda}\log\left(1+\mu\cos\left(2(\theta_{ij}-\phi)\right)\right) \\ \text{subject to} \quad & 0 \le \mu \le 1 \\ & -\pi/2 \le \phi < \pi/2, \end{aligned}$$



Improvement using NNs

- MDP drops as method gets more sophistocated
- NN with weights gives 20% smaller MDPs
- Equivalent to 40% more instrument area!
- Also gives better location of initial interaction point and estimate of event energy!

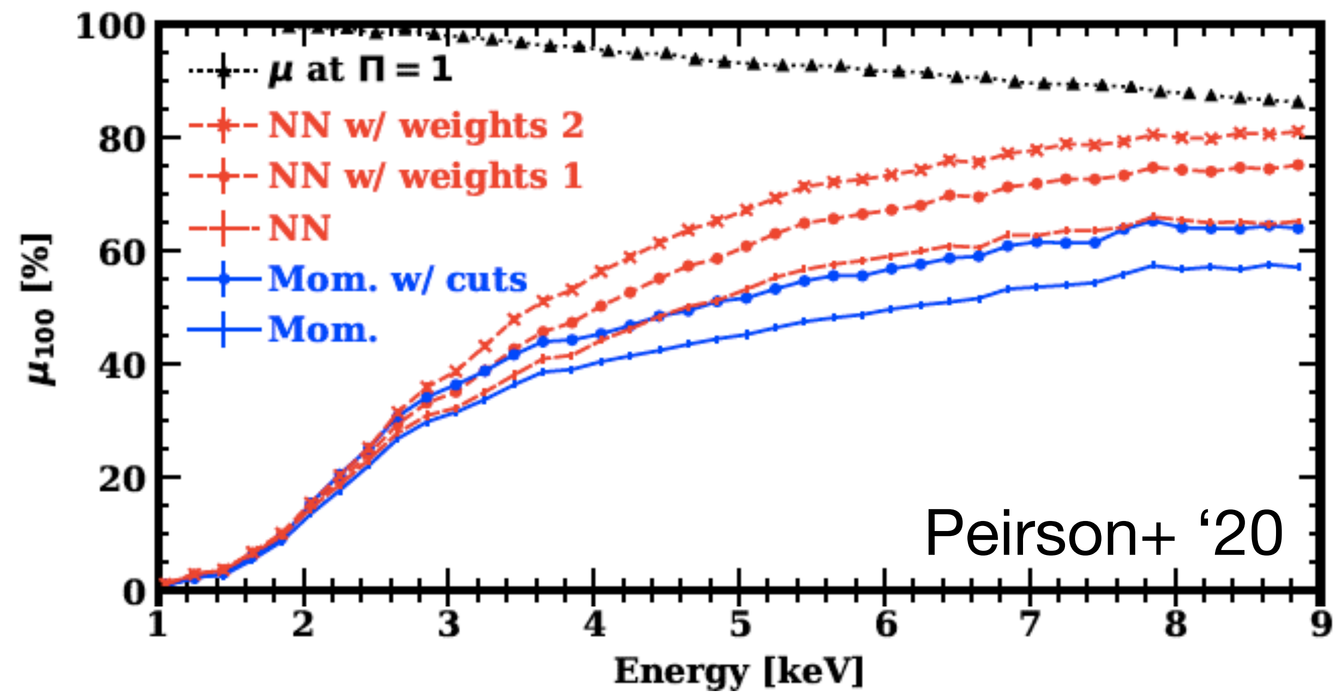
Peirson+ ‘20

Peirson+ ‘20

Method	μ_{100}(%)	$N_{\rm eff}/N$	λ	MDP$_{99}$(%)
Mom.	27.0	1.0	0	5.03 ± 0.02
Mom. w/ cut	31.3	0.80	–	4.88 ± 0.03
Mom. w/ weights	31.4	0.88	0.67	4.61 ± 0.01 ←
NN	28.7	1.0	0	4.72 ± 0.02
NN w/ weights	32.6	0.95	1	4.26 ± 0.02
NN w/ weights	36.8	0.81	1.83	4.09 ± 0.02 ←
NN w/ weights (bootstrap)	36.1	0.85	1.83	4.07 ± 0.02 ←
NN w/ weights	37.7	0.76	2	4.12 ± 0.02

Improvement using NNs

- MDP drops as method gets more sophistocated
- NN with weights gives 20% smaller MDPs
- Equivalent to 40% more instrument area!
- Also gives better location of initial interaction point and estimate of event energy!



Peirson+ '20

Method	μ_{100}(%)	N_{eff}/N	λ	MDP$_{99}$(%)
Mom.	27.0	1.0	0	5.03 ± 0.02
Mom. w/ cut	31.3	0.80	–	4.88 ± 0.03
Mom. w/ weights	31.4	0.88	0.67	4.61 ± 0.01 ←
NN	28.7	1.0	0	4.72 ± 0.02
NN w/ weights	32.6	0.95	1	4.26 ± 0.02
NN w/ weights	36.8	0.81	1.83	4.09 ± 0.02 ←
NN w/ weights (bootstrap)	36.1	0.85	1.83	4.07 ± 0.02 ←
NN w/ weights	37.7	0.76	2	4.12 ± 0.02

Why is $\mu_E \neq 1$?

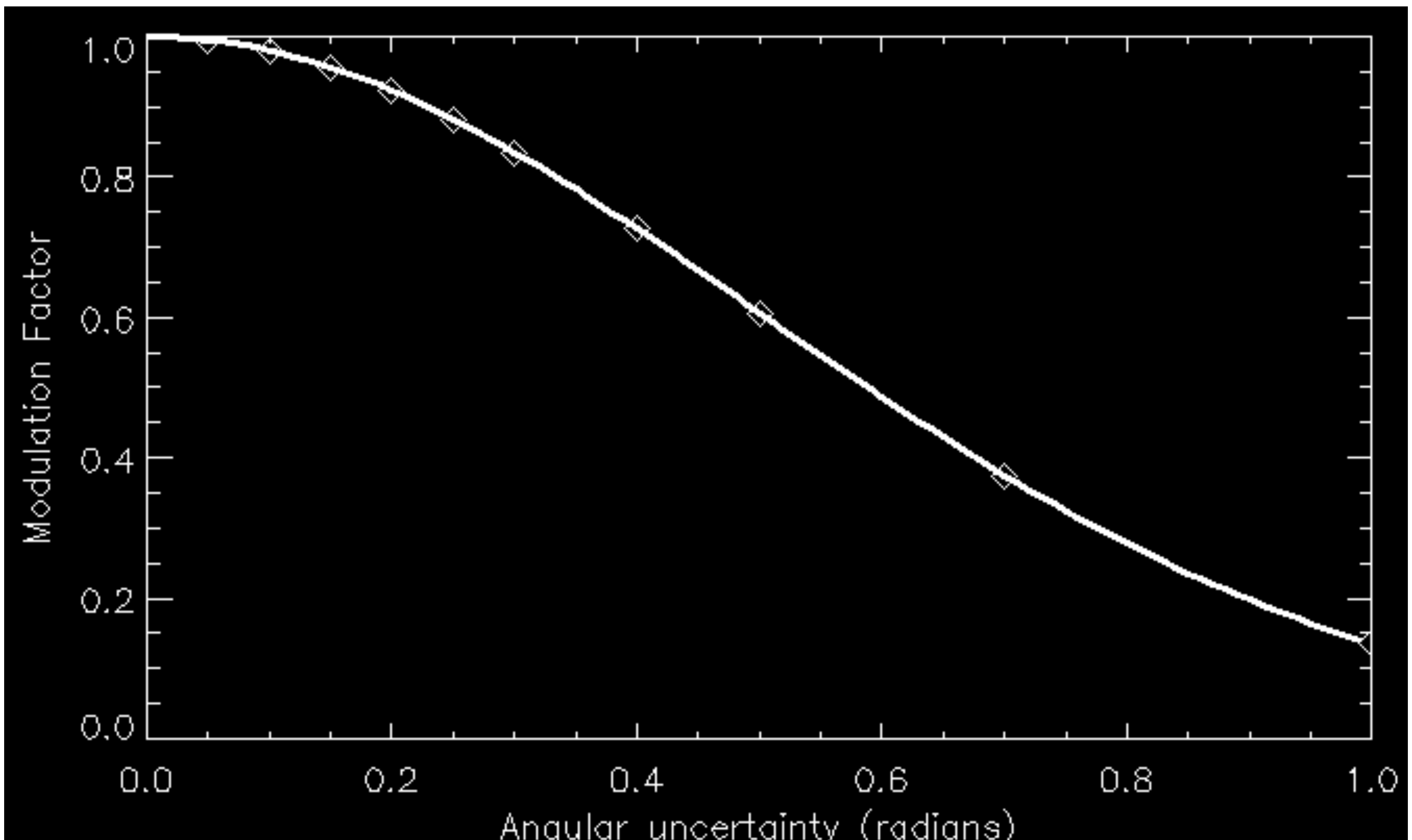
- The modulation factor reflects instrument reality
- For IXPE, the true photoelectron azimuth ψ', is imperfectly known: $\psi \neq \psi'$
  - If perfect, then $p(\psi) = p(\psi') \propto \cos 2(\psi' - \varphi)$ for 100% polarized source
- Hypothesis: $\mu < 1$ due to uncertainty in ψ
  - Assume Gaussian error distribution: $\psi \sim N(\psi', \sigma)$
  - Modulated counts are

$$C(\psi) = 1 + \int G(\psi; \psi', \sigma) \cos 2(\psi' - \varphi) \; d\psi' = 1 + \mu \cos 2(\psi - \varphi),$$

    where $\mu = [C(\varphi) - C(\varphi + \pi/2)]/[C(\varphi) + C(\varphi + \pi/2)] = e^{-2\sigma^2}$
  - Actual integration over $-m\pi \leq \psi' \leq m\pi$, for $m\pi \gg \sigma$
  - Modulation factor depends only on angle measurement uncertainty
  - Thus, $\mu(E)$ **encapsulates** Gaussian **measurement uncertainties**

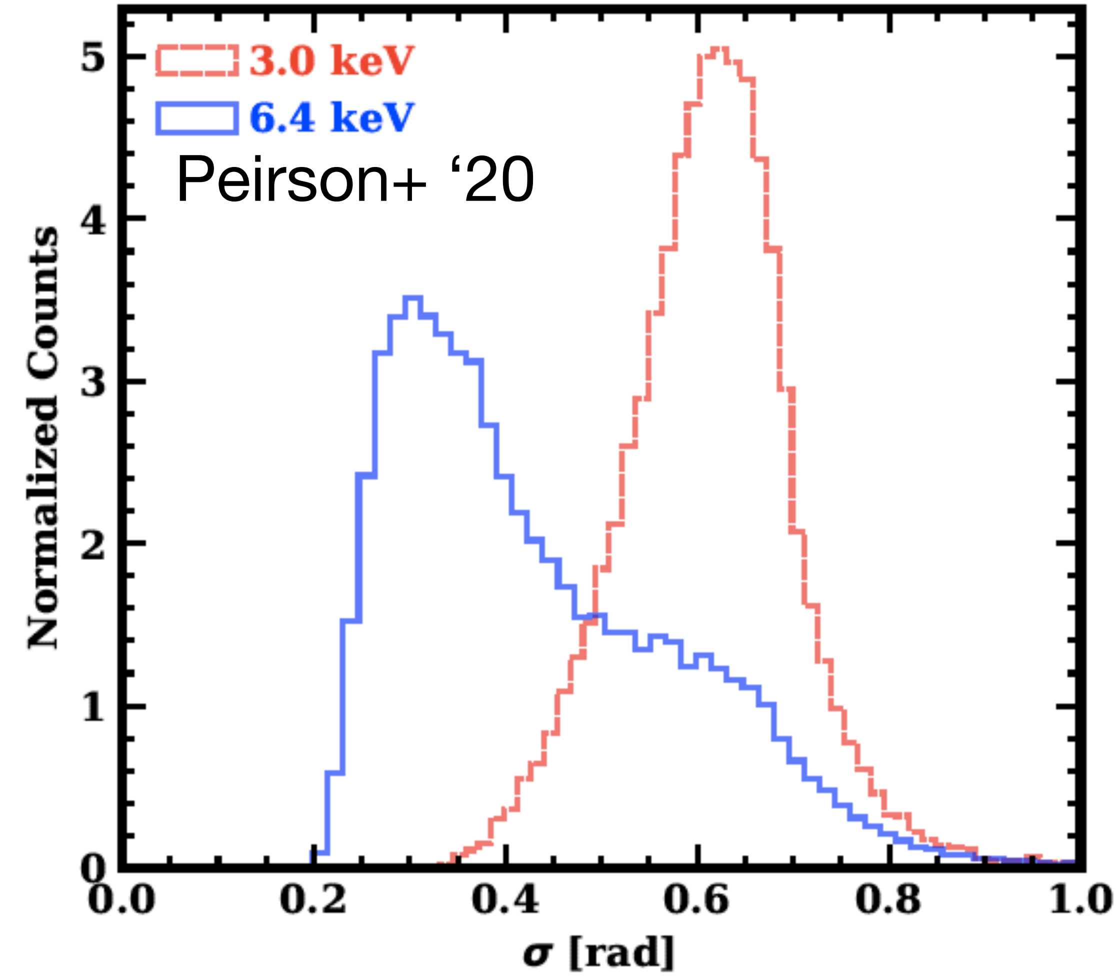
Modulation Drops with Blurring

- Computed convolution, obtaining $\lambda(\psi;\sigma)$
- Determined $\mu = (C_{\max} - C_{\min})/C_{\max} + C_{\min}) = e^{-2\sigma^2}$
- Individual track uncertainties derivable from CNN track analysis



Using the CNN Uncertainties

- Track σ distribution is not a delta function and not even unimodal



- Generally:

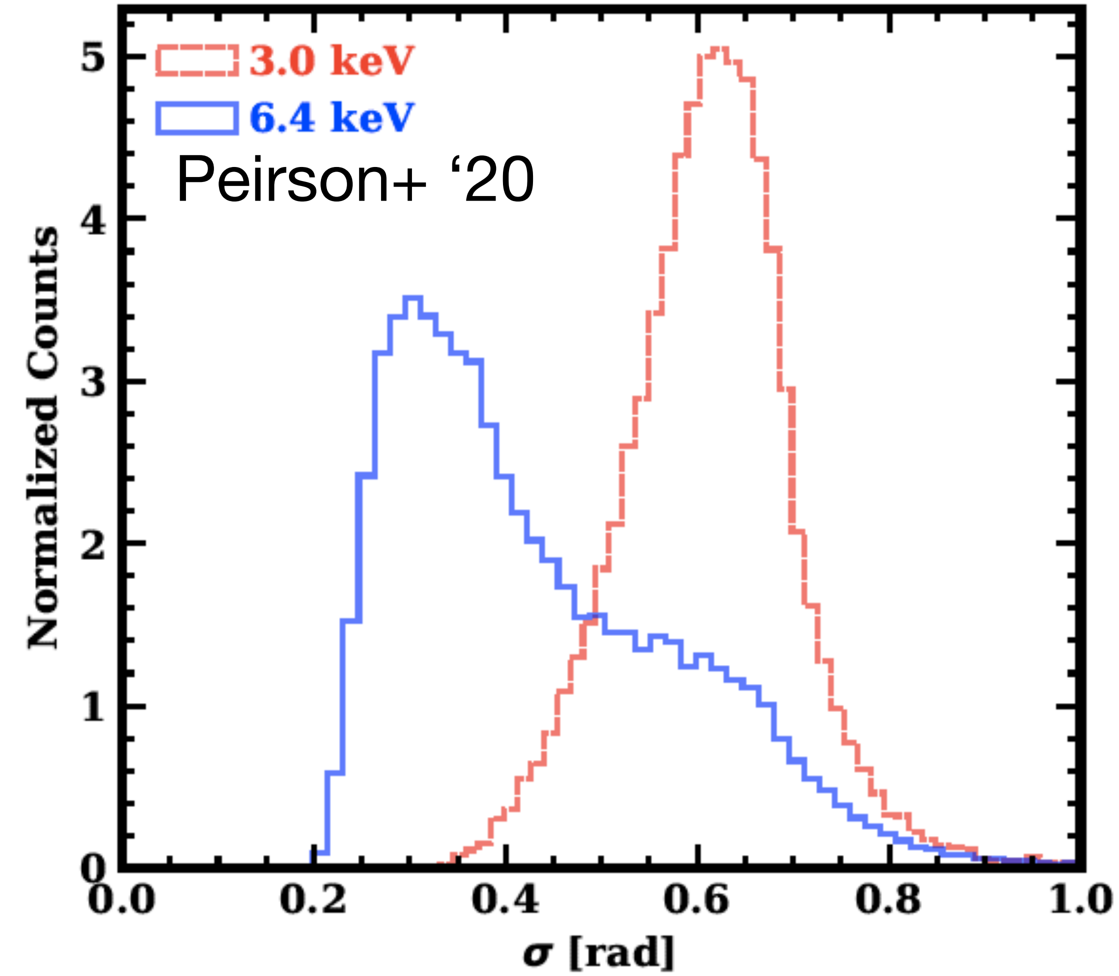
$$\mu = \int p(\sigma)\mu(\sigma)d\sigma = \int p(\sigma)e^{-2\sigma^2}d\sigma$$

- Simple case: a bimodal distribution of σ, single *E*
  - fraction f with $\sigma \ll 1$, $\mu_1 = 1$, $1-f$ have large σ, $\mu_2 = 0 \longrightarrow \mu = f$
  - fraction f have $\mu = \mu_1$, $1-f$ have large σ, $\mu_2 = 0 \longrightarrow \mu = f\mu_1$

- Summary: can predict μ(E) from CNN p(σ_E)

Using the CNN Uncertainties

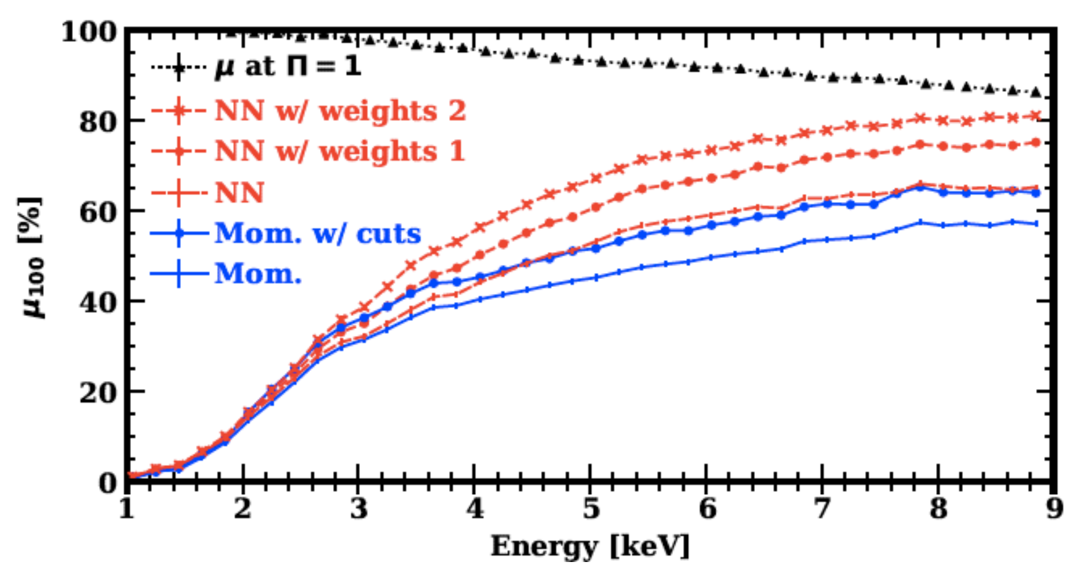
- Track σ distribution is not a delta function and not even unimodal



- Generally:

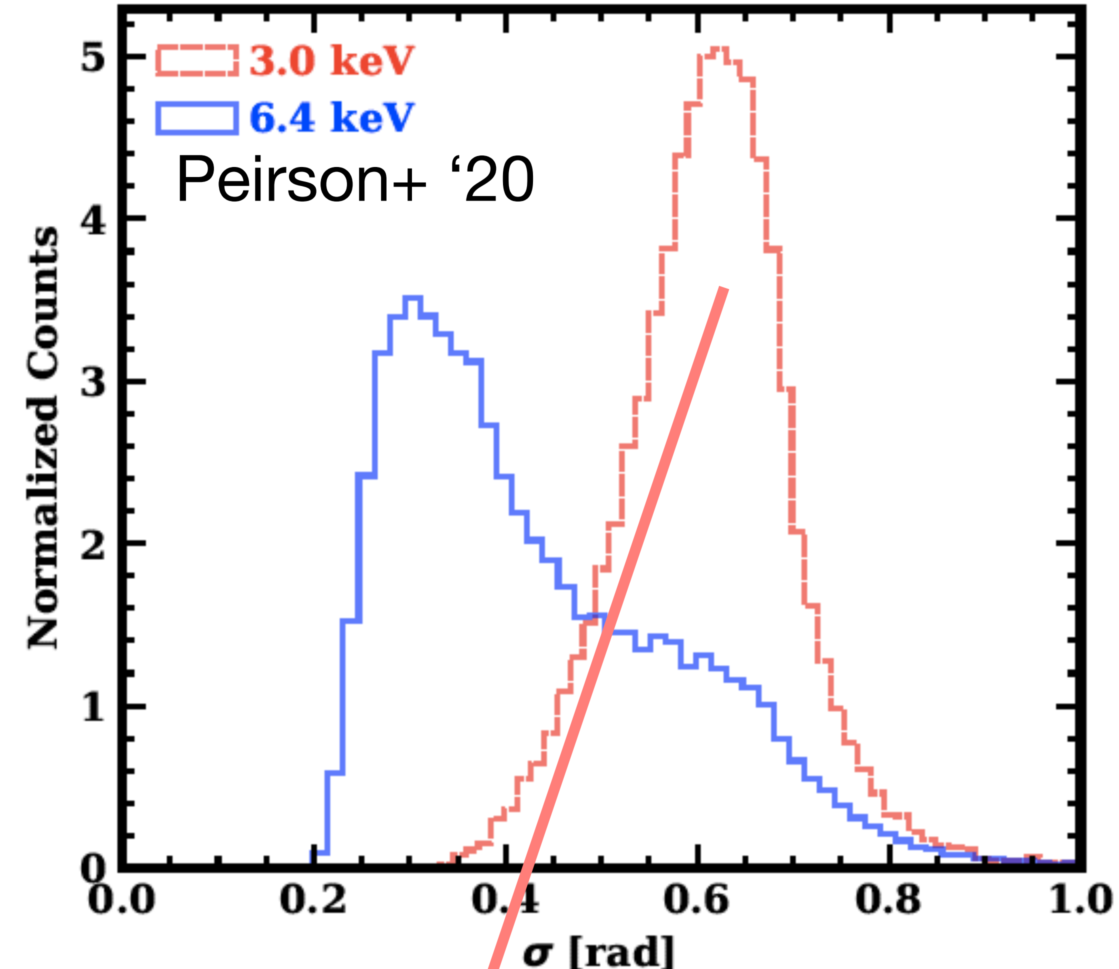
$$\mu = \int p(\sigma)\mu(\sigma)d\sigma = \int p(\sigma)e^{-2\sigma^2}d\sigma$$

- Simple case: a bimodal distribution of σ, single *E*
  - fraction f with $\sigma \ll 1$, $\mu_1 = 1$, $1 - f$ have large σ, $\mu_2 = 0$ –> $\mu = f$
  - fraction f have $\mu = \mu_1$, $1 - f$ have large σ, $\mu_2 = 0$ –> $\mu = f\mu_1$
- Summary: can predict μ(E) from CNN p(σ_E)



Using the CNN Uncertainties

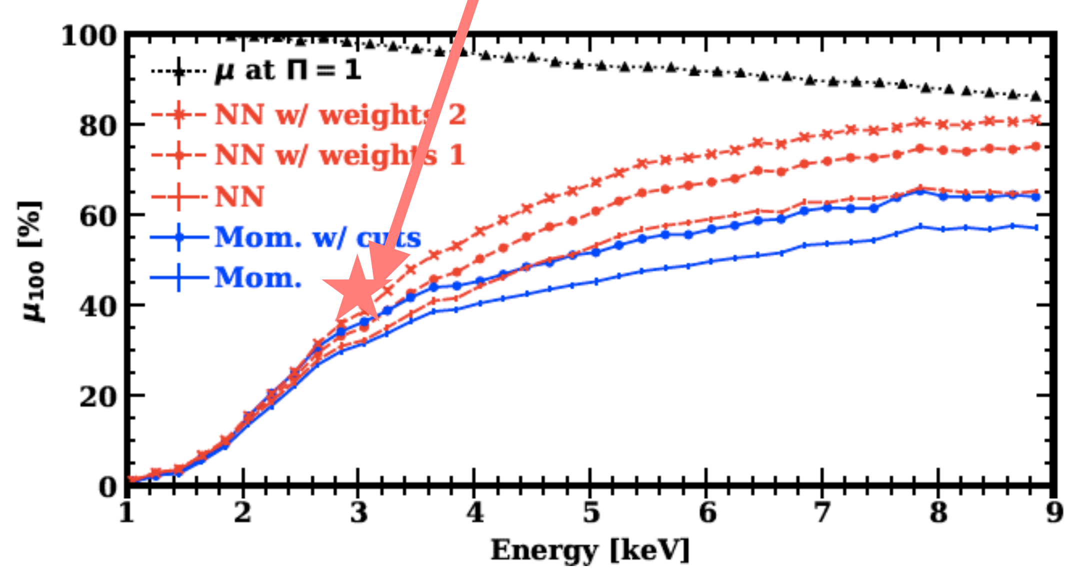
- Track σ distribution is not a delta function and not even unimodal



- Generally:

$$\mu = \int p(\sigma)\mu(\sigma)d\sigma = \int p(\sigma)e^{-2\sigma^2}d\sigma$$

- Simple case: a bimodal distribution of σ, single *E*
  - fraction f with $\sigma \ll 1$, $\mu_1 = 1$, $1-f$ have large σ, $\mu_2 = 0 \;{-}{>}\; \mu = f$
  - fraction f have $\mu = \mu_1$, $1-f$ have large σ, $\mu_2 = 0 \;{-}{>}\; \mu = f\mu_1$



- Summary: can predict μ(E) from CNN $p(\sigma_E)$

Using the CNN Uncertainties

- Track σ distribution is not a delta function and not even unimodal

- Generally:

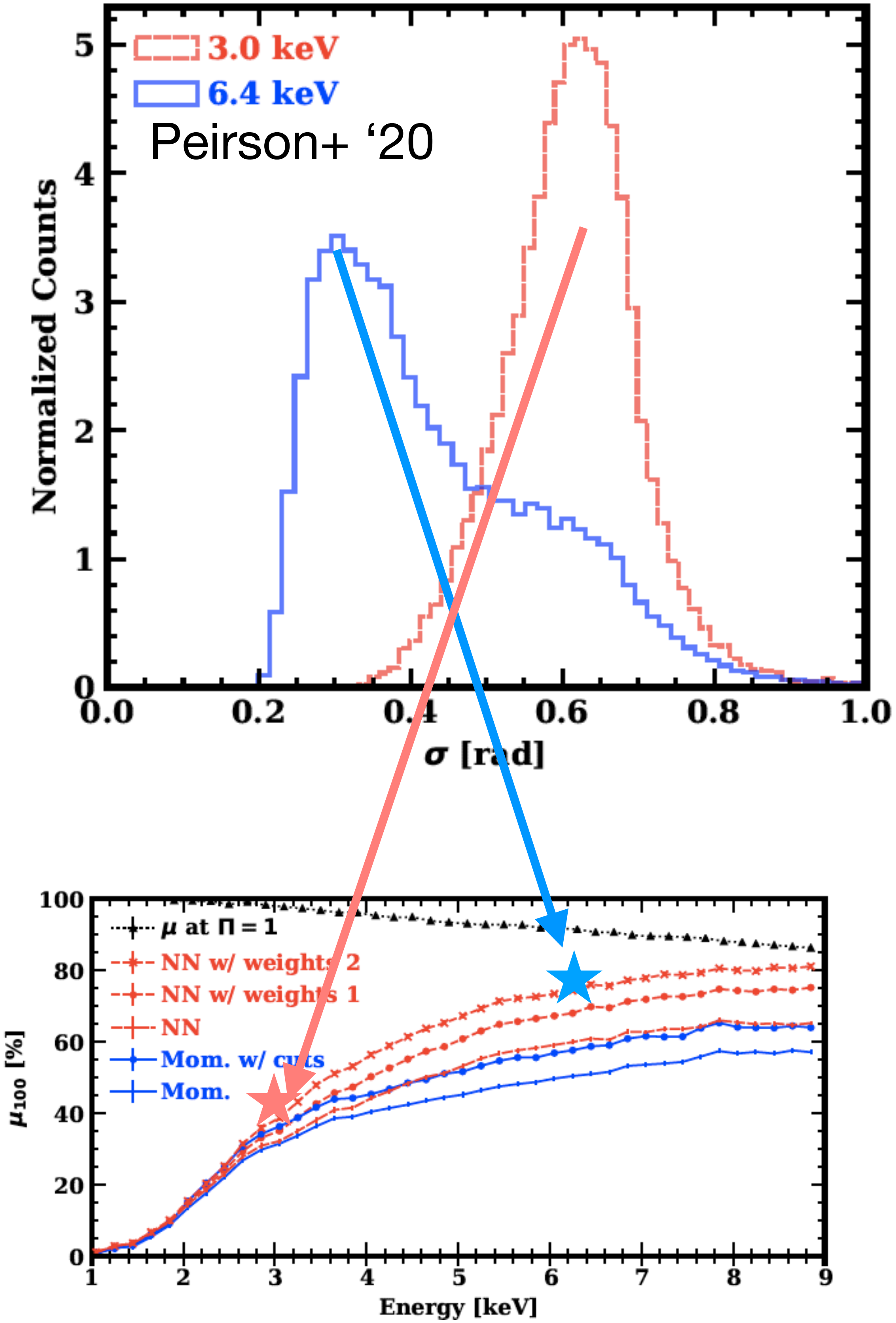
$$\mu = \int p(\sigma)\mu(\sigma)d\sigma = \int p(\sigma)e^{-2\sigma^2}d\sigma$$

- Simple case: a bimodal distribution of σ, single *E*

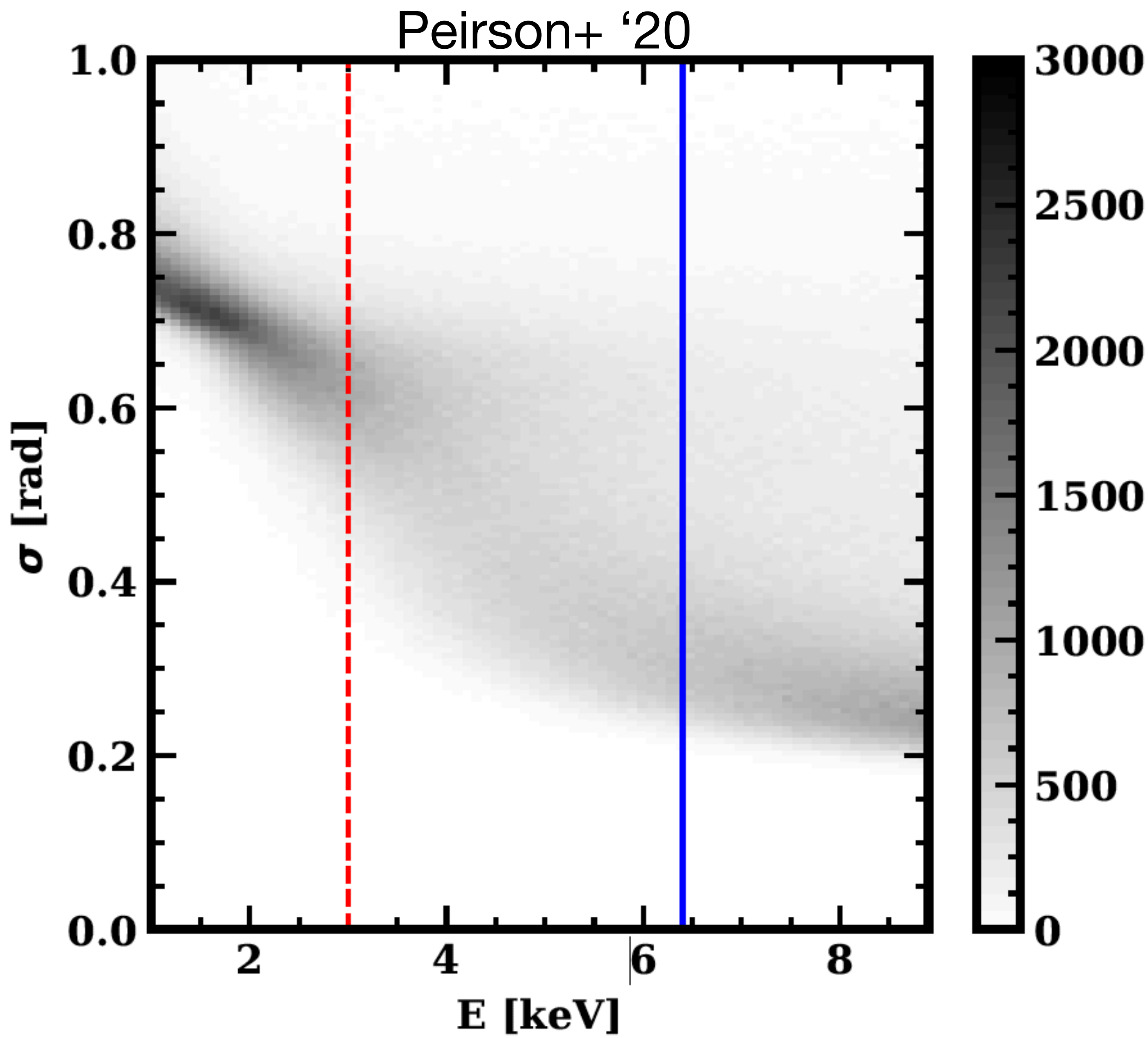
  - fraction f with $\sigma \ll 1$, $\mu_1 = 1$, $1-f$ have large σ, $\mu_2 = 0$ —> $\mu = f$

  - fraction f have $\mu = \mu_1$, $1-f$ have large σ, $\mu_2 = 0$ —> $\mu = f\mu_1$

- Summary: can predict μ(E) from CNN p(σ_E)



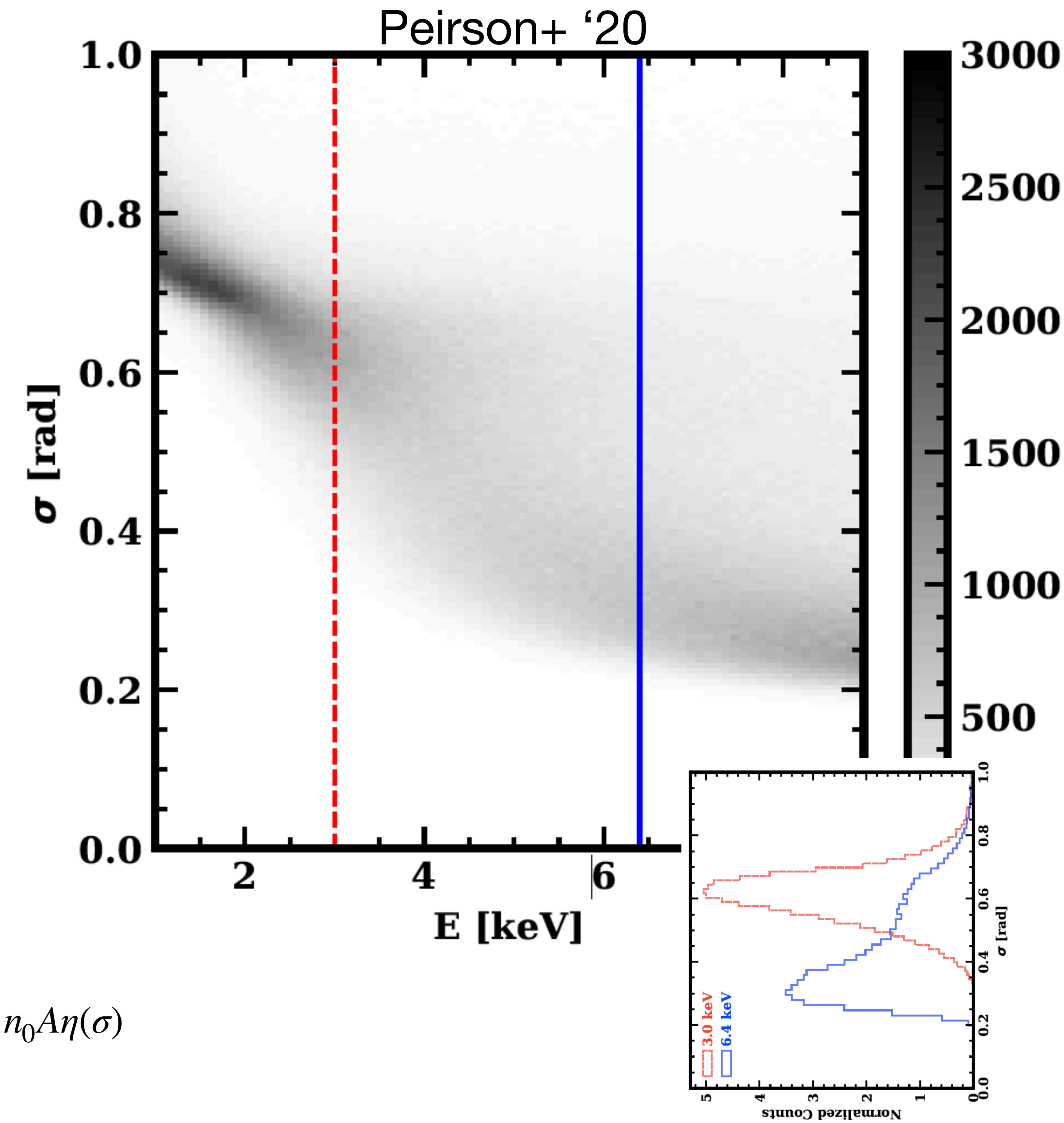
Broad-Band Analysis



- Modulation depends only on σ
- Likelihood depends only on modulation
- Thus, likelihood depends on distribution of σ
- New measure:

$$\frac{dN}{d\sigma} = \int n_E A_E p(\sigma; E) dE = n_0 A \int \phi(E) \alpha(E) p(\sigma; E) dE = n_0 A \eta(\sigma)$$

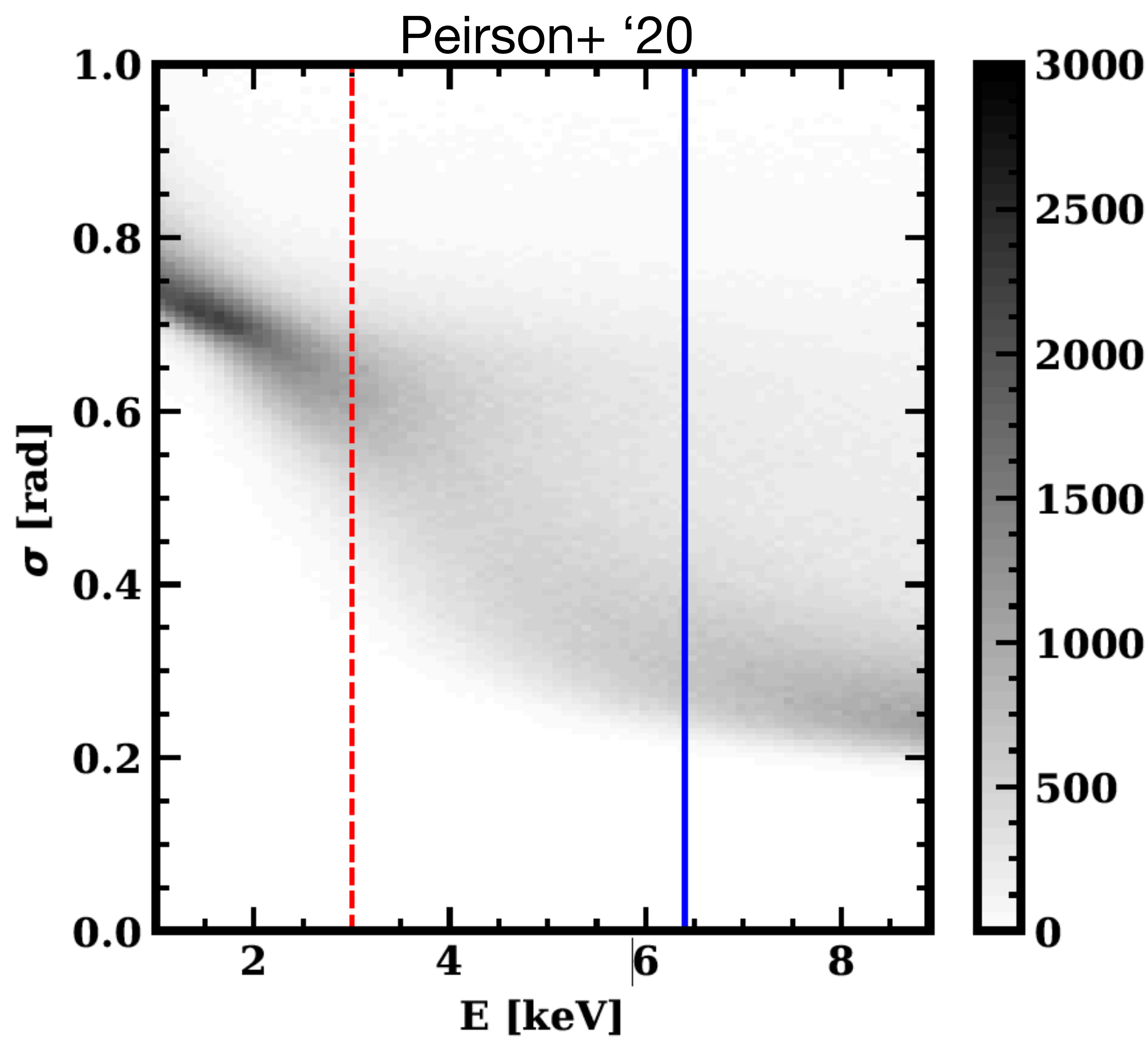
Broad-Band Analysis



- Modulation depends only on σ
- Likelihood depends only on modulation
- Thus, likelihood depends on distribution of σ
- New measure:

$$\frac{dN}{d\sigma} = \int n_E A_E p(\sigma; E) dE = n_0 A \int \phi(E) \alpha(E) p(\sigma; E) dE = n_0 A \eta(\sigma)$$

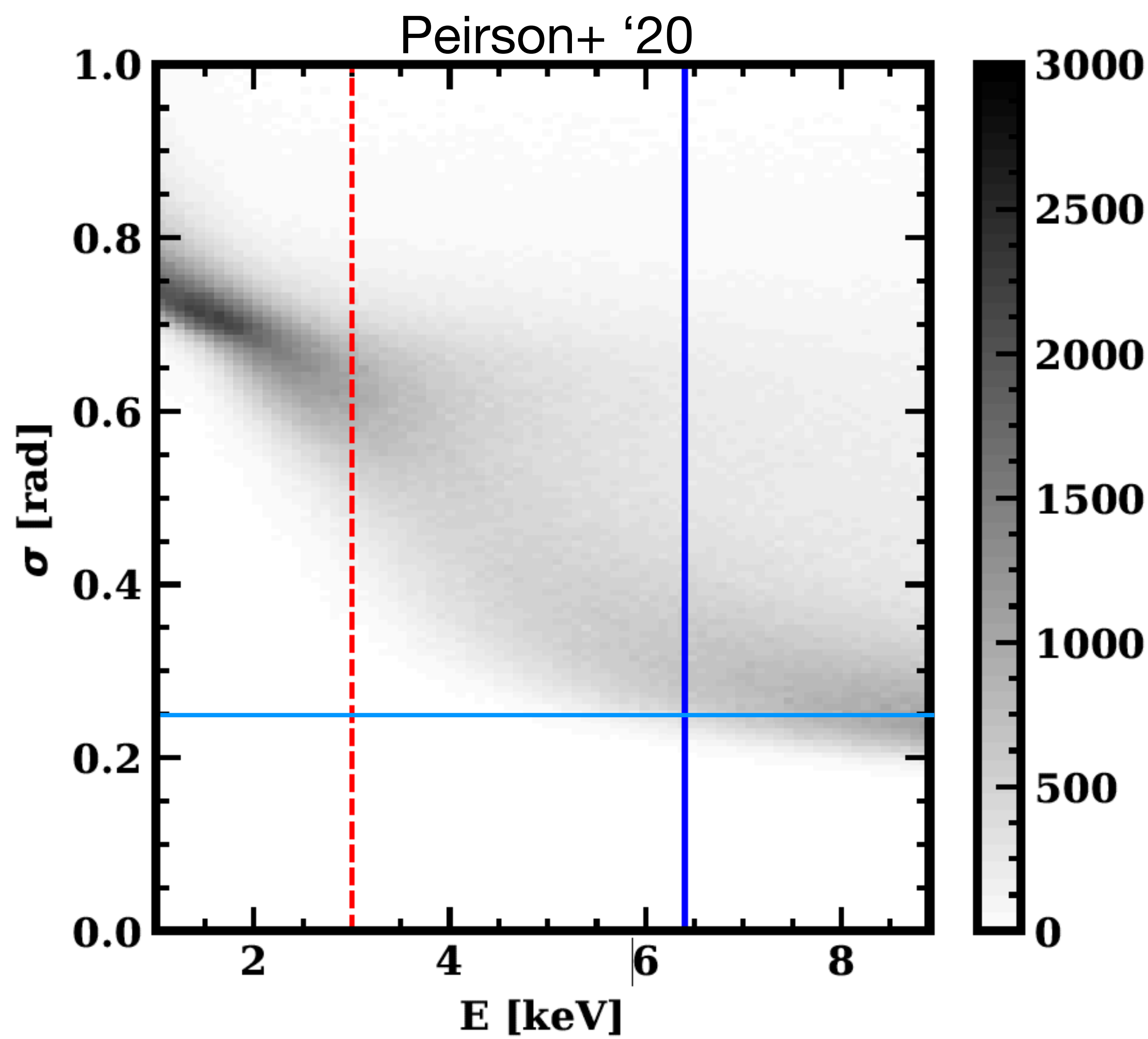
Broad-Band Analysis



- Modulation depends only on σ
- Likelihood depends only on modulation
- Thus, likelihood depends on distribution of σ
- New measure:

$$\frac{dN}{d\sigma} = \int n_E A_E p(\sigma; E) dE = n_0 A \int \phi(E) \alpha(E) p(\sigma; E) dE = n_0 A \eta(\sigma)$$

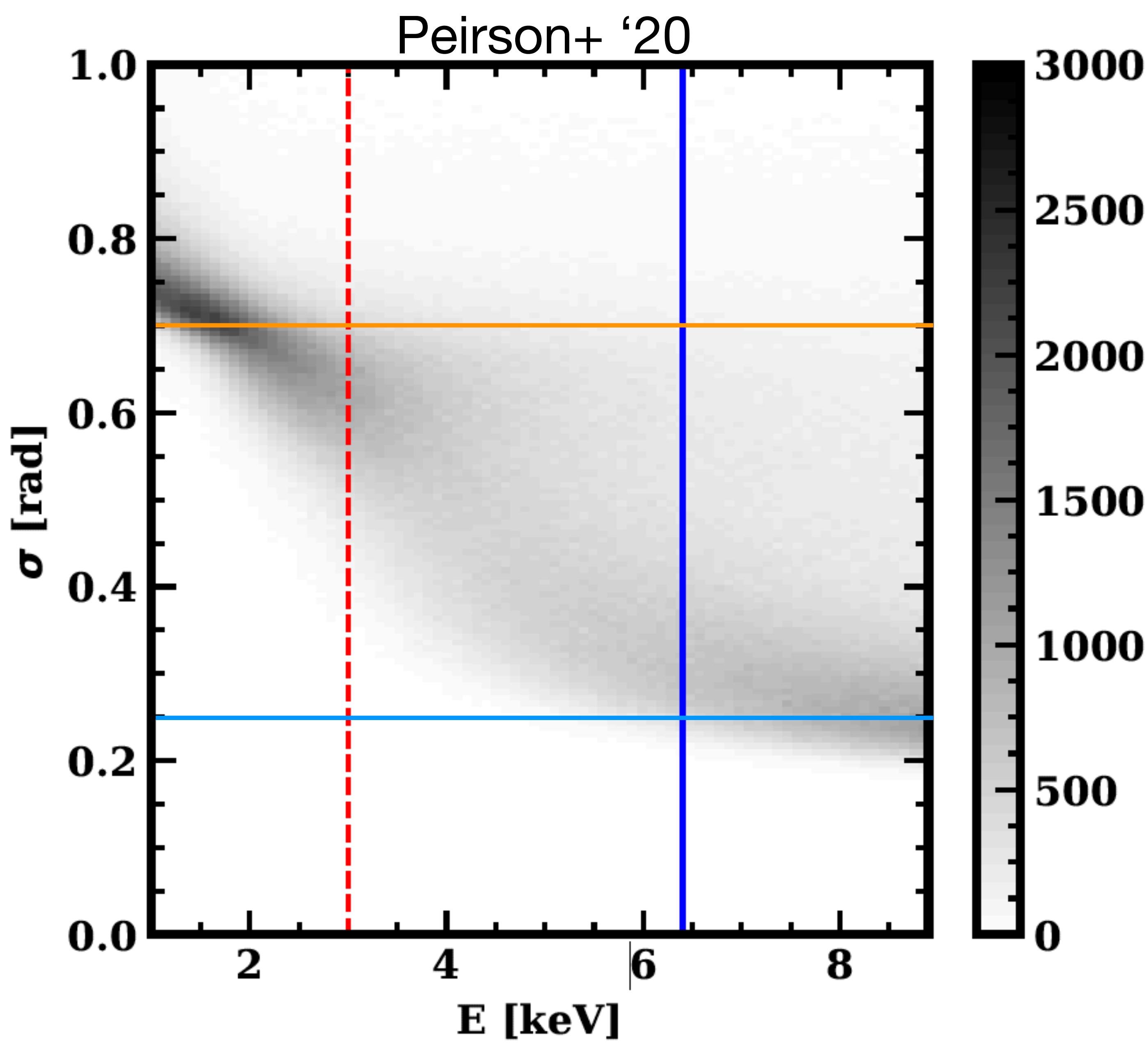
Broad-Band Analysis



- Modulation depends only on σ
- Likelihood depends only on modulation
- Thus, likelihood depends on distribution of σ
- New measure:

$$\frac{dN}{d\sigma} = \int n_E A_E p(\sigma; E) dE = n_0 A \int \phi(E) \alpha(E) p(\sigma; E) dE = n_0 A \eta(\sigma)$$

Broad-Band Analysis



- Modulation depends only on σ
- Likelihood depends only on modulation
- Thus, likelihood depends on distribution of σ
- New measure:

$$\frac{dN}{d\sigma} = \int n_E A_E p(\sigma; E) dE = n_0 A \int \phi(E)\alpha(E)p(\sigma; E)dE = n_0 A \eta(\sigma)$$

Likelihood Analysis

- Original: $\lambda(n_E, \Pi, \varphi; E, \psi) = [1 + \Pi\mu_E \cos(2\psi - 2\varphi)]n_E A_E T dE d\psi$
  - gives MDP$_{99}$ $= 4.29/\sqrt{\sum \mu_E^2 C(E)}$
- Update with:

$$\lambda(n_0, \Pi, \varphi; E, \psi, \sigma) = \int d\psi' \; [1 + \Pi \cos(2\psi' - 2\varphi)]G(\psi; \psi', \sigma)n_E A_E T p(\sigma; E)$$

$$= \; [1 + \Pi e^{-2\sigma^2} \cos(2\psi - 2\varphi)]n_E A_E T p(\sigma; E)$$

- Transform to σ space:

$$\tilde{\lambda}(n_0, q, u; \psi, \sigma) = \int \lambda dE = [1 + e^{-2\sigma^2}(q \cos 2\psi + u \sin 2\psi)]n_0 A T \eta(\sigma)$$

and $\tilde{S}(q, u) = -2\sum_i \ln(1 + qe^{-2\sigma_i^2} \cos 2\psi_i + ue^{-2\sigma_i^2} \sin 2\psi_i)$

- Result: $\mathrm{MDP}_{99} = 4.29/\sqrt{\sum e^{-4\sigma^2} C(\sigma)}$, for small Π (or large C[σ])