iDark
The intelligent dark matter survey

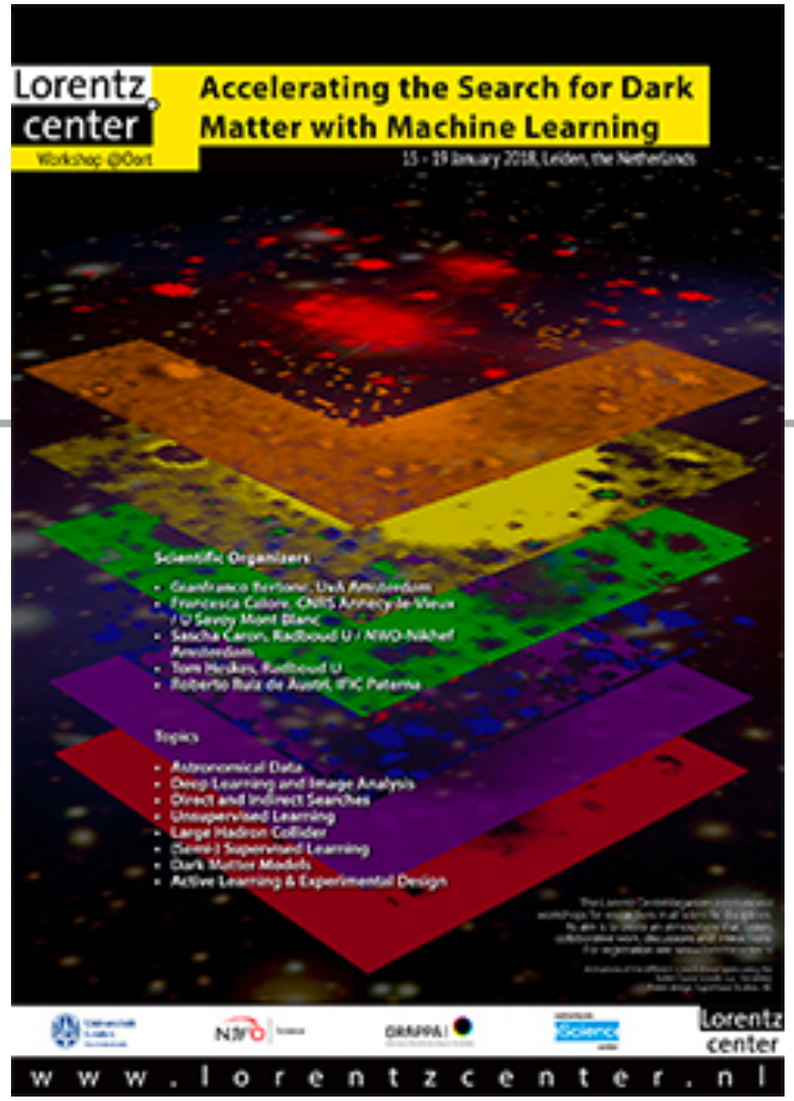
Radboud University

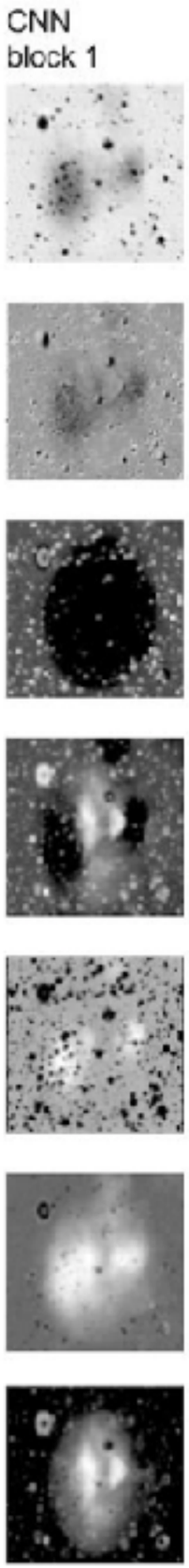
netherlands eScience center

INDIRECT DETECTION & MACHINE LEARNING

LUC HENDRIKS

RADBOUD UNIVERSITY NIJMEGEN



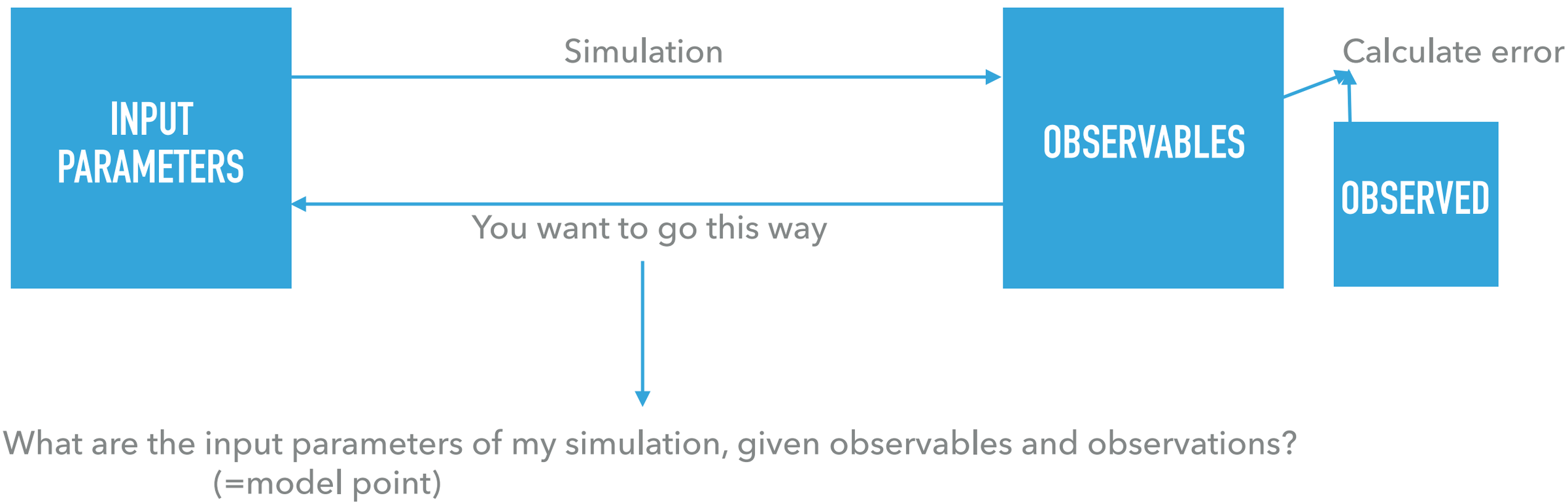


OUTLINE

- Inverse problem
- Different approaches to deal with γ-ray data for indirect detection
  - SkyFACT / D3PO
  - Deep learning
- Some ways to use deep learning with simulations
  - Inverting inverse problem
  - Simulating simulations

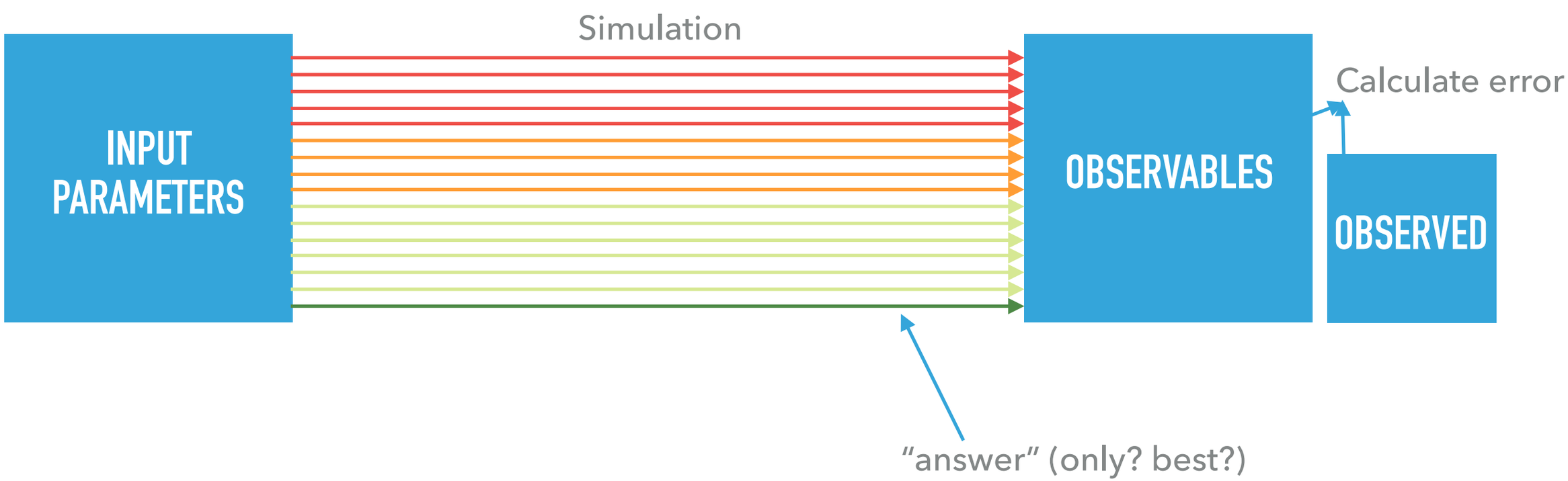
INVERSE PROBLEM

- Many problems in (astro)physics are an **inverse problem**



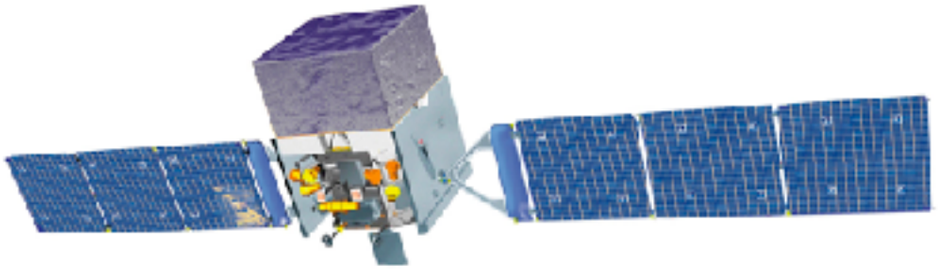
INVERSE PROBLEM

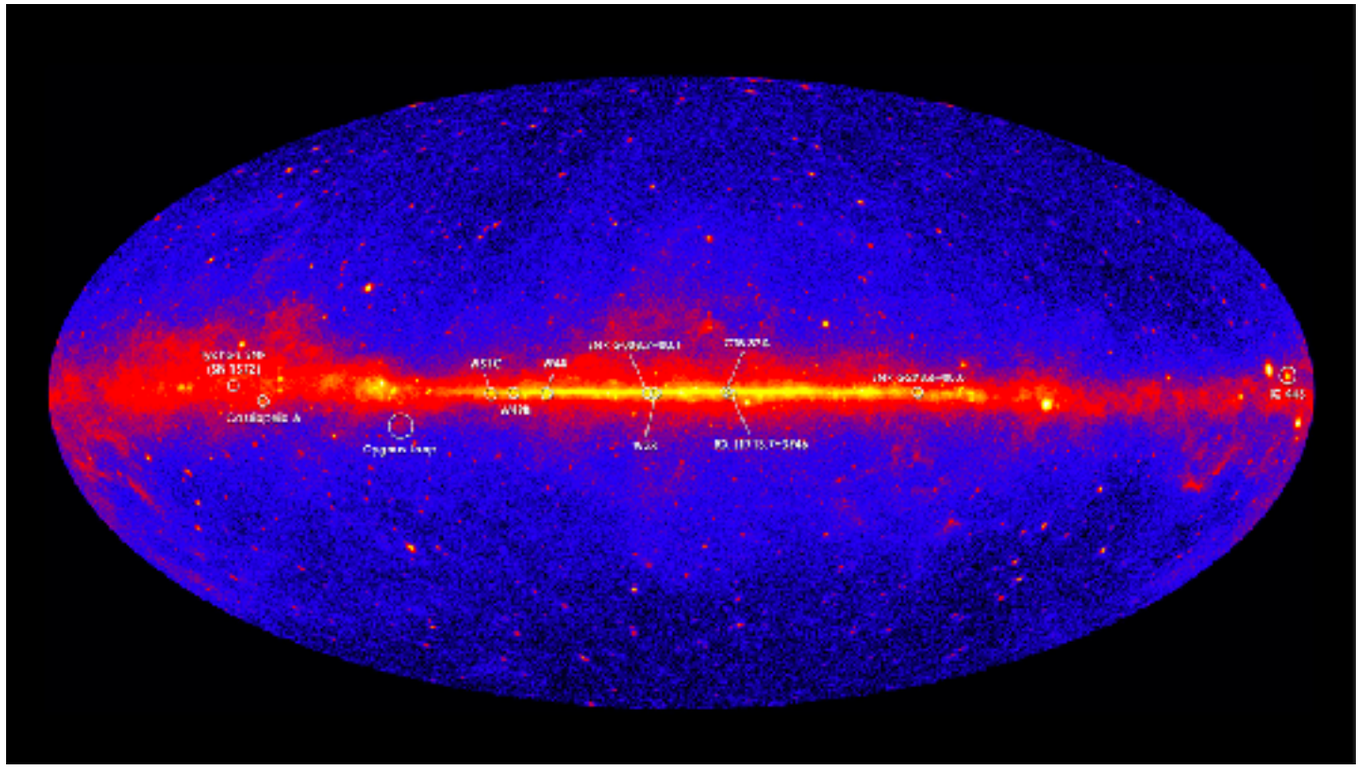
- You cannot go from observables to input parameters
- Need to scan the input parameter space for regions that predict observables that agree with observations
- Methods: random sampling, gradient descent, MCMC, nested sampling, …
- Computationally very expensive



INDIRECT DETECTION

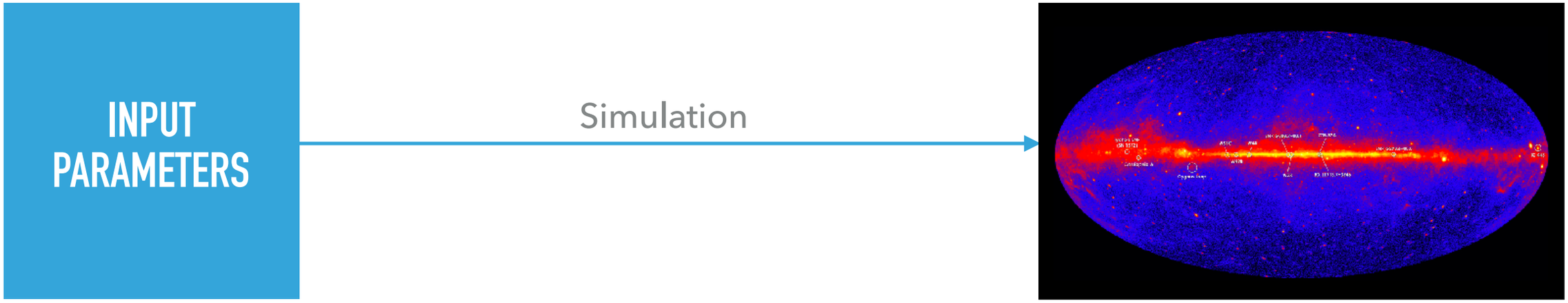
- Detection of decay or annihilation products of DM
- Focus on γ-rays
- Fermi-LAT is the go-to instrument





INDIRECT DETECTION — CHALLENGES

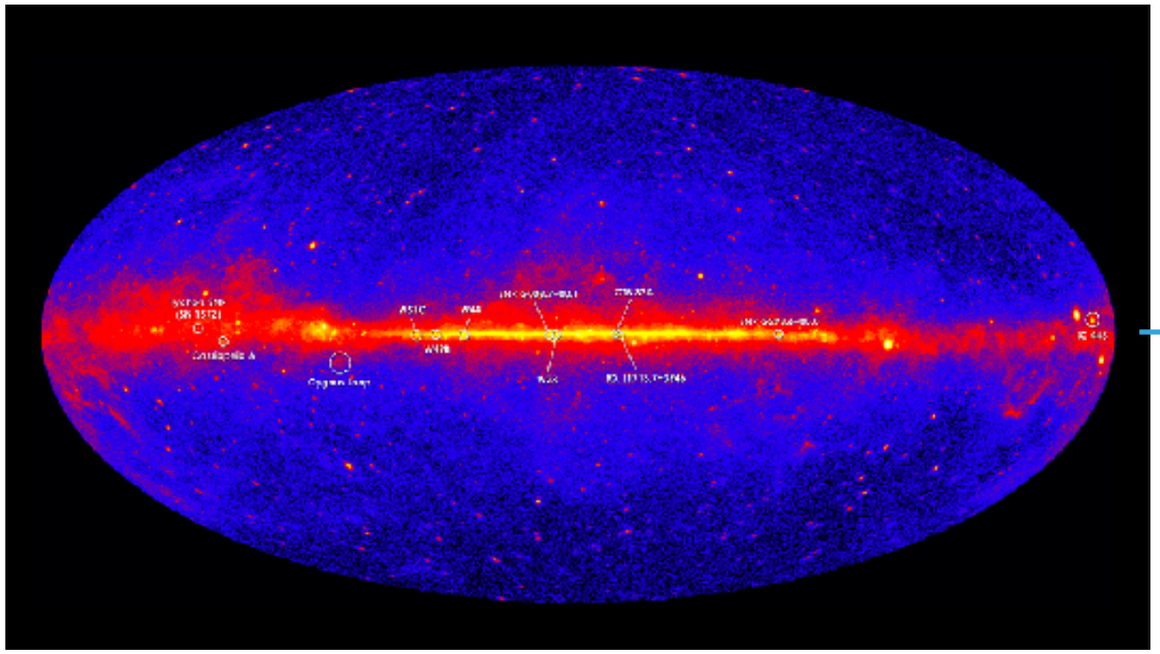
- You only have 1 image of the sky – no training data
- Develop models that reconstruct the γ-ray sky

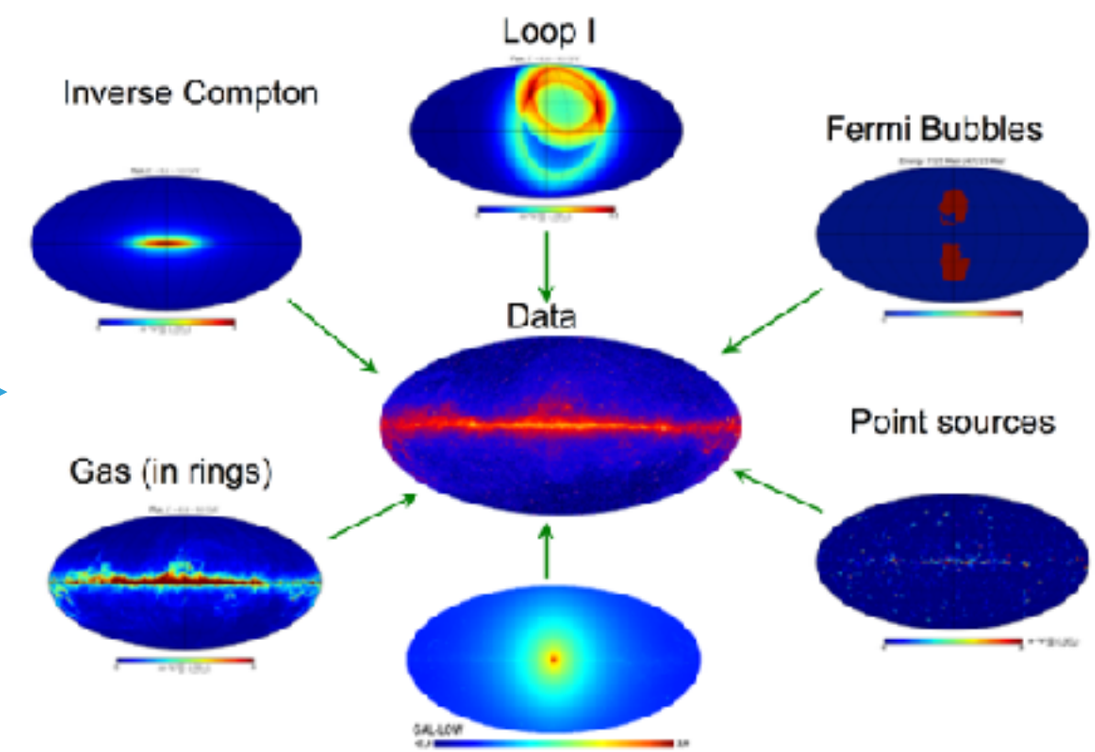


- What are all the input parameters? What are their values?
- Typical inverse problem

INDIRECT DETECTION

- Decompose gamma ray sky into different components
- Weighted sum different components
- Make fit. Add component / edit weights to improve fit
- However, the sum of components cannot perfectly match the sky…





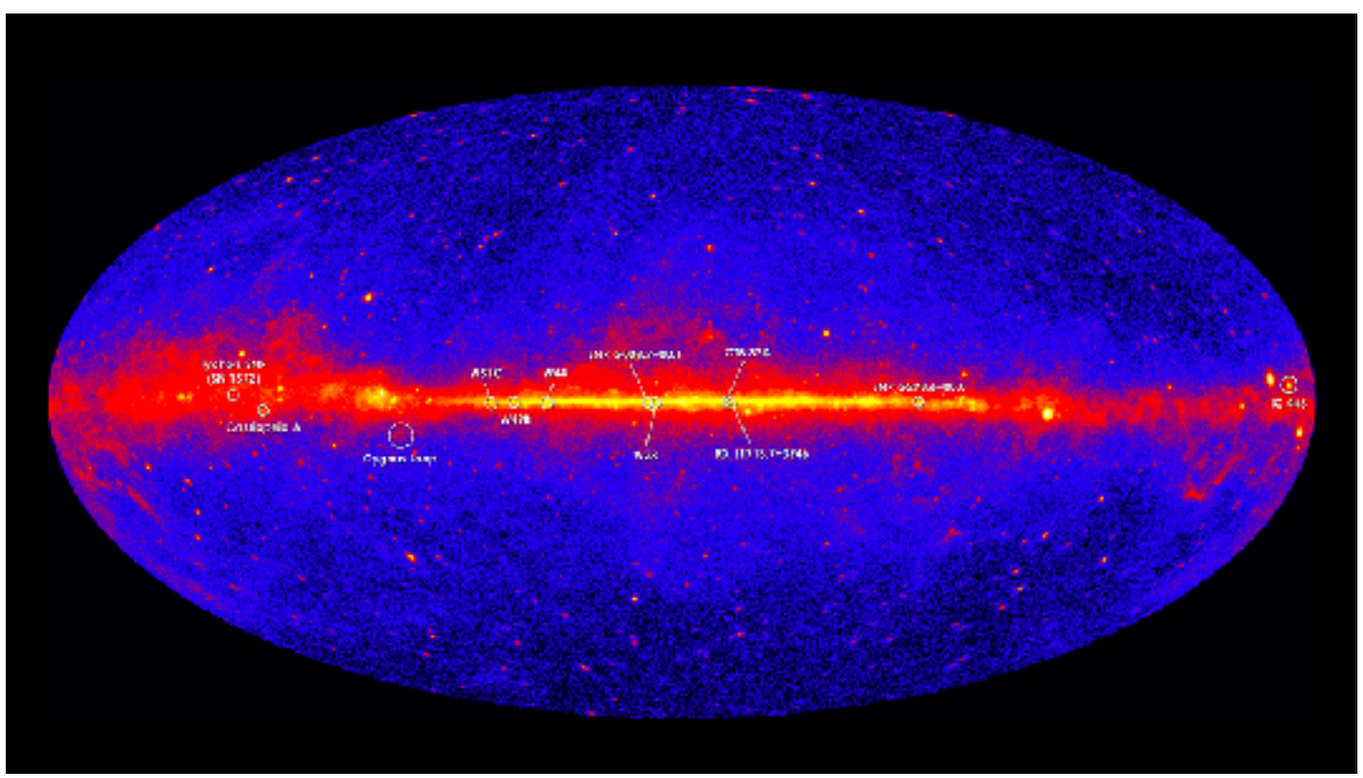
GC excess. DM???

SKYFACT

Add nuisance parameters
+ regularisation

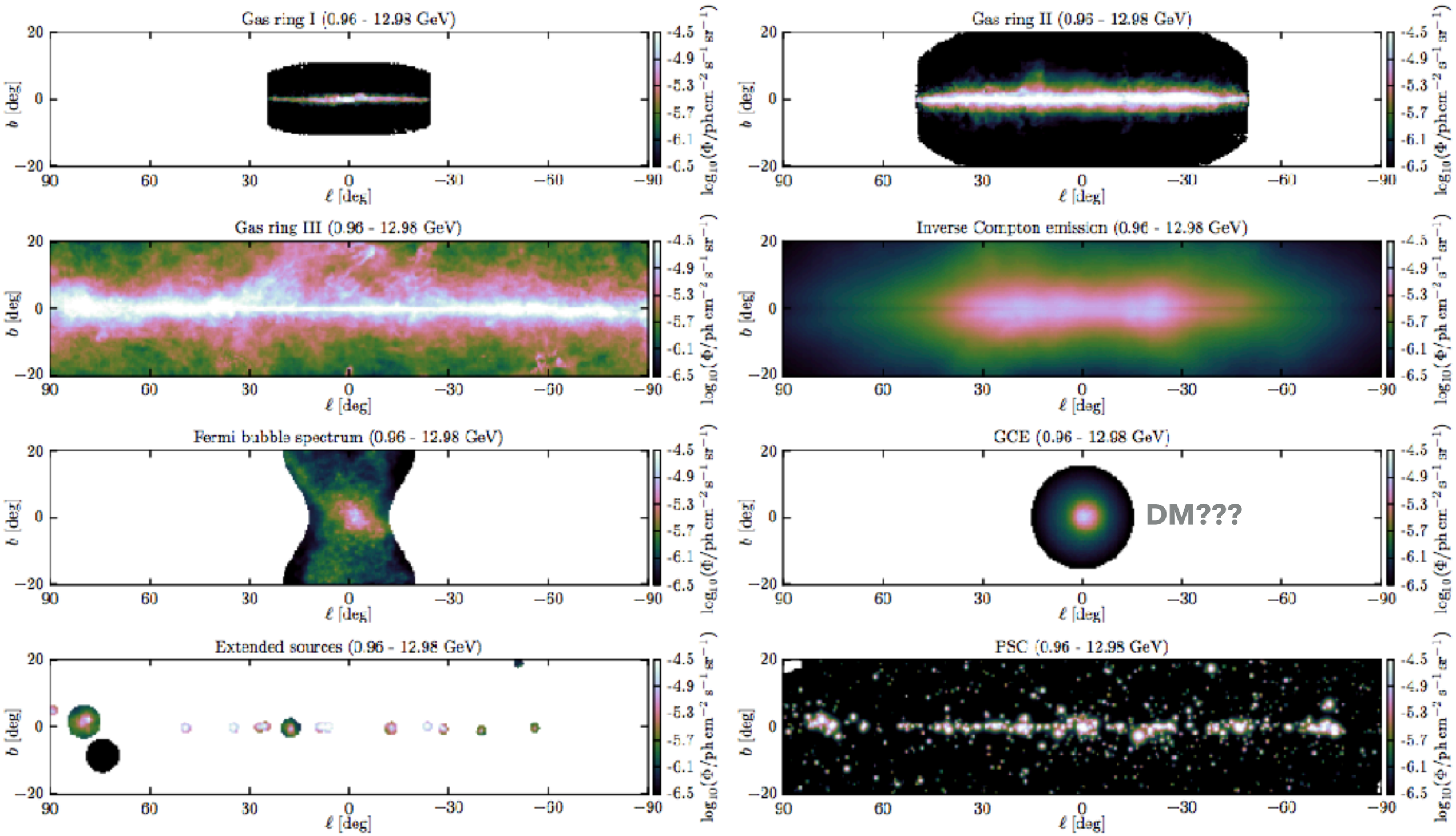
$$\phi_{pb} = \sum_k T_p^{(k)} \tau_p^{(k)} \cdot S_b^{(k)} \sigma_b^{(k)} \cdot \nu^{(k)}$$

- Assume uncorrelated error in pixels
  (lots of imperfections in models and measurements)
- Modify models (all energy bin / pixel values) to reduce errors
- You will find zones with high error that cannot be fixed with nuisance parameters
  - Add new component



https://arxiv.org/abs/1705.04065
Emma Storm, Christoph Weniger, Francesca Calore

SKYFACT


Gas ring I (0.96 - 12.98 GeV)
Gas ring II (0.96 - 12.98 GeV)
Gas ring III (0.96 - 12.98 GeV)
Inverse Compton emission (0.96 - 12.98 GeV)
Fermi bubble spectrum (0.96 - 12.98 GeV)
GCE (0.96 - 12.98 GeV)
DM???
Extended sources (0.96 - 12.98 GeV)
PSC (0.96 - 12.98 GeV)
b [deg]
ℓ [deg]
log10(Φ/ph cm−2 s−1 sr−1)

INFORMATION FIELD THEORY & D3PO

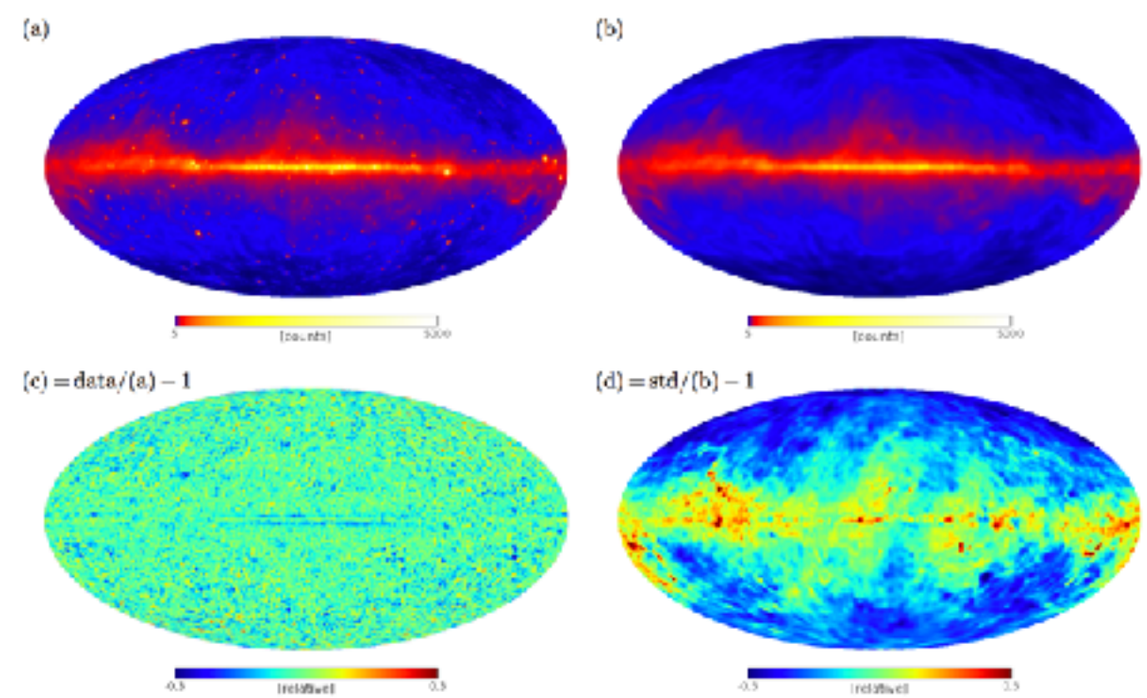
- Connect Bayesian analysis to statistical physics & QFT
- Deconvolve, denoise, decompose photon observations
  - Use IFT to get information out of photon observations
  - Open source code: http://wwwmpa.mpa-garching.mpg.de/ift/d3po/

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d|s)\,\mathcal{P}(s)}{\mathcal{P}(d)}$$

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(d,s)}{\mathcal{P}(d)} \equiv \frac{e^{-\mathcal{H}(d,s)}}{\mathcal{Z}(d)}$$

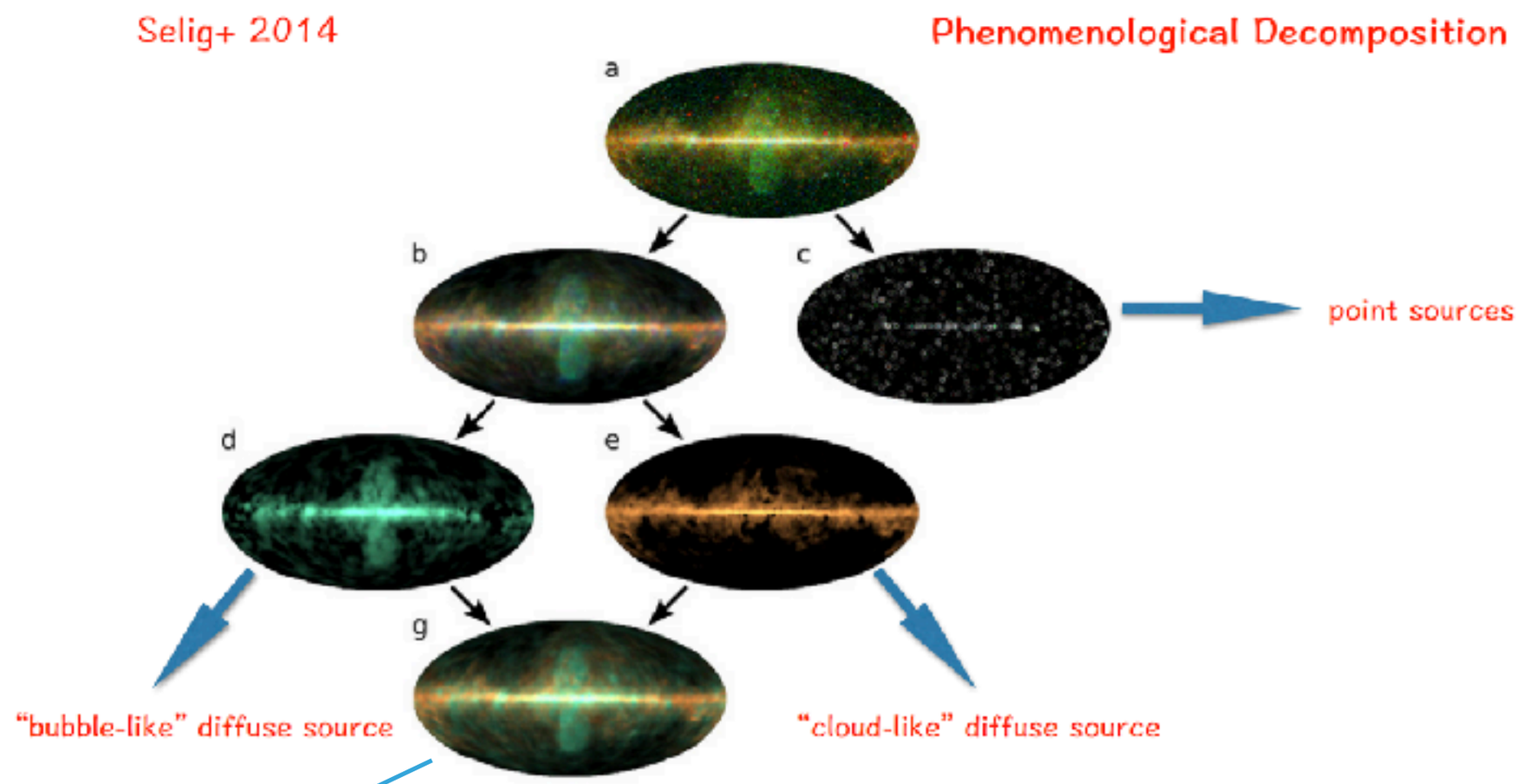
$$\mathcal{H}(d,s) \equiv -\ln \mathcal{P}(d,s) = -\ln \mathcal{P}(d|s) - \ln \mathcal{P}(s) \equiv \mathcal{H}(d|s) + \mathcal{H}(s)$$

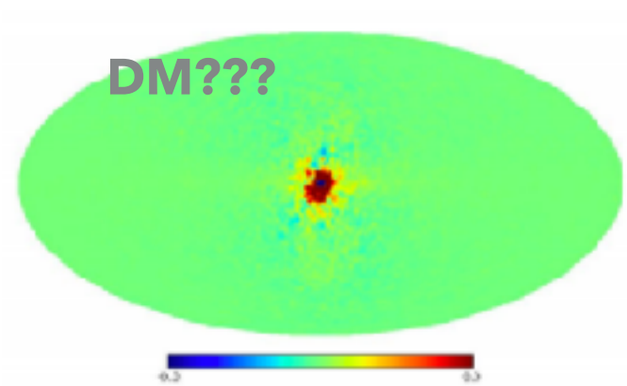
$$\mathcal{Z}(d) \equiv \mathcal{P}(d) = \int \mathcal{D}s\, \mathcal{P}(d,s)$$



https://arxiv.org/abs/1311.1888
Marco Selig and Torsten A. Enßlin

PHENOMENOLOGICAL DECOMPOSITION



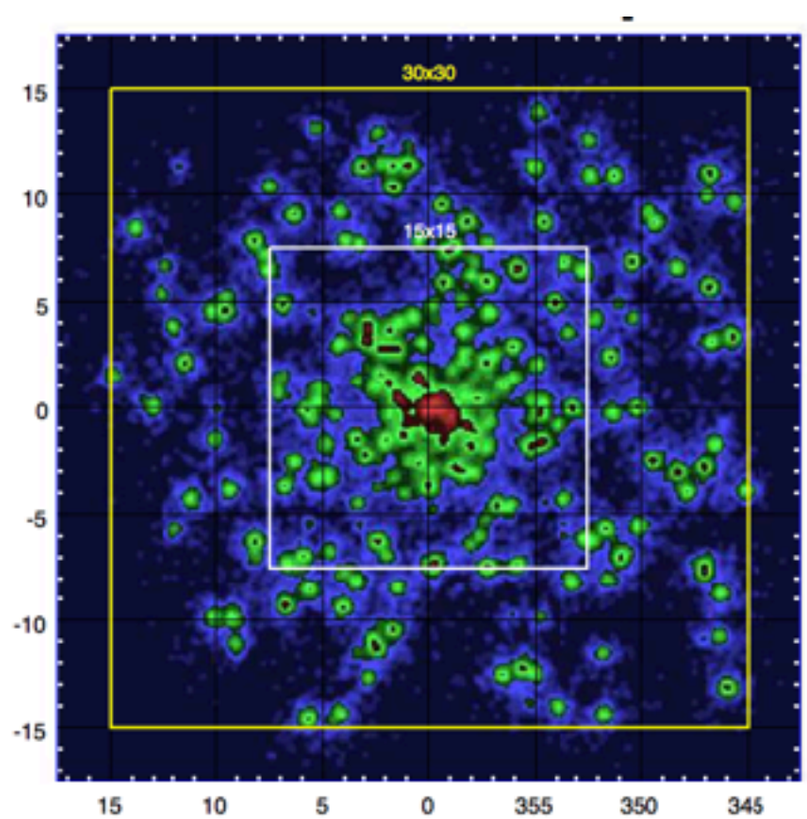


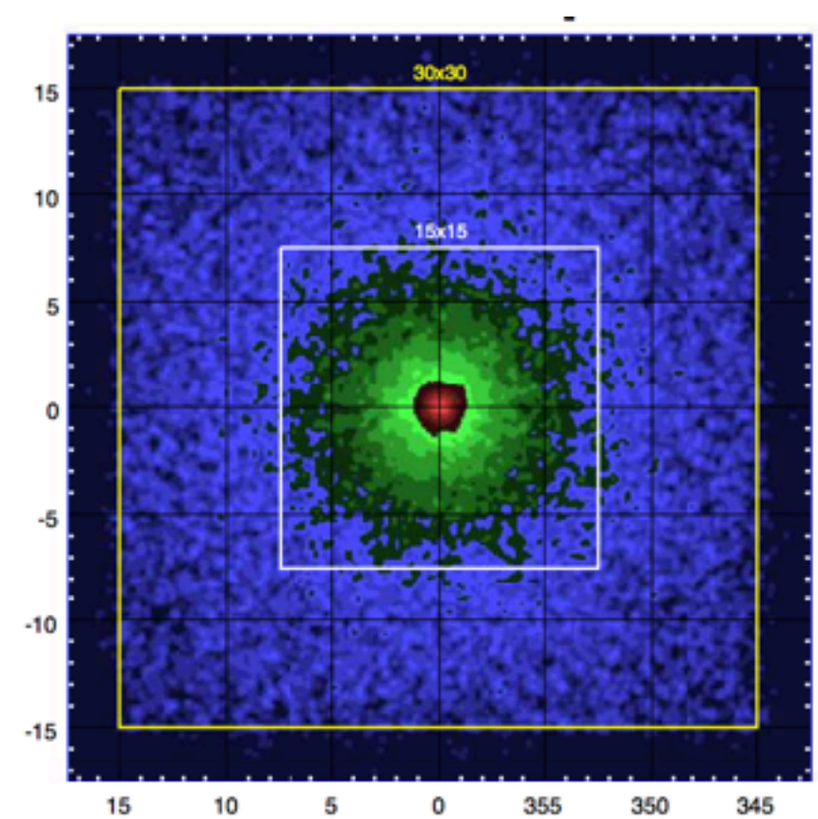
"GC excess" template

From (Xiaoyuan Huang)

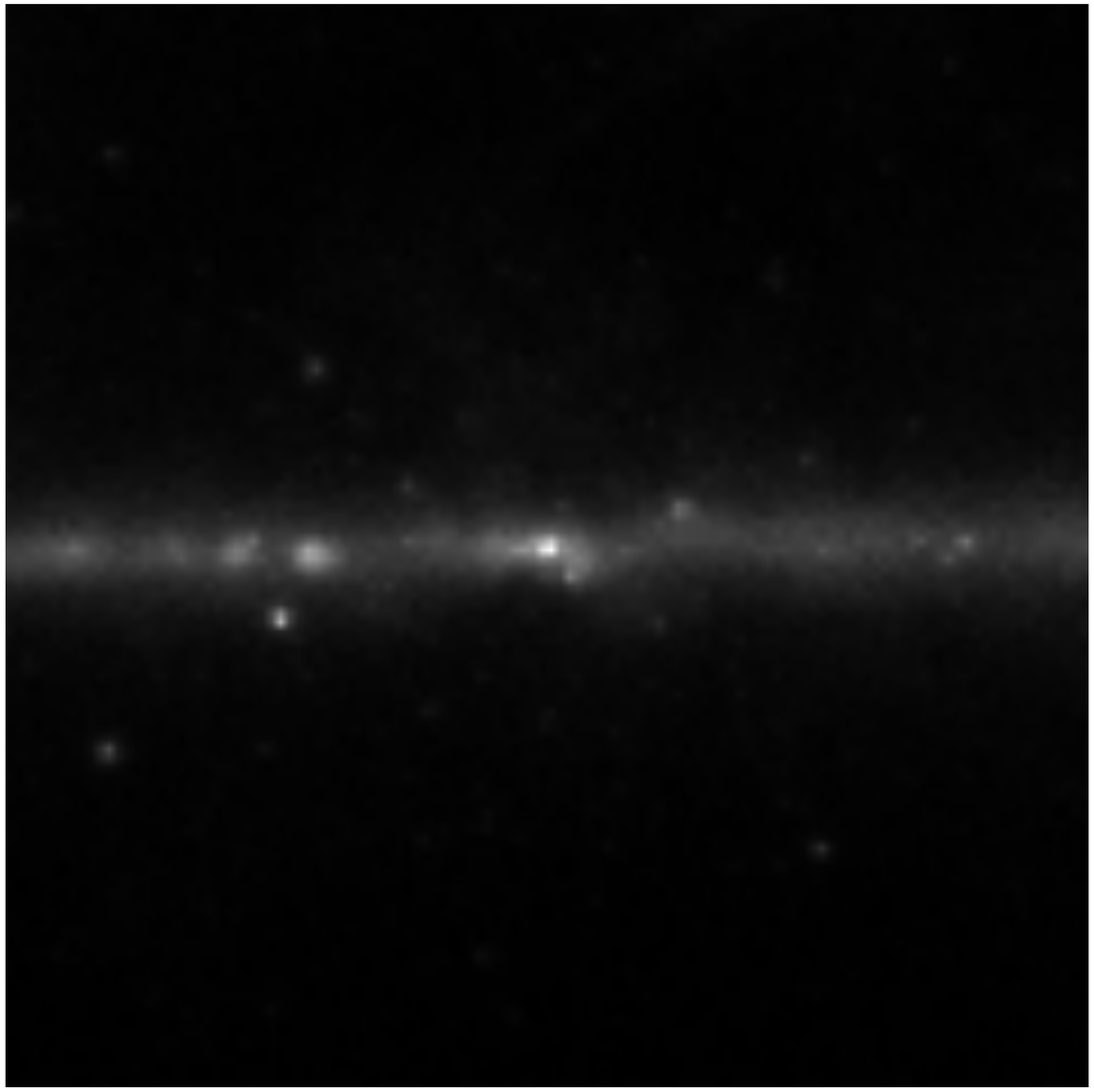
DEEP LEARNING

- Question: Is the GC excess from a diffuse source or collection of unresolved point sources?
- Can we answer this question with deep learning?
- Assume diffuse and point source component distributed by gNFW profile

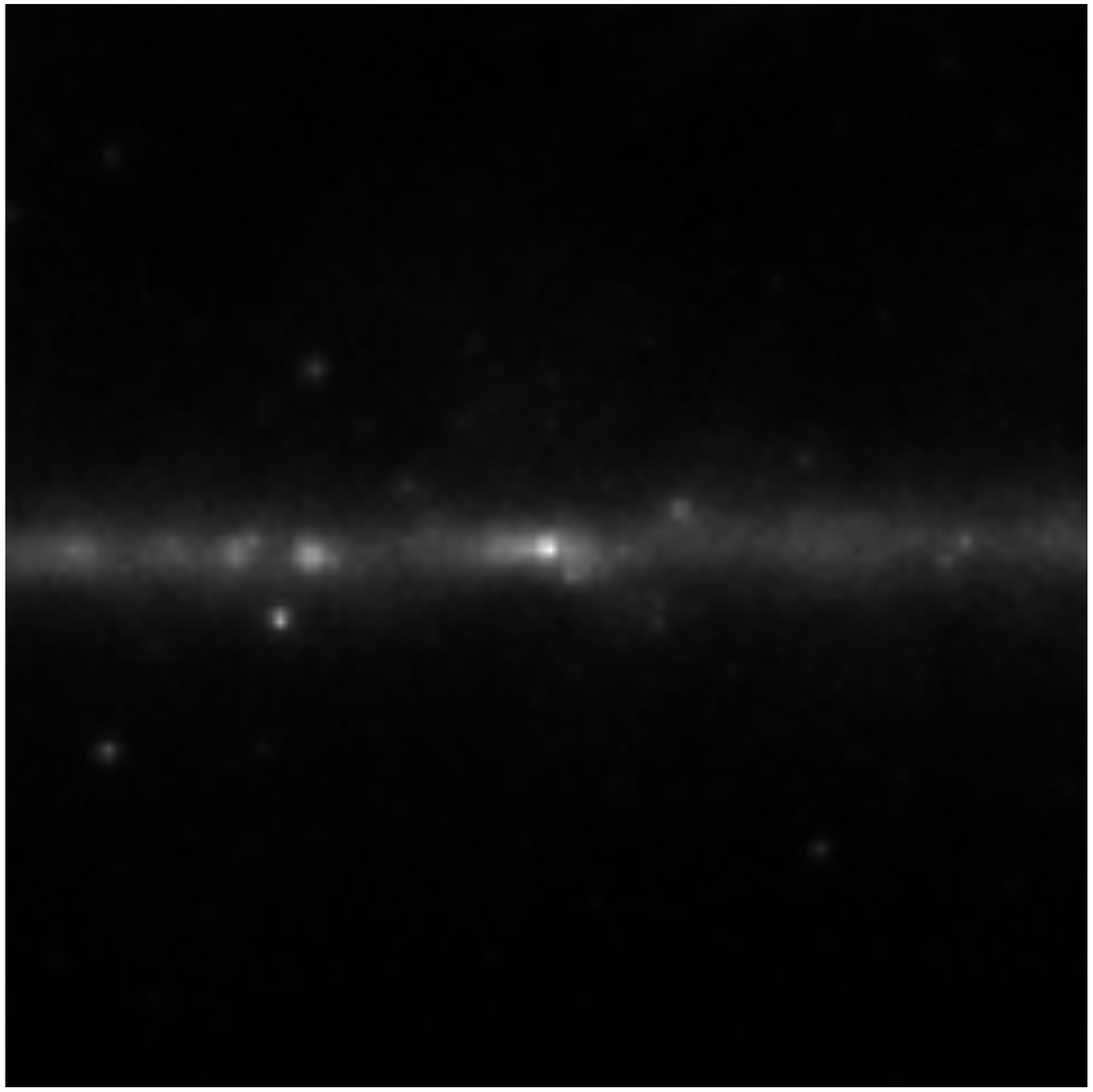




EXAMPLE OF TWO SIMULATIONS



$f_{src} = 0.9883$

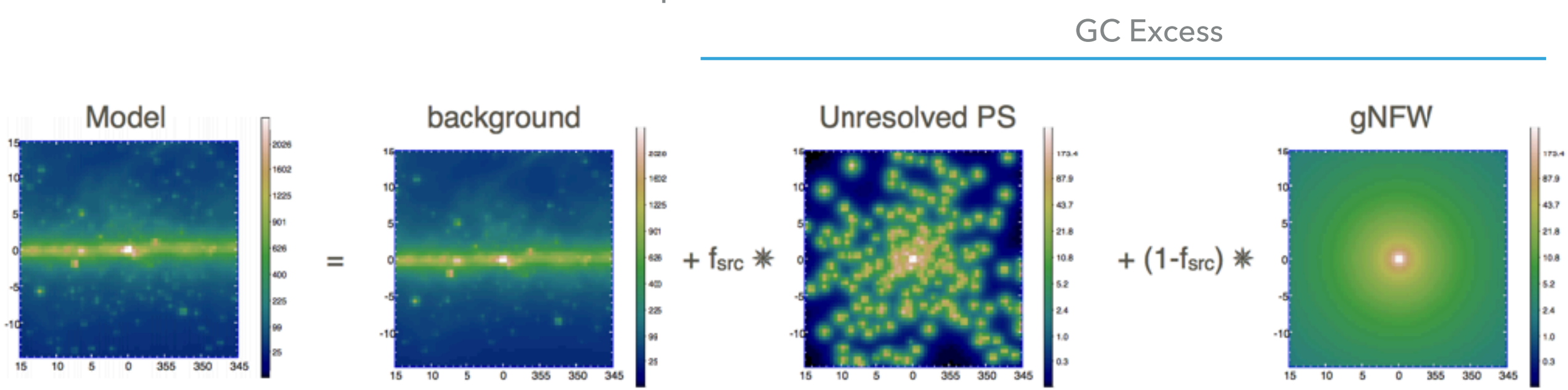


$f_{src} = 0.0275$

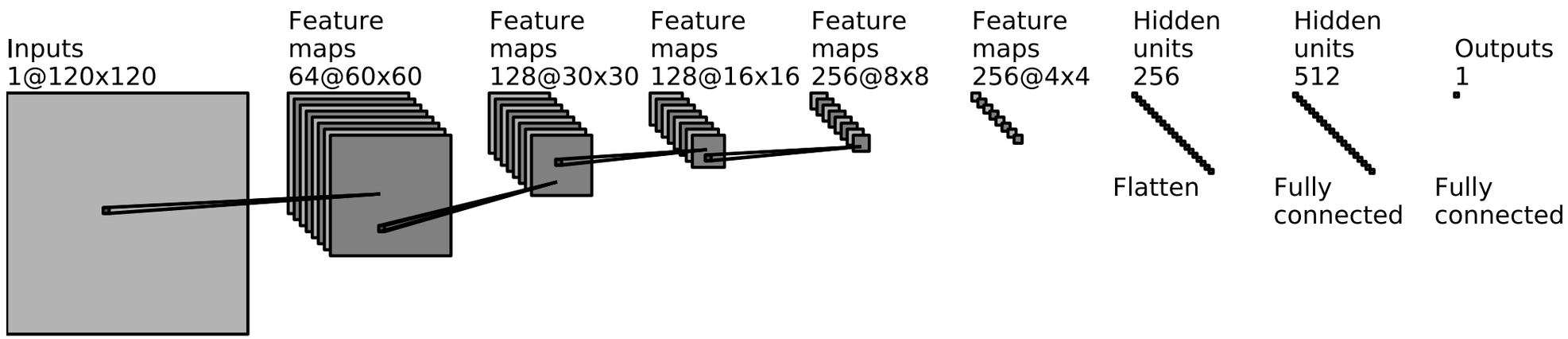
Try yourself! http://fermiai.s3-website-eu-west-1.amazonaws.com/

HOW TO APPLY DEEP LEARNING TO Y-RAY DATA?

- Goal: determine the component of point sources vs diffuse source of the GC excess – f_{src}
- Simulate GC using Fermi tools (5 parameters)
  - Output is photon count map of photons between 1-6GeV (no spectrum information, will be improved in new version)
- Sample from simulations in 5D parameter space
- Train network to predict f_{src} accurately in all scenarios of the other components
- Apply on real data – sample big enough so that reality is somewhere in 5D space
- Network trained on simulated data to predict f_{src}



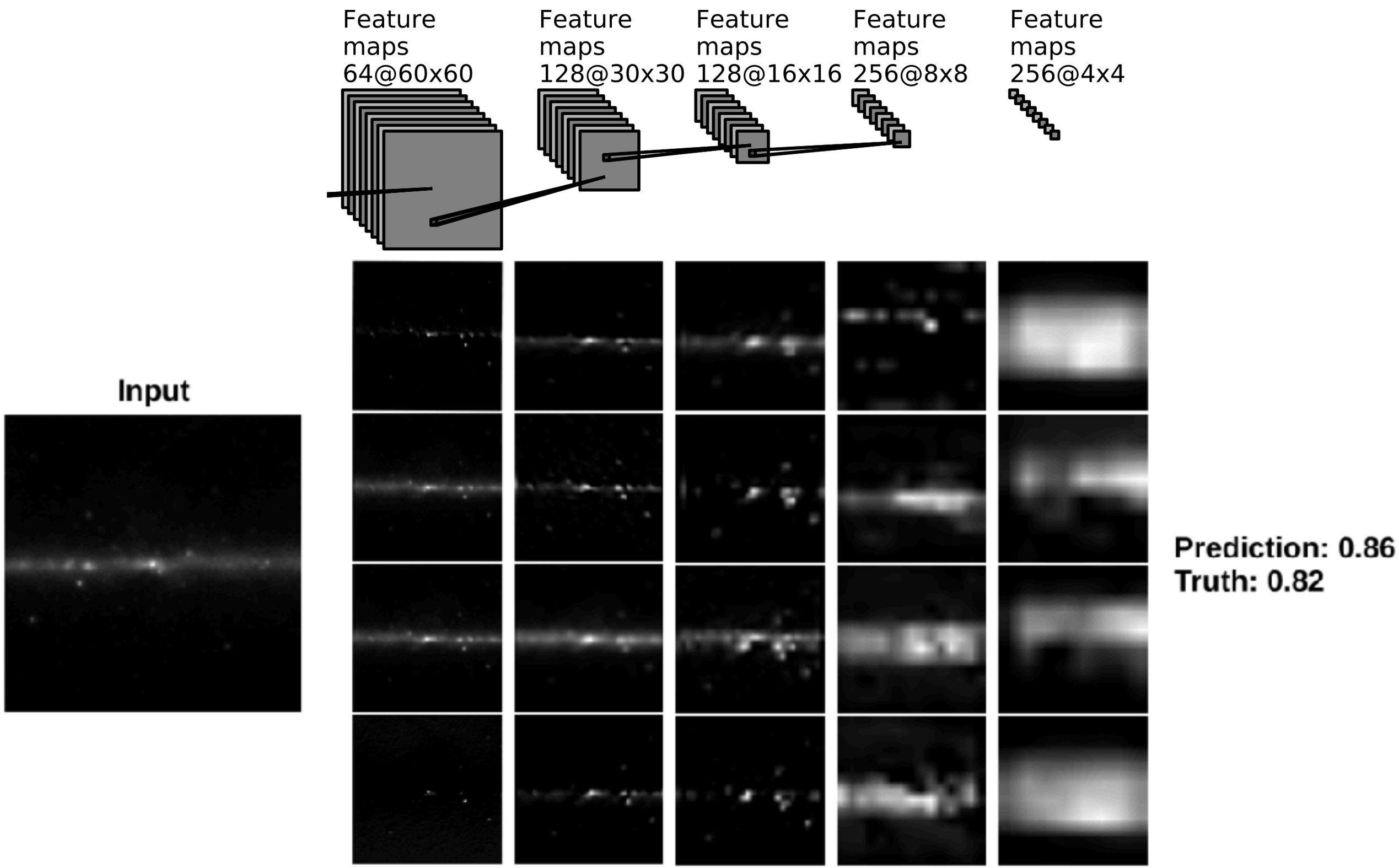
CONVOLUTIONAL NEURAL NETWORK



Max-pooling after every convolution
Local response normalization after every other convolution

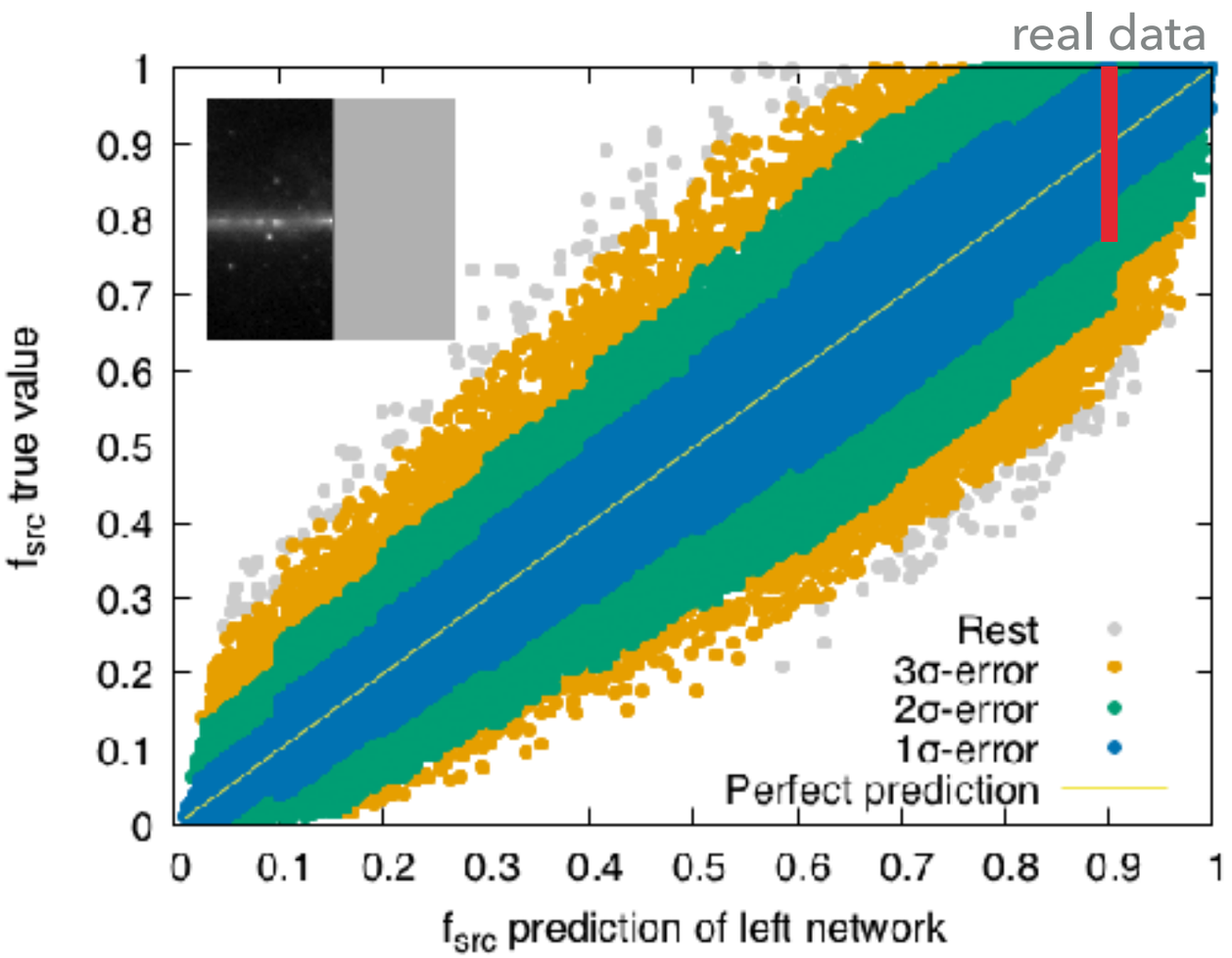
- Next version:
  - More input parameters in simulations
  - Output of network are all input parameters instead of only point source fraction

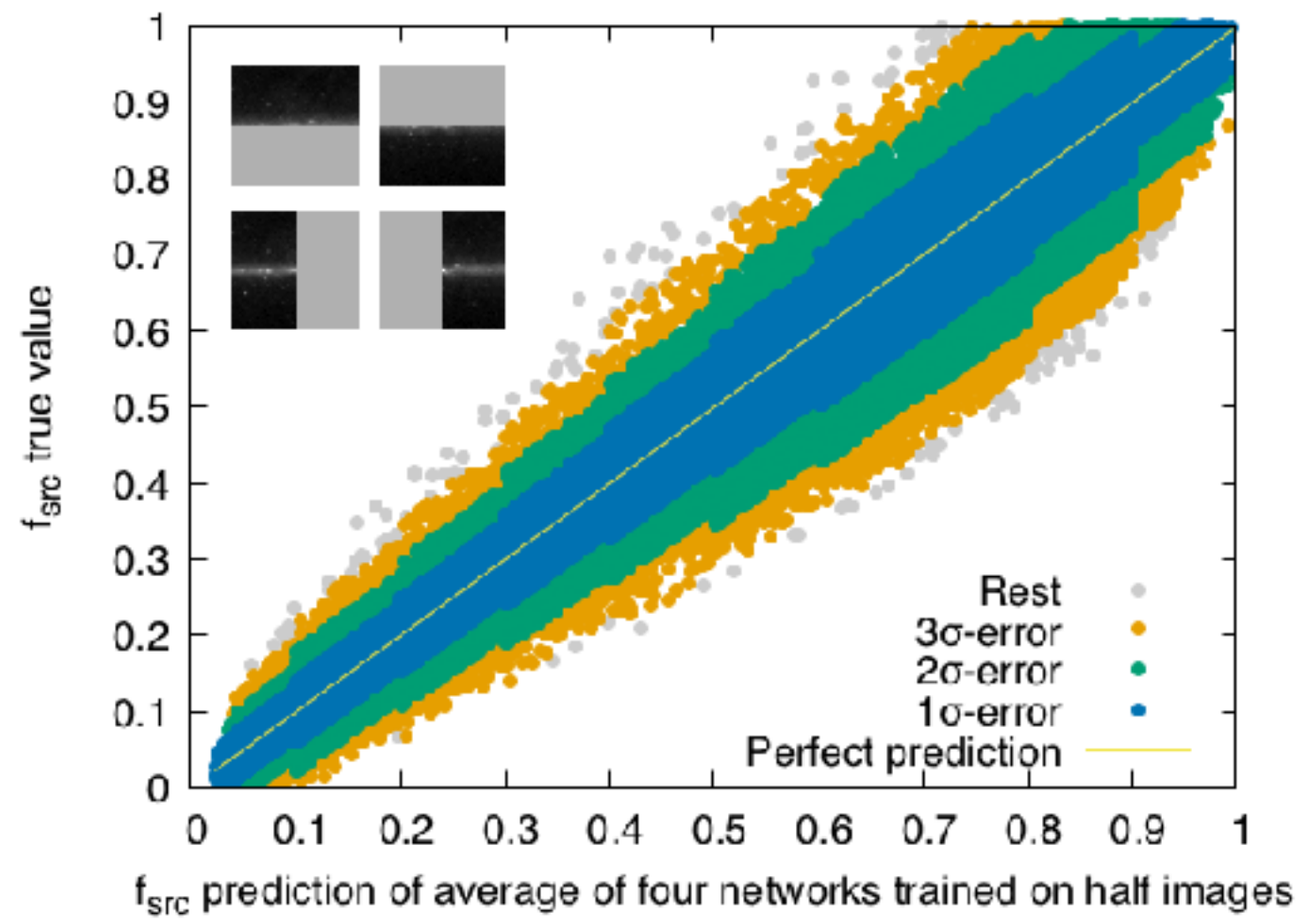
DEEP LEARNING



RESULTS

- Train using 3 background models, test on 2 others
- Test data: 2x30000 test points





Because we are doing a followup study, real result is only evaluated on the left network to not bias ourselves

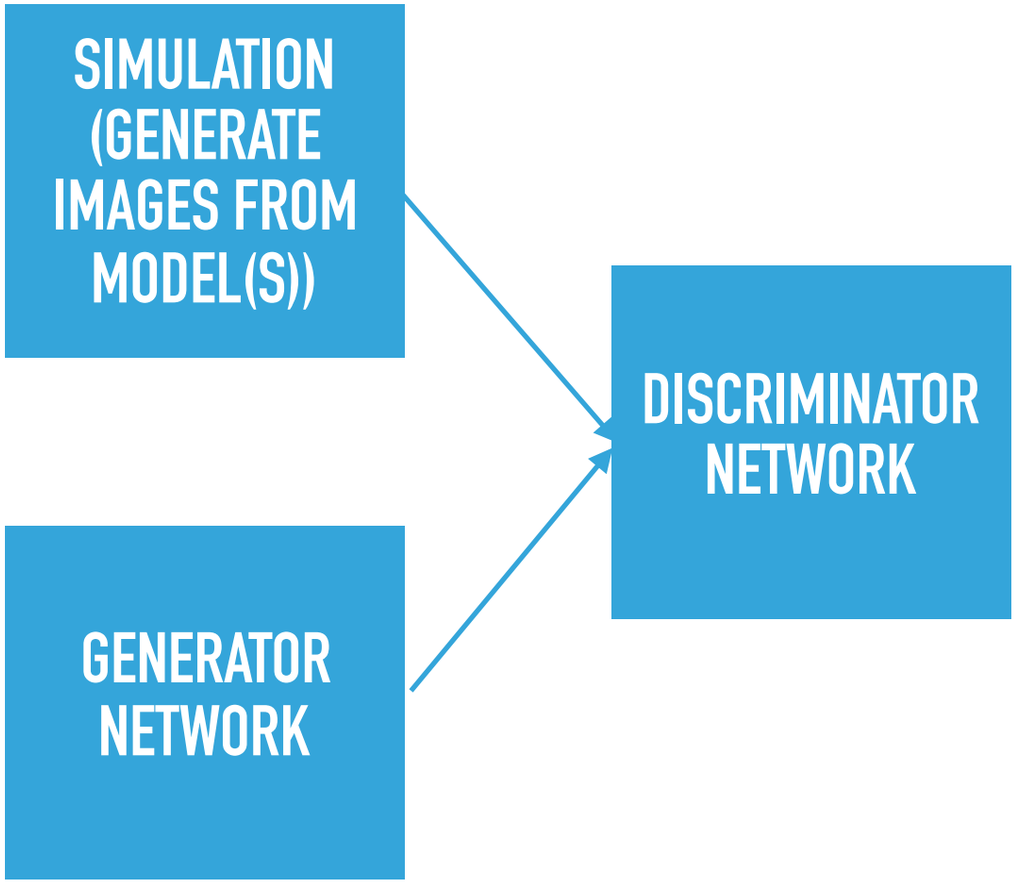
INVERTING INVERSE PROBLEM

- Use simulation to generate dataset
  - Observable = network input
  - Network output = model input
- No inverse problem anymore!
- However:
  - Network assumes simulation is reality
  - Even putting in a cat photo will return a value between 0 & 1, while we know it makes no sense
    - CNNs are known to be easily tricked
  - How to be able to trust this?

https://medium.com/@ageitgey/machine-learning-is-fun-part-8-how-to-intentionally-trick-neural-networks-b55da32b7196

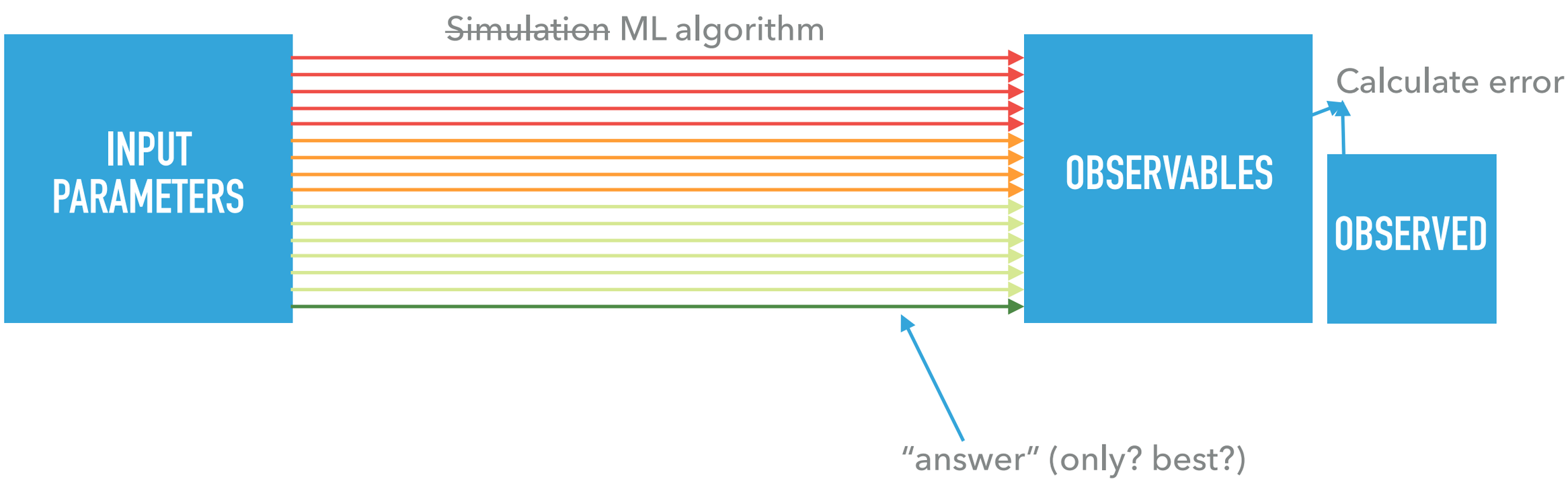
DEEP LEARNING APPROACH

- Adversarial model
- Generator network learns to generate images the model can generate
- Discriminator tries to distinguish
- When it cannot anymore, generator network is as good as your model
- Can the model generate images that are identical to real data?
- Which generator from which model can do this better / easier: compare models with each other



SIMULATING THE SIMULATION

- Another way to use deep learning to find input parameters is to simulate the simulation
- Use dataset from simulation to train a network
  - Network does the same as the simulation



SIMULATING THE SIMULATION

- Example from asteroseismology
- Train network on precomputed grid
- 10 input parameters
- 30 mins to find minimum

CONCLUSIONS

- Indirect detection methods can benefit greatly from ML methods
- Methods already used to deal with the Fermi-LAT data
  - SkyFACT
  - D3PO
  - ConvNets / deep learning
- Deep learning can be used to invert the inverse problem, or simulate simulations
  - Speed up simulations greatly