SIGGRAPH2007

A Gentle Introduction to Bilateral Filtering and its Applications

SIGGRAPH2007

How does bilateral filter relates with other methods?

Pierre Kornprobst (INRIA)

0:35

Many people worked on… *edge-preserving restoration*

Bilateral filter

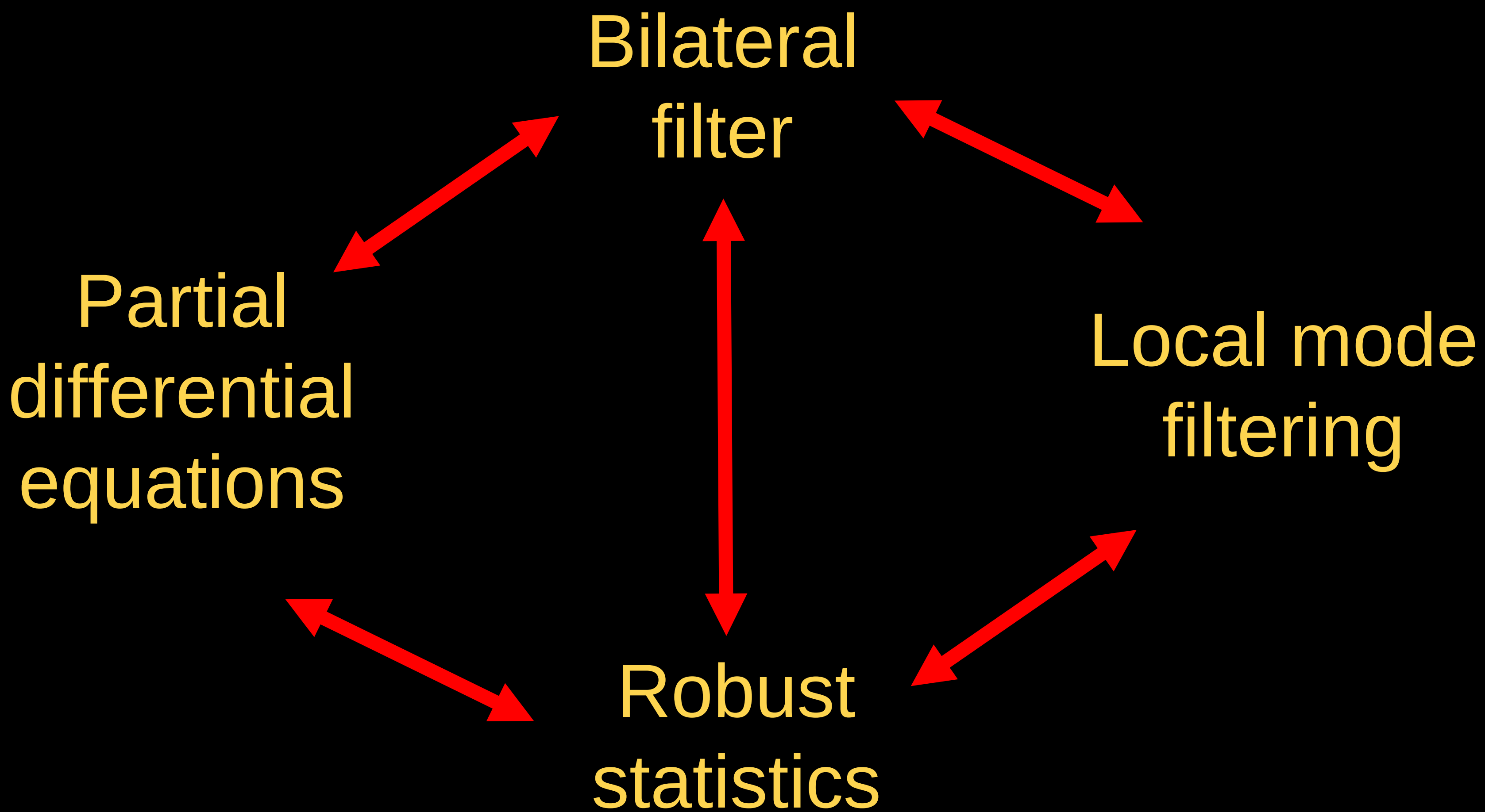
Partial differential equations

Anisotropic diffusion

Local mode filtering

Robust statistics

Goal: Understand how does bilateral filter relates with other methods



Goal: Understand how does bilateral filter relates with other methods

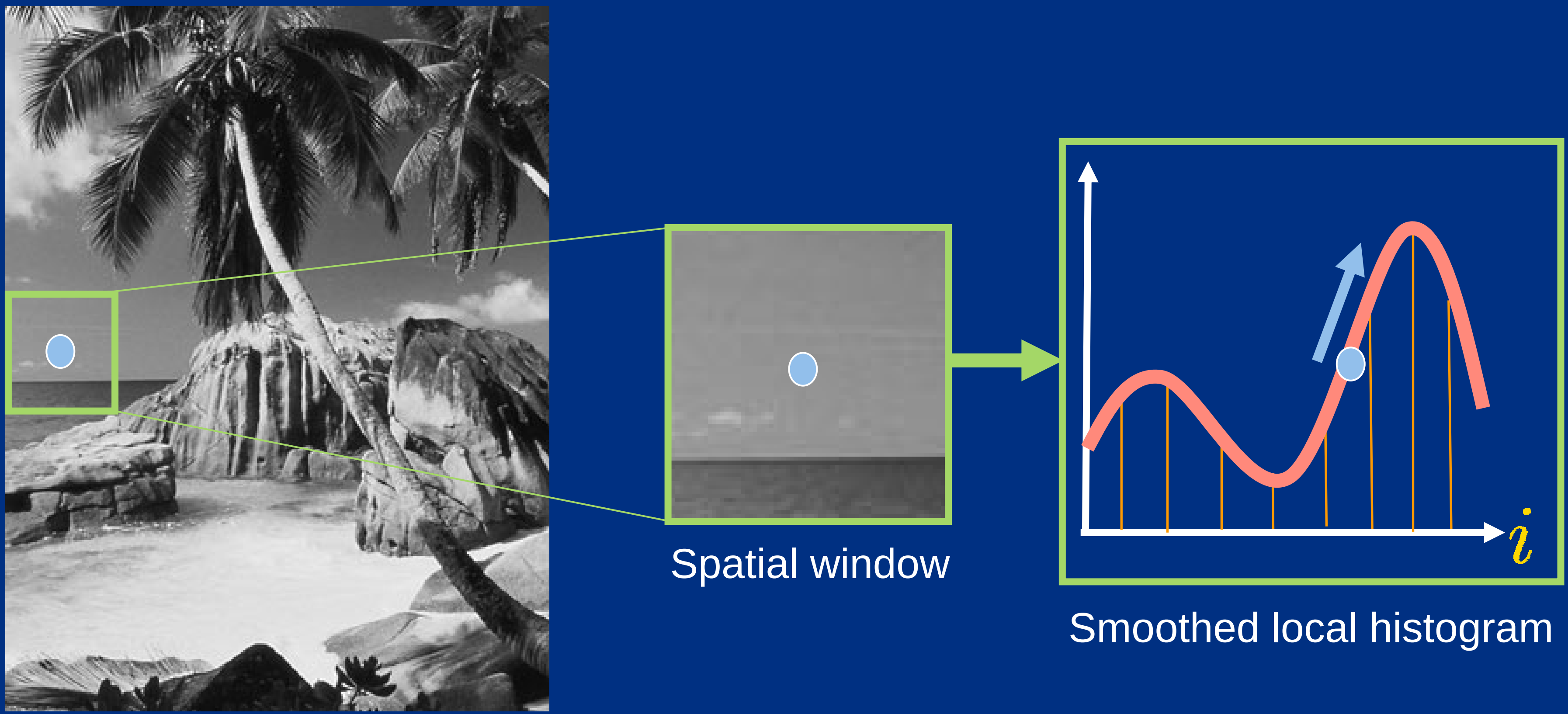
Bilateral filter

Local mode filtering

Partial differential equations

Robust statistics

Local mode filtering principle



You are going to see that BF has the same effect as local mode filtering

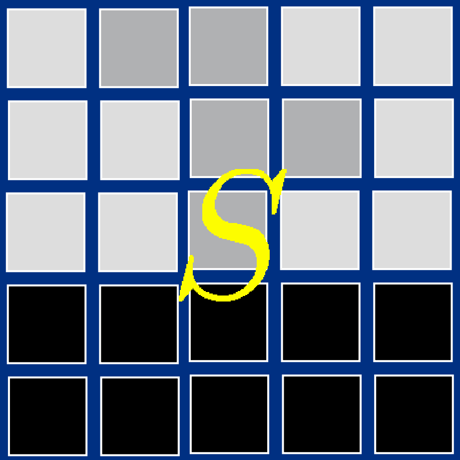
Let’s prove it!

- Define *global* histogram
- Define a *smoothed* histogram
- Define a *local* smoothed histogram
- What does it mean to look for *local modes*?
- What is the *link* with bilateral filter?

Definition of a *global* histogram

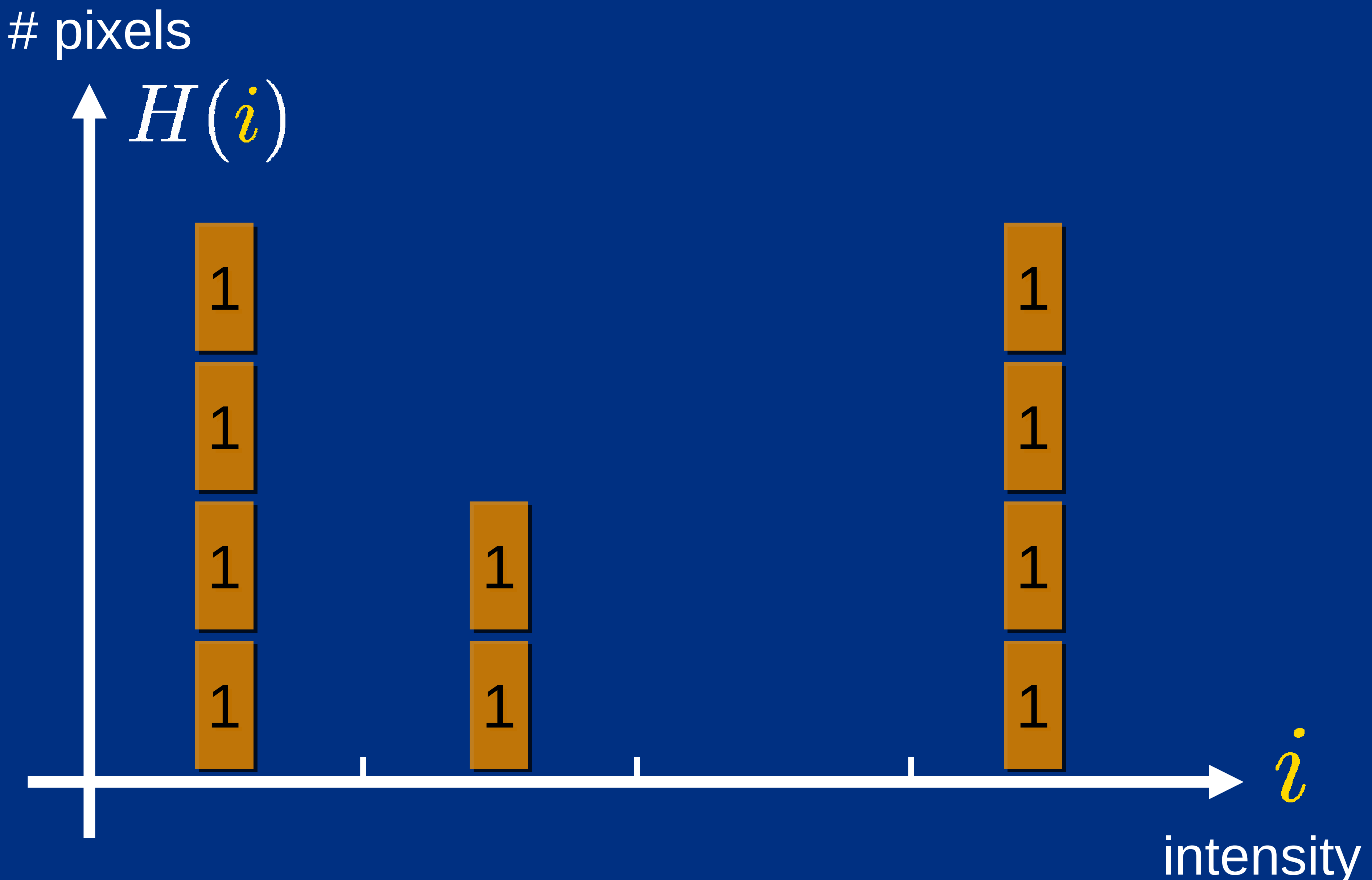
- Formal definition

$$H(i) = \sum_{p \in S} \delta(I_p - i)$$



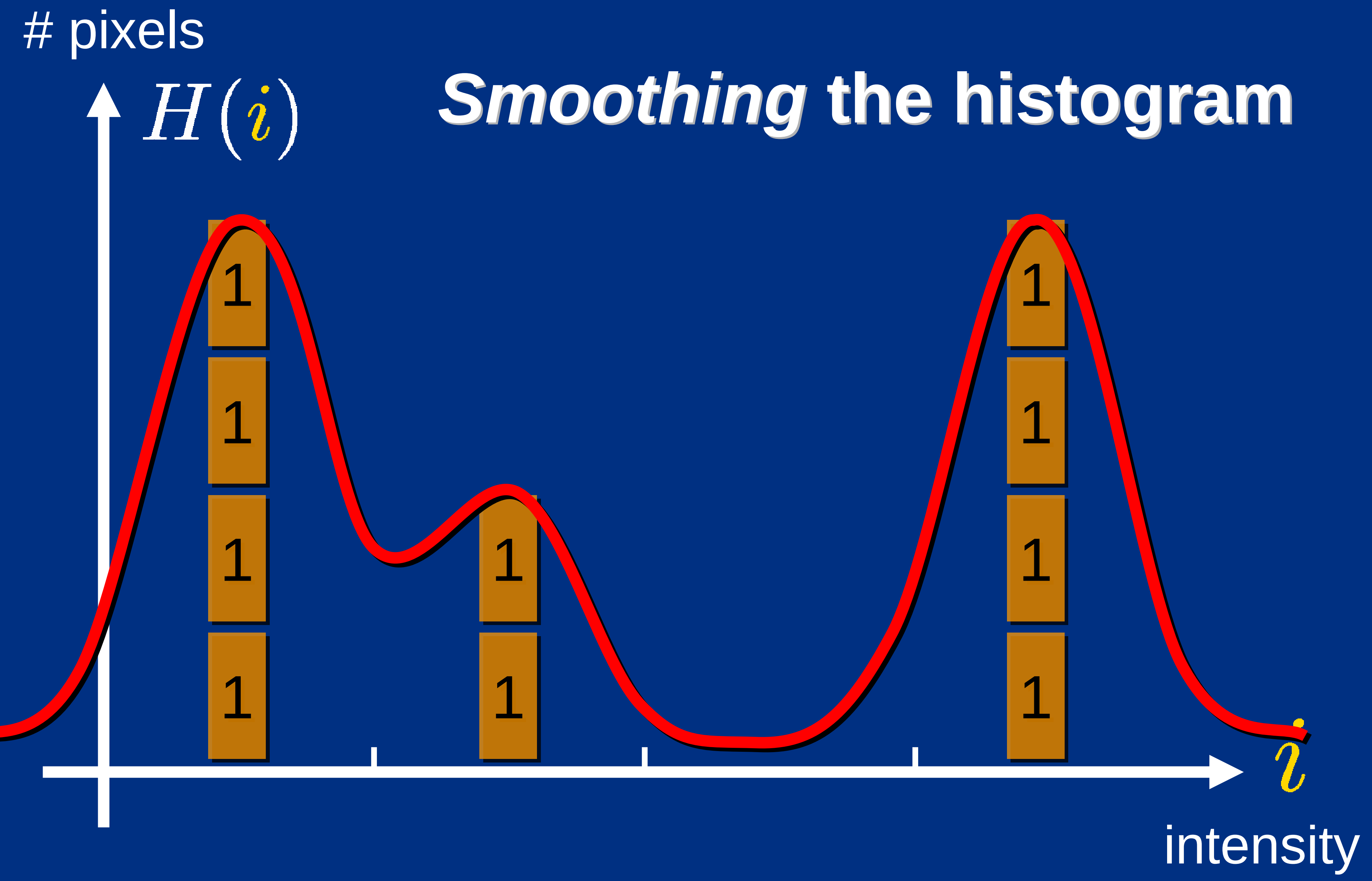
Where $\delta(.)$ is the dirac symbol (1 if t=0, 0 otherwise)

- A sum of dirac, « a sum of ones »



pixels
$H(i)$
1
1
1
1
1
1
1
1
1
1
i
intensity

Smoothing the histogram



Smoothing the histogram

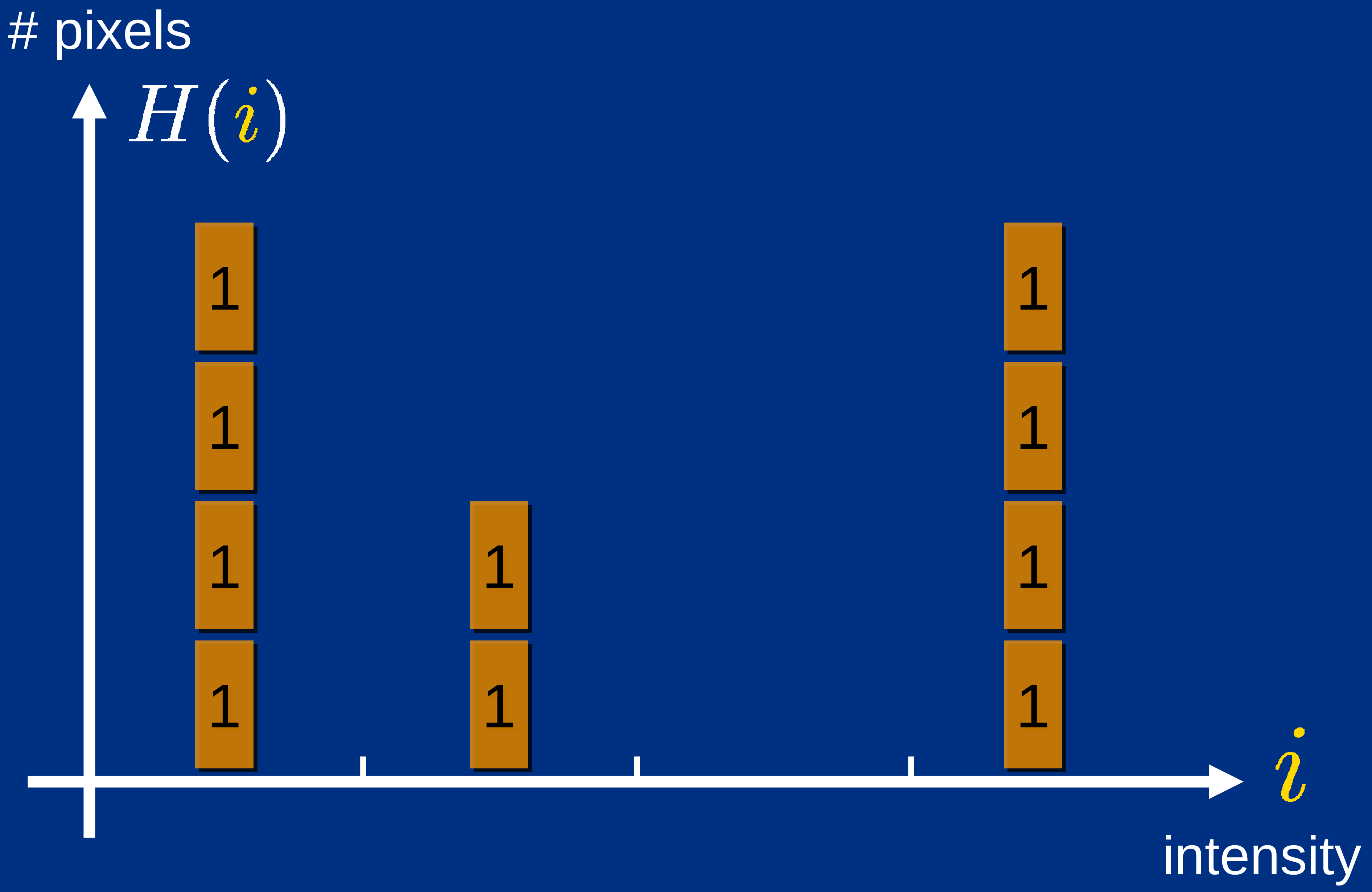
$$H \star G_{\sigma_r}(i) = \sum_{j \in \mathcal{I}} H(j) G_{\sigma_r}(i-j)$$

$$= \sum_{j \in \mathcal{I}} \sum_{p \in S} \delta(I(p) - j) G_{\sigma_r}(i-j)$$

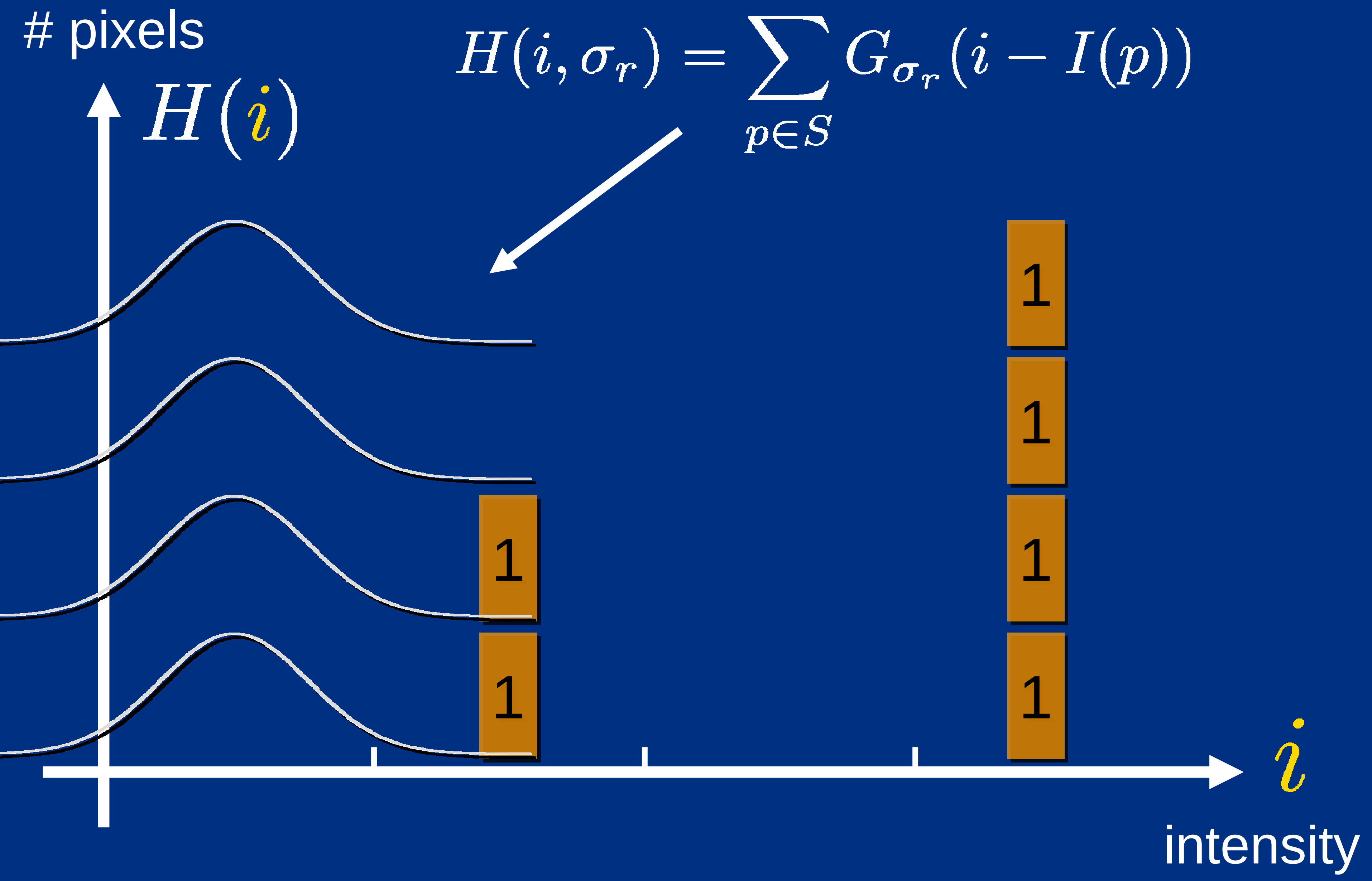
$$= \sum_{p \in S} \sum_{j \in \mathcal{I}} \delta(I(p) - j) G_{\sigma_r}(i-j)$$

$$= \sum_{p \in S} G_{\sigma_r}(i - I(p))$$

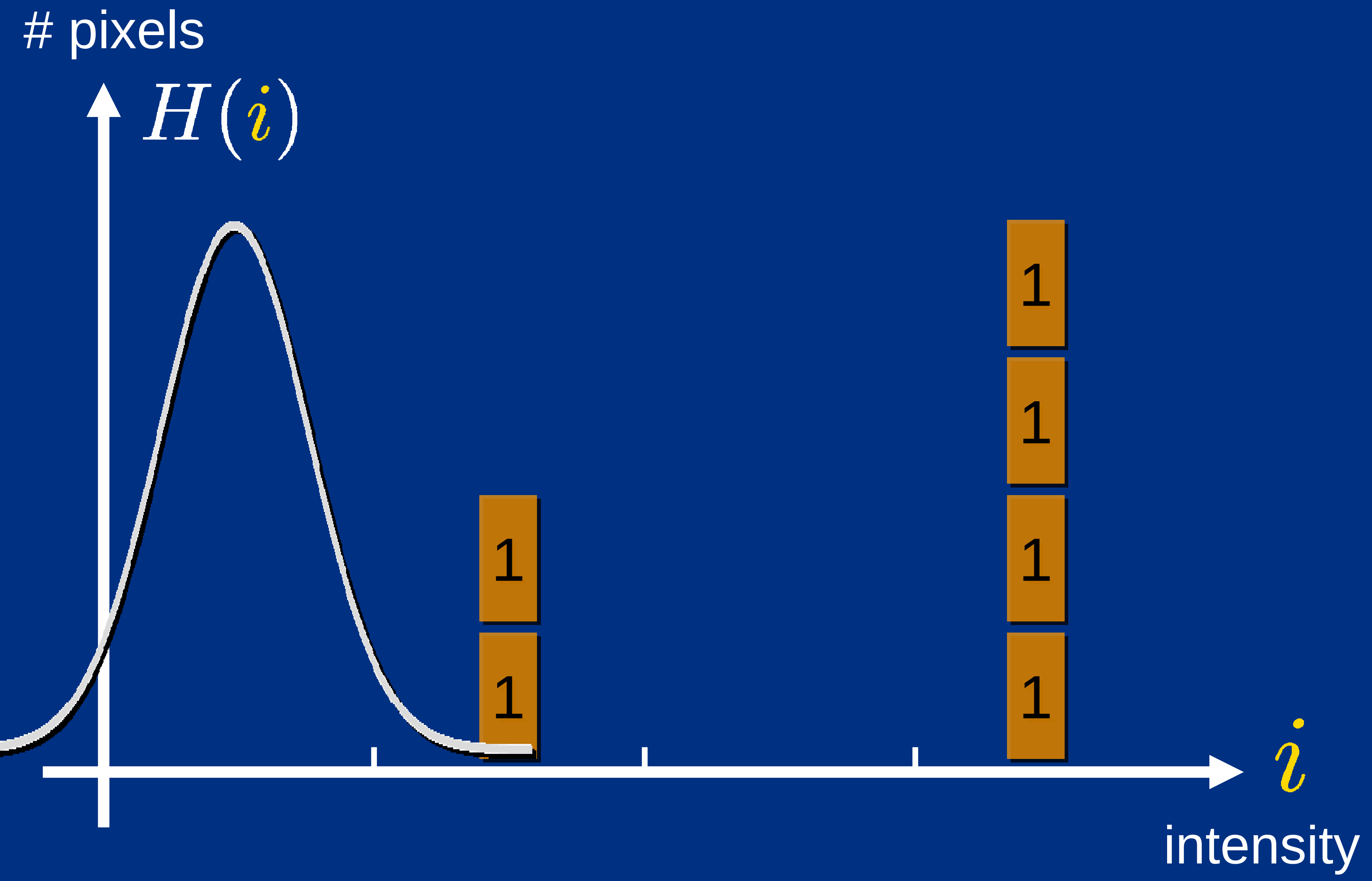
$\delta(I(p) - j) = 0$ unless $j = I(p)$

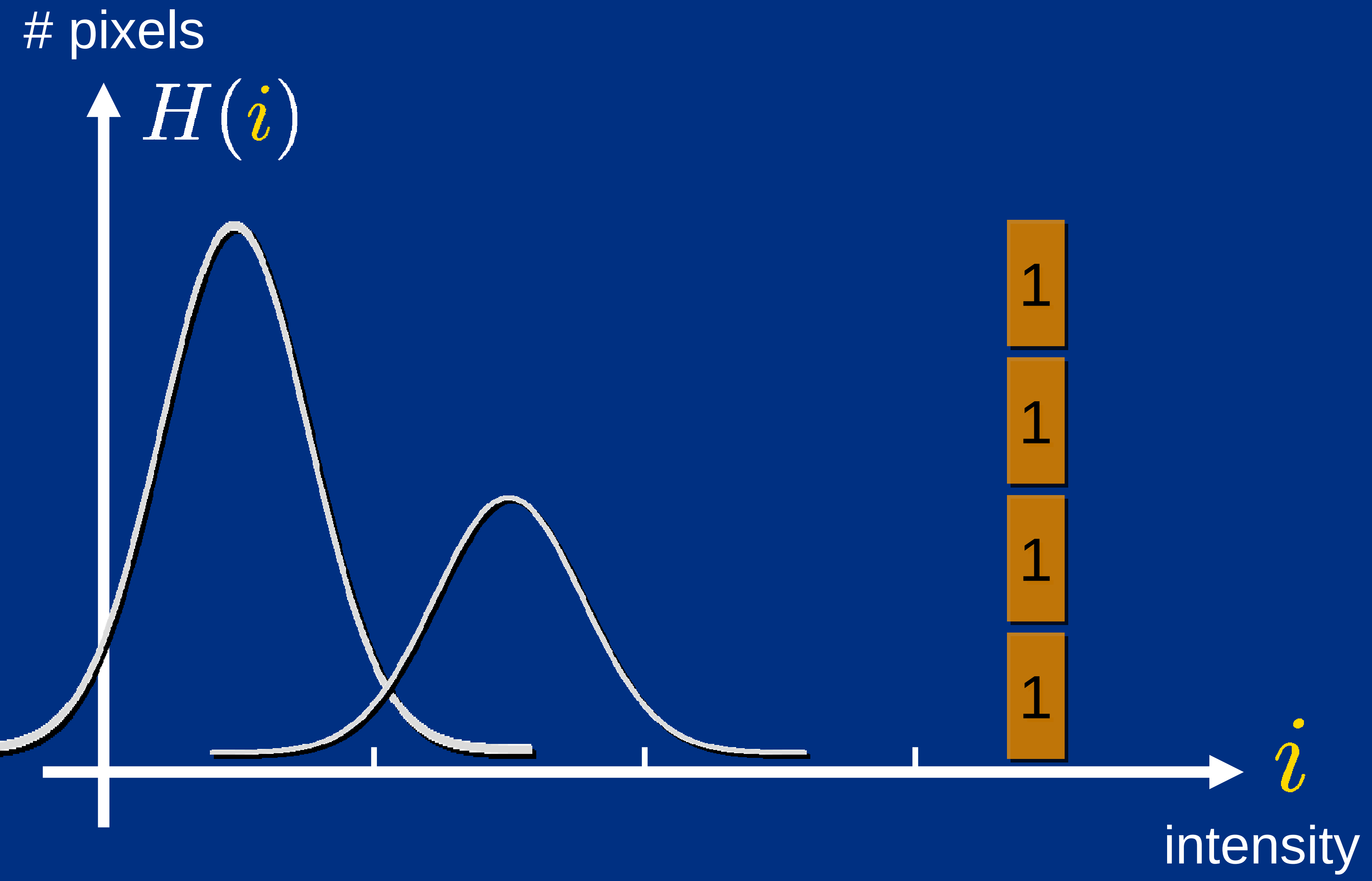


pixels
$H(i)$
1
1
1
1
1
1
1
1
1
1
i
intensity

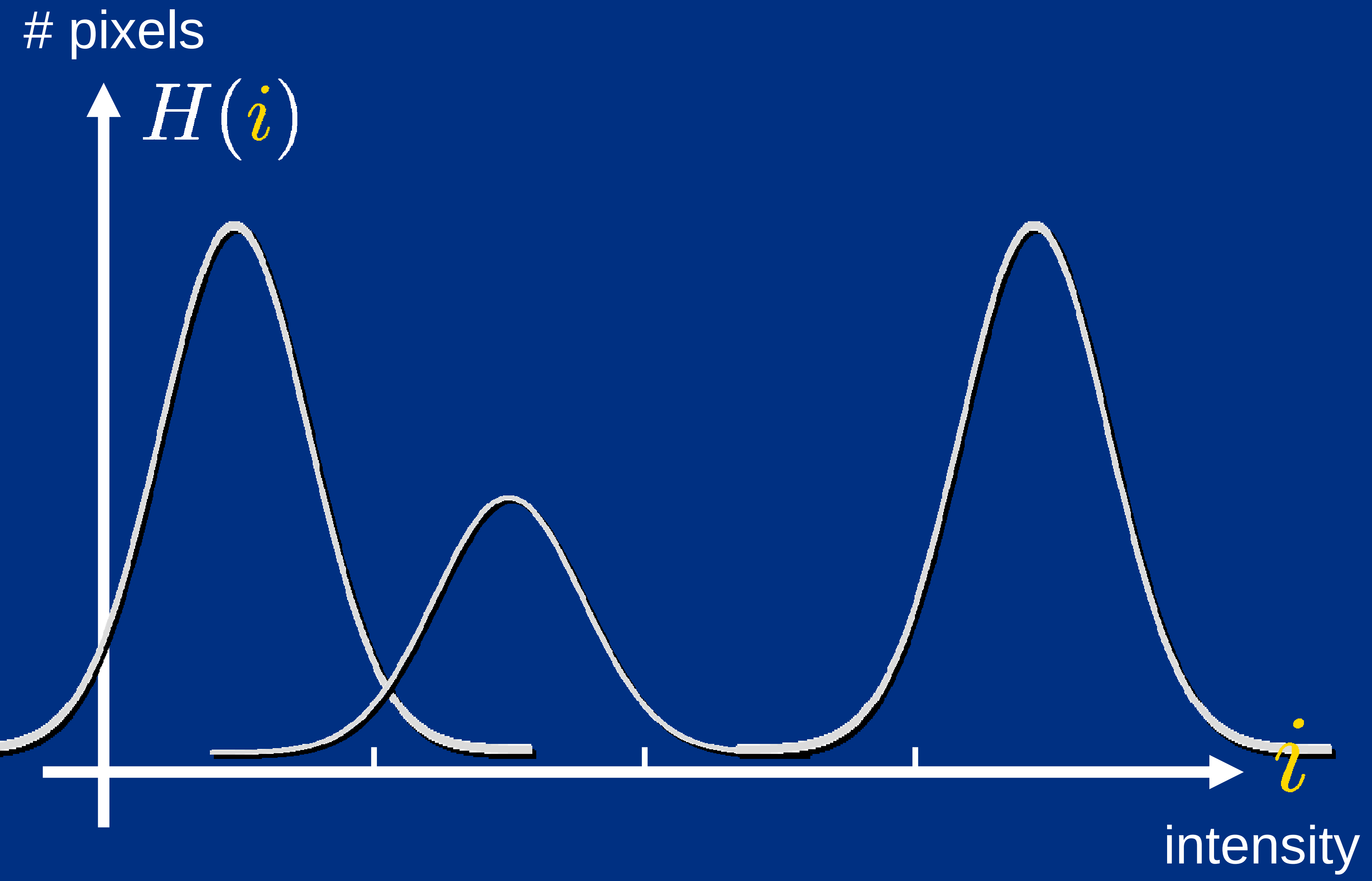


pixels
$H(i)$
$H(i,\sigma_r) = \sum_{p \in S} G_{\sigma_r}(i - I(p))$
1
1
1
1
1
1
i
intensity

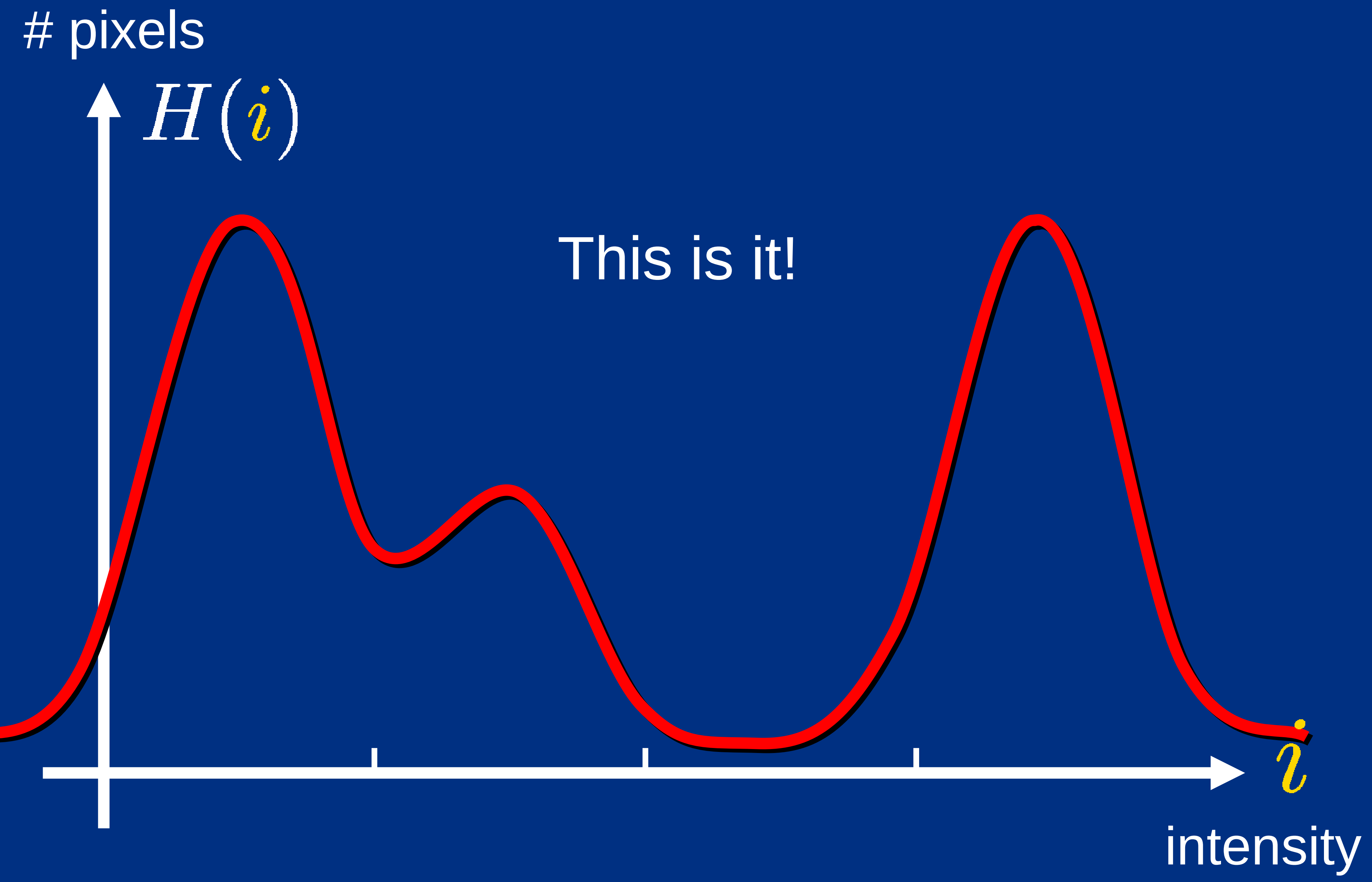

pixels
$H(i)$
1
1
1
1
1
1
i
intensity



pixels
$H(i)$
1
1
1
1
i
intensity



# pixels
H(i)
i
intensity



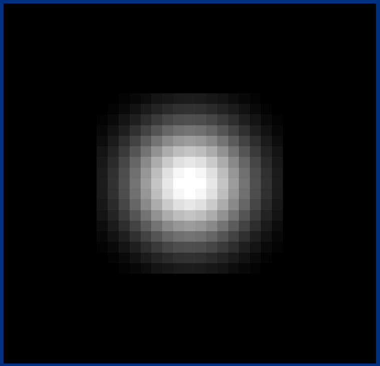
pixels
$H(i)$
This is it!
i
intensity

Definition of a *local* smoothed histogram

- We introduce a « smooth spatial window »

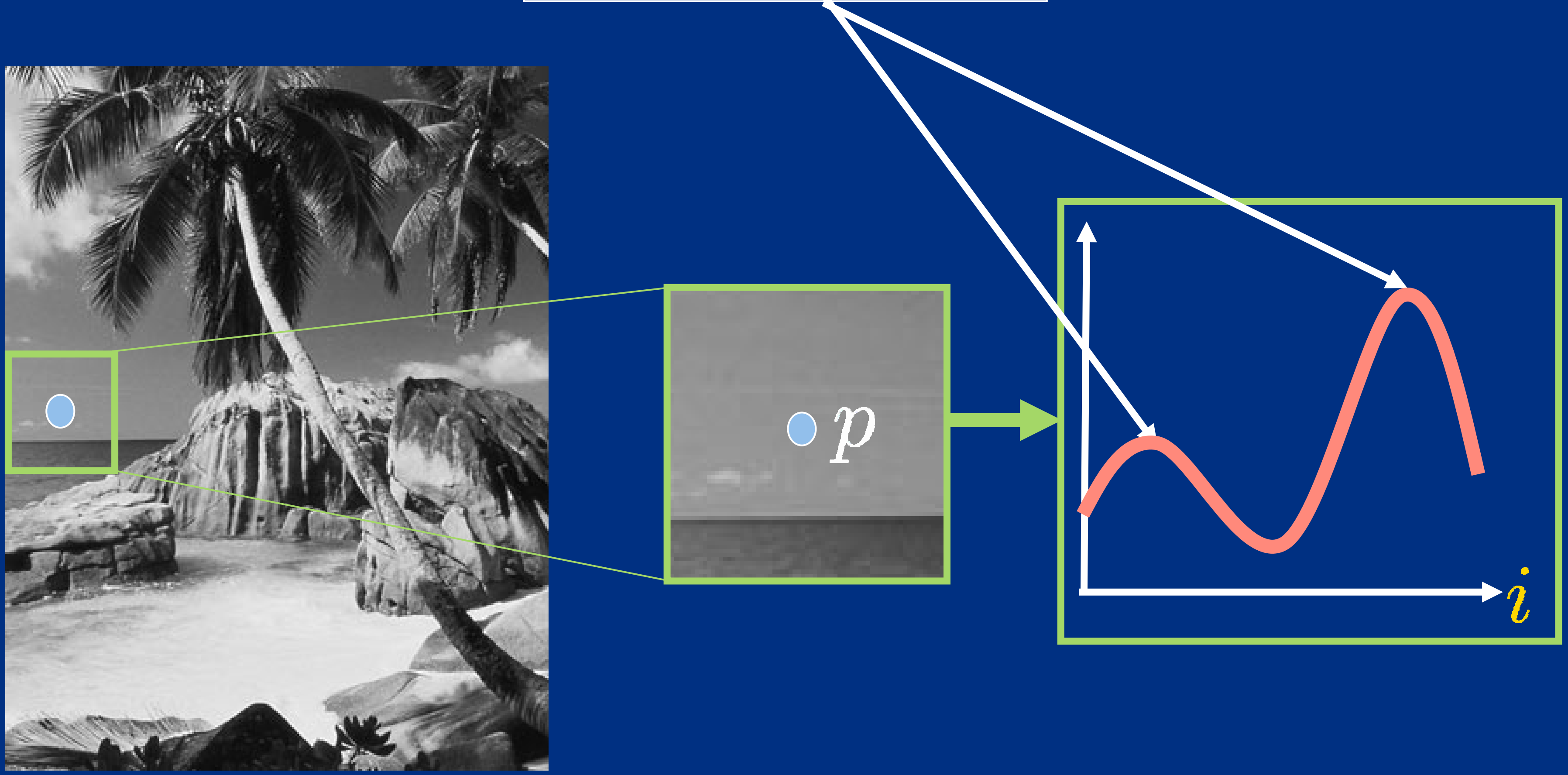
$$H(i, p, \sigma_r, \sigma_s) = \sum_{q \in \Omega} G_{\sigma_s}(p - q) G_{\sigma_r}(i - I(q))$$

where $\begin{cases} \sigma_r = \text{Smoothing of intensities} \\ \sigma_s = \text{Spatial window} \end{cases}$



And that's the formula to have in mind!

Definition of local modes



A local mode *i* verifies $$\frac{\partial}{\partial i} H(i, p, \sigma_r, \sigma_s) = 0$$

Local modes?

- Given

$$H(i, p, \sigma_r, \sigma_s) = \sum_{q \in \Omega} G_{\sigma_s}(p - q) G_{\sigma_r}(i - I_q)$$

- We look for $i \;/\; \frac{\partial}{\partial i} H(i, p, \sigma_r, \sigma_s) = 0$

- Result: $i = \frac{\sum_{q \in \Omega} G_{\sigma_s}(p - q) G_{\sigma_r}(i - I_q) I_q}{\sum_{q \in \Omega} G_{\sigma_s}(p - q) G_{\sigma_r}(i - I_q)}$

Local modes?

One iteration of the bilateral filter
amounts to converge to the local mode

I_p

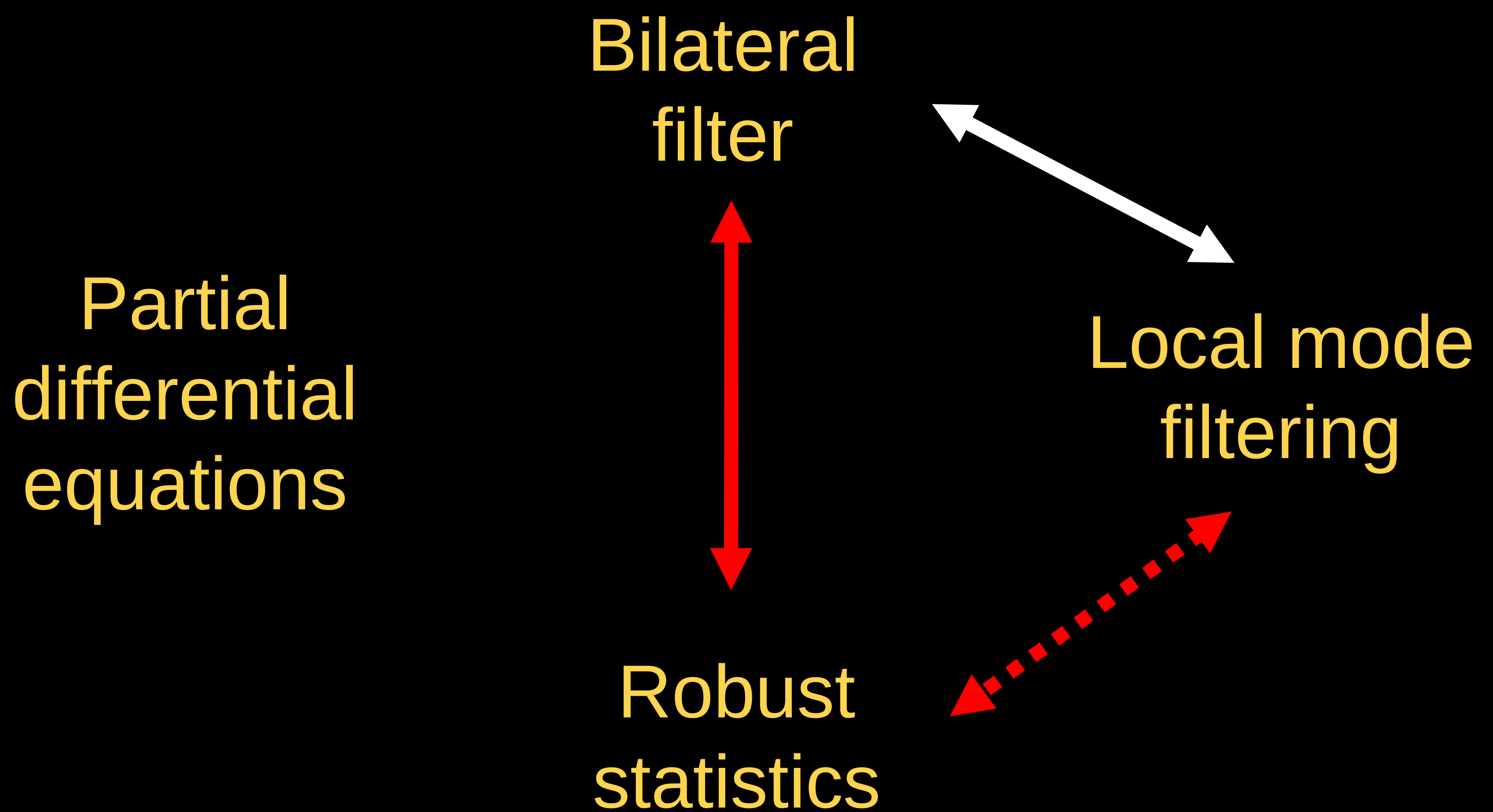
- Result: $i = \frac{\sum_{q \in \Omega} G_{\sigma_s}(p-q) G_{\sigma_r}(i - I_q) I_q}{\sum_{q \in \Omega} G_{\sigma_s}(p-q) G_{\sigma_r}(i - I_q)}$

Take home message #1

Bilateral filter is equivalent to mode filtering in local histograms

[Van de Weijer, Van den Boomgaard, 2001]

Goal: Understand how does bilateral filter relates with other methods



Robust statistics

- Goals: Reduce the influence of outliers, preserve discontinuities
- Minimizing a cost

Robust or not robust?

e.g., $$min_{I_*} \sum_{p \in S} \sum_{q \in \eta_p^4} \rho(I_q - I_p) + \mathrm{etc}$$

Penalizing differences between neighbors

Smoothing term

[Huber 81, Hampel 86]

Robust statistics

- Goals: Reduce the influence of outliers, preserve discontinuities
- Minimizing a cost (« local » formulation)

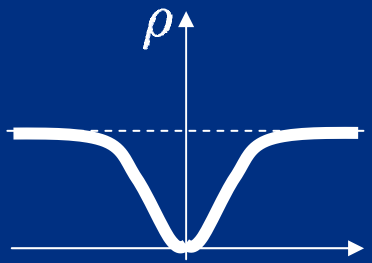
$$min_{I_*} \sum_{p \in S} \sum_{q \in \eta_p} G_{\sigma_s}(q-p)\rho(I_q - I_p)$$

- And to minimize it

$$I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|} \sum_{q \in \eta_p} G_{\sigma_s}(q-p)\rho'(I_q^t - I_p^t)$$

[Huber 81, Hampel 86]

If we choose $\rho(t) = 1 - G_{\sigma_r}(t)$



- The minimization of the error norm gives

$$I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|} \sum_q G_{\sigma_s}(q-p) G_{\sigma_r}(I_q^t - I_p^t)(I_q^t - I_p^t)$$

Iterated reweighted least-square

- The bilateral filter is

$$I_p^{t+1} = \frac{\sum_q G_{\sigma_s}(q-p) G_{\sigma_r}(I_q^t - I_p^t) I_q^t}{\sum_q G_{\sigma_s}(q-p) G_{\sigma_r}(I_q^t - I_p^t)}$$

Weighted average of the data

- So similar! They solve the same minimization problem! [Hampel etal., 1986]: The bilateral filter IS a robust filter!

Back to robust statistics…

Robust or not robust?

Error norm

$$min \sum_{p \in S} \sum_{q \in \eta_p} G_{\sigma_s}(p-q)\rho(I_p - I_q)$$

Influence function

$$I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|} \sum_{q \in \eta_p} G_{\sigma_s}(q-p)\rho'(I_q^t - I_p^t)$$

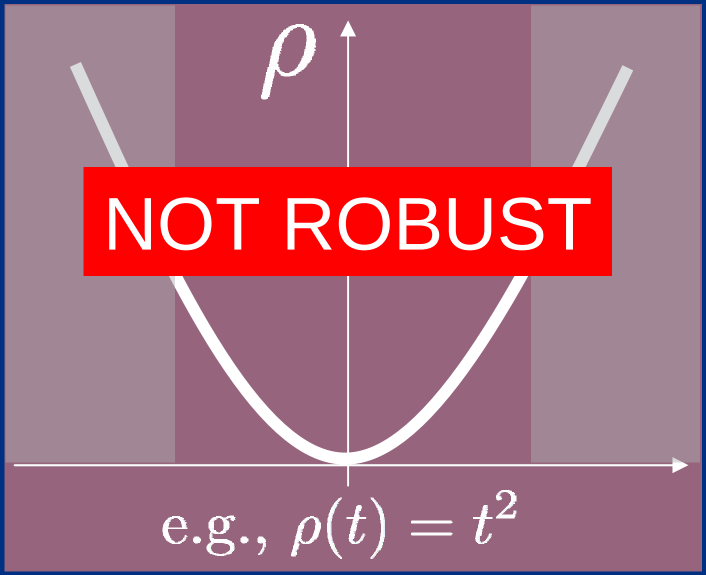
How to choose the error norm? How is the shape related to the anisotropy of the diffusion? *What's the graphical intuition?*

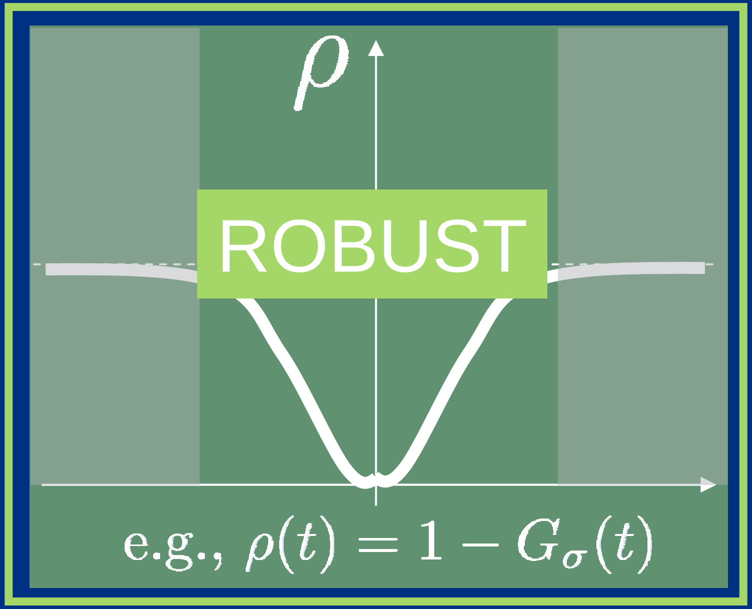
Graphical intuition

From the energy

Error norm

$$\min \sum_{p \in S} \sum_{q \in \eta_p} G_{\sigma_s}(p - q)\rho(I_p - I_q)$$





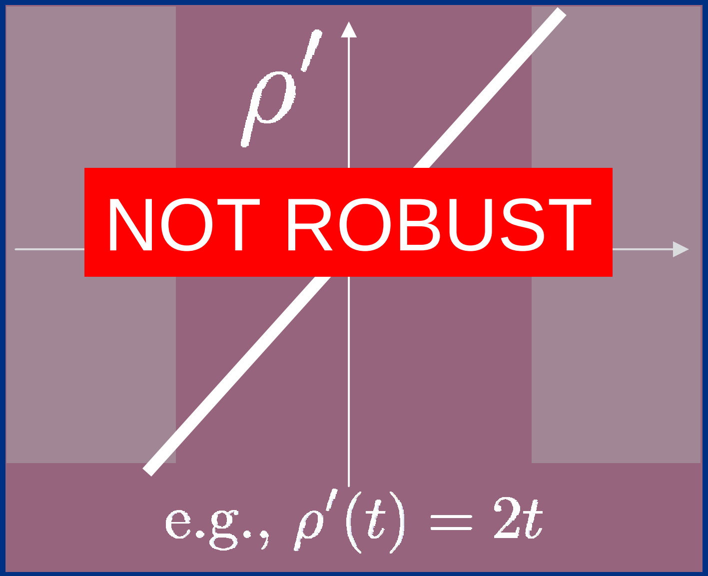
The error norm should not be too penalizing for high differences

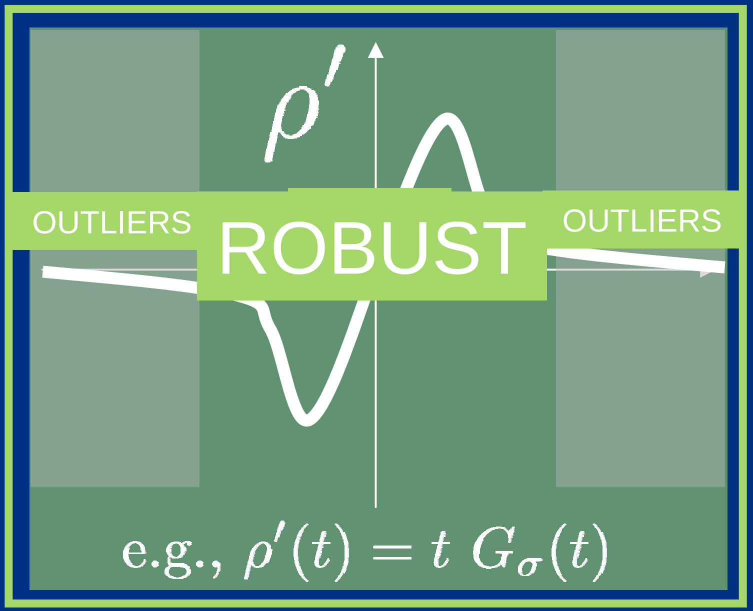
Graphical intuition

From its minimization

Influence function

$$I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|} \sum_{q \in \eta_p} G_{\sigma_s}(q-p)\,\rho'(I_q^t - I_p^t)$$

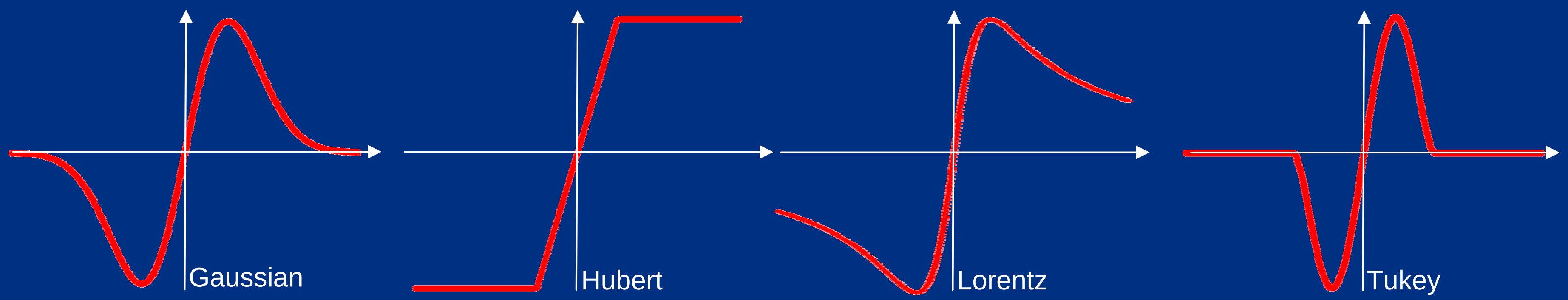




The influence function in the robust case reveals two different behaviors for inliers versus outliers

What is important here?

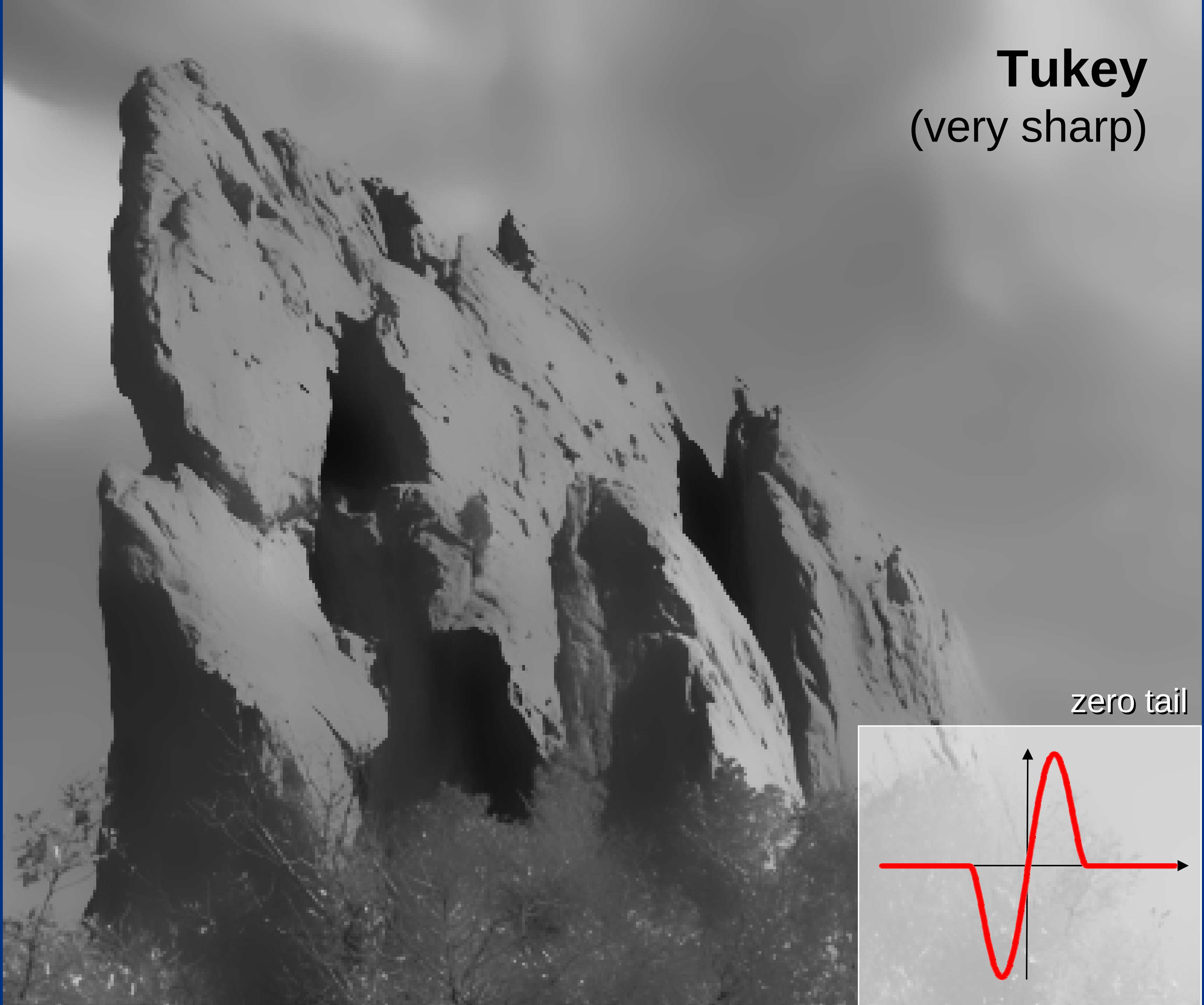
- The qualitative properties of this influence function, distinguishing inliers from outliers.
- In robust statistics, many influence functions have been proposed



Let’s see their difference on an example!

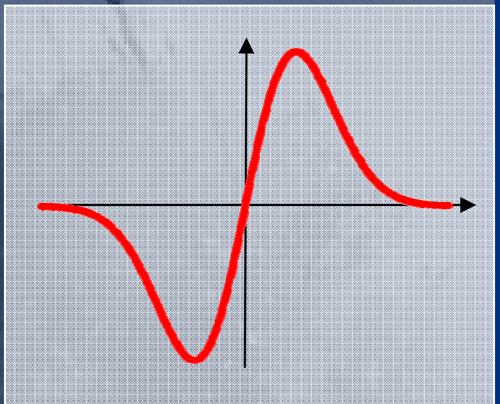

input

Tukey
(very sharp)

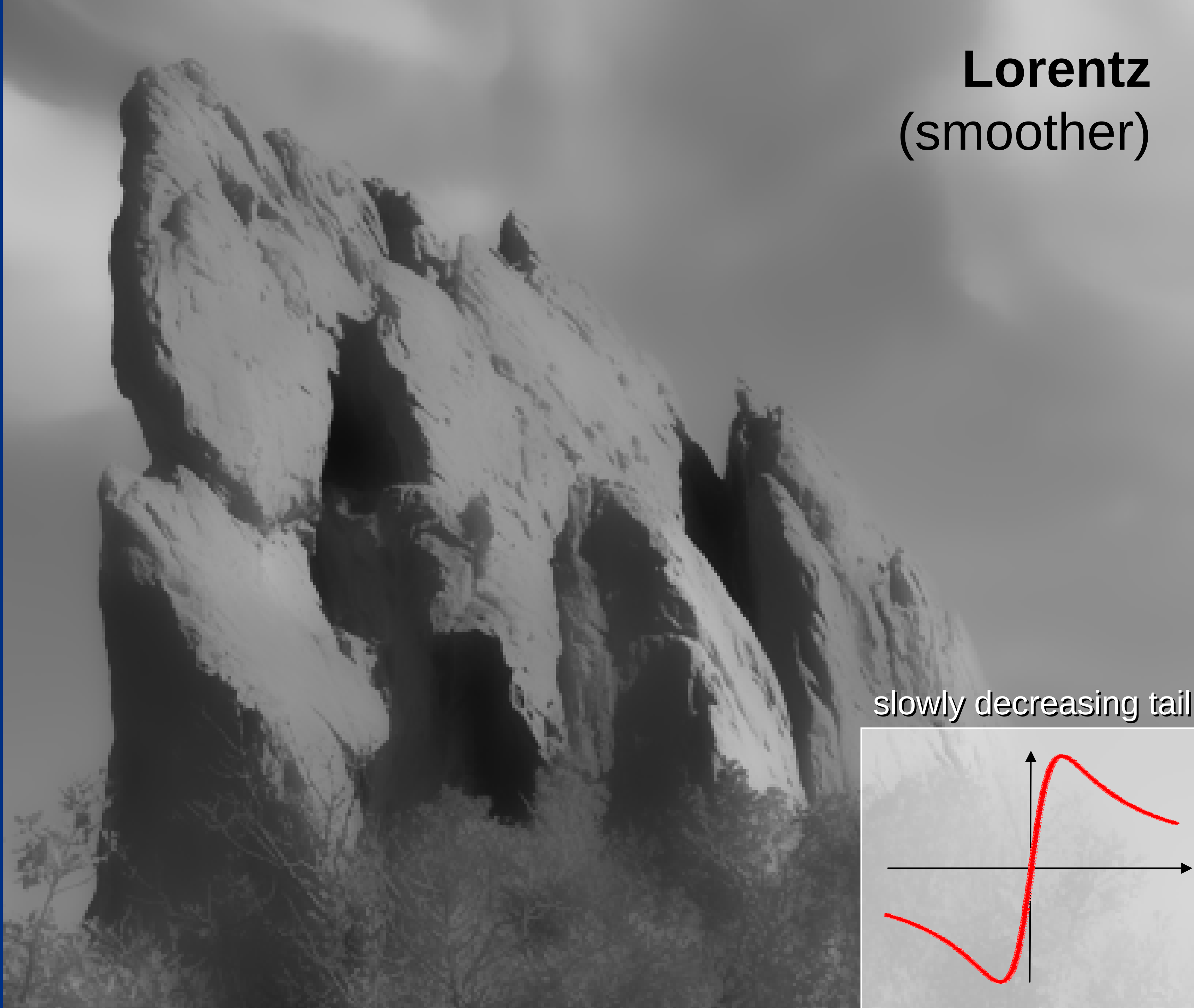


Gauss
(very sharp,
similar to Tukey)

fast decreasing tail

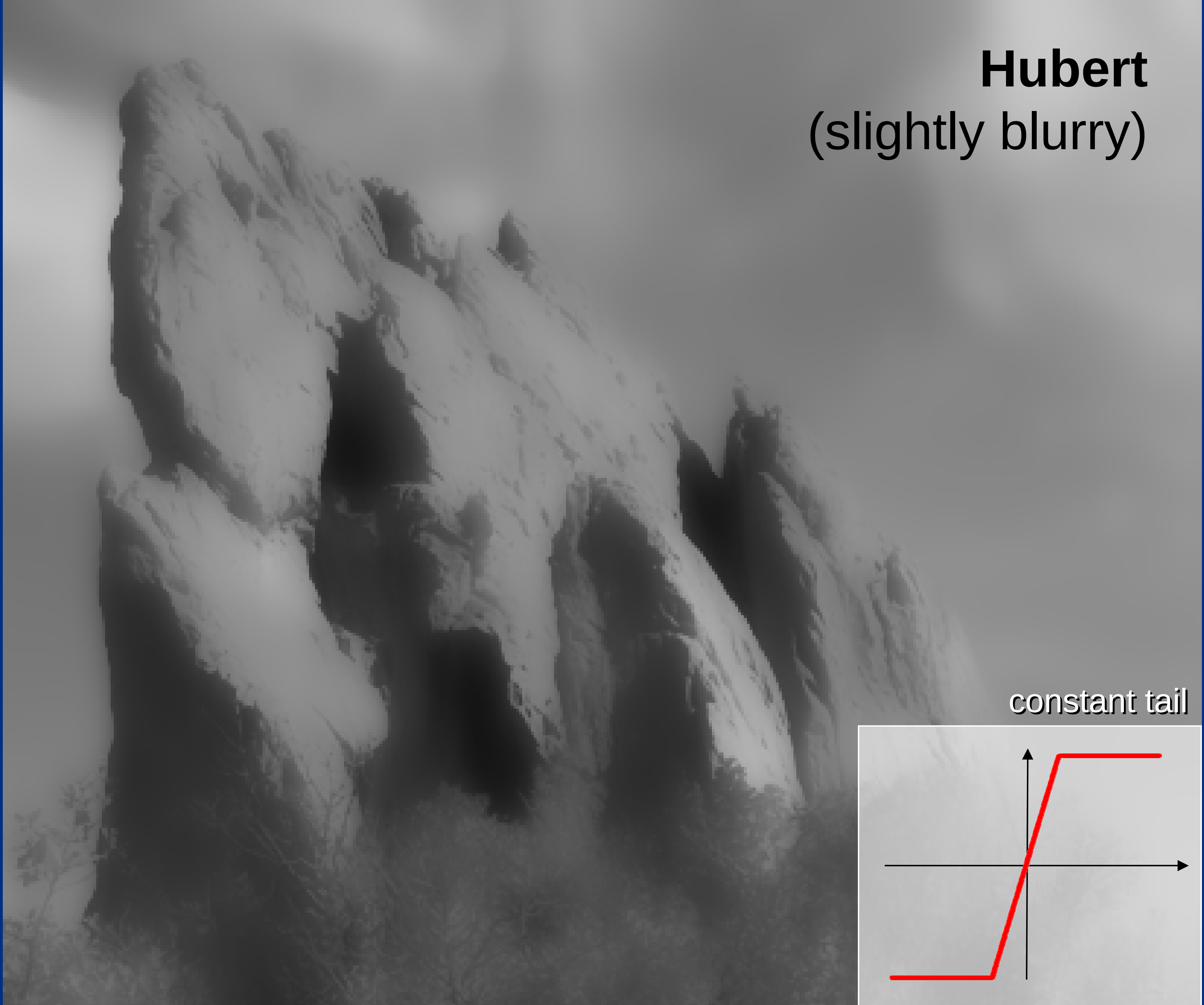


Lorentz
(smoother)



slowly decreasing tail

Hubert
(slightly blurry)



Take home message #2

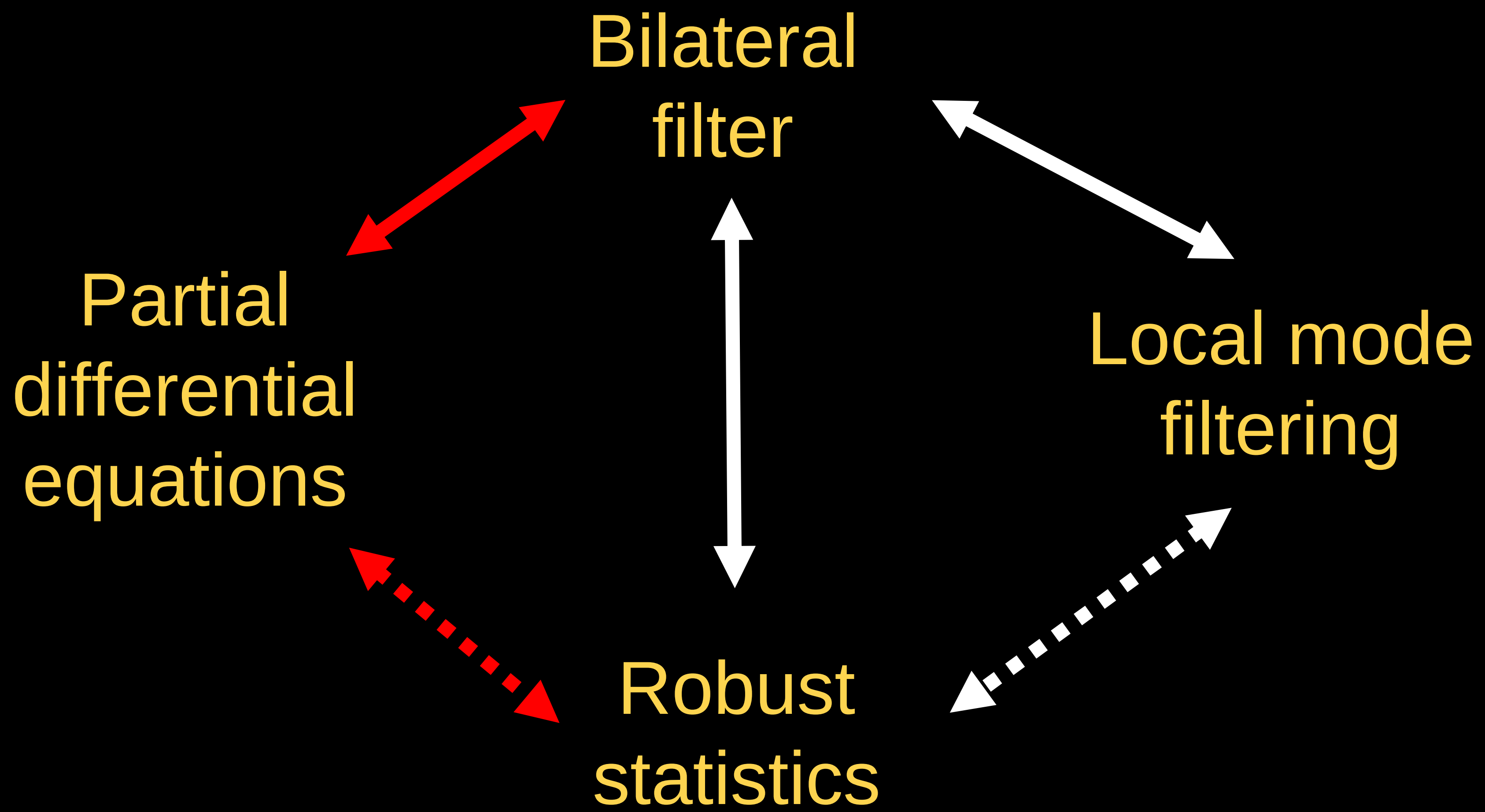
The bilateral filter is a robust filter.

Because of the range weight, pixels with different intensities have limited or no influence. They are *outliers*.

Several choices for the range function.

[Durand, 2002, Durand, Dorsey, 2002, Black, Marimont, 1998]

Goal: Understand how does bilateral filter relates with other methods



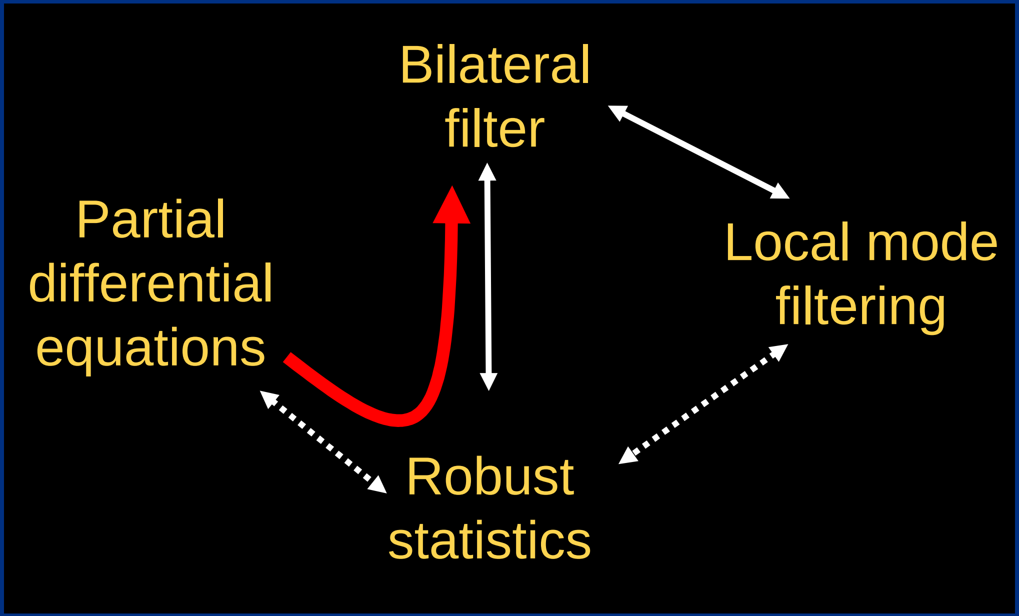
What do I mean by PDEs?

- **Continuous** interpretation of images
- Two kinds of formulations
  - Variational approach $\inf_{I} \int_{p \in \Omega} F(p, I, \nabla I) dp$
  - Evolving a partial differential equation

$$\frac{\partial I}{\partial t} = G(p, I, \nabla I, \ldots)$$

Two ways to explain it

- The « simple one » is to show the link between PDEs and robust statisitcs



Black, Marimont, 1998, etc]

continuous

$$\inf_I \int_{p\in\Omega} \rho(\nabla I)\,dp$$

$$\frac{\partial I}{\partial t} = \mathrm{div}\left(g(\nabla I)\nabla I\right)$$

discrete

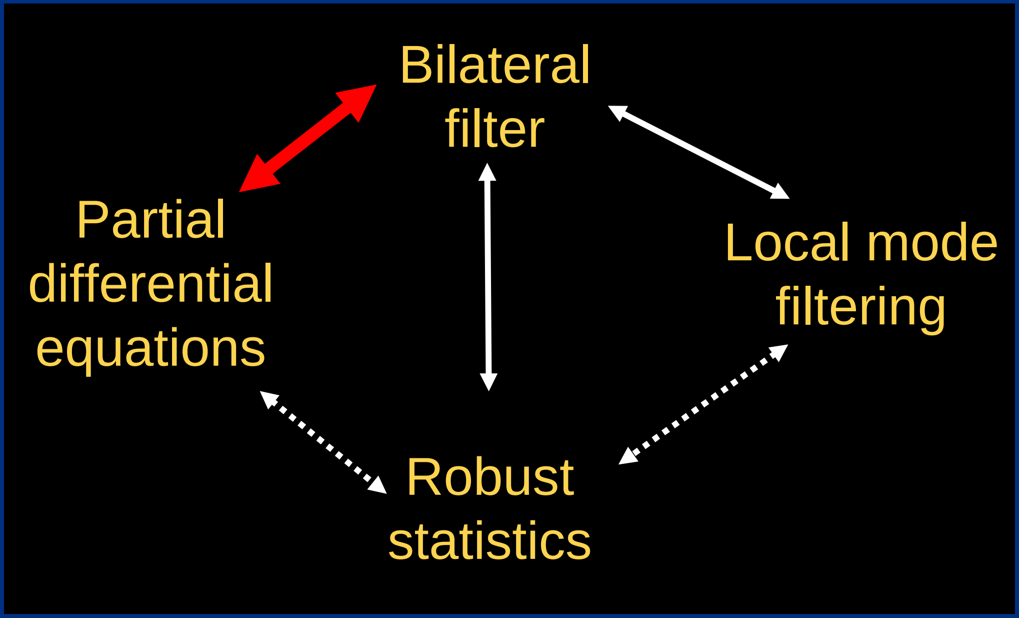
$$\inf_I \sum_{p\in\Omega}\sum_{q\in\eta_p^4} \rho(I_p - I_q)$$

$$I_p^{t+1} - I_p^t = \frac{\lambda}{|\eta_p|}\sum_{q\in\eta_s} g(\nabla I_{p,q})\nabla I_{p,q}$$

$$\text{with...}\ \nabla I_{p,q} = I_q - I_p$$

Two ways to explain it

- The « more rigorous one » is to show directly the link between a differential operator and an integral form



Gaussian solves heat equation

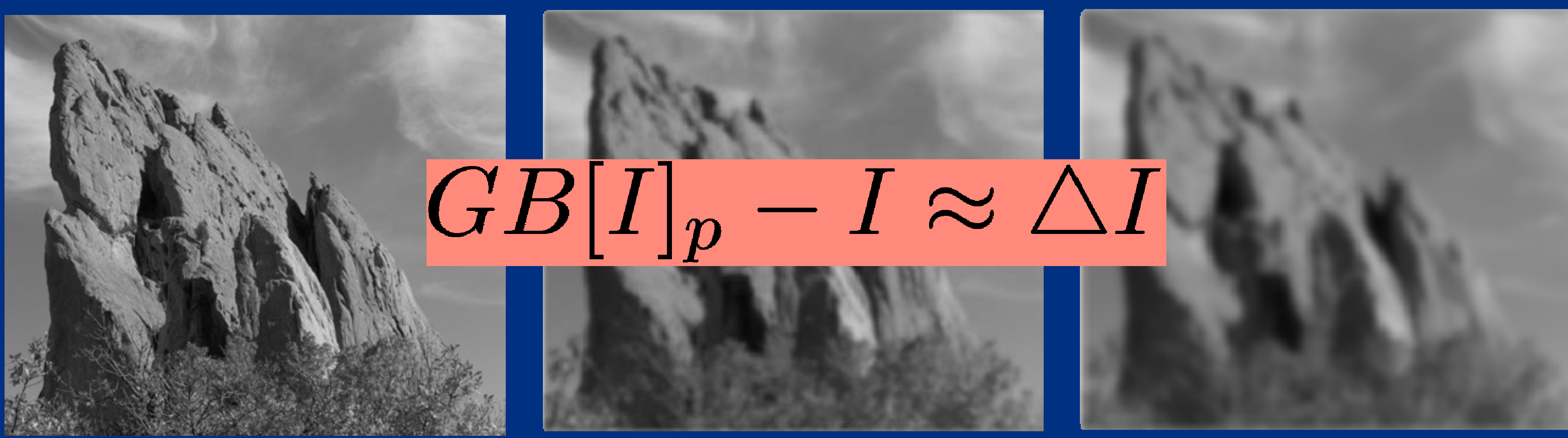
$$\frac{\partial I}{\partial t} = \triangle I = I_{xx} + I_{yy} \qquad t$$

- Linear diffusion
- When time grows, diffusion grows
- Diffusion is isotropic

Gaussian solves heat equation

$$\frac{\partial I}{\partial t} = \triangle I = I_{xx} + I_{yy}$$

t



$$GB[I]_p - I \approx \triangle I$$

σ_s

$$GB[I]_p = \int_S G_{\sigma_s}(q-p) I_q \, dq$$

$GB[I]_p$ Is a solution of the heat equation when $\sigma_s = \sqrt{2t}$

And with the range?

[Buades, Coll, Morel, 2005]

- Considering the Yaroslavsky Filter

$$Y[I]_p = \frac{1}{C(p)} \int_{B_{\sigma_s}(p)} G_{\sigma_r}(I_q - I_p) I_q dq$$

Integral representation
Space range is in the domain

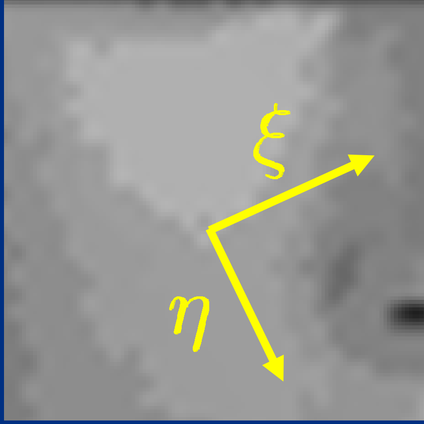
- When $\sigma_s, \sigma_r \to 0$

$$Y[I]_p - I_p \approx \text{nonlinear diffusion operators}$$

(operation similar to M-estimators)

At a very local scale, the asymptotic behavior of the integral operator corresponds to a diffusion operator

More precisely



- We have

$$Y[I]_p - I_q \approx \sigma_s^2 [\, \blacksquare \, I_{\eta\eta} + \blacksquare \, I_{\xi\xi}]$$

- And then we enter a large class of anisotropic diffusion approaches based on PDEs

$$\frac{\partial I}{\partial t} = [\, \blacksquare \, I_{\eta\eta} + \blacksquare \, I_{\xi\xi}]$$

New idea here: It is not only a matter of smoothing or not, but also to take into account the local structure of the image

Take home message #3

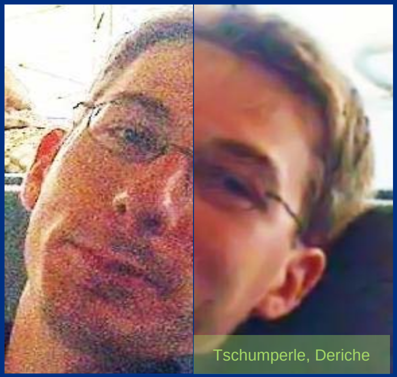
Bilateral filter is a discretization of a particular kind of a PDE-based anisotropic diffusion.

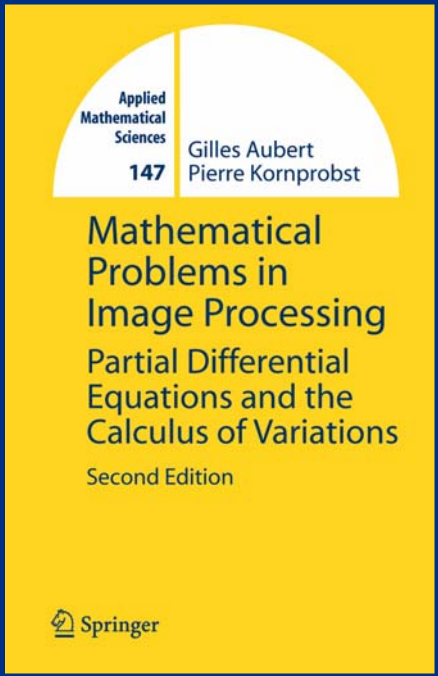
[Barash 2001, Elad 2002, Durand 2002, Buades, Coll, Morel, 2005]

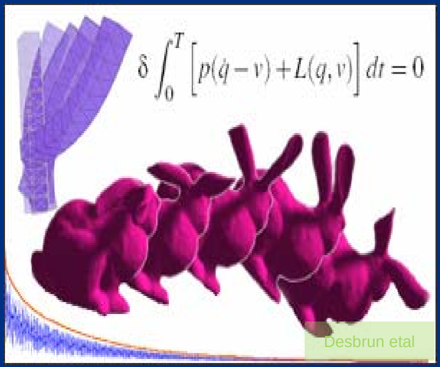
Welcome to the PDE-world!

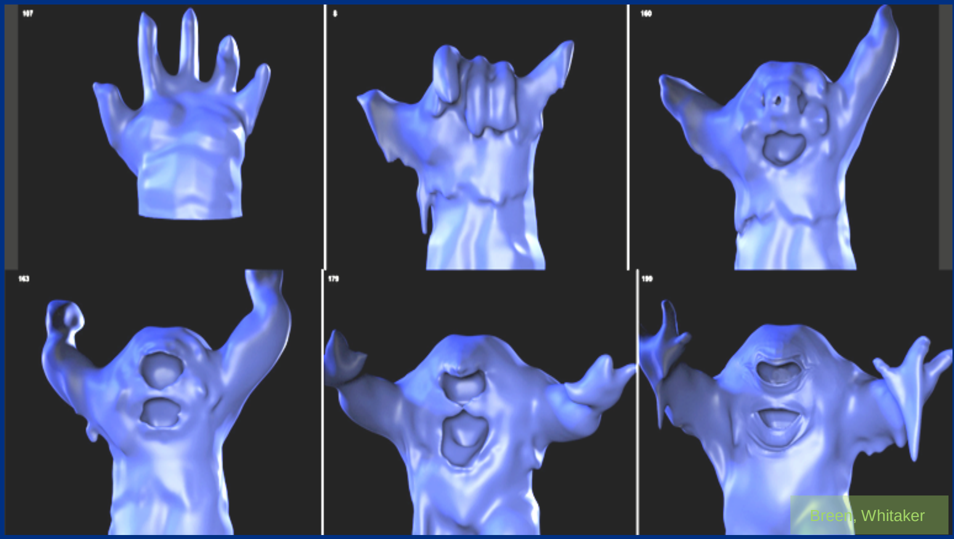
[Kornprobst 2006]

The PDE world at a glance



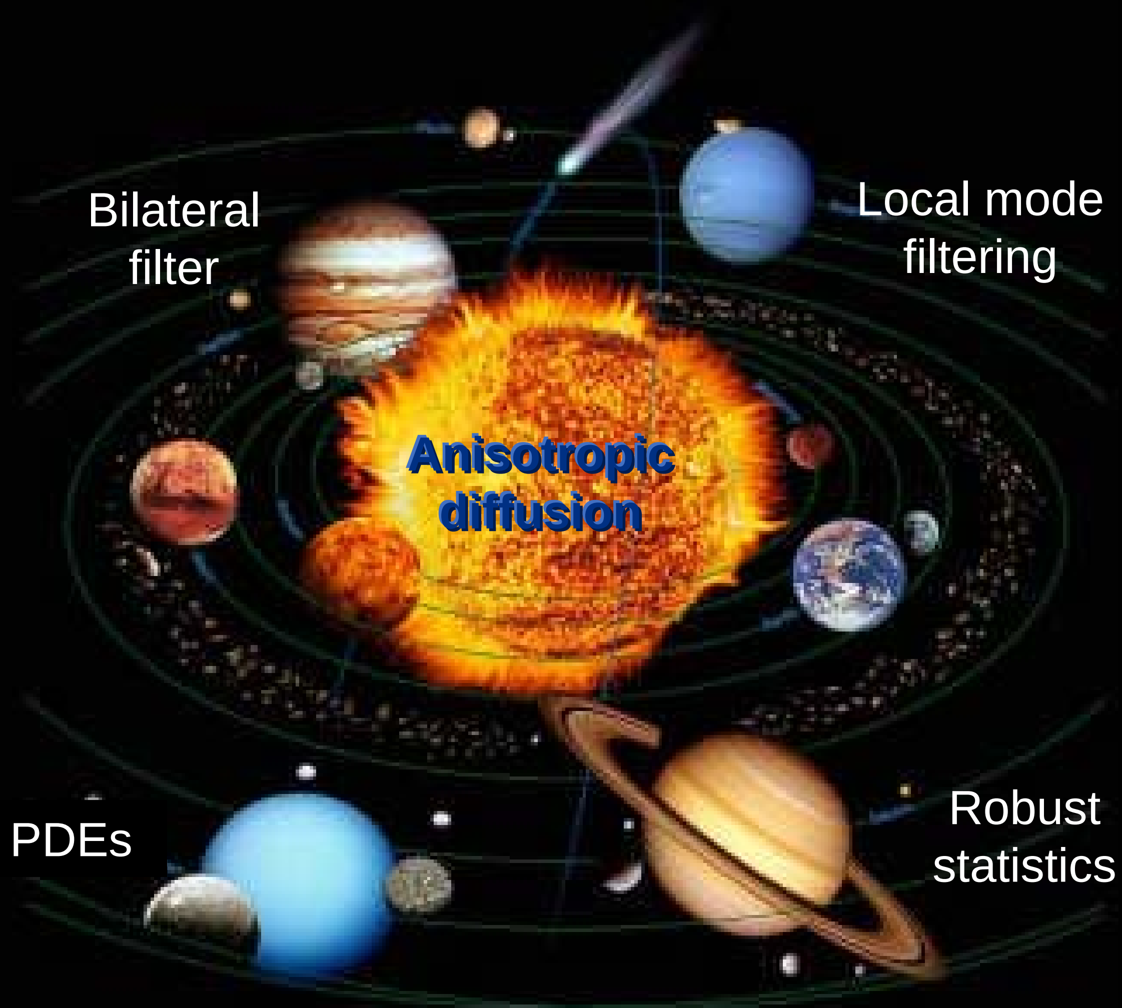






Summary

Bilateral filter is one technique for anisotropic diffusion and it makes the bridge between several frameworks. From there, you can explore news worlds!



Questions?

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche SOPHIA ANTIPOLIS - MÉDITERRANÉE

Pierre.kornprobst@inria.fr
http://pierre.kornprobst.googlepages.com/