

Lecture 20: Reweighting, and the One-Step Optimal Line Search

Chapter 8.3–8.4

Charles A. Bouman

Purdue University

Fall 2026

Today's Big Ideas

- Reweighting recasts the prior. A prior that is the *same everywhere* but *not Gaussian* becomes a *Gaussian* prior whose weights *vary across the image*.
- The new pairwise weight $\tilde{b}_{s,r}$ is *small at an edge*. So edges get weak smoothing and are preserved.
- Because the cost is now Gaussian (quadratic), the best step size is one exact formula, $\alpha^* = -\theta_1/\theta_2$. No search.
- That one step is the engine of a practical reconstruction.

Reweighting recasts the prior

The true MAP cost has a fixed weight $b_{s,r}$ and a non-Gaussian potential ρ :

$$f(x) = \frac{1}{2} \|y - Ax\|_{\Lambda}^2 + \sum_{\{s,r\} \in \mathcal{P}} b_{s,r} \rho(x_s - x_r) .$$

The surrogate keeps the data term and replaces ρ by a plain square, with a new weight $\tilde{b}_{s,r}$:

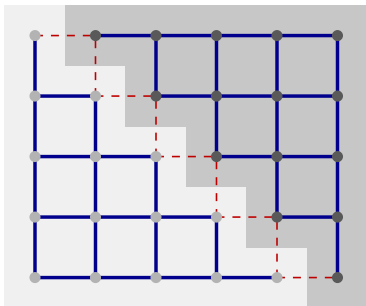
$$Q(x; x') = \frac{1}{2} \|y - Ax\|_{\Lambda}^2 + \sum_{\{s,r\} \in \mathcal{P}} \tilde{b}_{s,r} (x_s - x_r)^2 .$$

The only change is $b_{s,r} \rho(\cdot) \rightarrow \tilde{b}_{s,r} (\cdot)^2$. The new weight, computed at the current image x' , is

$$\tilde{b}_{s,r} = b_{s,r} \frac{\rho'(\Delta')}{2 \Delta'} , \quad \Delta' = x'_s - x'_r .$$

- **Same everywhere** $\rightarrow$ **varies with the image.**
- **Non-Gaussian** $\rightarrow$ **Gaussian.**

In the image: the weights drop along an edge



Two neighbors that straddle an edge differ a lot, so their surrogate weight $\tilde{b}_{s,r}$ is small.

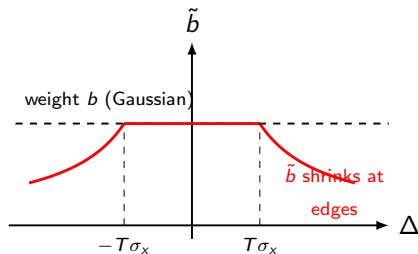
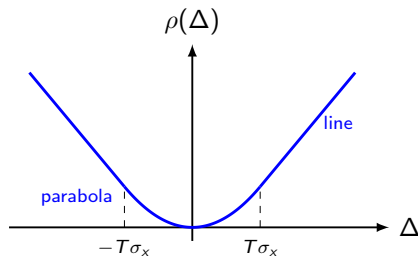
- **Strong** bonds inside a smooth region: full smoothing.
- **Weak** bonds across the edge: the edge is kept.

The prior is the same everywhere. It is only *this* image's edges, through the differences $x_s - x_r$, that make the surrogate a **non-homogeneous** Gaussian model.

Example: the Huber potential

The influence is clipped, so the weight is flat near 0 and shrinks at edges:

$$\frac{\tilde{b}_{s,r}}{b_{s,r}} = \frac{\rho'(\Delta')}{2\Delta'} = \begin{cases} \frac{1}{2\sigma_x^2} & |\Delta'| \leq T\sigma_x, \\ \frac{T}{2\sigma_x |\Delta'|} & |\Delta'| > T\sigma_x. \end{cases}$$



Weight starts equal to the Gaussian weight, then collapses at edges, so the edge is preserved.

Reference: surrogate weight for each potential

Potential	Weight factor $\tilde{b}_{s,r}/b_{s,r}$
Quadratic	$\frac{1}{2\sigma_x^2}$
Total Variation	$\frac{1}{2\sigma_x \Delta' }$ (infinite at $\Delta' = 0$)
Huber	$\frac{\text{clip}\{\Delta', [-T\sigma_x, T\sigma_x]\}}{2\sigma_x^2 \Delta'}$
log cosh	$\left \frac{T \tanh(\Delta'/T\sigma_x)}{2\sigma_x \Delta'} \right $
GGMRF	$\frac{ \Delta' ^{p-2}}{2\sigma_x^p}$
QGGMRF ($p = 2$)	$\frac{1}{2\sigma_x^2} \cdot \frac{ \Delta'/T\sigma_x ^{q-2} (q/2 + \Delta'/T\sigma_x ^{q-2})}{(1 + \Delta'/T\sigma_x ^{q-2})^2}$

Every weight shrinks as $|\Delta'|$ grows, so every one preserves edges. Only Total Variation blows up at $\Delta' = 0$.

The line search along a direction

Recall the surrogate we are minimizing, the Gaussian cost with weights $\tilde{b}_{s,r}$:

$$Q(x; x') = \frac{1}{2} \|y - Ax\|_{\Lambda}^2 + \sum_{\{s,r\} \in \mathcal{P}} \tilde{b}_{s,r} (x_s - x_r)^2 .$$

Given the current image x and a search direction d , the line search steps along d to the bottom of Q :

$$\alpha^* = \arg \min_{\alpha} Q(x + \alpha d; x) .$$

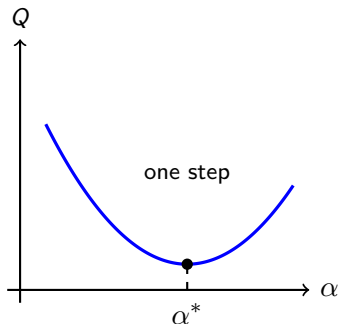
Because Q is quadratic in x , it is an exact quadratic in the scalar α ,

$$Q(x + \alpha d; x) = c + \theta_1 \alpha + \frac{1}{2} \theta_2 \alpha^2 .$$

One exact step to the bottom

The minimizer is one closed-form expression:

$$\alpha^* = \frac{-\theta_1}{\theta_2} = \frac{-\text{slope}}{\text{curvature}}$$



- $\theta_1 = d^t \nabla f(x)$ is the *true* slope. The surrogate matches f in value and slope at x .
- $\theta_2 = d^t H_Q d$ is the *surrogate* curvature along d . Here H_Q is the second-derivative (Hessian) matrix of the surrogate cost Q :
 $H_Q = A^t \Lambda A + [\text{prior term}]$.
- The minus sign steps downhill.

Computing the two terms cheaply

Expand $Q(x + \alpha d; x)$ and collect the powers of α . The coefficient of α is θ_1 , of $\frac{1}{2}\alpha^2$ is θ_2 . Each splits into a data part and a prior part:

$$\theta_1 = \underbrace{-(y - Ax)^t \Lambda A d}_{\text{data slope}} + \underbrace{\sum_{\{s,r\} \in \mathcal{P}} 2\tilde{b}_{s,r}(x_s - x_r)(d_s - d_r)}_{\text{prior slope}}$$

$$\theta_2 = \underbrace{\|Ad\|_{\Lambda}^2}_{\text{data curvature}} + \underbrace{\sum_{\{s,r\} \in \mathcal{P}} 2\tilde{b}_{s,r}(d_s - d_r)^2}_{\text{prior curvature}}$$

The surrogate Hessian H_Q is never built

$\theta_2 = d^t H_Q d$, where $H_Q d$ has a data part $A^t \Lambda A d$ and a prior part built from the $2\tilde{b}_{s,r}$ pairwise differences. The matrix H_Q is never formed — only the scalar above is computed.

One formula, two directions

The step formula $\alpha^* = -\theta_1/\theta_2$ does not depend on how the direction d is chosen. Two common choices give the same step:

Steepest descent: $d = -\nabla f(x)$

With $g = \nabla f(x)$, $\alpha^* = \frac{g^t g}{g^t H_Q g}$, where

$$g^t H_Q g = \|Ag\|_\Lambda^2 + \sum_{\{s,r\} \in \mathcal{P}} 2 \tilde{b}_{s,r} (g_s - g_r)^2 .$$

Coordinate descent (ICD): $d = \varepsilon_s$

The direction is one pixel: ε_s is one at pixel s and zero elsewhere. Then α^* is the single-pixel update.

Steepest descent on the surrogate

Each iteration recomputes the weights $\tilde{b}_{s,r}$, takes the gradient, and steps to the bottom of the surrogate. This is the Lab 4 algorithm.

Steepest Descent with Majorization of the Prior:

Initialize x

For K iterations {

$$g \leftarrow \nabla f(x) = -A^t \Lambda(y - Ax) + p, \quad \text{with } p_s = \sum_{r \in \partial s} 2\tilde{b}_{s,r}(x_s - x_r)$$

$$\alpha^* \leftarrow \frac{g^t g}{g^t H_Q g}$$

$$x \leftarrow x - \alpha^* g$$

}

The critical step — write $g^t H_Q g$ out in full

$$g^t H_Q g = \|Ag\|_{\Lambda}^2 + \sum_{\{s,r\} \in \mathcal{P}} 2\tilde{b}_{s,r}(g_s - g_r)^2.$$

Summary

- Reweighting turns a prior that is the same everywhere but non-Gaussian into a Gaussian prior whose weights vary across the image.
- The weight is small at an edge, so edges are preserved.
- Because the cost is Gaussian, the best step is one exact formula, $\alpha^* = -\theta_1/\theta_2$, with H_Q never formed.
- Steepest descent and ICD are this same step with different d .

The big idea

A spatially-varying Gaussian prior, minimized one exact step at a time — the engine of a practical MAP reconstruction.