Practical X/γ-ray spectroscopy
(basic analysis techniques in high-energy astrophysics)

Mariano Mendez

University of Groningen, The Netherlands

11th Cospar Capacity-building Workshop, Bangalore, India, 2010.

Practical X/γ-ray spectroscopy

Most X-ray/γ-ray instruments have low or moderate spectral resolution, typically $R = E/\Delta E$ ~ 3 – 20.

Compare with typical spectral resolution in optical spectrographs, R ~ 2000 – 5000 (up to R ~70000 possible).

However, band-passes of high-energy instruments extend over a factor ~30 – 100 in energy.

Fermi LAT extends over a factor ~15000 In energy! Compare with typical values of only ~3 – 5 in the optical.

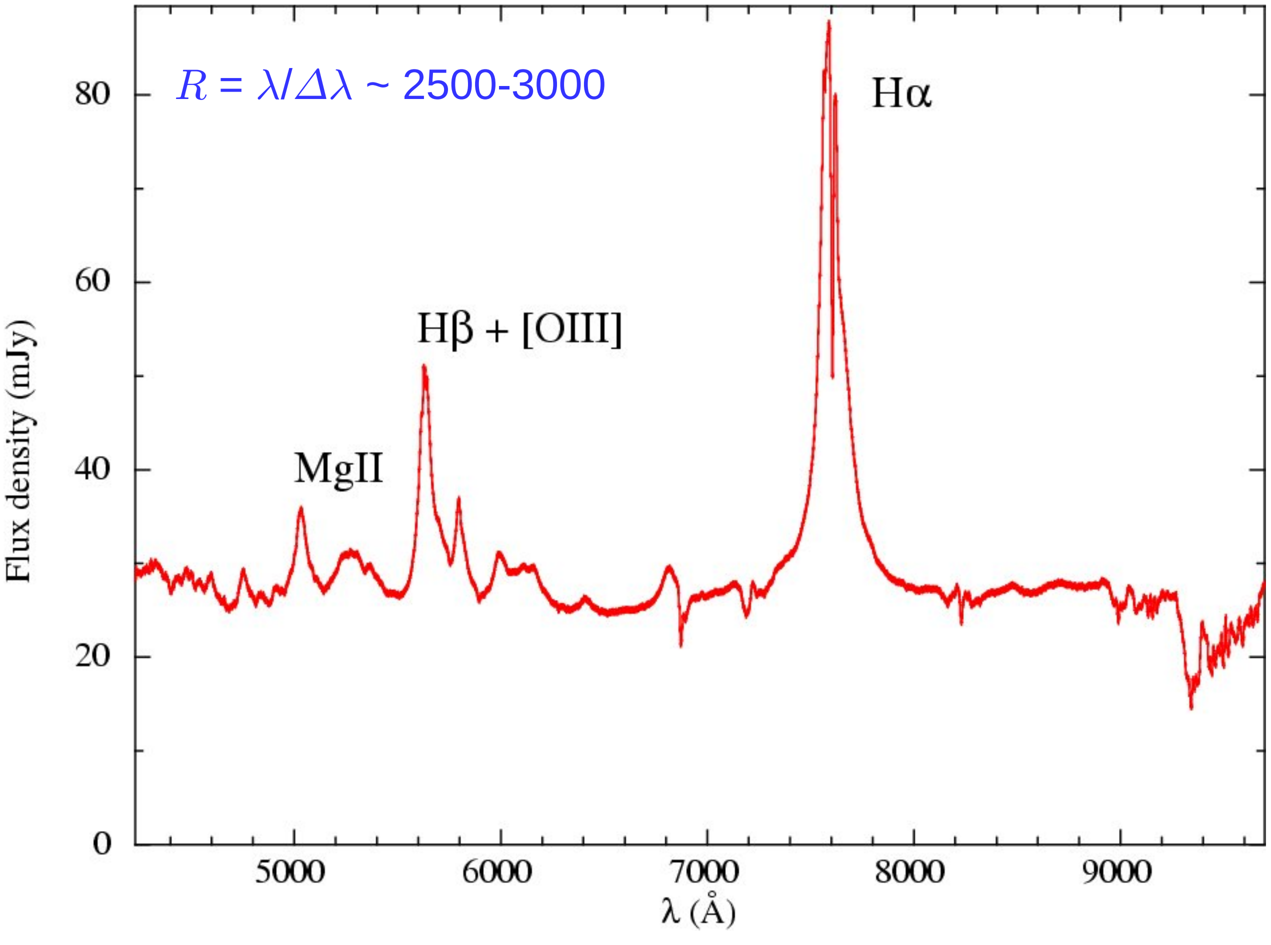
Practical X/γ-ray spectroscopy

Differently from what happens in the optical, the continuum shape of the spectrum often provides important physical information (cf. lectures on radiative processes).

Therefore, unlike in the optical, most uses of high-energy spectra usually involve the simultaneous analysis of the entire spectrum rather than an attempt to measure individual line strengths.

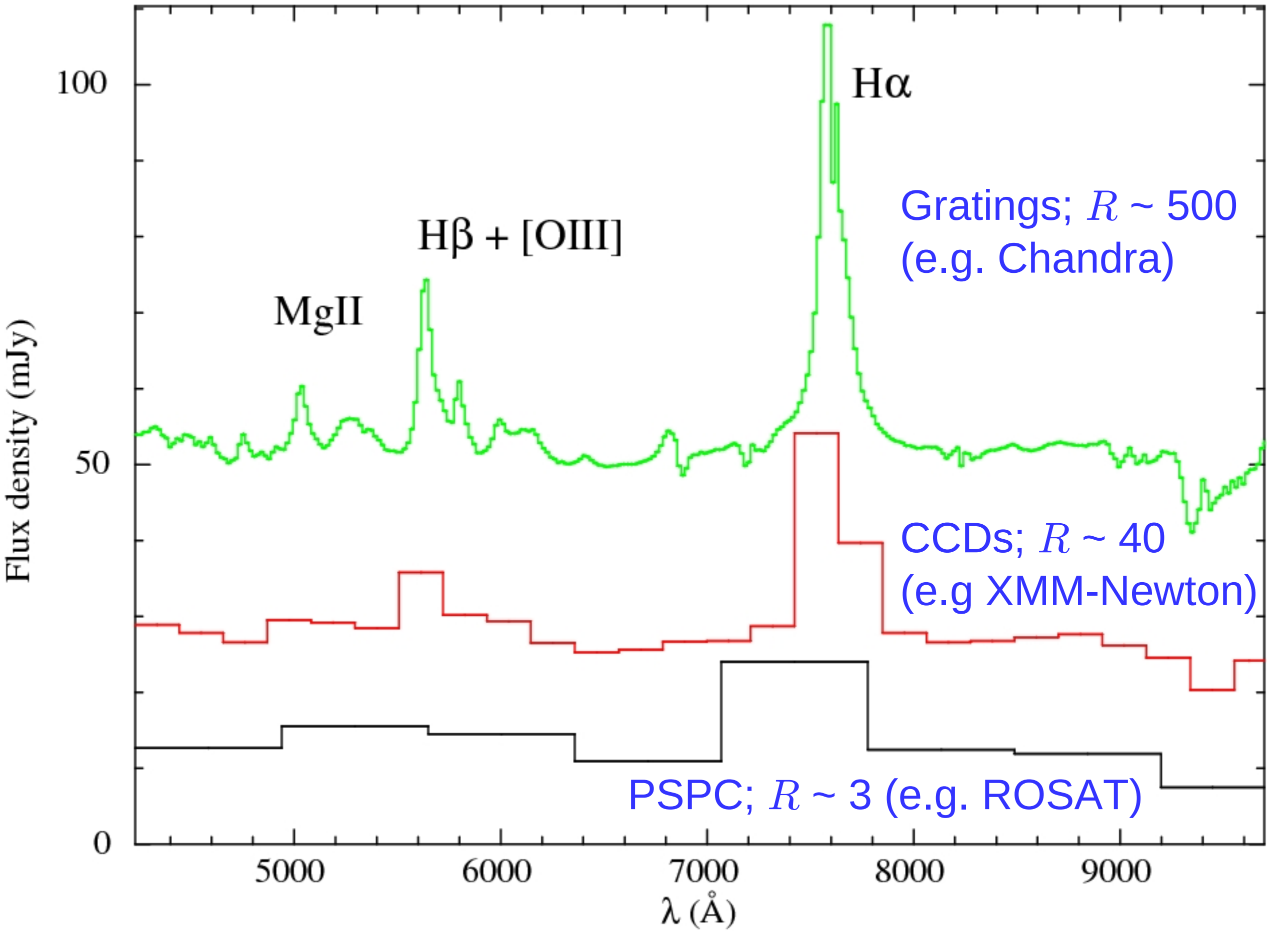
X-ray spectroscopy compared to optical spectroscopy

3C 273 Optical Spectrum (NTT/EMMI)



X-ray spectroscopy compared to optical spectroscopy

3C 273 Optical Spectrum with typical X-ray instruments



Practical X/γ-ray spectroscopy

The limited spectral resolution of the X-ray/γ-ray detectors means that there is a non-negligible probability that a photon of energy E that enters the telescope is assigned an energy $E' \neq E$ within a certain energy range around E.

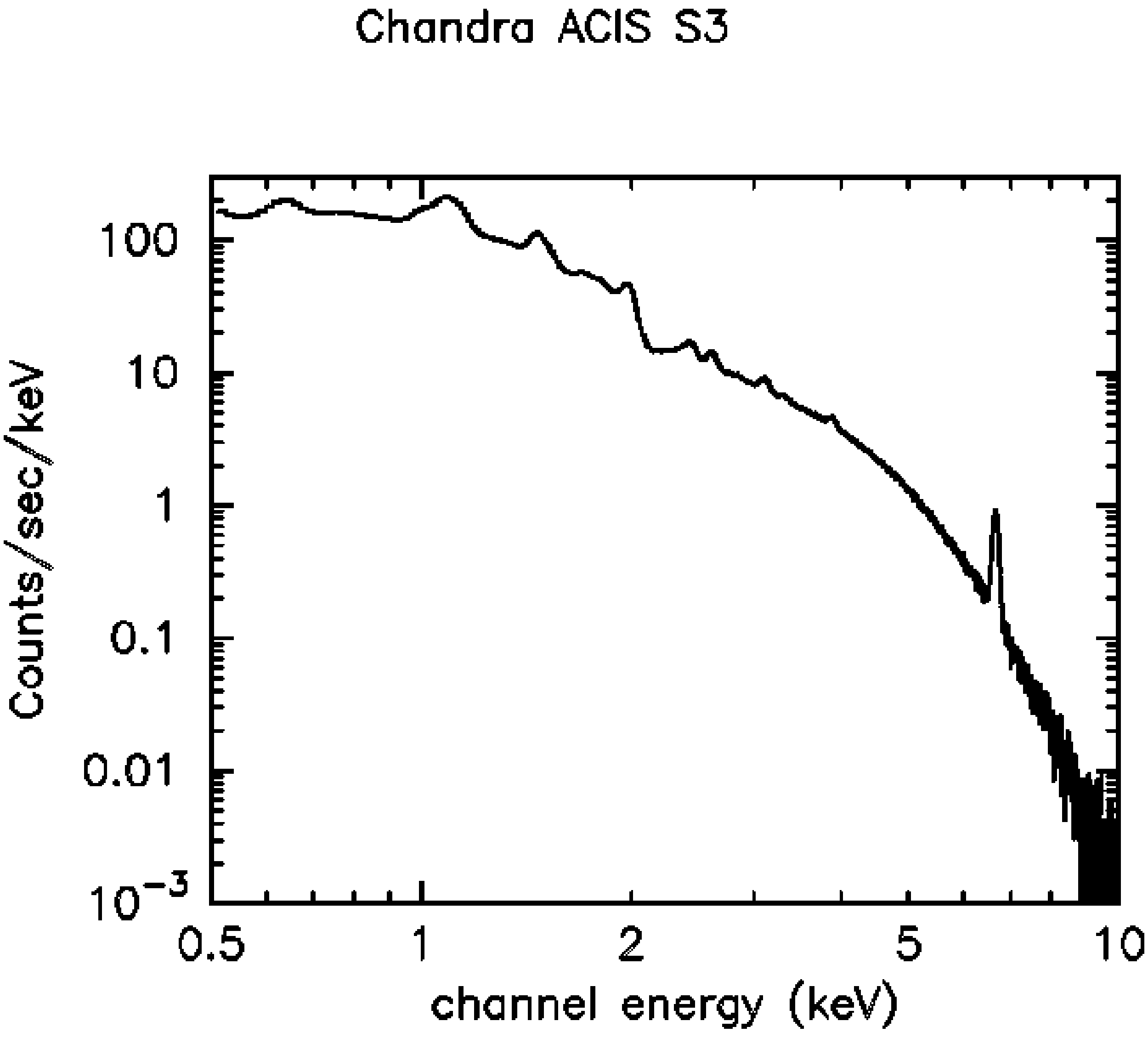
This can be expressed in terms of a function, $R(E,E')$, that gives the probability that a photon of energy E is assigned an energy E' in the detector.

The Basic Problem (1)

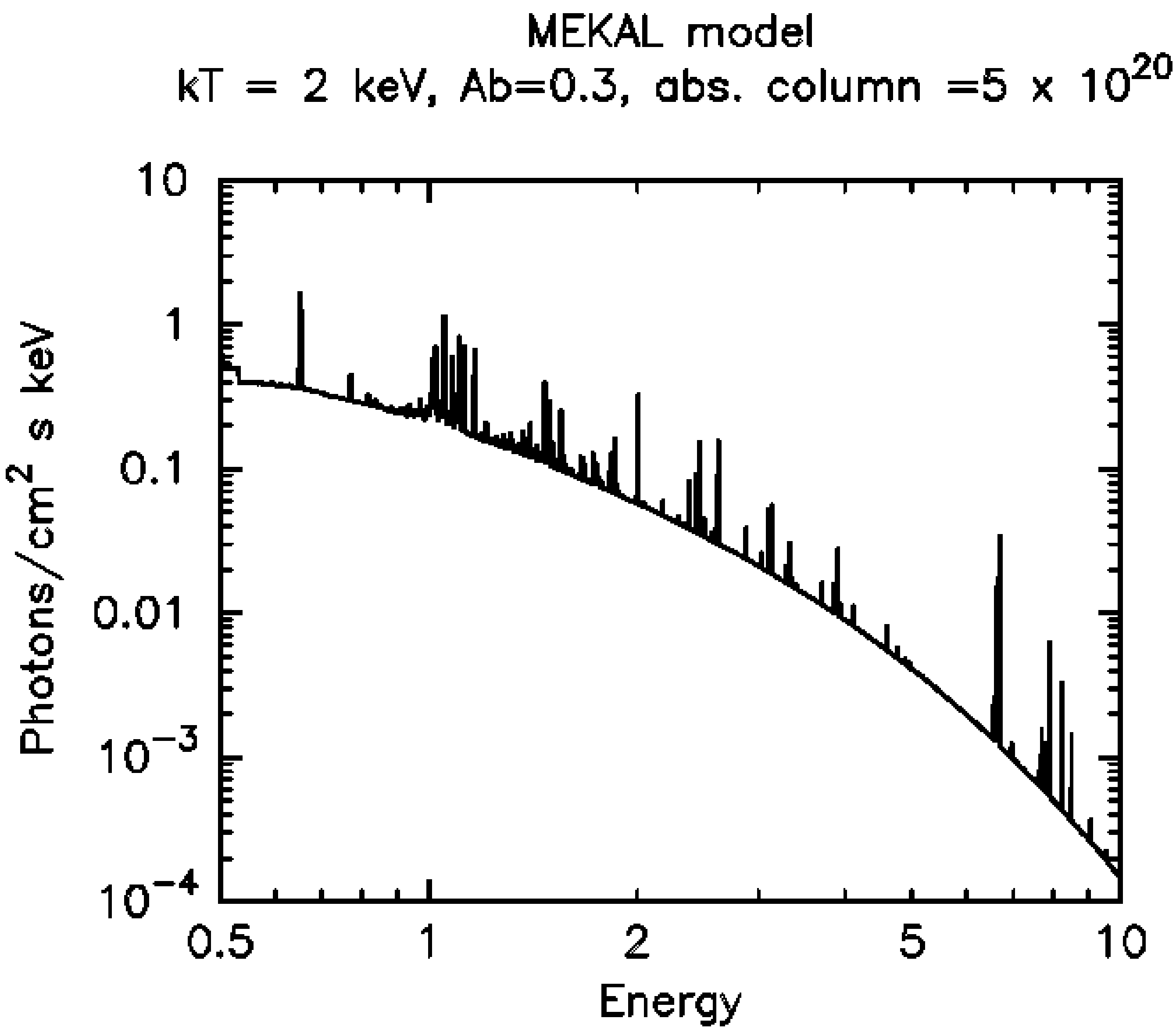
Suppose we observe a source and detect $D(E')$ photons per unit area, per unit time, per unit energy as a function of the **measured** energy E' of the photons, how can we infer $S(E)$, the number of photons per unit area, per unit time, per unit energy **emitted** by that source?

(Remember that the measured energy of each photon is not necessarily, and most likely is not, equal to the energy of the emitted photon!)

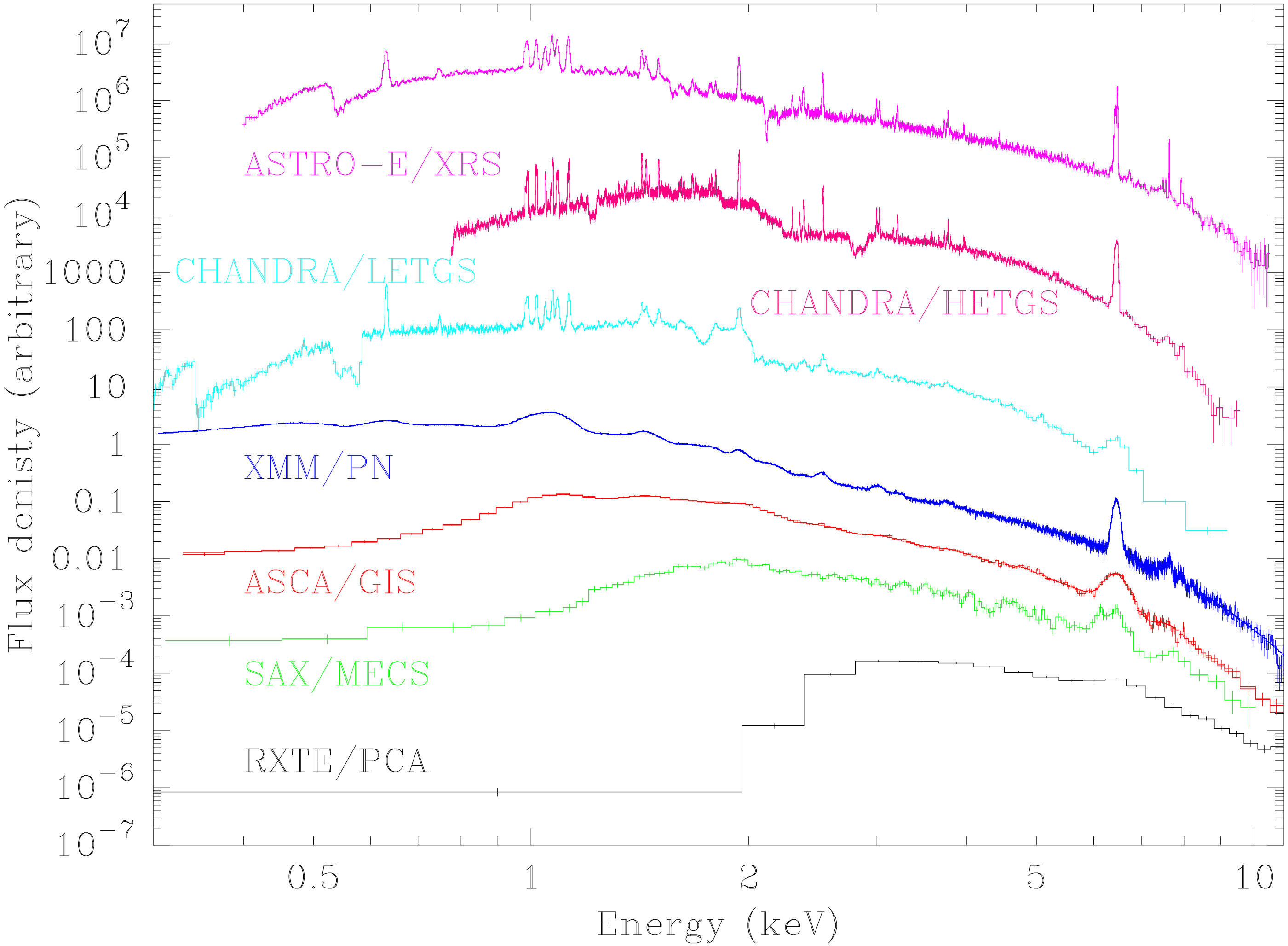
Can we start with this…



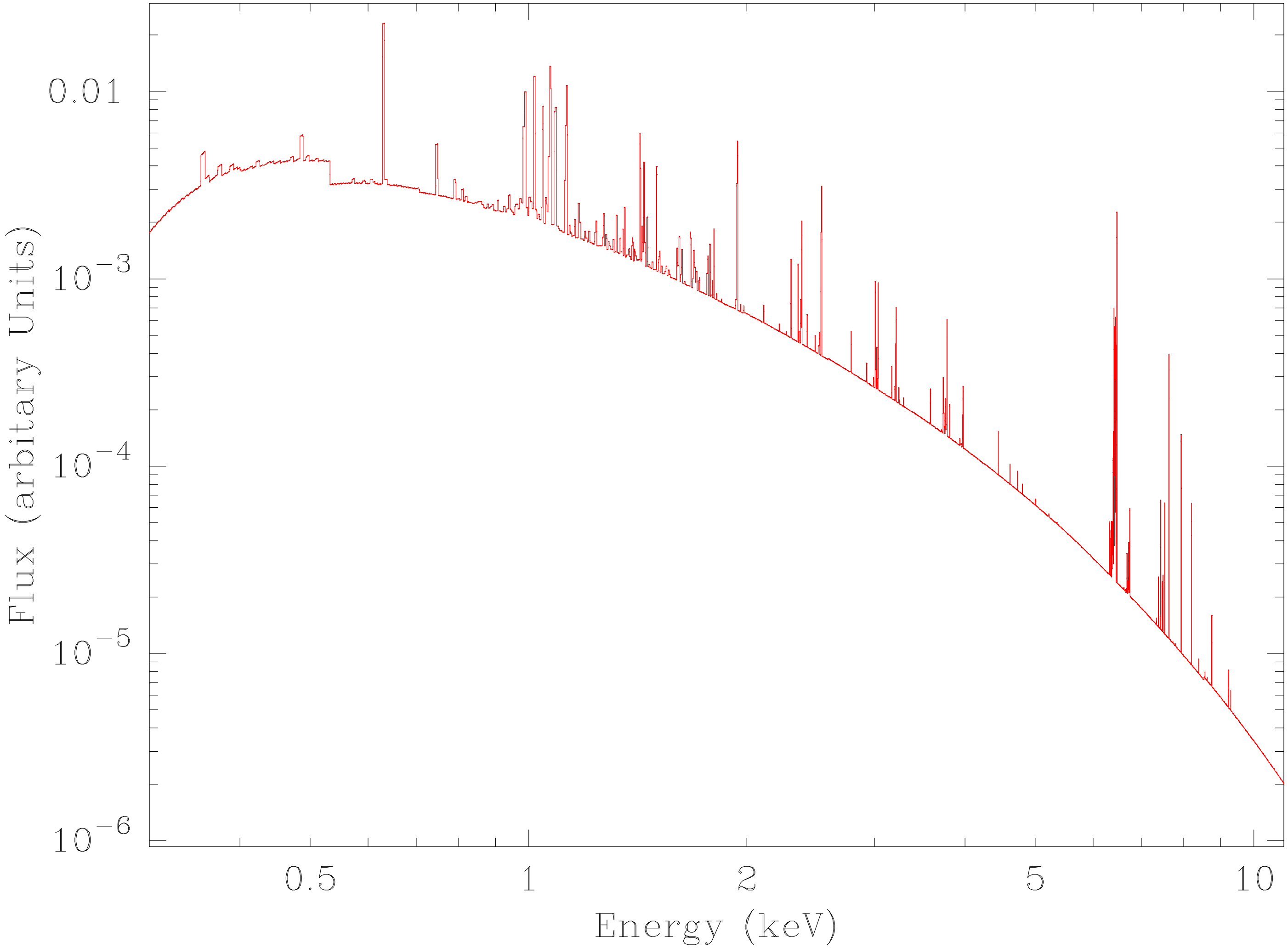
… and deduce this

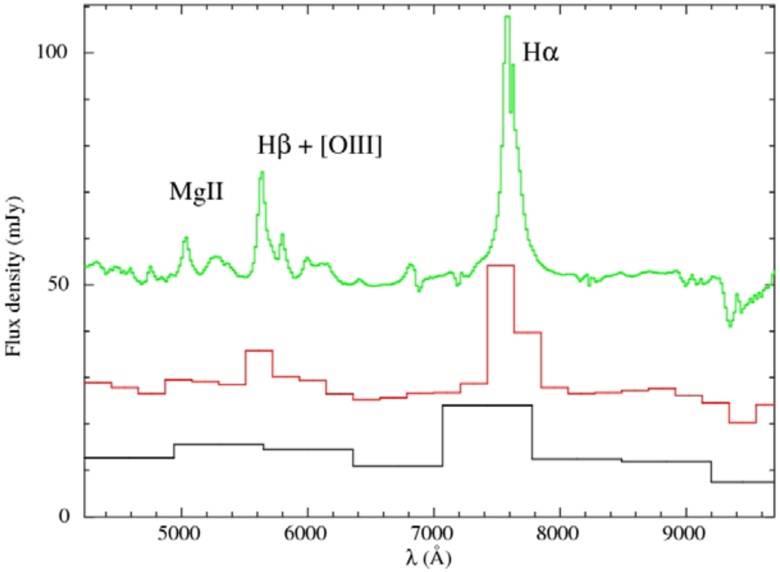


Can we start with any of these…

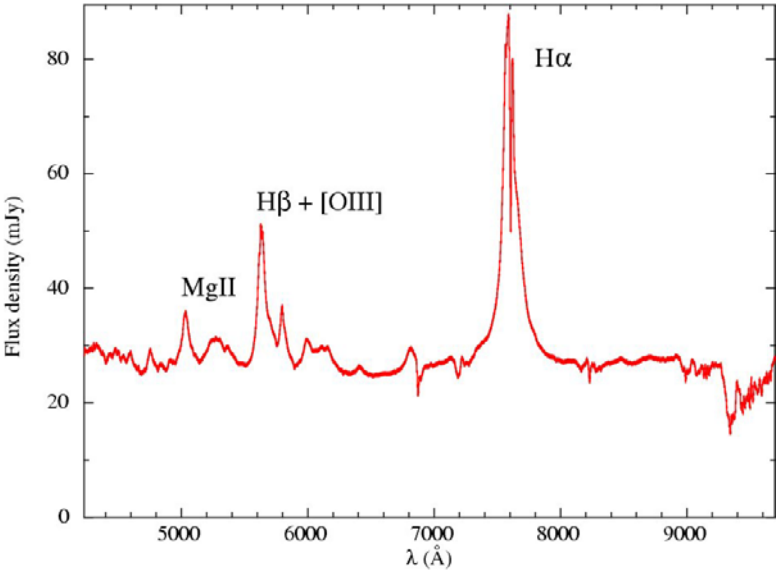


… and deduce this



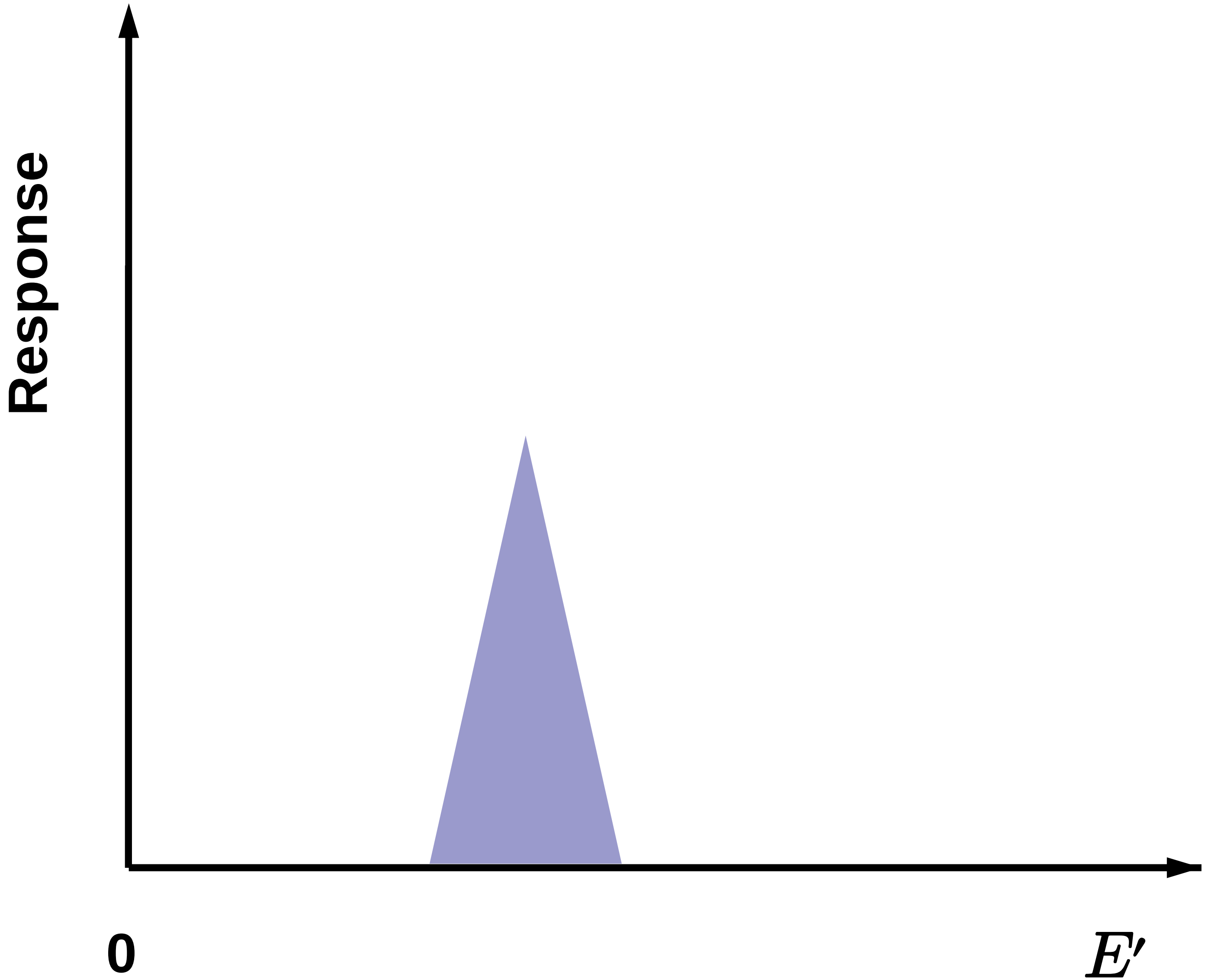


Can we start with these…

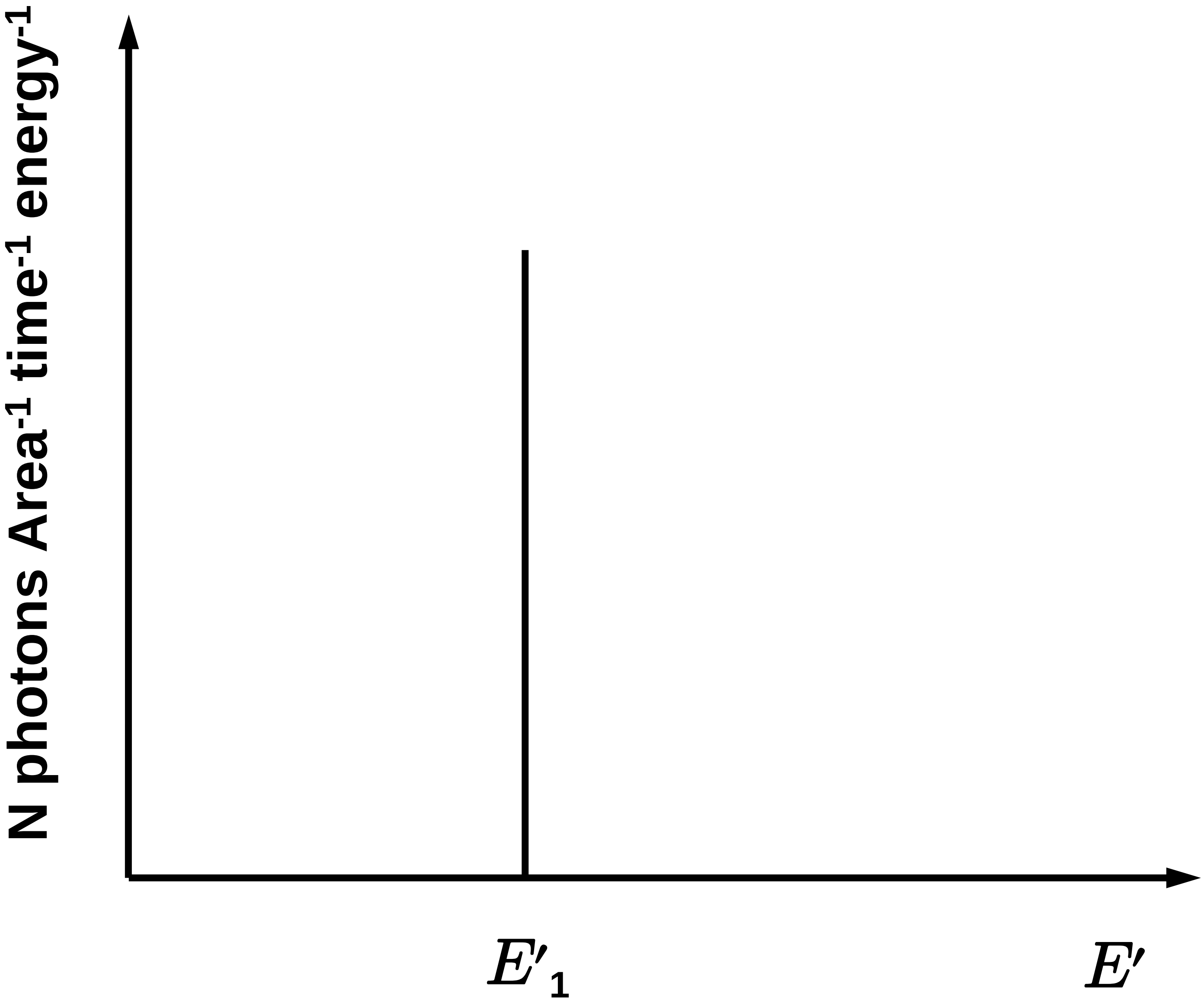


and deduce this?

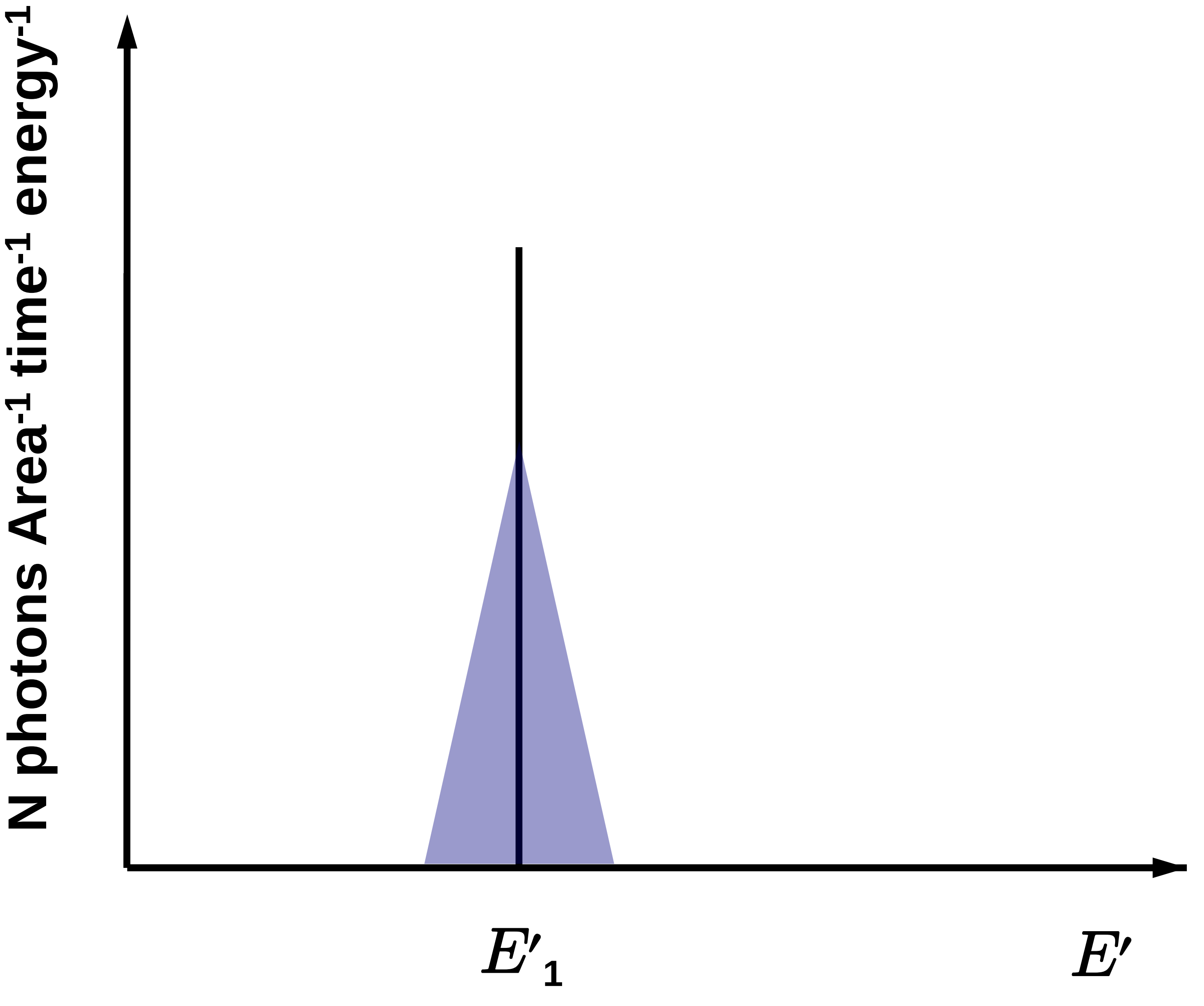
A simple graphical example



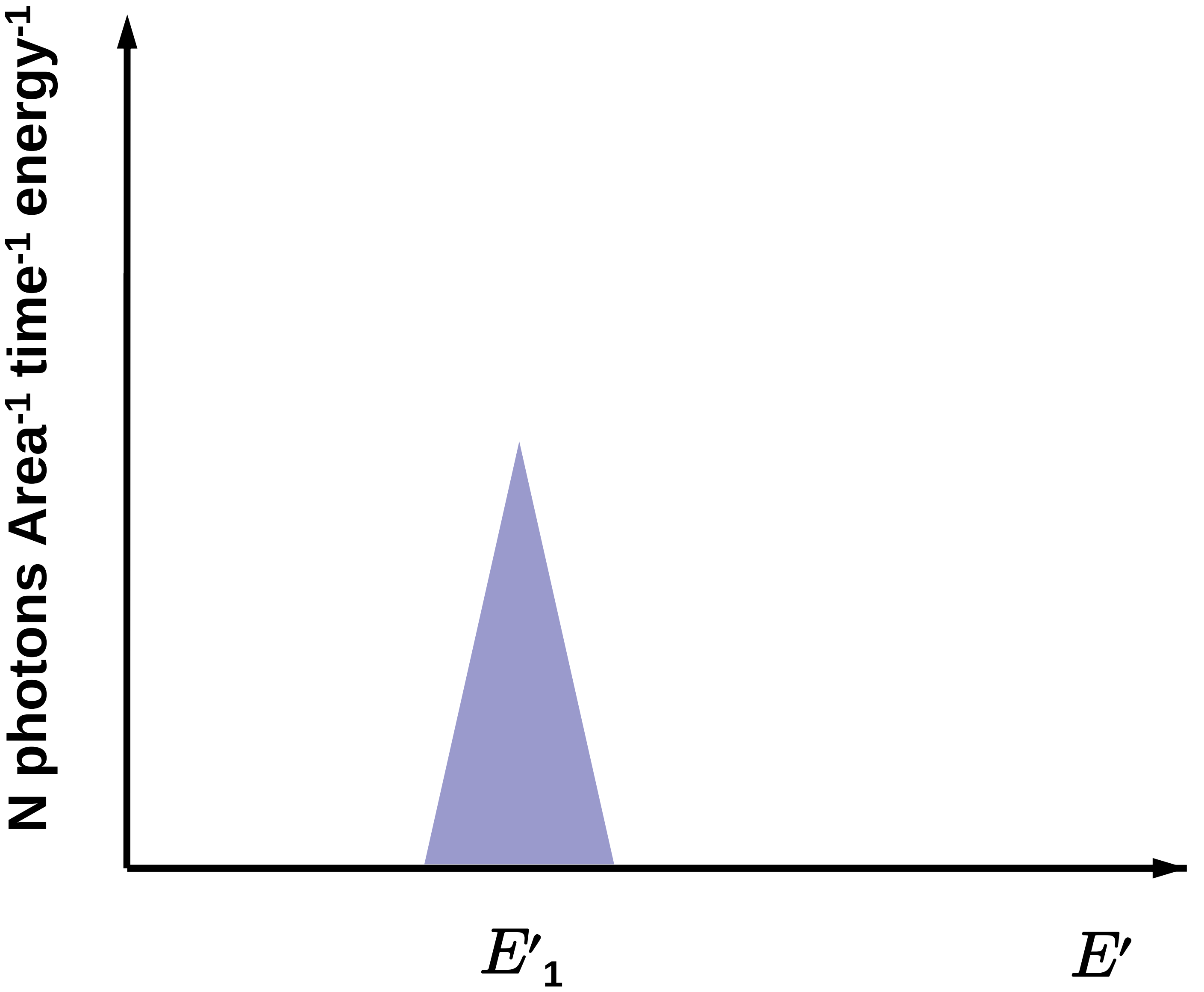
A simple graphical example



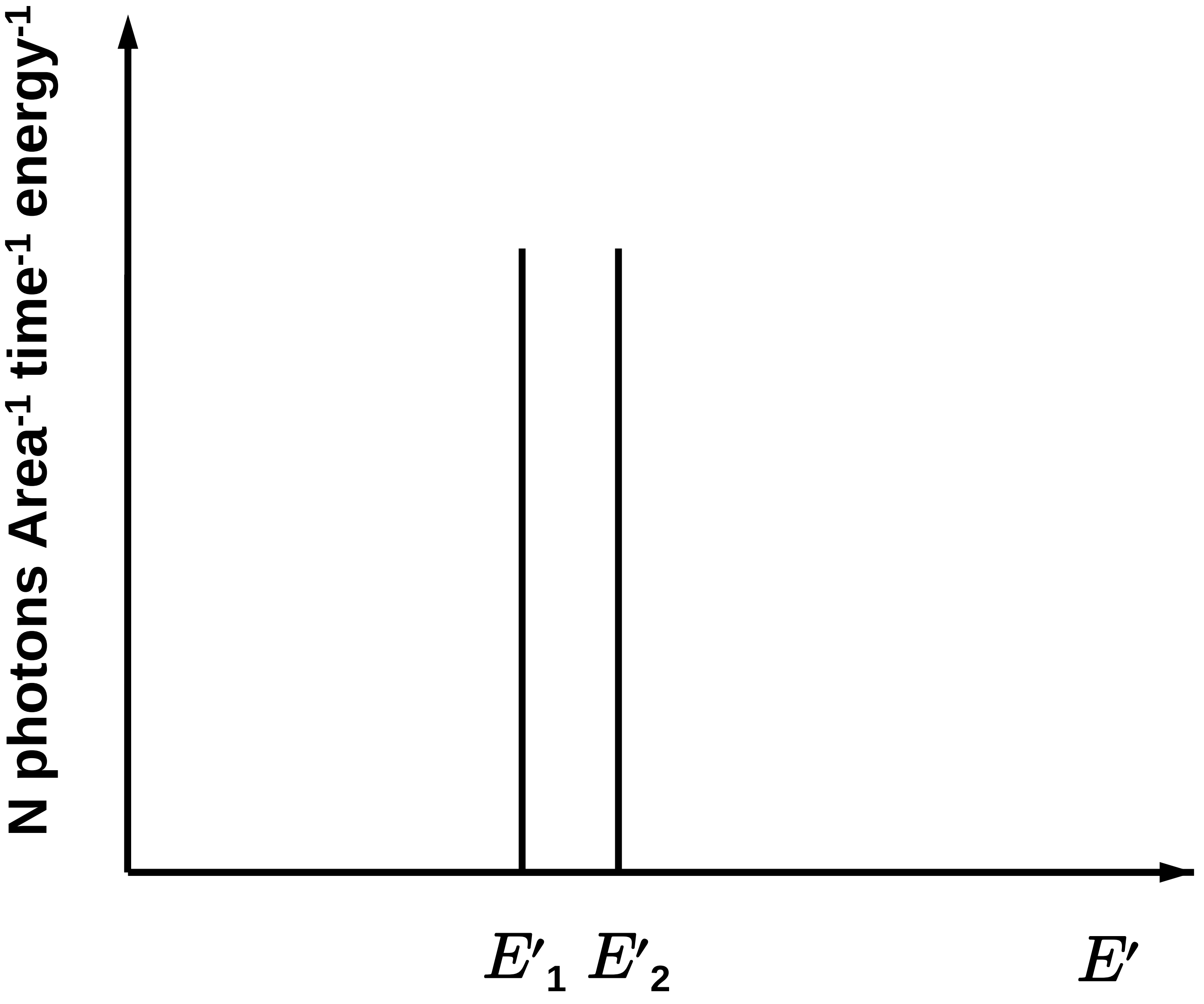
A simple graphical example



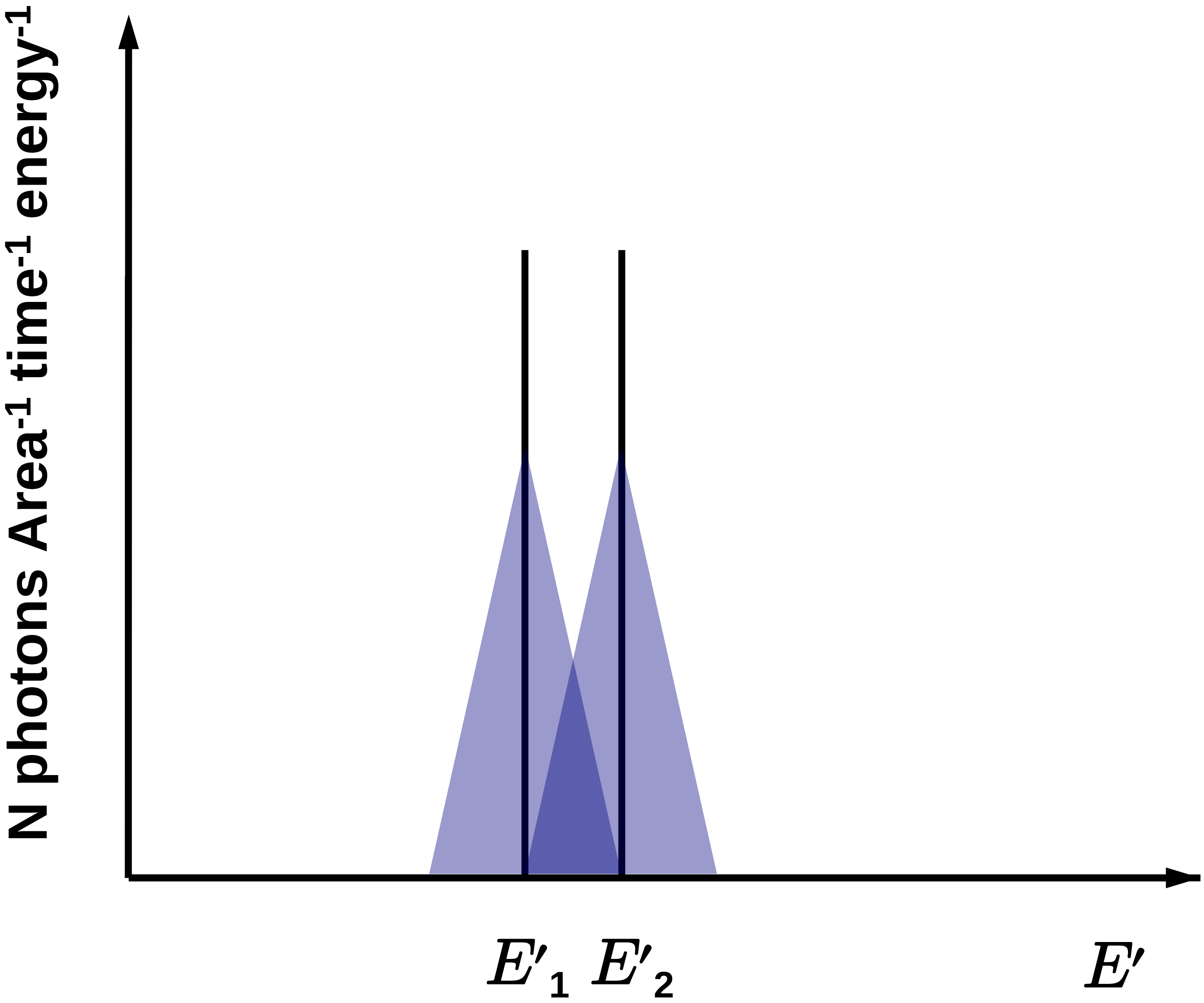
A simple graphical example



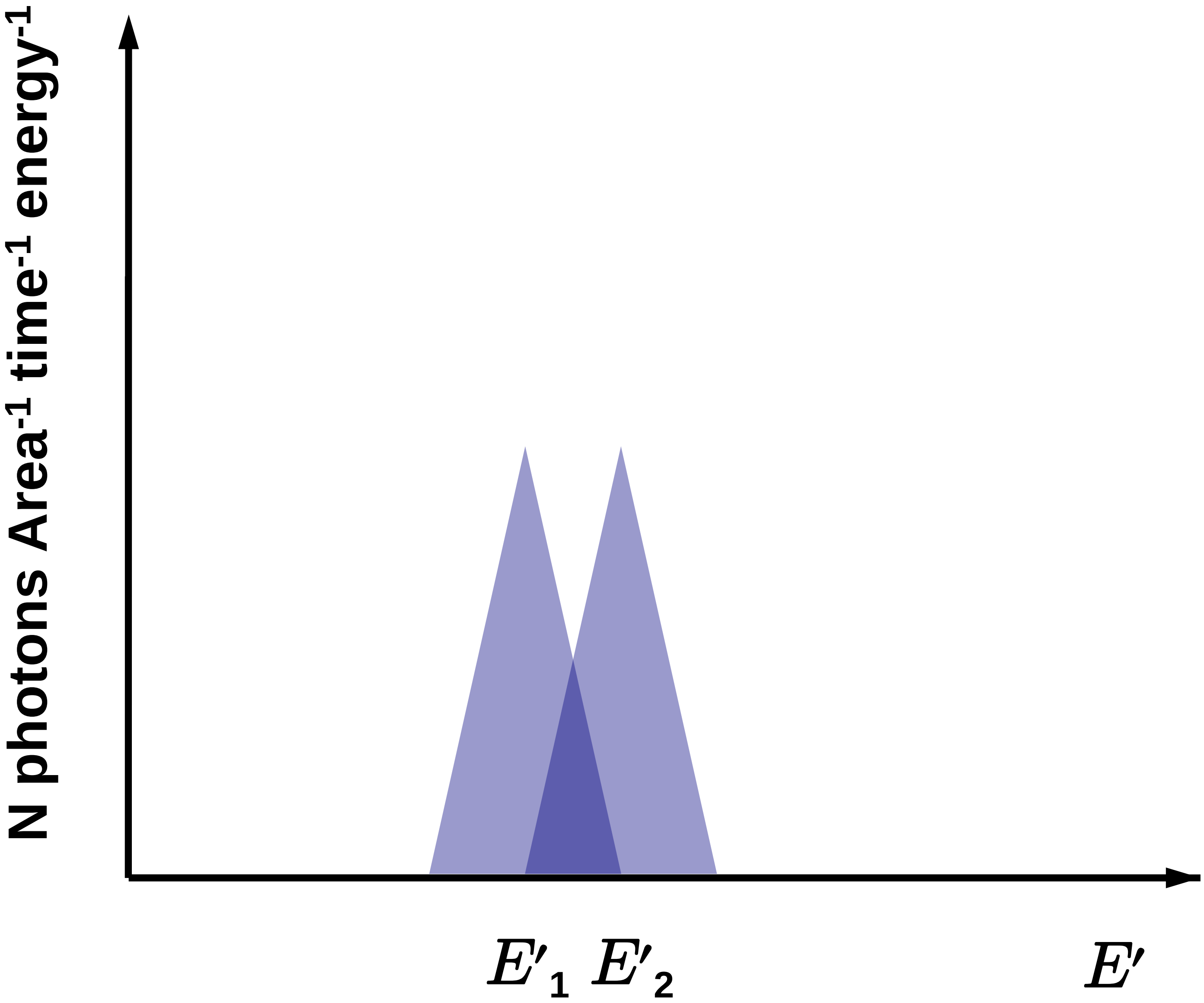
A simple graphical example



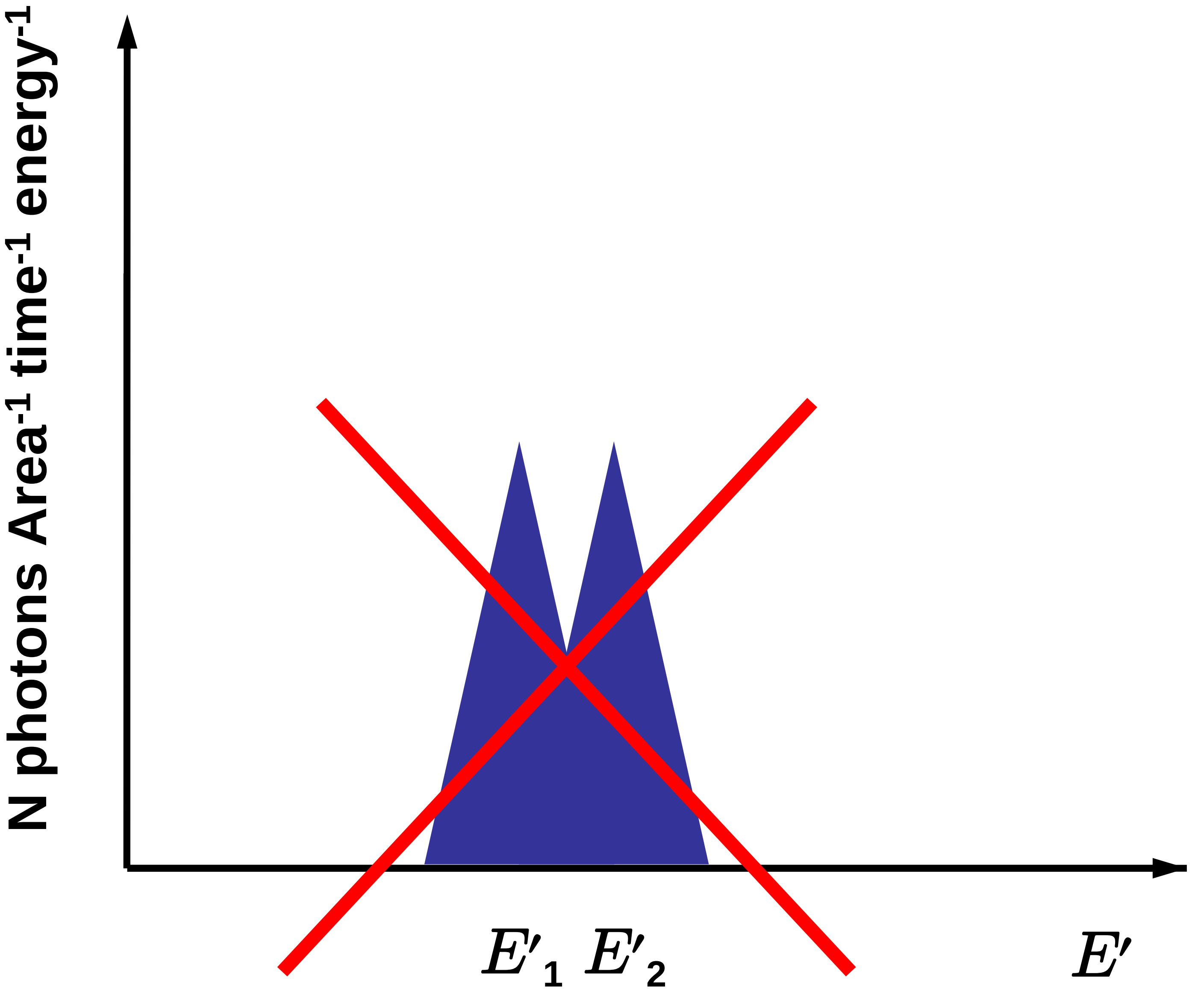
A simple graphical example



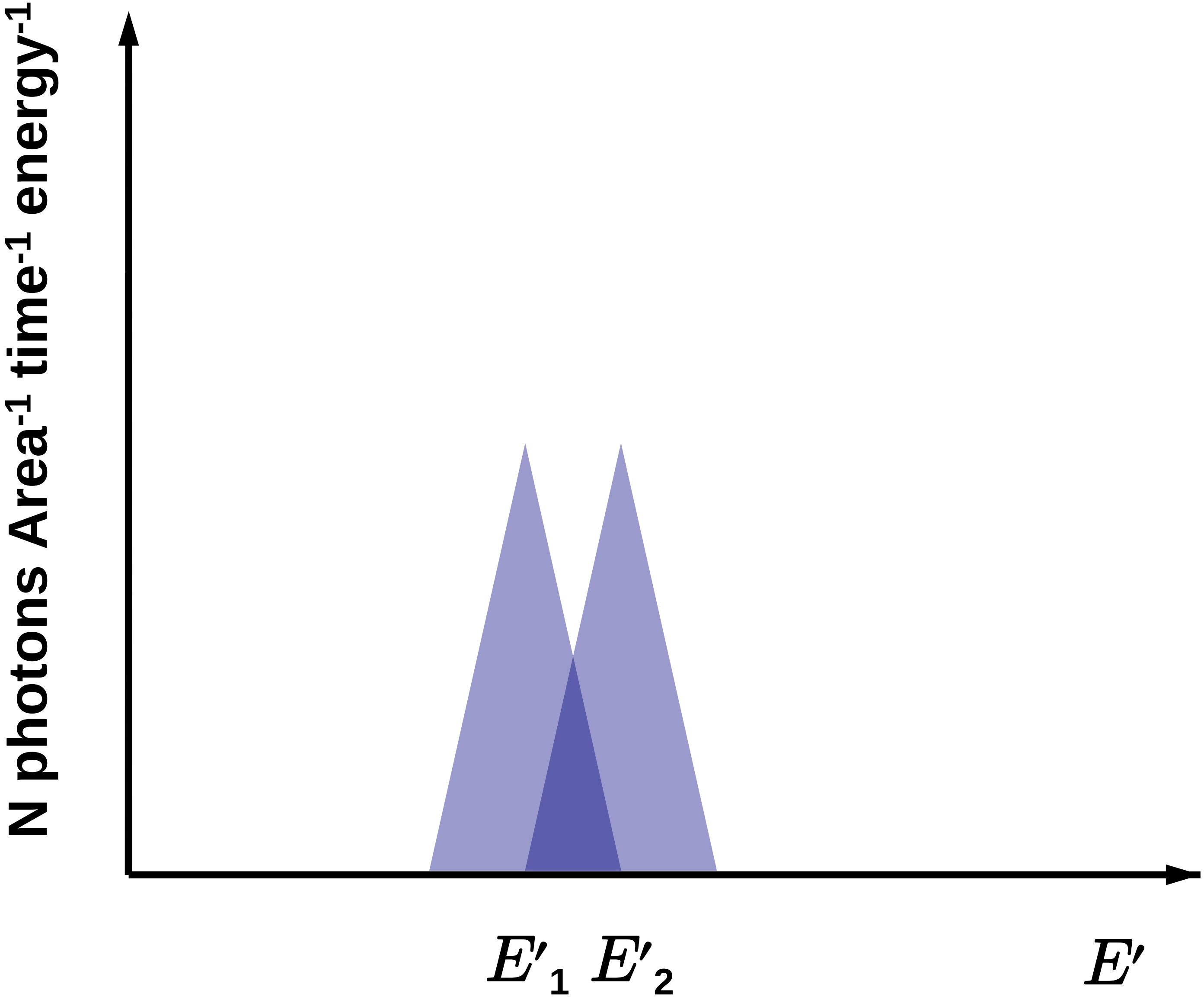
A simple graphical example



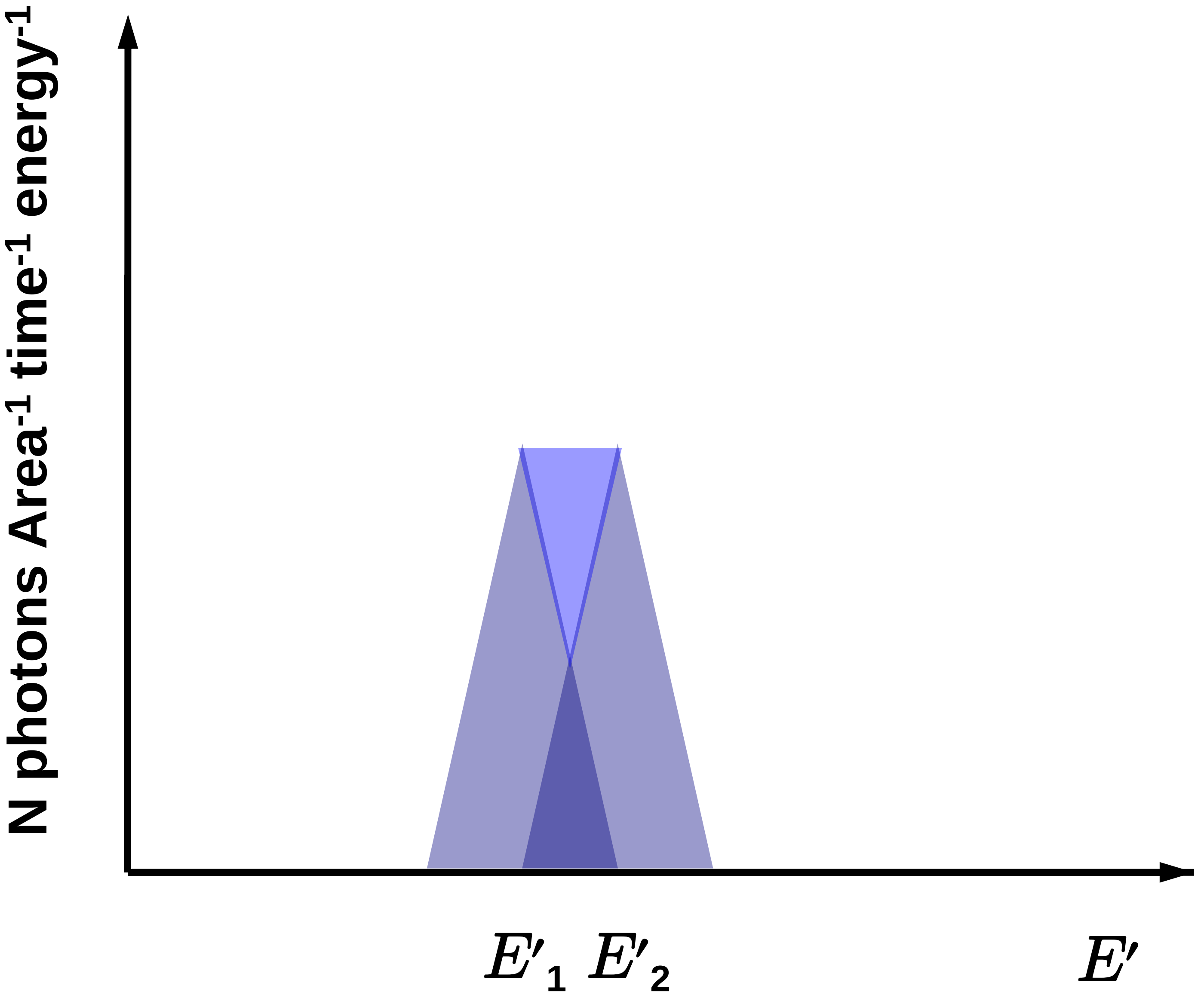
A simple graphical example



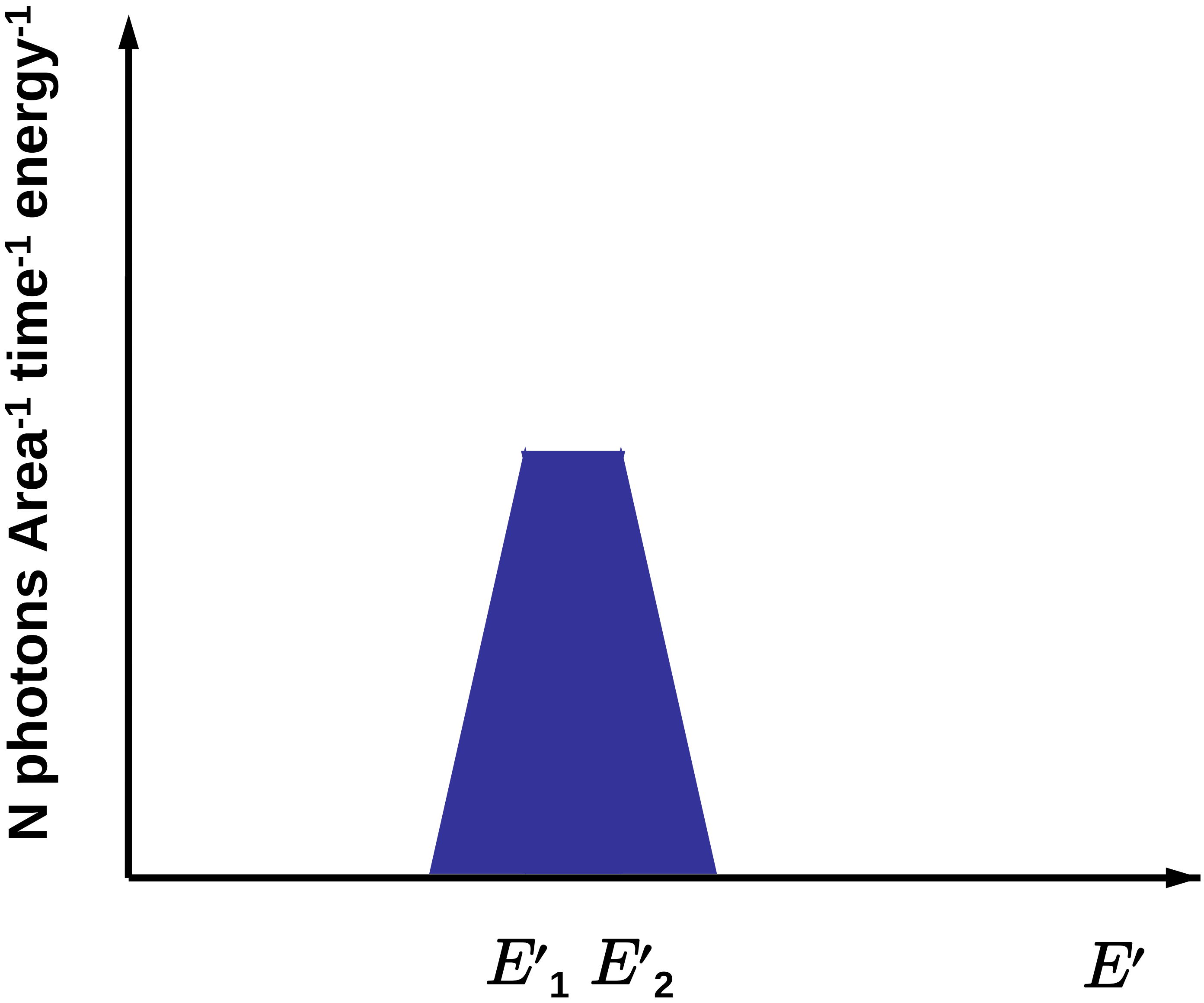
A simple graphical example



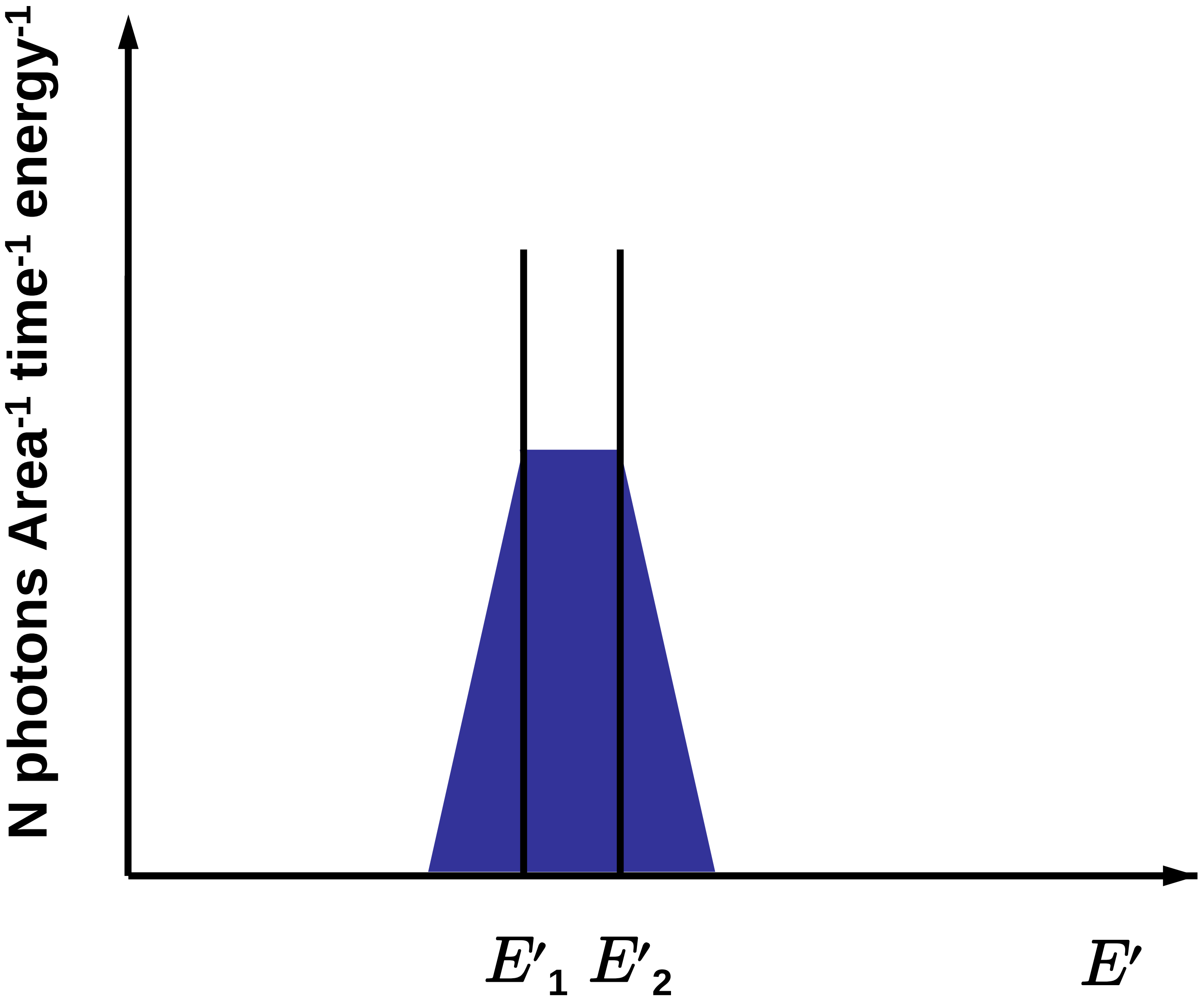
A simple graphical example



A simple graphical example

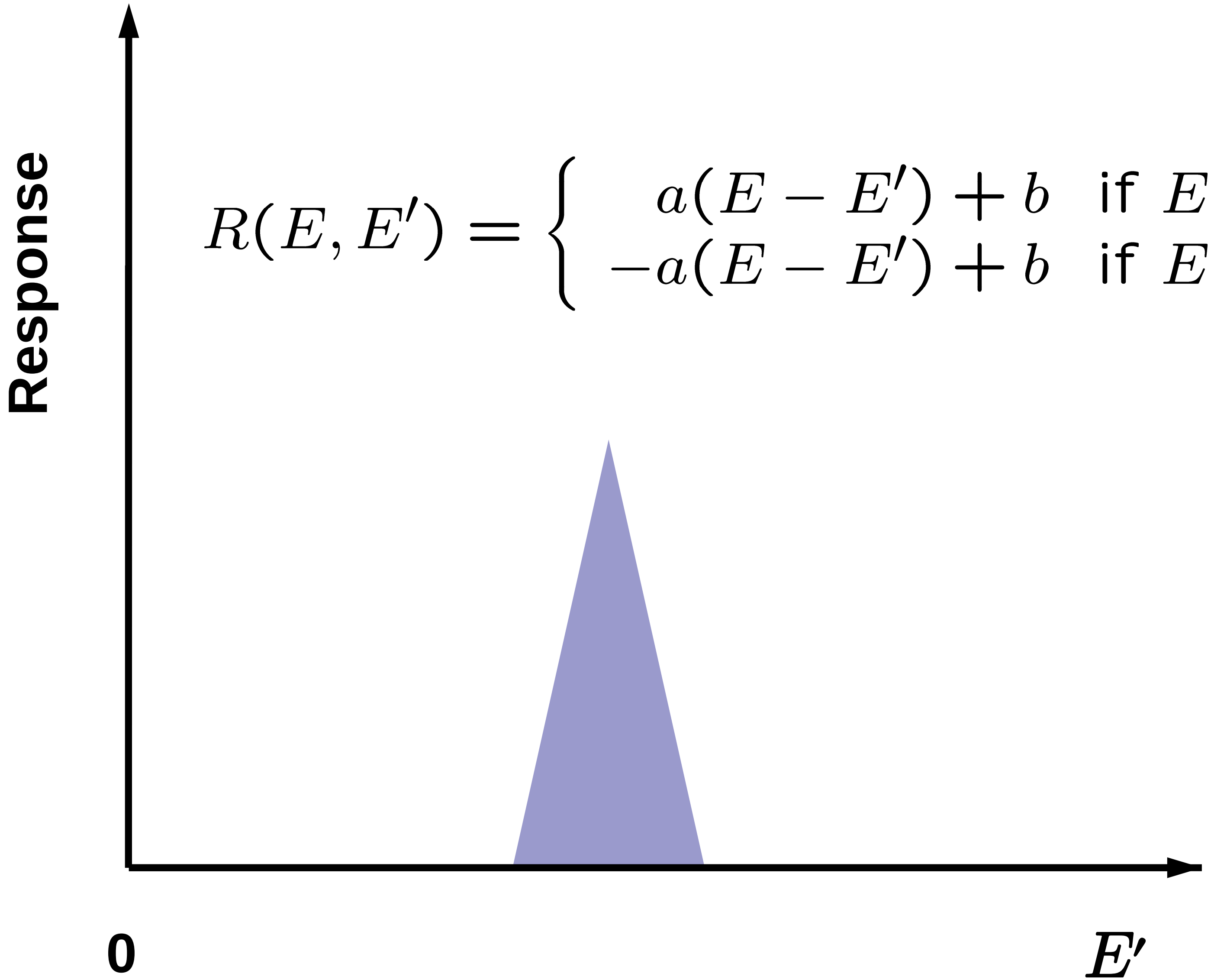


A simple graphical example

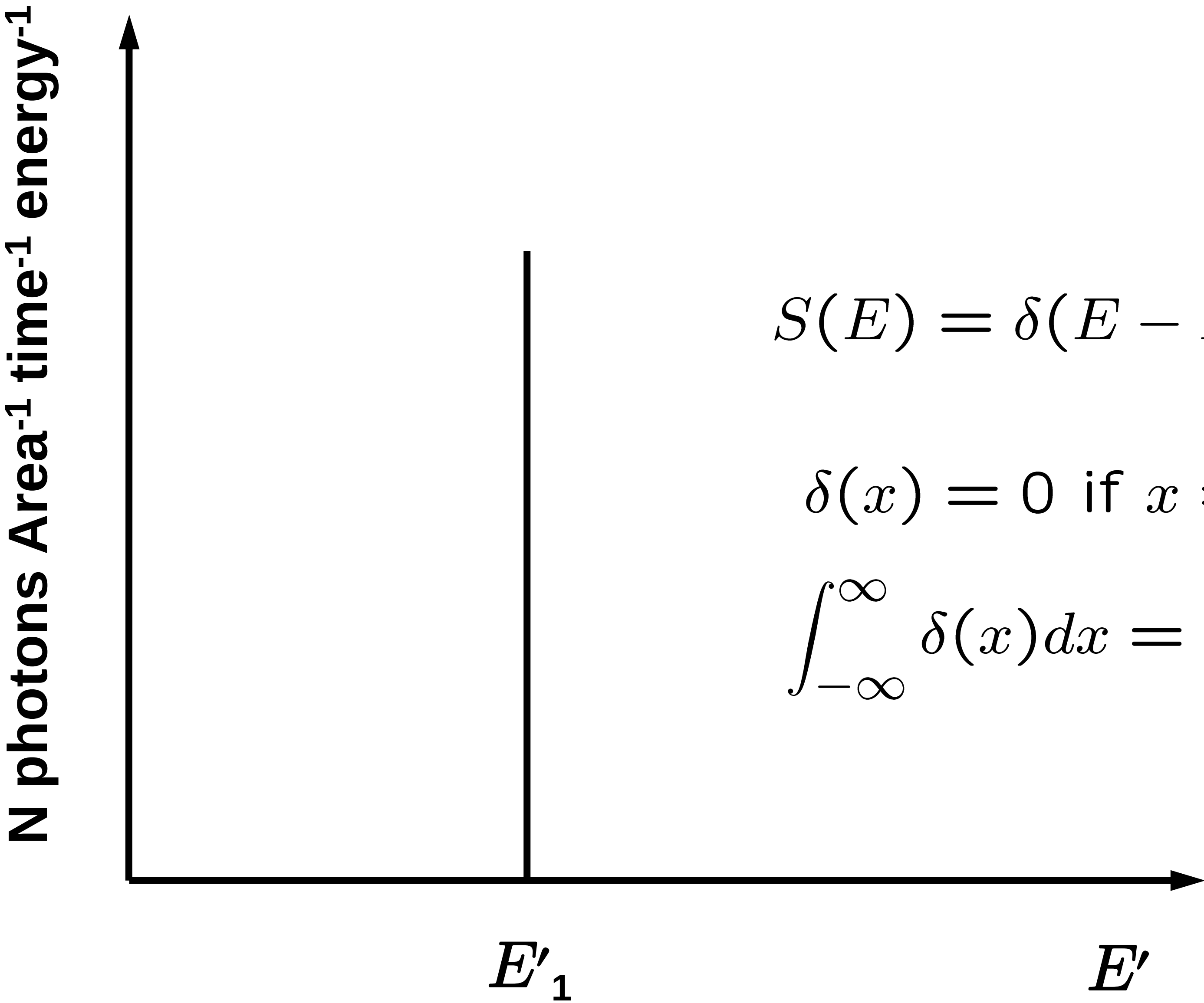


A simple graphical example

$$R(E, E') = \begin{cases} a(E - E') + b & \text{if } E \leq E' \\ -a(E - E') + b & \text{if } E \geq E' \end{cases}$$



A simple graphical example



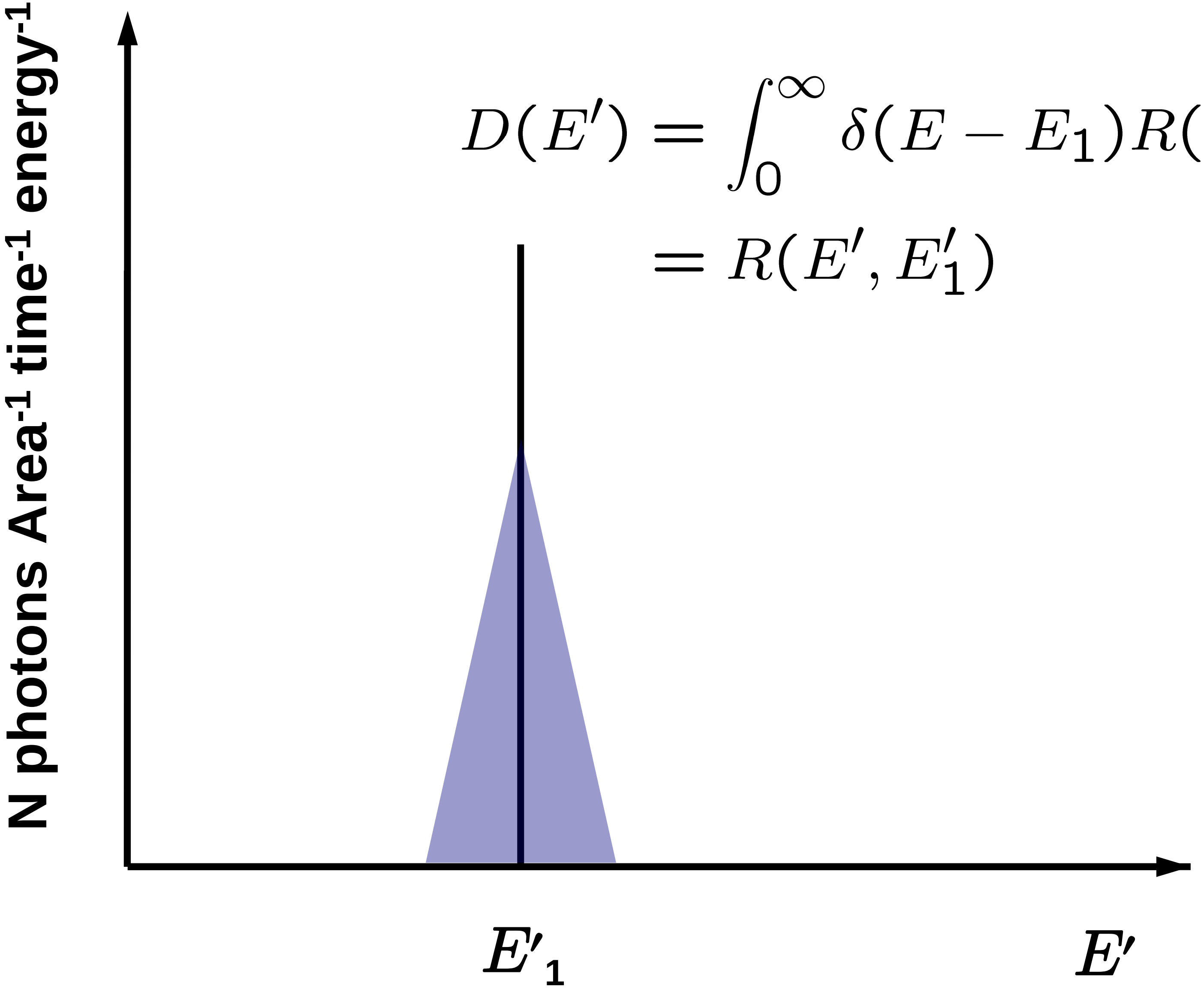
$$S(E) = \delta(E - E_1)$$

$$\delta(x) = 0 \text{ if } x \neq 0$$

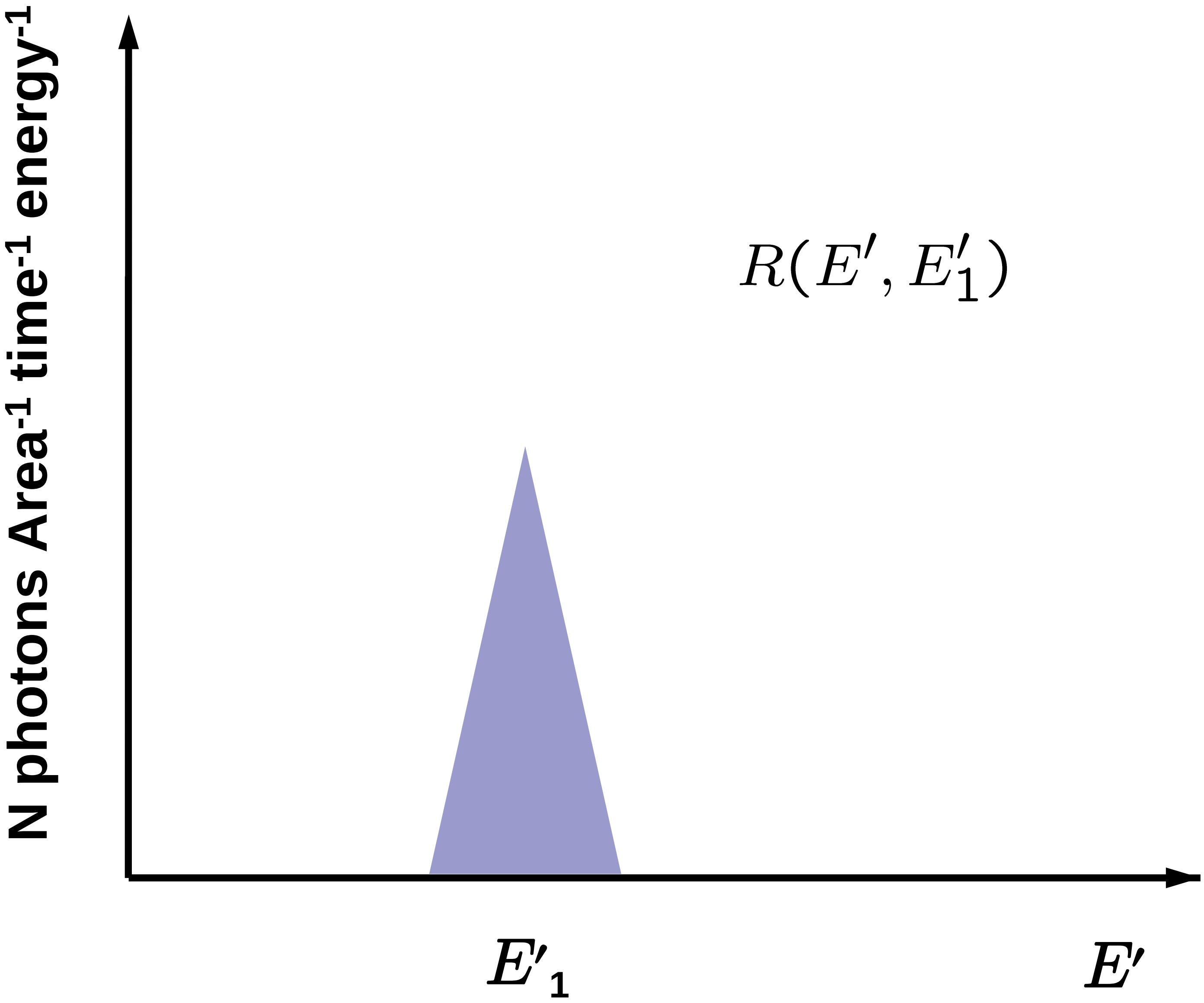
$$\int_{-\infty}^{\infty} \delta(x)dx = 1$$

A simple graphical example

$$D(E') = \int_0^\infty \delta(E - E_1) R(E, E') dE$$
$$= R(E', E'_1)$$



A simple graphical example



The Basic Problem (1)

Suppose we observe a source and detect $D(E')$ photons per unit area, per unit time, per unit energy as a function of the **measured** energy E' of the photons, how can we infer $S(E)$, the number of photons per unit area, per unit time, per unit energy **emitted** by that source?

The answer is that what we see is the convolution of the incident spectrum, $S(E)$, and the instrument response, $R(E,E')$, describing the way the instrument responds to the incident photon.

The Basic Problem (2)

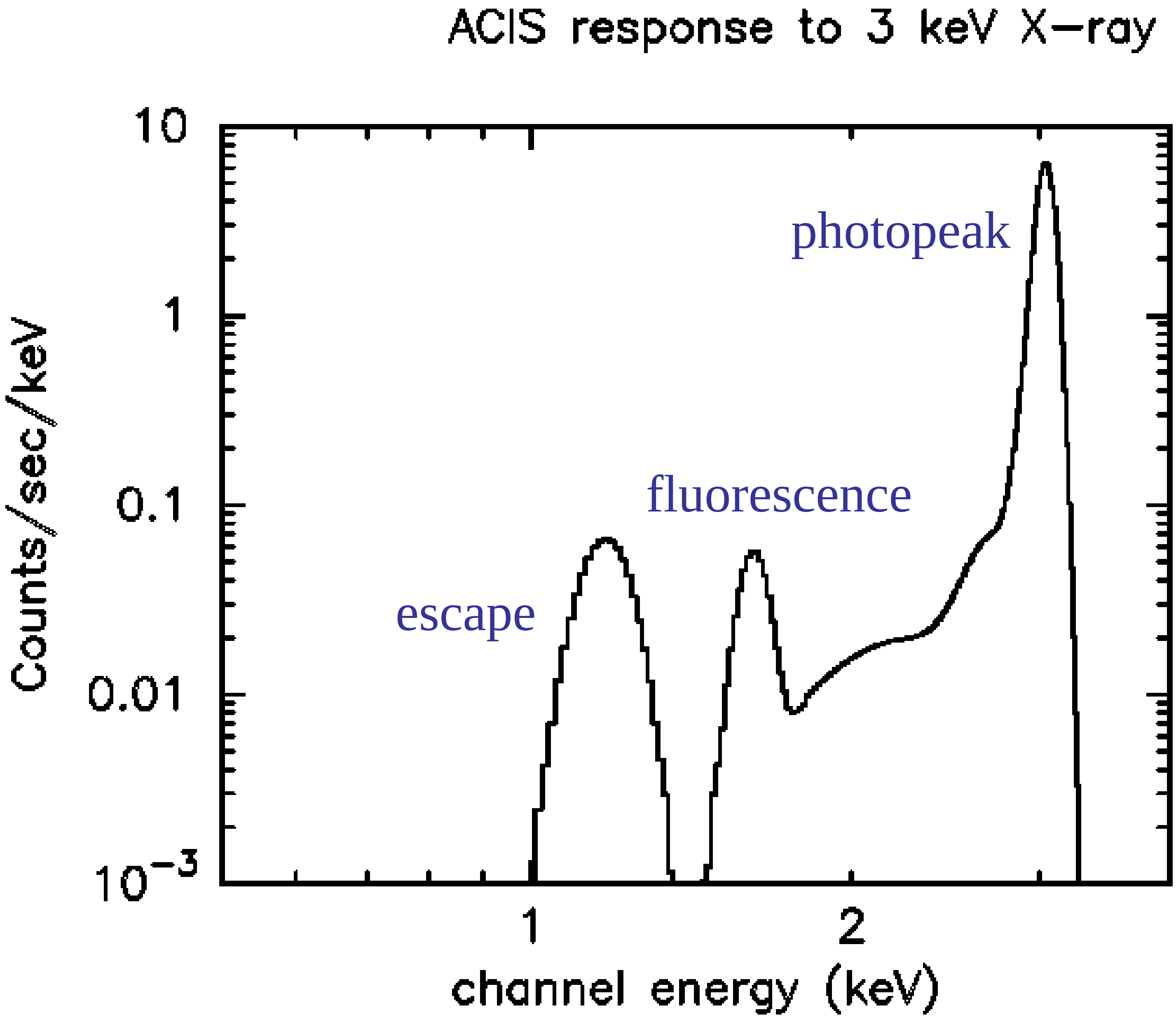
In general, the incident and observed spectrum will be related by the integral relation:

$$D(E') = T \int R(E, E') A(E) S(E) dE$$

- T is the exposure time (*e.g.*, in seconds)
- $R(E',E)$ is the probability of an incoming photon of energy E being measured at an energy E', with

$$\int R(E', E) dE' = 1.$$

A real example of $R(E',E)$



The Basic Problem (2)

In general, the incident and observed spectrum will be related by the integral relation:

$$D(E') = T \int R(E, E') A(E) S(E) dE$$

- $A(E)$ is the energy-dependent effective area of the telescope and detector system (*e.g.*, in cm^2).

- $S(E)$ is the source flux at the front of the telescope (*e.g.*, in photons/cm^2/s/keV)

The Inverse problem

This is known as the remote sensing problem (or inverse problem) and arises in many areas of astronomy as well as, *e.g.*, geophysics and medical imaging.

In mathematics the integral equation is known as a Fredholm equation of the first kind.

The Basic Problem (3)

$$D(E') = T \int R(E, E') A(E) S(E) dE$$

We know T from the observation, and we assume we know $A(E)$ and $R(E',E)$ from the instrument calibration; we want to solve this integral equation for $S(E)$.

In reality we have discrete measurements over finite energy intervals. Let's say we have divided the energy range of interest into M intervals (also known as bins).

The Basic Problem (3)

We then have the data measured in each bin, D_i, where i runs from **1** to M.

Similarly, the emitted spectrum and effective area of the detector will now be S_j and A_j, respectively, where j runs from **1** to N. The integral equation

$$D(E') = T \int R(E, E') A(E) S(E) dE$$

will then become

$$D_i = T \sum_{j=1}^{N} R_{ij} A_j S_j,$$

The Basic Problem (3)

$$D_i = T \sum_{j=1}^{N} R_{ij} A_j S_j,$$

where, as mentioned, D_i is the number of photons observed in channel i, and S_j is the flux in photons/cm^2/s in energy bin j.

We measure D_i and we want to find S_j.

The Basic Problem (3)

$$D_i = T \sum_{j=1}^{N} R_{ij} A_j S_j.$$

We can rewrite the above using matrix notation as

$$D = T \ \mathbf{R} \ A \ S,$$

The vector D is the product of the **matrix** **R** times the vectors A and S, and the scalar T.

R is usually called the *Redistribution Matrix Function (RMF)*, or if A is absorbed into **R**, the *Response Matrix (RSP)*.

The Basic Problem (3)

$$D = T\ \mathbf{R}\ S$$

The obvious tempting solution is to invert the above:

$$S = \frac{1}{T} D \mathbf{R}^{-1}.$$

Where $\mathbf{R}^{-1}$ is the inverse, in matrix sense, of **R**.

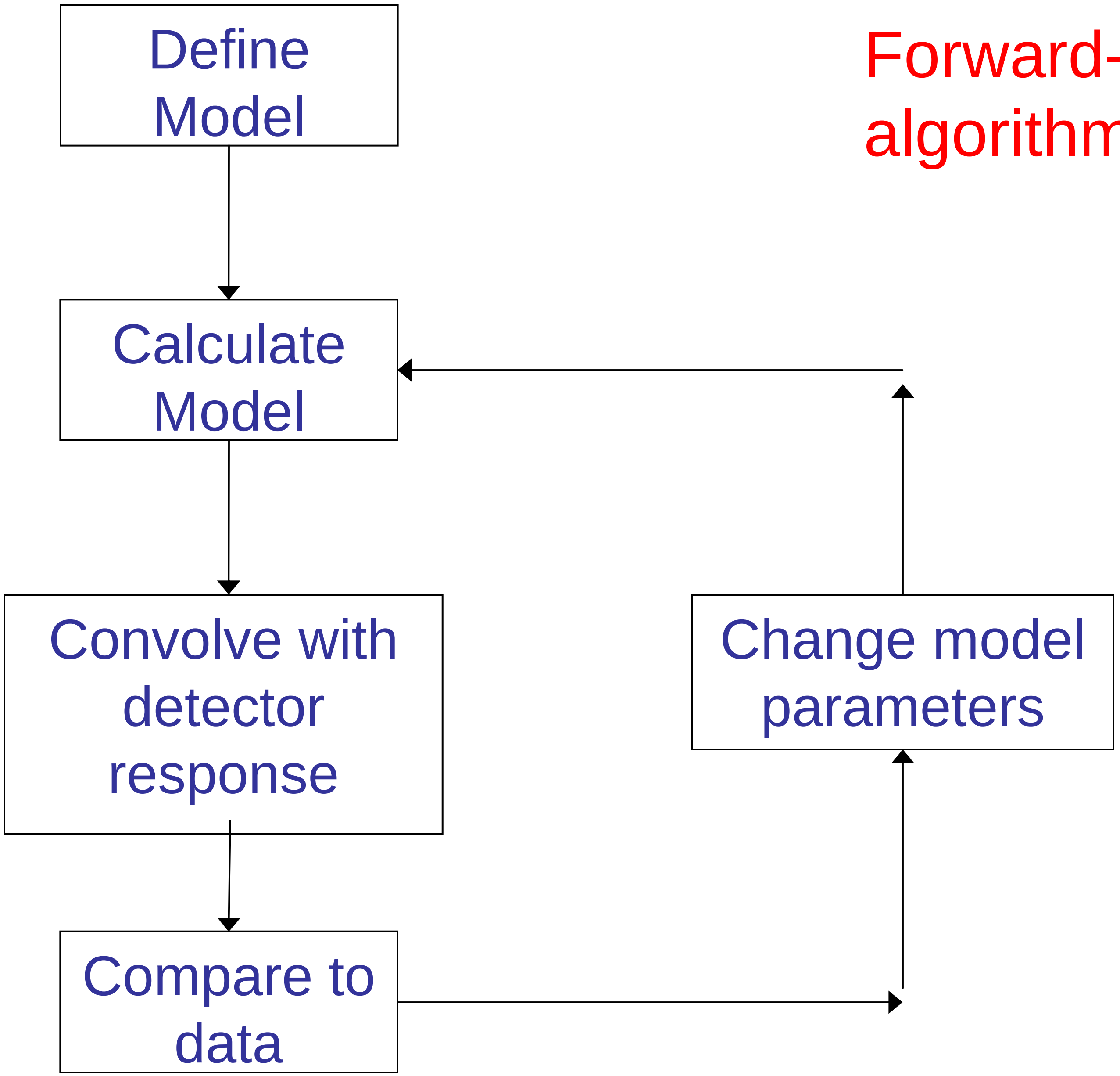
This does not work! The S_j derived in this way are very sensitive to slight changes in the data, D_i, and the response matrix, **R**. This is a great method for amplifying noise!

Forward fitting

The standard method of analyzing X-ray/γ-ray spectra is “forward-fitting”. This comprises the following steps…

- Calculate a model spectrum.
- Multiply the result by an instrumental response matrix.
- Compare the result with the actual observed data by calculating some statistic.
- Modify the model spectrum and repeat till the best value of the statistic is obtained.

Forward-fitting algorithm



Forward fitting

This only works if the model spectrum can be expressed in a reasonably small number of parameters so that the model can be varied in some sensible fashion.

The aim of the forward-fitting is then to obtain the best-fit and confidence ranges of these parameters.

The solution is NOT unique. Only some extra physical knowledge helps to get an answer.

Spectral fitting programs

- XSPEC - part of HEAsoft. General spectral fitting program with many models available.
- Sherpa - part of CIAO. Multi-dimensional fitting program which includes the XSPEC model library and can be used for spectral fitting.
- SPEX - from SRON in the Netherlands. Spectral fitting program specializing in collisional plasmas and high resolution spectroscopy.
- ISIS - from the MIT Chandra HETG group. Mainly intended for the analysis of grating data.

Models

Models are usually built up from individual components. These can be thought of as two basic types: additive (an emission component e.g. blackbody, line,…) or multiplicative (something which modifies the spectrum e.g. absorption).

Model = $M = M_1 \times M_2 \times (A_1 + A_2 + M_3 \times A_3) + A_4$

Additive Models

Basic additive (emission) models include :

- blackbody
- thermal bremsstrahlung
- power-law
- collisional plasma
- Gaussian or Lorentzian lines

Spectral packages have many more models available covering specialized topics such as accretion disks, comptonized plasmas, non-equilibrium ionization plasmas, multi-temperature collisional plasmas…

Multiplicative and other models

Multiplicative models include :

- photoelectric absorption due to the neutral ISM
- photoelectric absorption due to ionized material
- high energy exponential roll-off.
- edge with $1/E^3$ roll-off.

There are also other types of model components (convolution, mixing) which are used like a multiplicative model but perform more complicated operations on the current model.

χ^2 - fit

Given the data, D_i, and the model values at the same bins as the data, M_i, calculate:

$$X^2 = \sum_{i=1}^{n} \frac{(D_i - M_i)^2}{\sigma_i^2},$$

where $\sigma_i = M_i^{1/2}$ (expected error).

χ^2 - fit

However, having not yet fit the data, we do not know what is the expected value in channel i, M_i, and hence we cannot calculate the expected error.

It is usual to take the observed error as a proxy to the expected one:

$$X^2 = \sum_{i=1}^{n} \frac{(D_i - M_i)^2}{\sigma_{D_i}^2}$$

where $\sigma_{D_i} = D_i^{1/2}$ (observed error).

χ^2 - fit

Notice that this error is “biased” (it is not a proper representation of the true error. More about bias in the Statistics lectures next week). For instance, if one channel happens to have fewer photons than expected, the error will also be smaller than expected, and that channel will have a strong influence in the value of X^2. In the extreme, if a channel has 0 counts (not unthinkable in γ-ray spectra), X^2 goes to infinity!

$$X^2 = \sum_{i=1}^{n} \frac{(D_i - M_i)^2}{\sigma_{D_i}^2}$$

χ^2 - fit

If M is the correct model, the function X^2 follows a χ^2 distribution with $n - m$ degrees of freedom (*dof*), where n is the number of data points (data channels) available, and m is the number of parameters in the model.

In that case, the average of X^2 is $\langle X^2 \rangle = n - m$, and the standard deviation is $\sigma_{X^2} = \sqrt{2(n-m)}$.

In general the model M will be rejected if X^2 is larger than $n - m + f \times \sqrt{2(n-m)}$, where f is determined by the confidence limit that one chooses.

χ^2 - fit

If the correct model is M' instead of M, the expected value of X^2 is no longer $n - m$, but $n - m + r$, with:

$$r = N \int \left[\frac{f'(E) - f(E)}{f(E)} \right]^2 f(E) dE,$$

where f and f' are the probability distributions for the photons in models M and M', respectively, and N is the total number of photons in the spectrum. As can be seen, r does not depend on the number of bins in the data, n.

χ^2 - fit

It can happen that one has chosen the wrong model to describe the data, but if n is too large one still ends up accepting it because

$$X^2 = [n - m + r] < [n - m + f \times \sqrt{2(n-m)}].$$

If you use χ^2 to fit your data, never oversample your spectra.

Notice also that χ^2 assumes all measurements are independent, but if you oversample, channels are going to be correlated, and you are also going to overestimate the number of degrees of freedom.

χ^2 - fit

The fits will be dominated by bins that, due to statistical fluctuations, have small number of counts.

If you use χ^2 to fit your data, you should rebin your spectra until you have enough counts in each channel so that the observed variance is approximately equal to the expected one (random fluctuations will not bias it).

Alternatively you may want not to assume errors are Gaussian (one of the assumptions behind χ^2), given that they are in fact Poissonian!