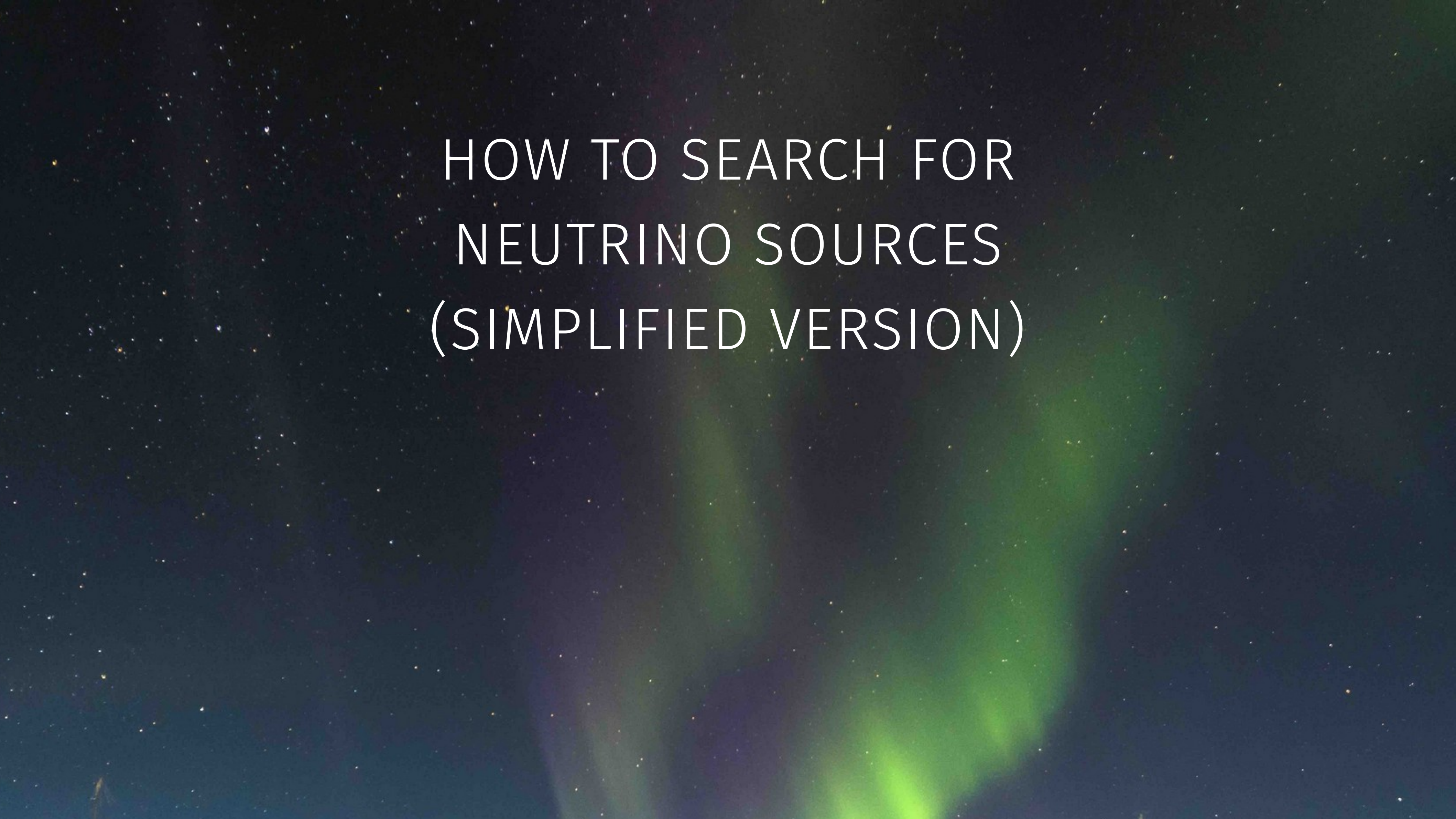
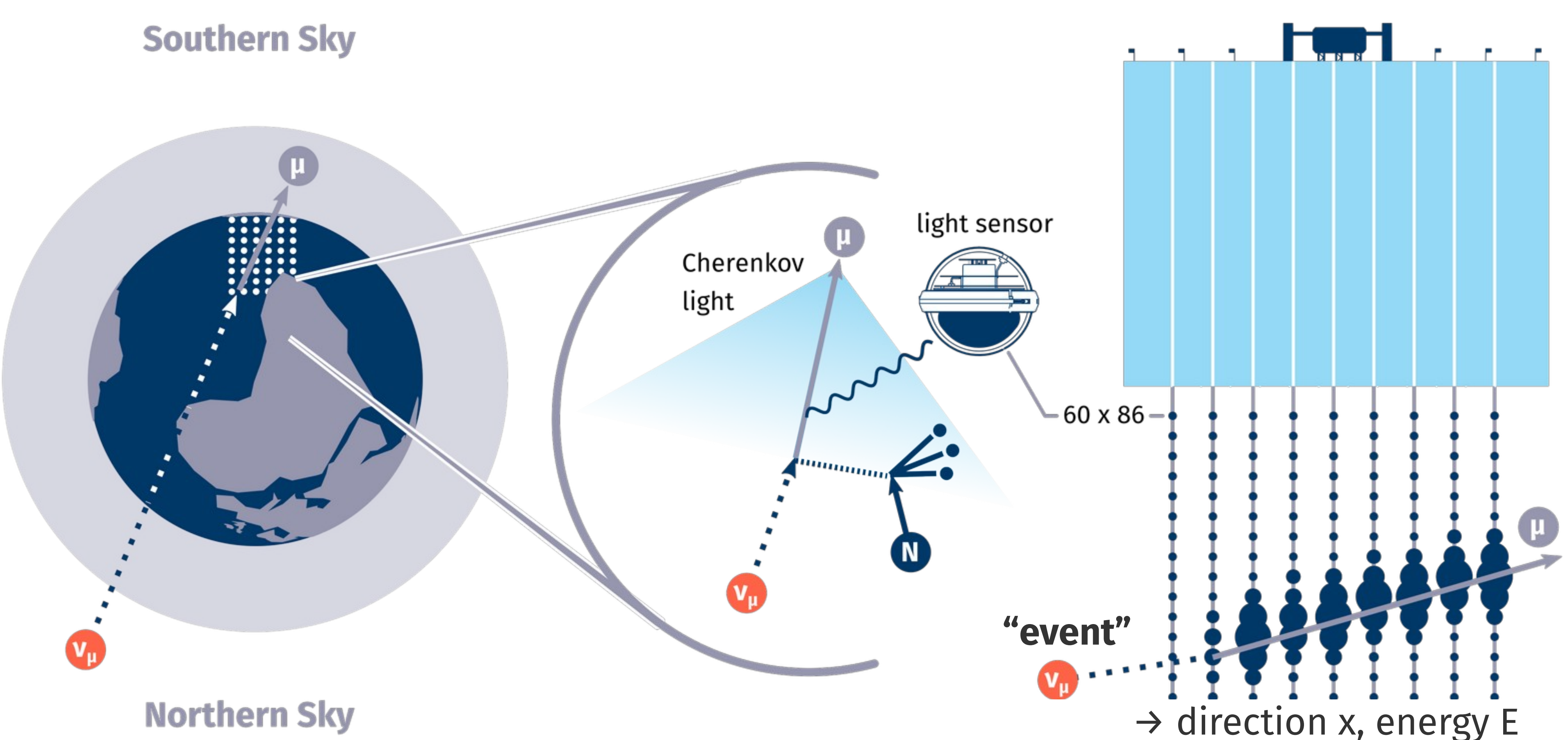
HOW TO SEARCH FOR NEUTRINO SOURCES (SIMPLIFIED VERSION)

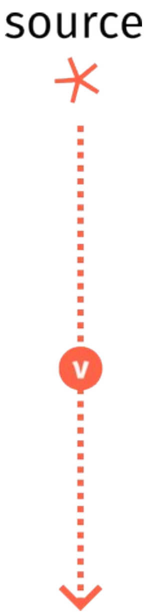


Detecting a neutrino signal with IceCube



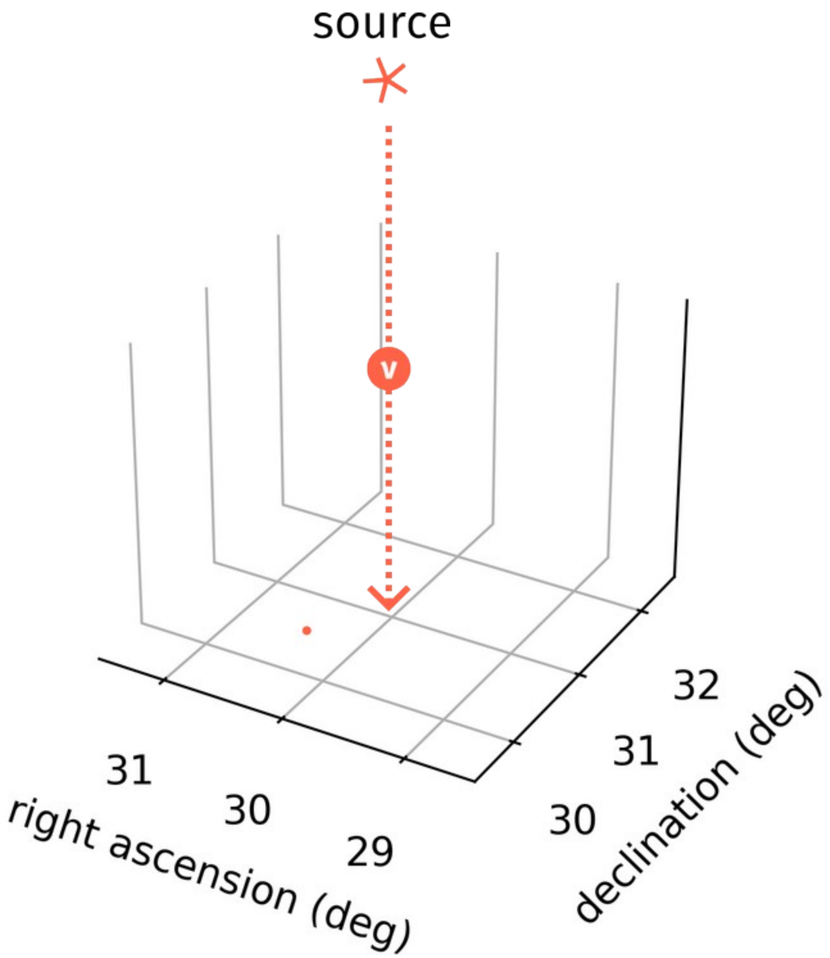
→ direction x, energy E

Arrival directions from a source



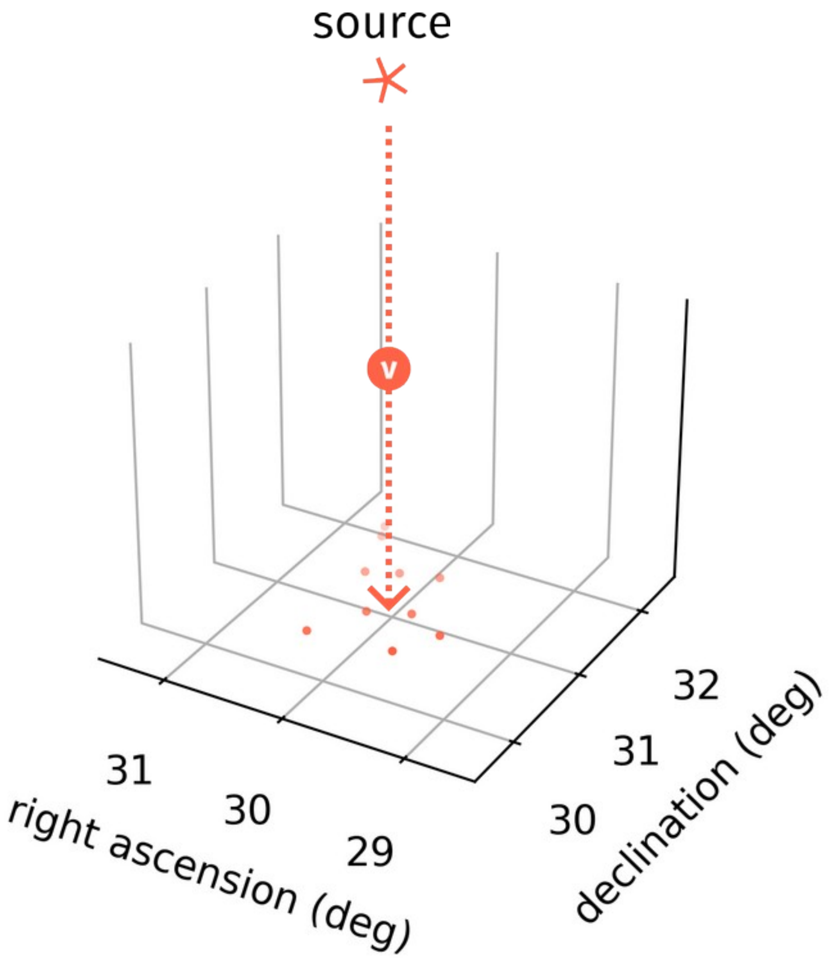
- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



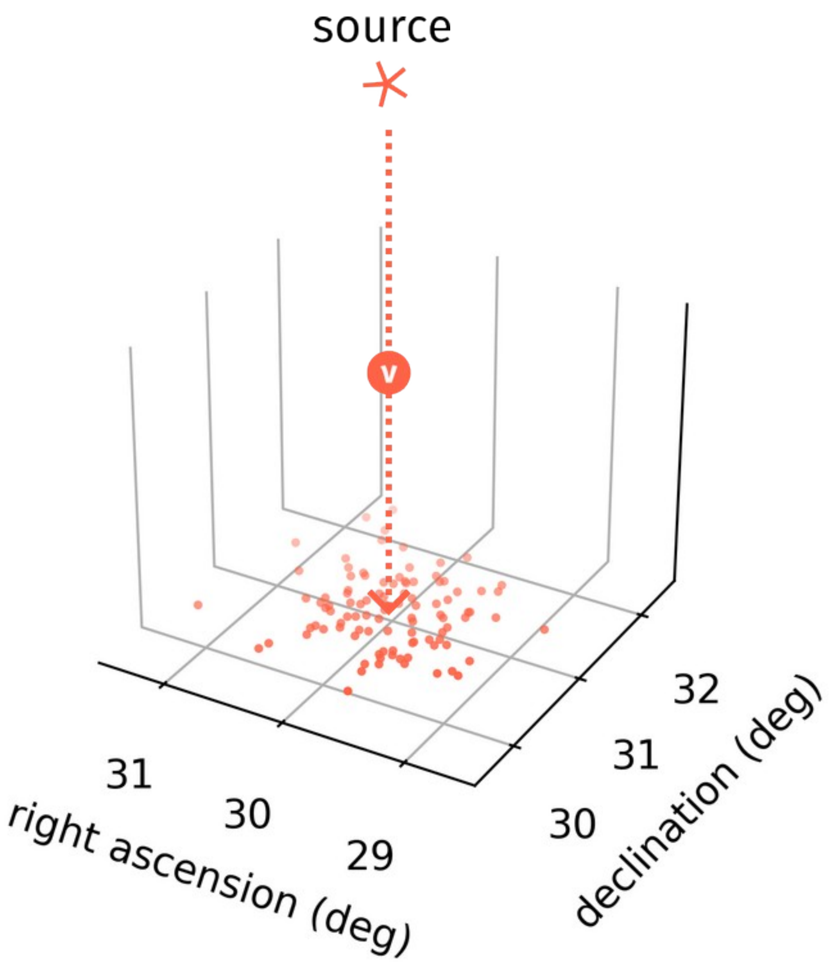
- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



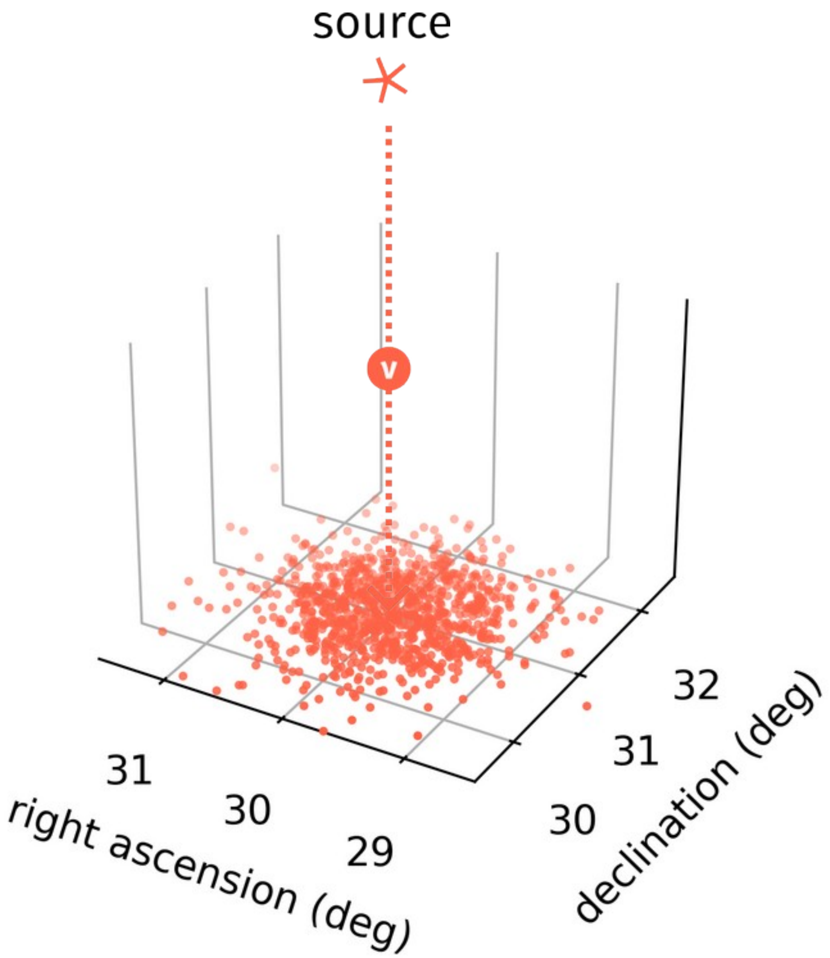
- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



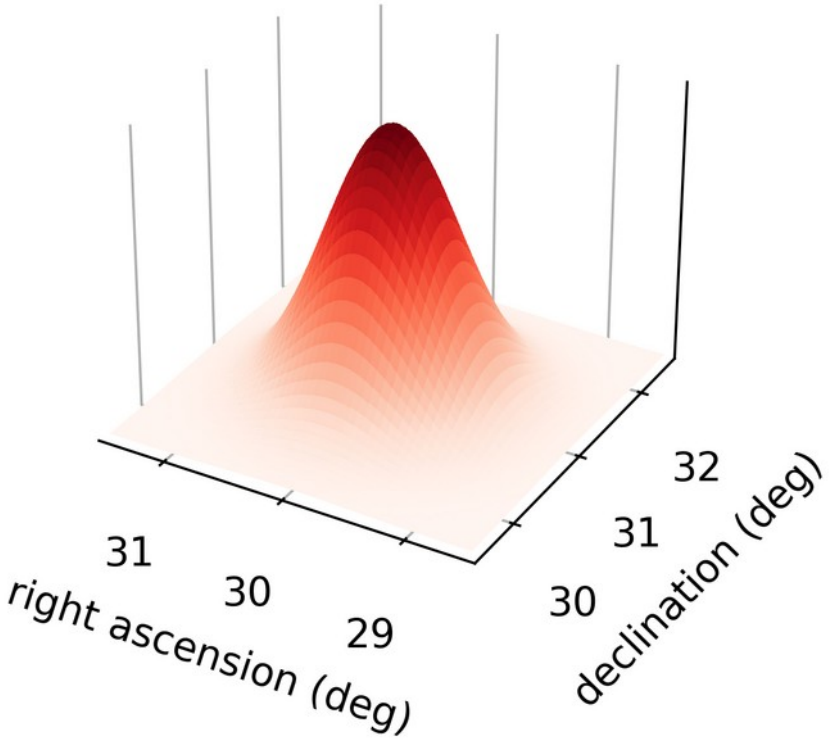
- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



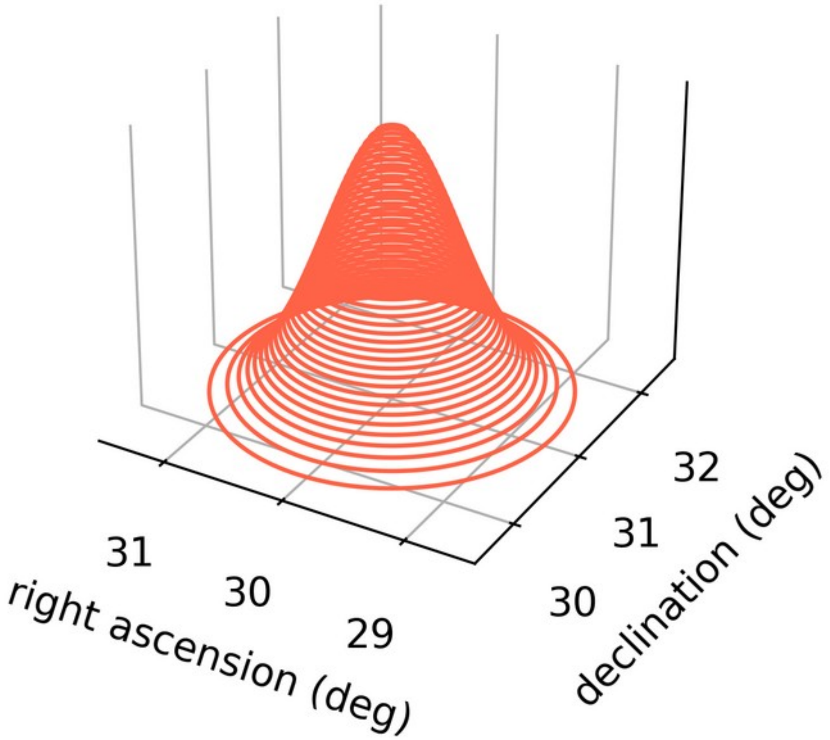
- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



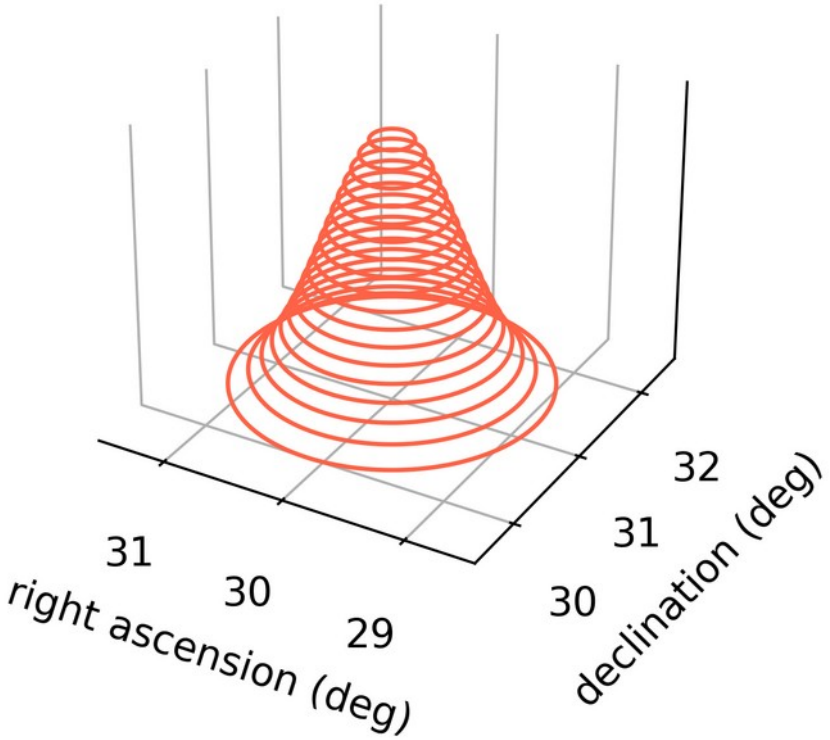
- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



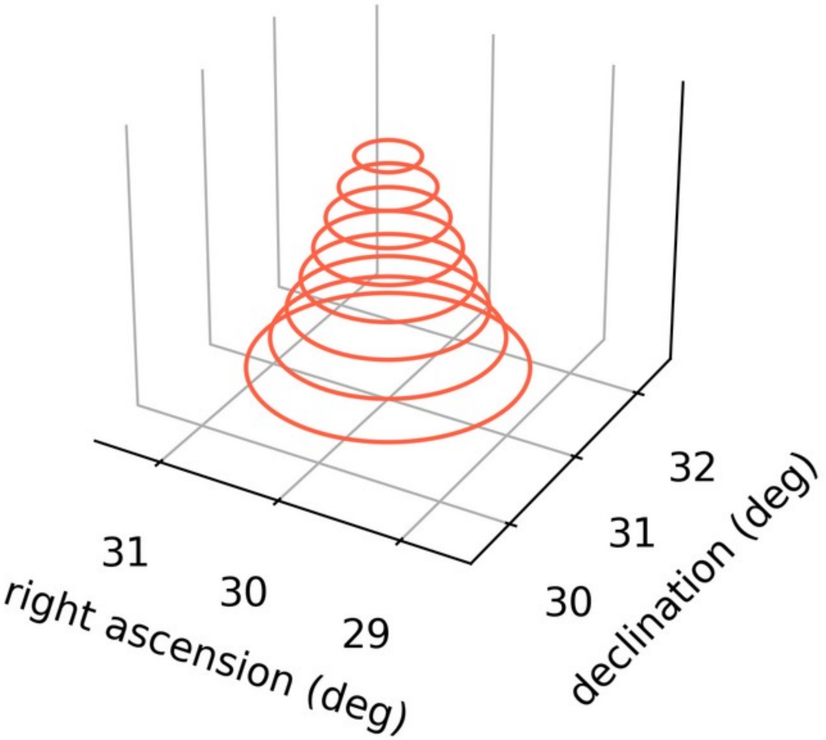
- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



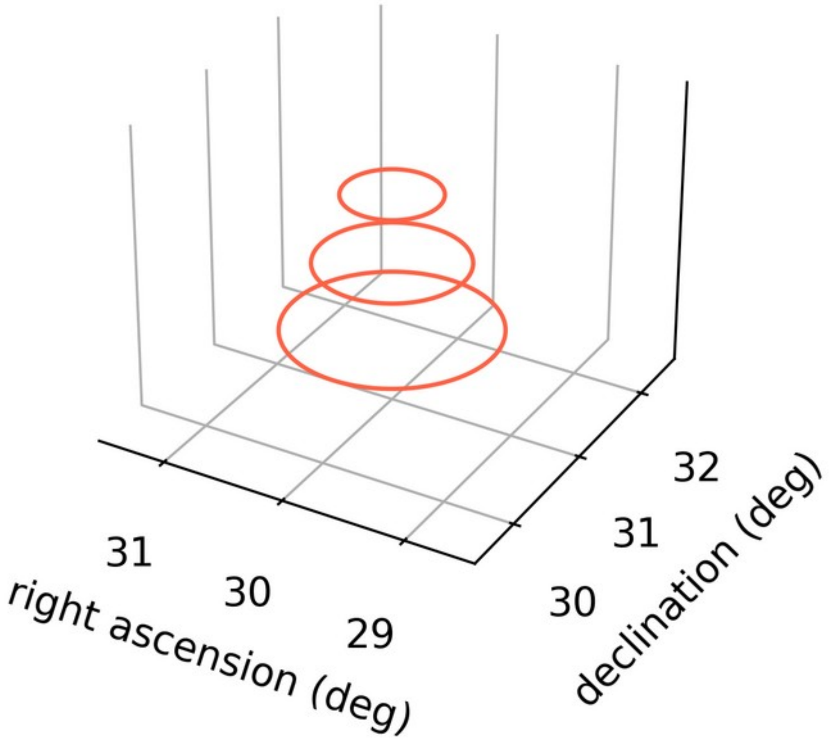
- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



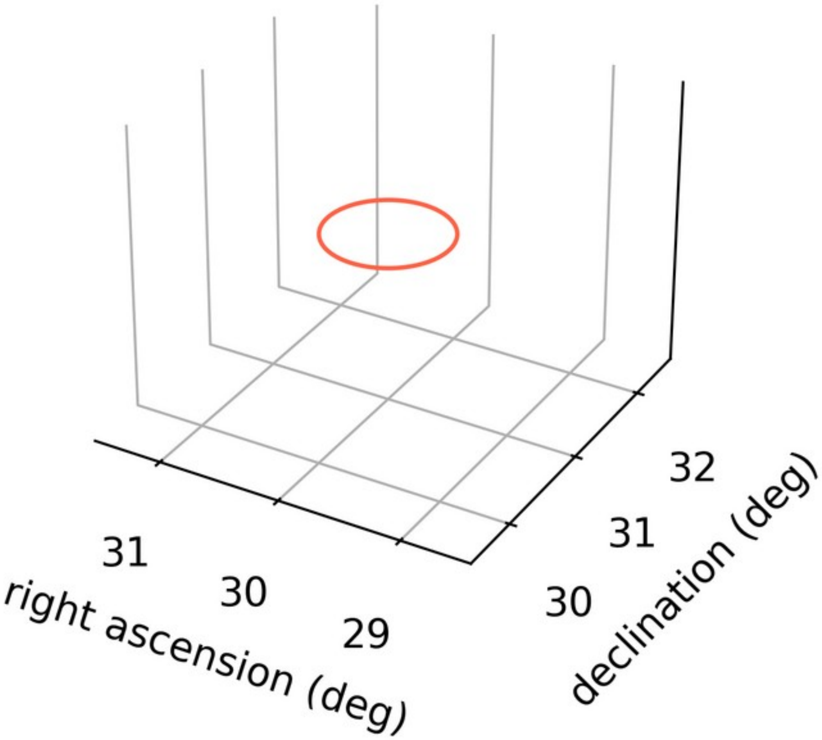
- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

Arrival directions from a source



- Simulate (neutrino) events from a source
- Reconstructed directions are random
- More neutrinos
  → denser around the source
  → density as function of direction
  → reveals the **point spread function** (PSF)
- Can be approximated with
  - Kent distribution
    (Gaussian on a sphere)
  - and a width parameter σ
    (actually per event not per source)

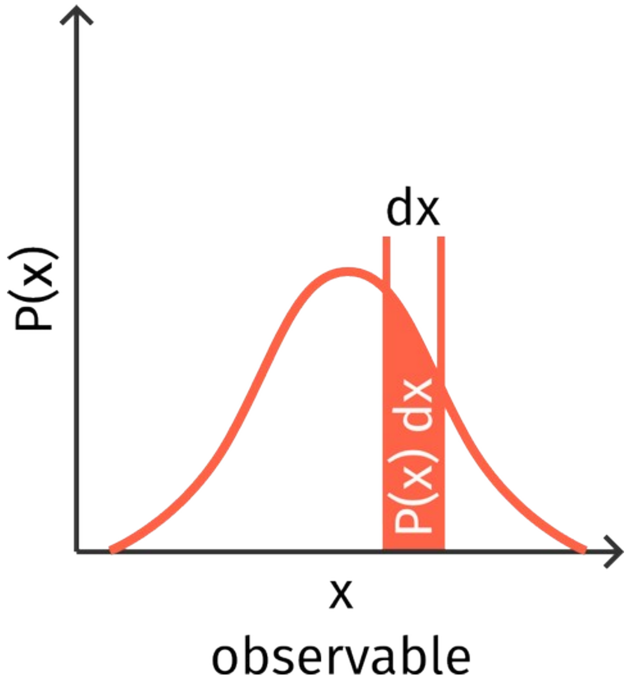
Signal PDF

- A probability density function (PDF) P
  - describes the outcome of observable x
  - (during one event)
  - Prob(“event arrives in dx”) = P(x) dx
  - ∫ P(x) dx = 1
- Normalize PSF over direction **x = (RA, dec)**

  → centre on a source

  → signal PDF **S(x)** for that source
  (probability density to be reconstructed at x)
- S(x; σ) = Kent(Ψ; σ)

  with Ψ = angle(event x, source)



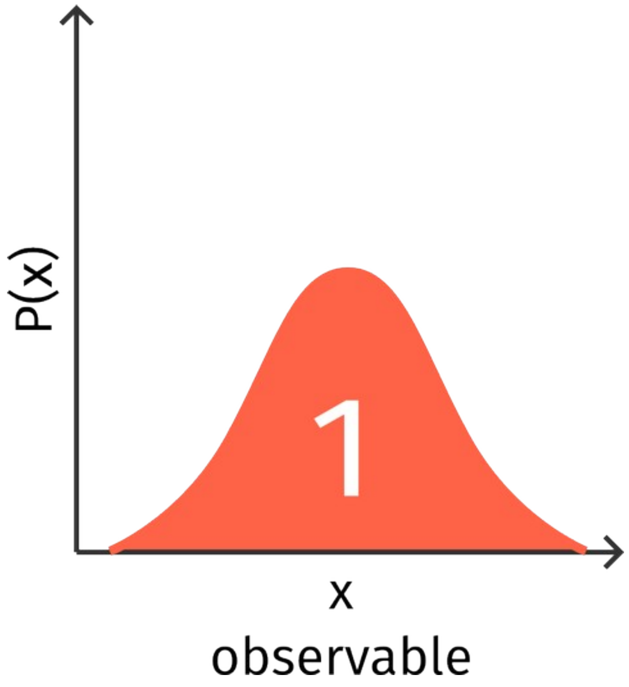
Signal PDF

- A probability density function (PDF) P
  - describes the outcome of observable x
  - (during one event)
  - Prob("event arrives in dx") = P(x) dx
  - $\int P(x)\,dx = 1$
- Normalize PSF over direction **x = (RA, dec)**

  → centre on a source

  → signal PDF **S(x)** for that source
  (probability density to be reconstructed at x)
- S(x; σ) = Kent(Ψ; σ)

  with Ψ = angle(event x, source)



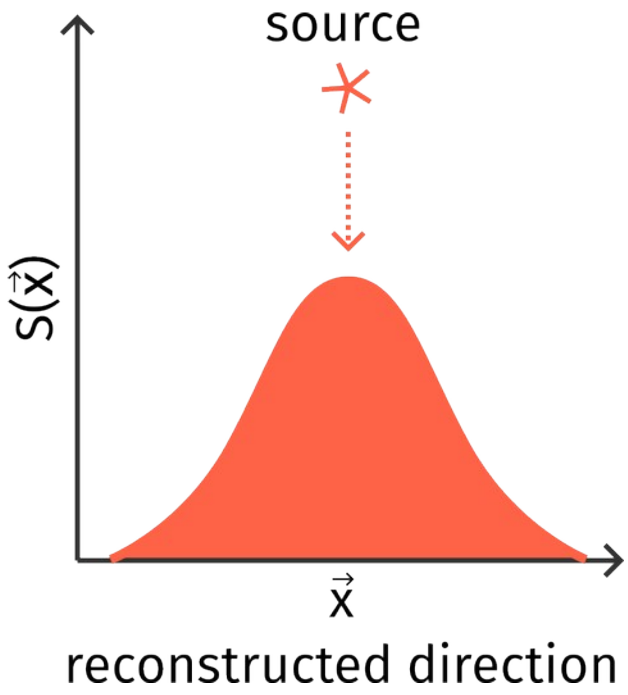
Signal PDF

- A probability density function (PDF) P
  - describes the outcome of observable x
  - (during one event)
  - Prob(“event arrives in dx”) = P(x) dx
  - ∫ P(x) dx = 1
- Normalize PSF over direction **x = (RA, dec)**

  → centre on a source

  → signal PDF **S(x)** for that source
  (probability density to be reconstructed at x)
- S(x; σ) = Kent(Ψ; σ)

  with Ψ = angle(event x, source)



Signal PDF

- A probability density function (PDF) P
  - describes the outcome of observable x
  - (during one event)
  - Prob("event arrives in dx") = P(x) dx
  - $\int P(x)\, dx = 1$
- Normalize PSF over direction **x = (RA, dec)**

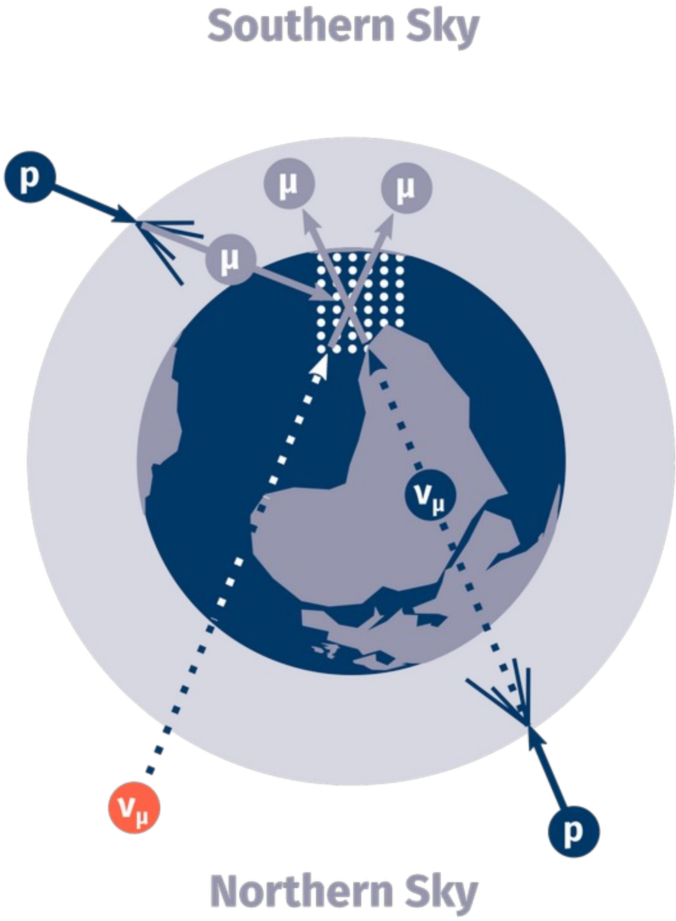
  → centre on a source

  → signal PDF **S(x)** for that source
  (probability density to be reconstructed at x)
- S(x; σ) = Kent(Ψ; σ)

  with Ψ = angle(event x, source)

Backgrounds (for neutrino telescopes)

Cosmic Rays
from all directions

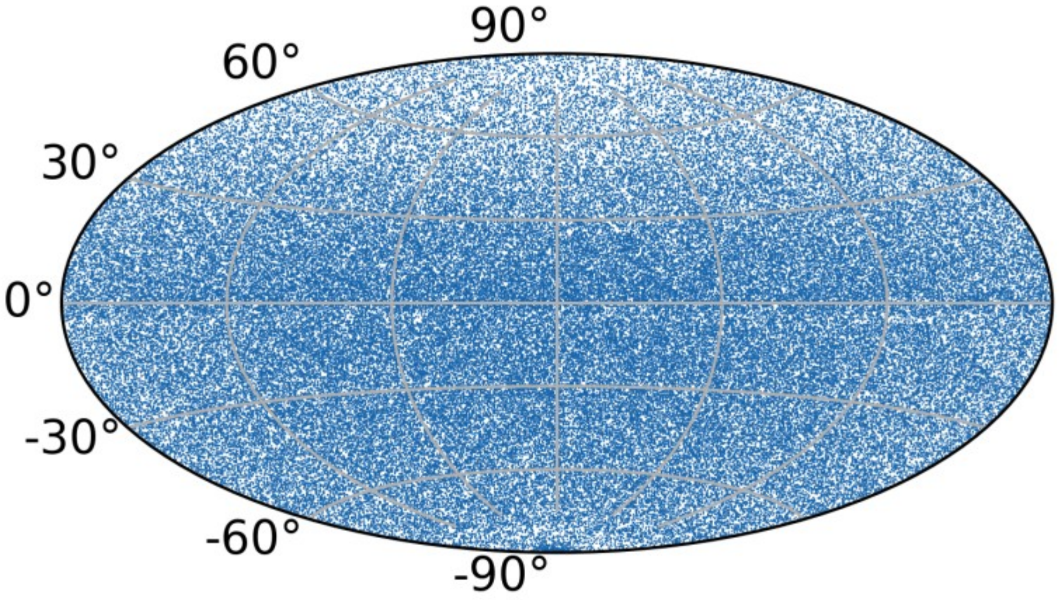


- atmospheric muons
  (can reduce with event selection)
- atmospheric neutrinos
  (look like signal events)

→ individual events look like signal
this is our **background**

- atmospheric neutrinos
- (muons blocked by the Earth)

IceCube sky

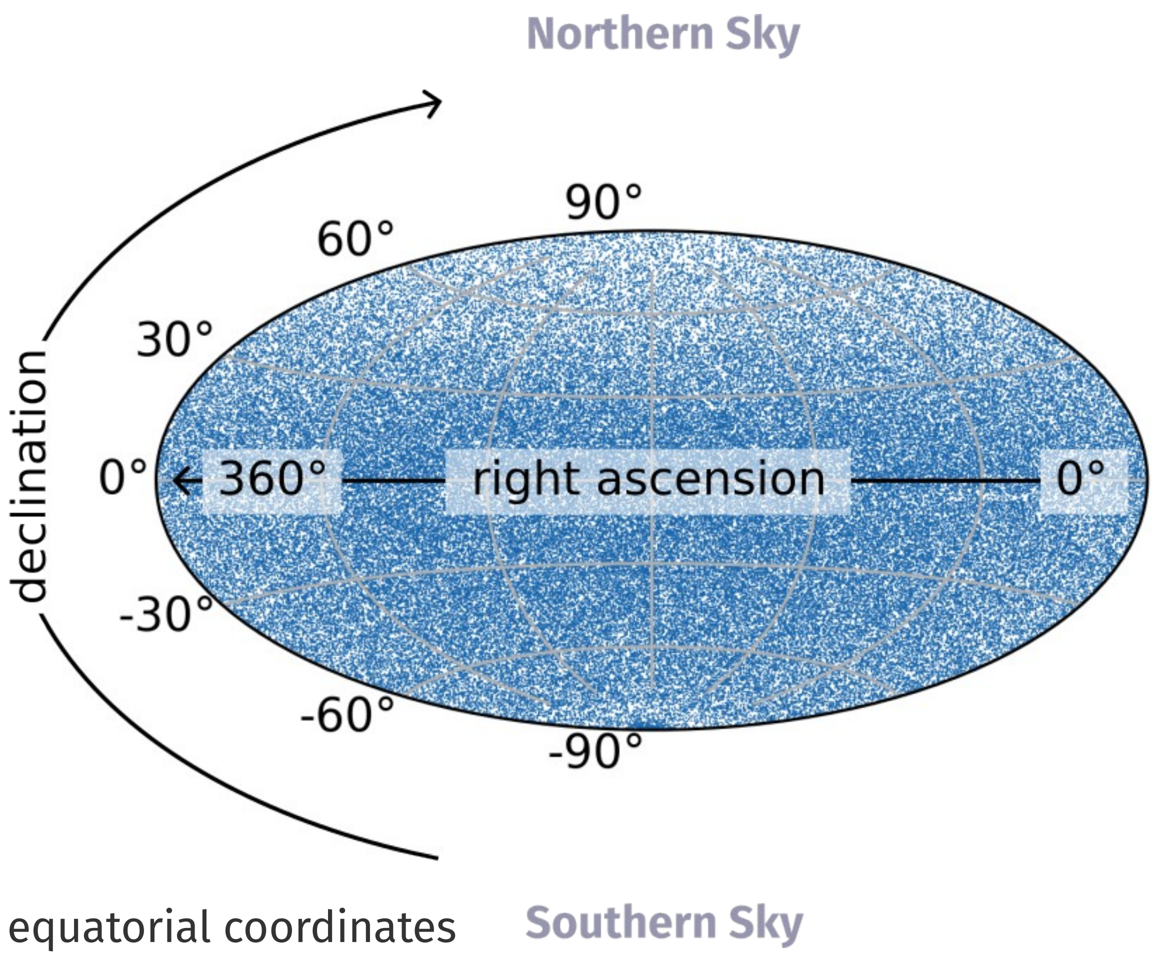


- 1 dot = 1 event
- **reconstructed** direction
- equatorial coordinates

- One year of data
- Selected to reduce muons
- Still mostly background!
- Can use data to estimate background PDFs

equatorial coordinates

IceCube sky



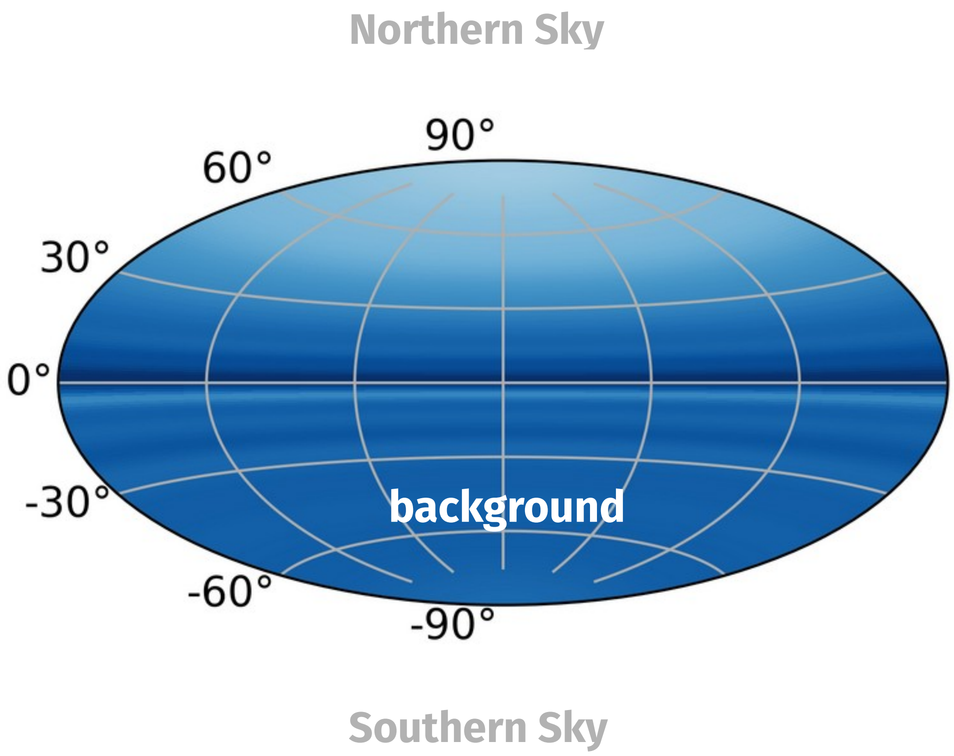
equatorial coordinates

- 1 dot = 1 event
- **reconstructed** direction
- equatorial coordinates

- One year of data
- Selected to reduce muons
- Still mostly background!
- Can use data to estimate background PDFs


Geograp ic South Pole

Background PDF

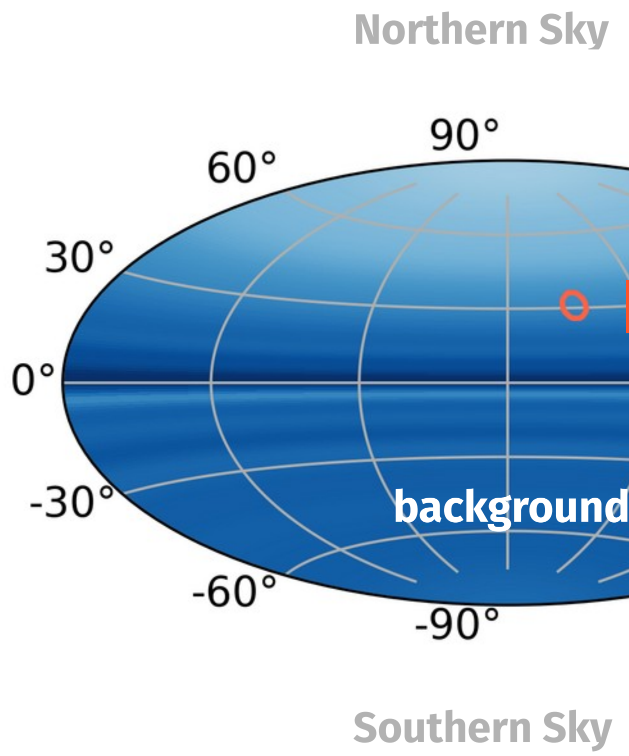


- distribution of data events in declination
- assume no dependence on right ascension (which is true for IceCube, **not** others)

  → background PDF **B(x)**
- shown here as sky map, darker = higher

Background PDF

FZU

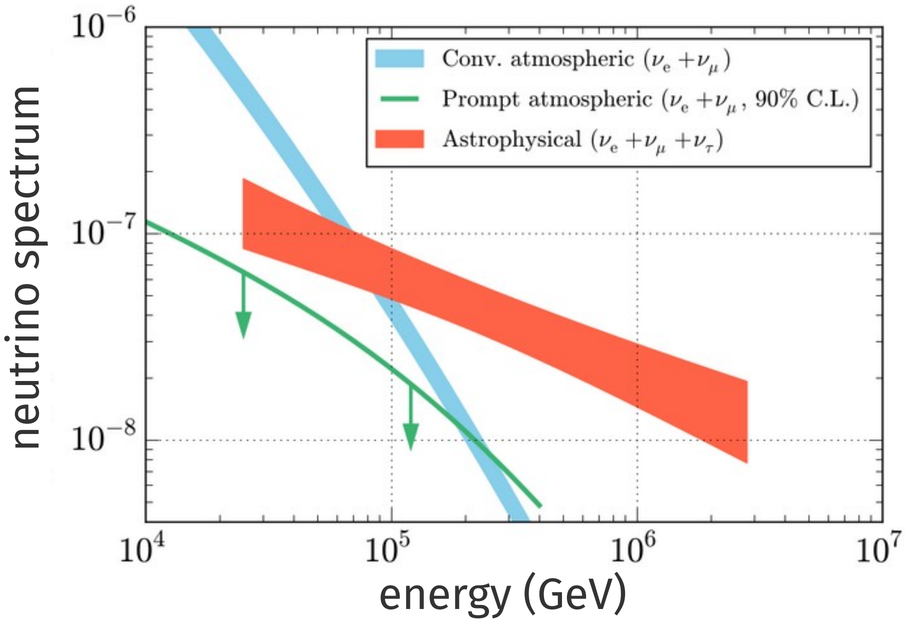


- distribution of data events in declination
- assume no dependence on right ascension (which is true for IceCube, **not** others)

  → background PDF **B(x)**
- shown here as sky map, darker = higher

Neutrino energy spectra

Low energies more atmospheric

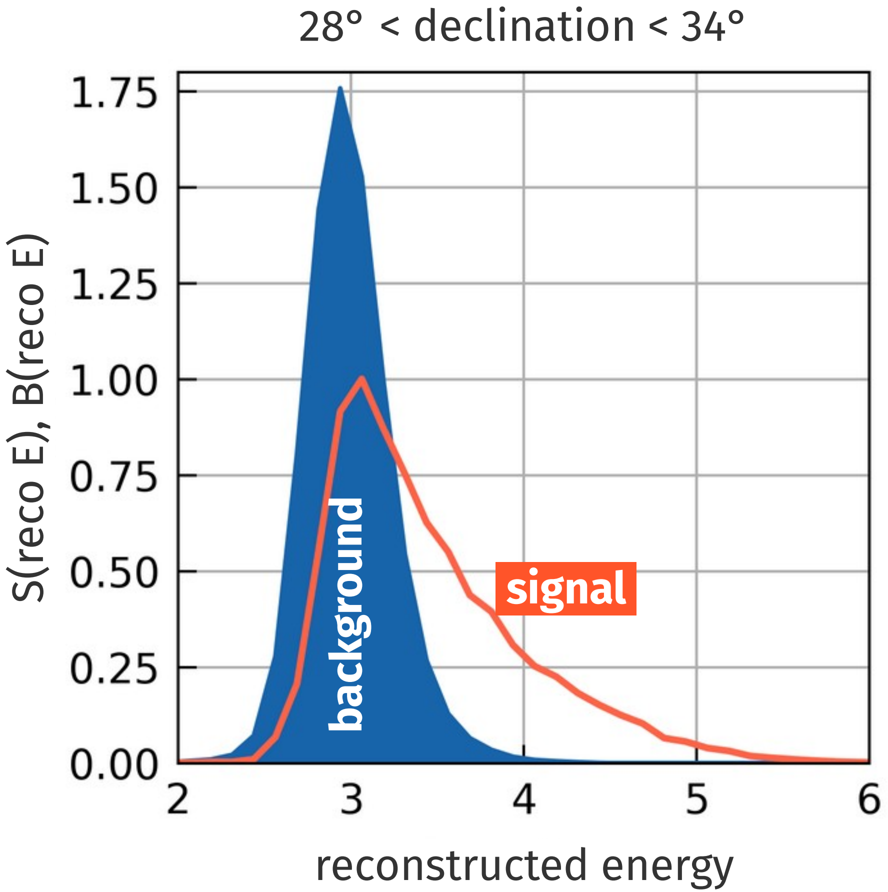


High energies more astrophysical

Reconstructed energy PDFs

- Reconstructed energy ≠ true energy
- Signal and background are different
- Example shown here

  → additional PDF parameter:
- S(x) → **S(x, E)** = S(x) S(E; declination)
- B(x) → **B(x, E)** = B(x) B(E; declination)



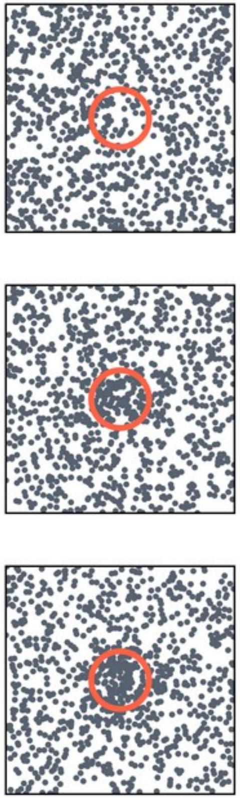
Signal + background

- Data sample of N events
- N is large → $N \approx N_{expected}$
- Our hypothesis:
  - n_S signal events
  - $(N - n_S)$ background events
- PDF to describe this data:

$$\left(\frac{n_S}{N}\right) S(\vec{x}, E) + \left(1 - \frac{n_S}{N}\right) B(\vec{x}, E)$$

- Given the data, how can we tell if $n_S > 0$?

N = 1000



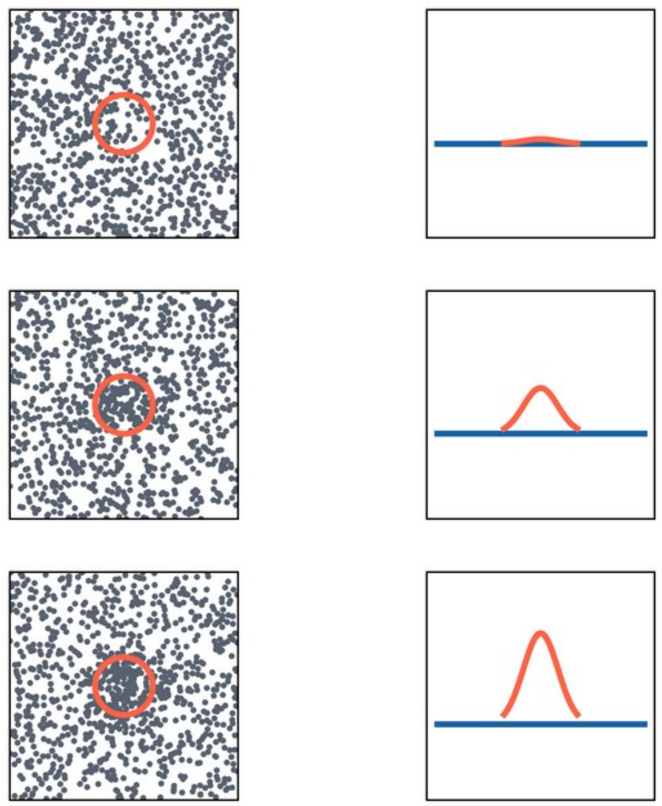
Signal + background

- Data sample of N events
- N is large → $N \approx N_{expected}$
- Our hypothesis:
  - n_S signal events
  - $(N - n_S)$ background events
- PDF to describe this data:

$$\left(\frac{n_S}{N}\right) S(\vec{x}, E) + \left(1 - \frac{n_S}{N}\right) B(\vec{x}, E)$$

- Given the data, how can we tell if $n_S > 0$?

N = 1000



$n_S = 10$

$n_S = 100$

$n_S = 200$

Ideas?

Likelihood

- Mathematicians came up with this recipe:

  1. include the unknown parameters
  of the hypothesis in a PDF

  2. use PDF & data to define a “likelihood”
  (“probability to obtain this data”)

  3. find parameter values where L is largest
  → that is the “maximum likelihood estimate”

$$L(\vec{a}) = \prod_{i=1}^{N} P(\vec{x}_i; \vec{a})$$

- In our case:

  1. Use the “signal + background” P(x, E)

  2. Multiply one term per IceCube event

  3. Find $n_S \geq 0$ where L is largest → $\hat{n}_S$

  4. Interpret the result?

$$L(n_S) = \prod_{i=1}^{N} \left[\left(\frac{n_S}{N} \right) S(\vec{x}_i, E_i) + \left(1 - \frac{n_S}{N} \right) B_i(\vec{x}_i, E_i) \right]$$

$$L(\hat{n}_S) \geq L(n_S) \forall n_S > 0$$

Test statistic

- Easier to calculate sums than products
  → logarithm of the likelihood

  → complete to Wilks' test statistic
- Properties:
  - TS ≥ 0
  - the higher the TS, the more signal
  - If $\hat{n}_S = 0$ → TS = 0
- But can get TS > 0 from pure background!

$$\mathrm{TS} = 2\log\frac{L(\hat{n}_S)}{L(n_S = 0)}$$

$$L(n_S = 0) = \prod_{i=1}^{N} B(\vec{x}_i, E_i)$$

$$\Rightarrow \mathrm{TS} = 2\sum_{i=1}^{N}\log\left[\left(\frac{n_S}{N}\right)\frac{S}{B}(\vec{x}_i, E_i) + \left(1 - \frac{n_S}{N}\right)\right]$$

Background fluctuations

- Background can fluctuate to TS > 0
- How can we quantify it?
- Need to simulate that we had multiple trials
- IceCube: randomize right ascension ("scrambling")

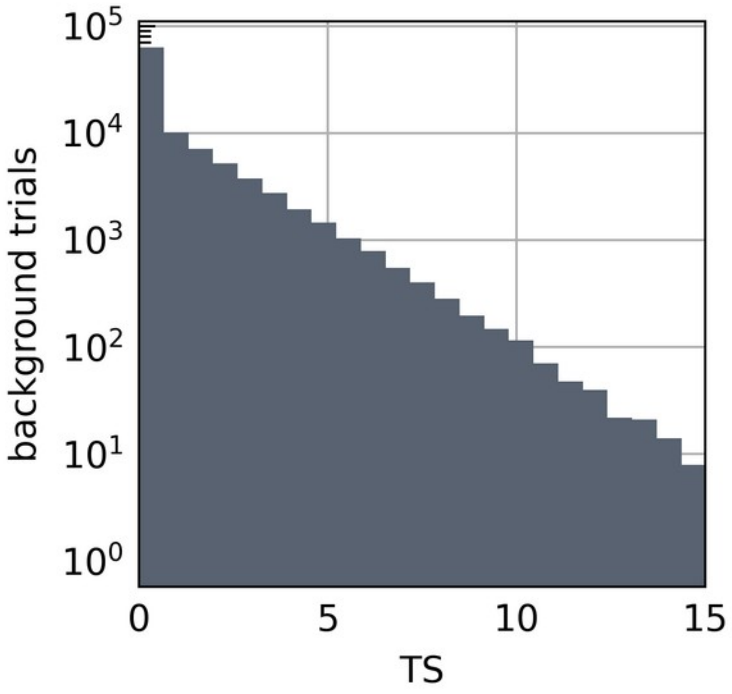
  → different events end up near the source

  → these can not come from the source

  → equivalent to a pure-BG event sample (same detector and event selection)
- For each trial:

  → maximize TS for each "trial"

  → repeat to reveal TS distribution



Background fluctuations

- Background can fluctuate to TS > 0
- How can we quantify it?
- Need to simulate that we had multiple trials
- IceCube: randomize right ascension ("scrambling")

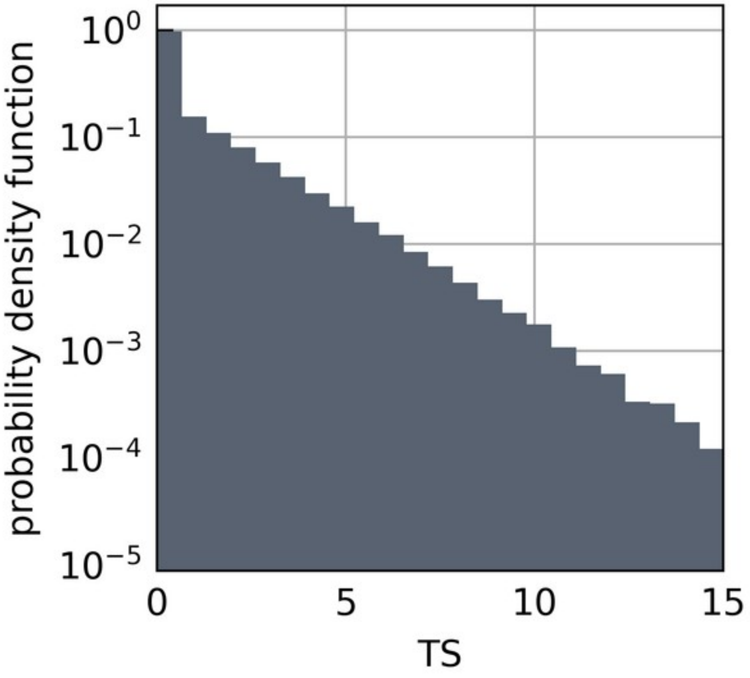
  → different events end up near the source

  → these can not come from the source

  → equivalent to a pure-BG event sample (same detector and event selection)
- For each trial:

  → maximize TS for each "trial"

  → repeat to reveal TS distribution



Significance and p-values

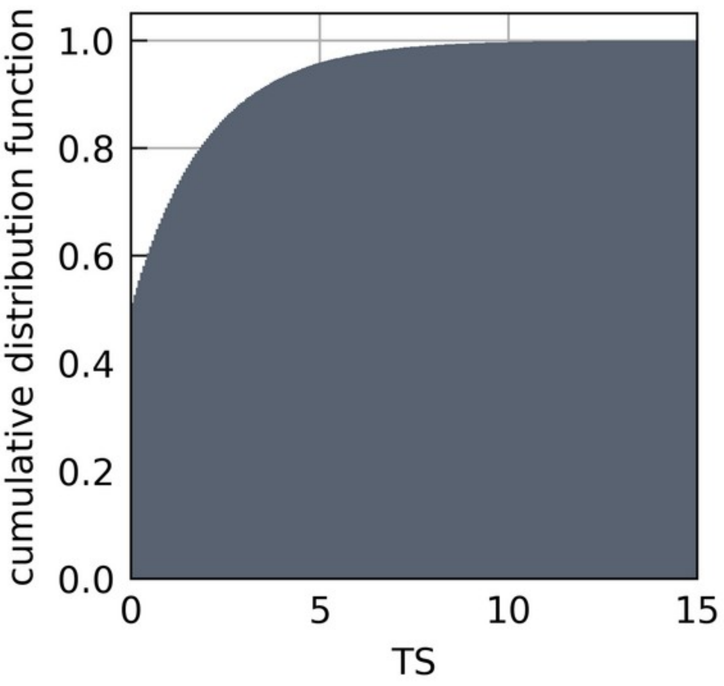
- Reject "pure BG" hypothesis for TS > threshold
- Probability to **falsely** reject it is the significance **α**
- A common (but arbitrary) choice:
  - $\alpha = 2.87 \times 10^{-7}$
  - "5 sigma"
- Set threshold = real data TS

  → combine with background trials

  → compute survival function

  → **p-value**
- Most of the time you don't have such a clear result
- So instead we try to interpret

$$\mathrm{CDF(TS)} = \int_0^{\mathrm{TS}} \mathrm{PDF(TS')}\,d\mathrm{TS'}$$



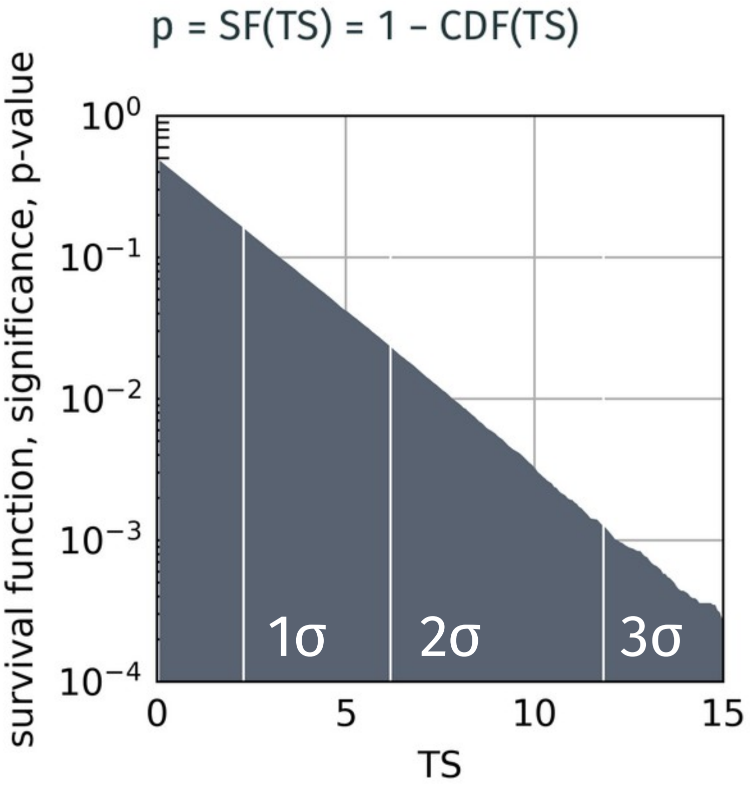
Significance and p-values

- Reject “pure BG” hypothesis for TS > threshold
- Probability to **falsely** reject it is the significance **α**
- A common (but arbitrary) choice:
  - $\alpha = 2.87 \times 10^{-7}$
  - “5 sigma”
- Set threshold = real data TS

  → combine with background trials

  → compute survival function

  → **p-value**
- Most of the time you don’t have such a clear result
- So instead we try to interpret



Do not be overimpressed by the attention and mathematical detail that goes into significance. This is largely due to the fact that it is under control and mathematically tractable. The power is usually far more important, but you can often say very little about it as the alternative hypothesis tends to be vague and unquantifiable. When overawed by a discussion of the significance of some test, it is salutary to remember that you can make a valid but meaningless significance test by choosing a random number r between 0 and 1 and rejecting the hypothesis if r is less than α.

R. J. Barlow, “Statistics – A Guide to the Use of Statistical Methods in the Physical Sciences”

Less verbose: the work does not end at the p-value.

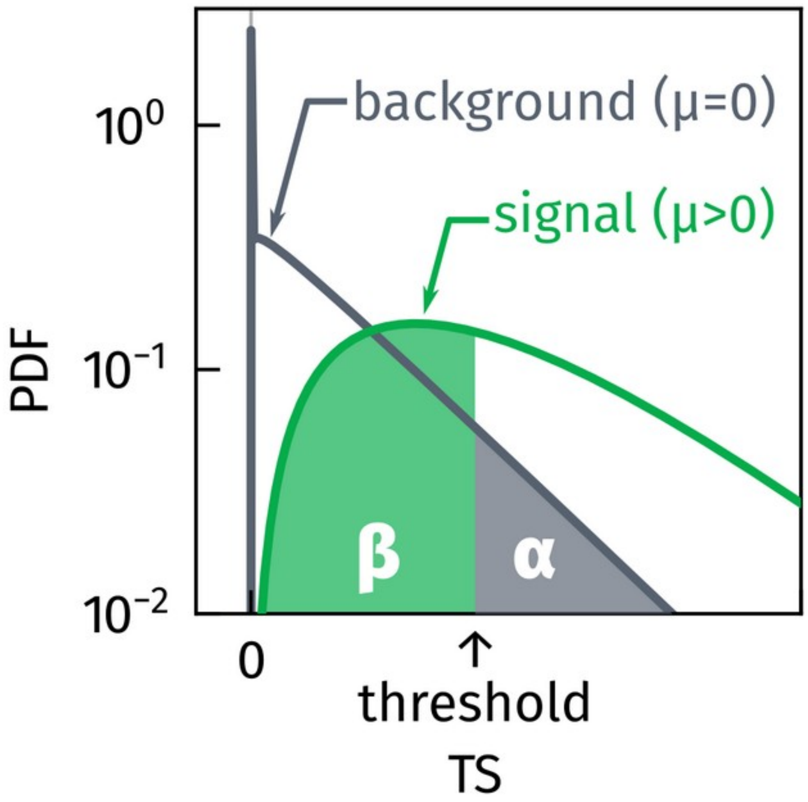
Background + signal trials

- Generate fake event samples:

  background + signal × **μ**

  i.e. scrambled data + simulation of signal
- Maximize test statistic for each event sample
  → distributions of TS given μ

  → measure power **β** of a hypothesis test

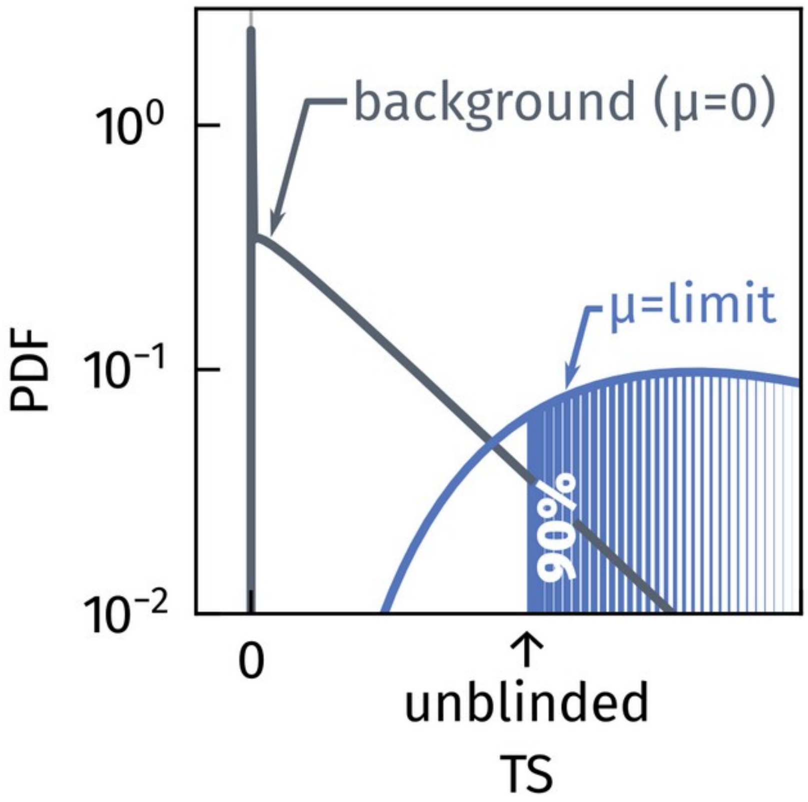


Upper limits

- Frequentist upper limit (for our case):

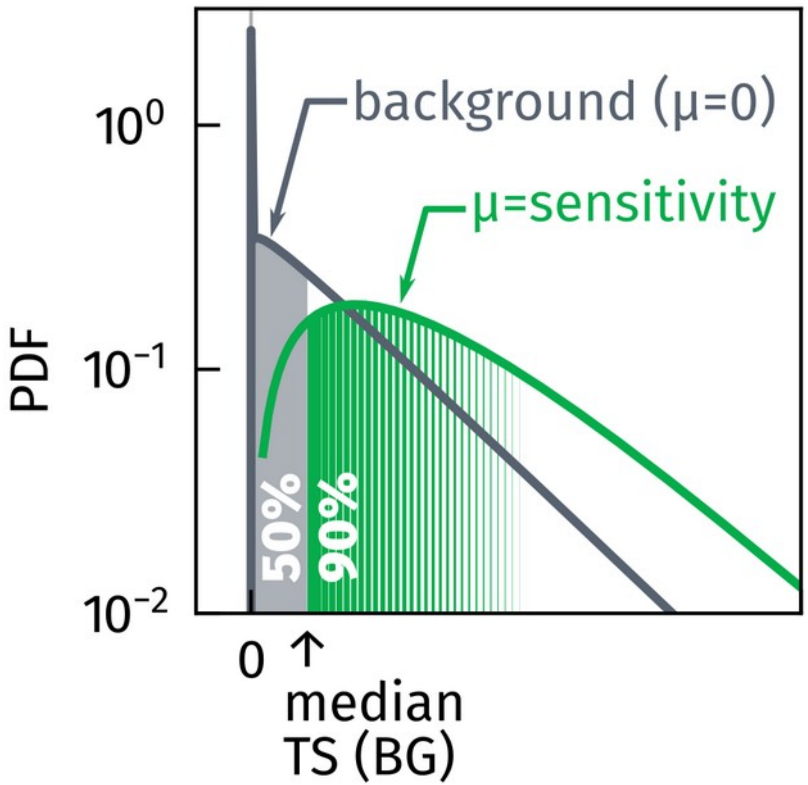
  "If the true μ exceeds the upper limit, the probability to get a smaller TS than observed is 10% or less"

- Produce signal trials for a "$\mu^{(90)}$"
- where 10th percentile of TS = TS(data)
- "$\mu^{(90)}$" = the upper limit
  - on that hypothesis
  - at 90% "confidence level" (C.L.)



Sensitivities

- If we could repeat the experiment,
- and there really was no signal,
- → sensitivity = median limit obtained
- Easy to construct in Neyman:
- Find “$\mu^{(90)}$” where 90% of TS > median TS(μ=0)
- Characterizes the analysis
  → can develop it blindly



Stacking

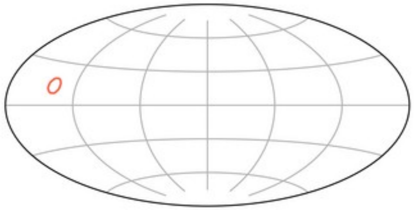
- Signal still indiscernable?

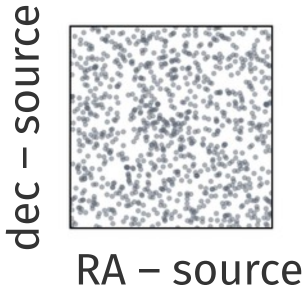
  → Collect more events

  - Wait for more data

    or nearly equivalently:

  - Combine multiple sources





- In mathematical terms:
  - add more PDFs
  - single signal PDF replaced by weighted sum

  → can become clearer

  (for the right choice of sources and weights)

$$S(\vec{x}, E)$$

Stacking

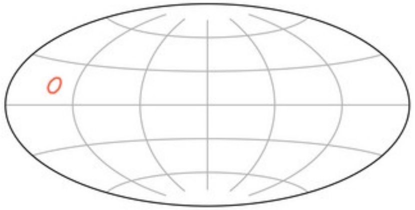
- Signal still indiscernable?

  → Collect more events

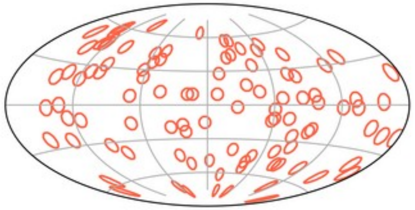
  - Wait for more data

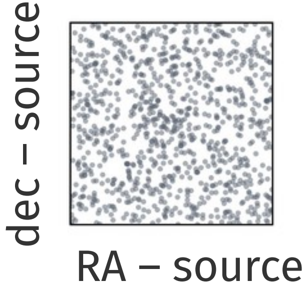
    or nearly equivalently:

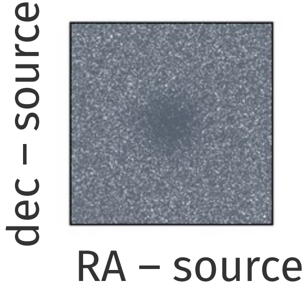
  - Combine multiple sources



→







- In mathematical terms:
  - add more PDFs
  - single signal PDF replaced by weighted sum

  → can become clearer

  (for the right choice of sources and weights)

$$S(\vec{x}, E) \rightarrow \sum_{k \in \text{sources}} w_k S_k(\vec{x}, E)$$

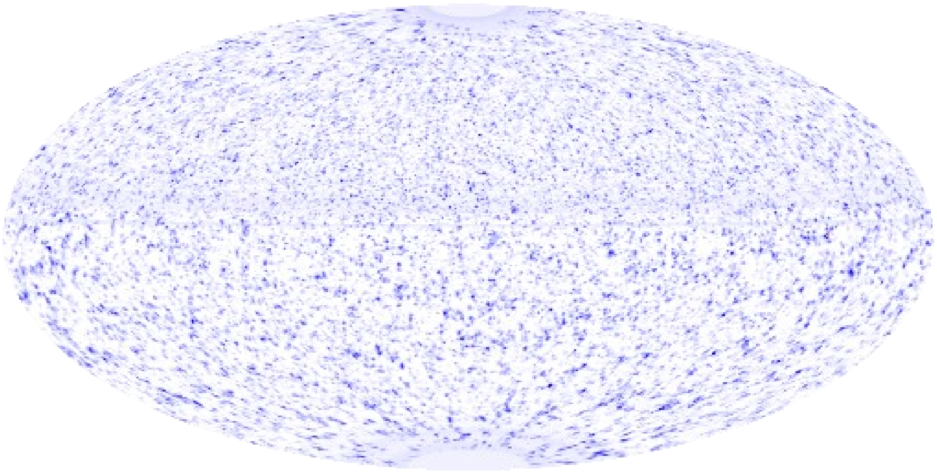
$$\sum_{k \in \text{sources}} w_k = 1$$

All sky scan

- Could also make fewer assumptions
  - no choice of source
  - no choice of source class
- Try all directions, calculate p-value each
- Getting back to a single result:

  → choose smallest p_{min}
- Repeat background trials with this extra step
- Probability to obtain a lower p_{min}

  → p_{post}, larger
- multiplying a trial factor – the price we pay



All sky scan

- Could also make fewer assumptions
  - no choice of source
  - no choice of source class
- Try all directions, calculate p-value each
- Getting back to a single result:

  → choose smallest p_{min}
- Repeat background trials with this extra step
- Probability to obtain a lower p_{min}

  → p_{post}, larger
- multiplying a trial factor – the price we pay