

RECOVERY OF EDGES FROM SPECTRAL DATA WITH NOISE—A NEW PERSPECTIVE

SHLOMO ENGELBERG AND EITAN TADMOR

ABSTRACT. We consider the problem of detecting edges in piecewise smooth functions from their N -degree spectral content, which is assumed to be corrupted by noise. There are three scales involved: the “smoothness” scale of order $1/N$, the noise scale of order η and the $\mathcal{O}(1)$ scale of the jump discontinuities. We use concentration factors which are adjusted to the noise variance, $\eta \gg 1/N$, in order to detect the underlying $\mathcal{O}(1)$ -edges, which are separated from the noise scale, $\eta \ll 1$.

CONTENTS

1. Introduction and statement of main results	1
2. Detection of edges - concentration kernels	3
3. Noisy data – a new perspective	7
3.1. Constrained minimization	7
3.2. Balancing the different types of errors	9
4. Numerical results. I	10
5. Noisy data and smoothness — concentration kernels revisited	11
6. Numerical results. II	14
References	15

1. INTRODUCTION AND STATEMENT OF MAIN RESULTS

We consider the detection of edges in piecewise smooth data from its Fourier projection,

$$S_N f(x) = \sum_{|k| \leq N} \widehat{f}(k) e^{ikx}, \quad \widehat{f}(k) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\xi) e^{-i\xi x} d\xi.$$

Our approach is based on the technique of *concentration kernels* advocated in [4, 5] and the closely related techniques described in [3]. The technique

Date: February 1, 2008.

1991 Mathematics Subject Classification. Piecewise smoothness, edge detection, noisy data, concentration kernels, constrained optimization, separation of scales.

Key words and phrases. 42A10, 42A50, 65T10.

Acknowledgment. E.T. was supported in part by NSF grant 04-07704 and ONR grant N00014-91-J-1076. Part of this research was carried out while the second author was visiting the Weizmann Institute and he is grateful for their hospitality.

presented makes use of the fact that if f is discontinuous, its Fourier coefficients contain a slowly decaying part associated with the jumps of f . This part decays much more slowly than the rapidly decaying smooth part of f . For example, if f is smooth except for a single jump discontinuity at $x = z$ of size $[f](z) := f(z+) - f(z-)$, the jump discontinuity is associated with slowly decaying Fourier coefficients,

$$\widehat{f}(k) = [f](z) \frac{e^{-ikz}}{2\pi ik} + \widehat{s}(k), \quad |\widehat{s}(k)| \lesssim \frac{1}{|k|^2}, \quad [f](x) := f(x+) - f(x-),$$

where $\widehat{s}(k)$ are the Fourier coefficients of s – the smooth part of f ; the smoother s is, the faster is the decay of $\widehat{s}(k)$. Concentration kernels succeed in separating the two sets of coefficients. To this end one computes

$$\mathcal{K}_N^\sigma * (S_N f)(x) = \pi i \sum_{|k| \leq N} \operatorname{sgn}(k) \sigma_N \left(\frac{|k|}{N} \right) \widehat{f}(k) e^{ikx}$$

Here, $\sigma_N(\xi)$ can be drawn from a large family of properly normalized concentration factors that are at our disposal. The resulting function tends to zero in regions in which f is smooth and tends to the amplitude of the jumps at points where the function has a jump discontinuity,

$$\mathcal{K}_N^\sigma * (S_N f)(x) = [f](x) + \mathcal{O}\left(\frac{\log N}{N}\right).$$

Thus, edges are detected by *separation of scales*.

In this paper we utilize concentration kernels to detect edges from spectral information which is corrupted by *white noise*. In this context we observe that there are three scales involved – edges of order $\mathcal{O}(1)$, noise with variance $\mathcal{O}(\eta)$ and the smooth part of f which is resolved within order $\mathcal{O}(1/N)$ or smaller. Here, we can separate the noisy part, $n(x)$, from the smooth part,

$$s(x) \rightarrow s(x) + n(x), \quad E(|\widehat{n}(k)|^2) = \eta;$$

if $\eta \lesssim 1/N$ then the noisy part could be identified with (or below) the $\mathcal{O}(1/N)$ -variation of the smooth part of f . In this case, there are essentially two scales and edges can be detected using the usual framework of concentration kernels advocated in [4, 5]. Thus, our main focus in this paper is when the smoothness scale is dominated by the scale of the noise which is still well-separated from the $\mathcal{O}(1)$ -scale of the jumps,

$$\frac{1}{N} \lesssim \eta \ll 1.$$

The spectral information is now corrupted by white noise, affecting both low and high frequencies,

$$\widehat{f}(k) = [f](z) \frac{e^{-ikz}}{2\pi ik} + \widehat{s}(k) + \widehat{n}(k)$$

In order to separate edges from the noisy scale, the edge detector, $\mathcal{K}_N^\sigma f(x)$, must be properly adapted to the presence of white noise. We show how to

design edge detectors that optimally compensate for noise and for the effects of the smooth part of the signal.

The paper is organized as follows. In section 2 we discuss the general framework of edge detection based on concentration kernels. We revisit the results of [5, 7], providing a simpler proof for the concentration property for a large family of concentration factors. In particular, we trace the precise dependence of the error on the regularity of the associated concentration factor. This will prove useful when we deal with noisy data in section 3. Here, we introduce our new perspective, where concentration factors are derived by a constrained minimization while taking into account the two main ingredients of our data — jump discontinuities and the noisy parts of the data. Numerical results are demonstrated in sections 4. In section 5 we extend our construction of concentration factors to include the *three* ingredients of the data — taking into account the smooth part of the data in addition to the edges and noisy parts; numerical results are presented in 6.

2. DETECTION OF EDGES - CONCENTRATION KERNELS

Consider an f which is *piecewise smooth* in the sense that it is sufficiently smooth except for finitely many jump discontinuities, say at $x = z_1, z_2, \dots, z_J$, where

$$[f](z_j) := f(z_j+) - f(z_j-) \neq 0, \quad j = 1, 2, \dots, J.$$

Given the Fourier coefficients, $\{\widehat{f}(k)\}_{k=-N}^N$, we are interested in detecting the edges of the underlying piecewise smooth f , namely, to detect their location, z_1, \dots, z_J and their amplitudes, $[f](z_1), \dots, [f](z_J)$.

We utilize edge detection based on *concentration kernels*

$$(2.1a) \quad \mathcal{K}_N^\sigma(y) := - \sum_{k=1}^N \sigma_N \left(\frac{k}{N} \right) \sin ky, \quad \frac{\sigma_N(\xi)}{\xi} \in C^1[0, 1].$$

We shall need the kernel $-\mathcal{K}_N^\sigma$ to have (approximately) unit mass,

$$\int_0^\pi \mathcal{K}_N^\sigma(y) dy \approx -1.$$

To this end, we require that $\sigma(\xi) \equiv \sigma_N(\xi)$ be a properly normalized concentration factor; that $\sigma(\xi)$ satisfy

$$(2.1b) \quad \int_0^1 \frac{\sigma(\xi)}{\xi} d\xi = 1.$$

Indeed, the rectangular quadrature rule yields

$$\begin{aligned} \int_0^\pi \mathcal{K}_N^\sigma(y) dy &= \sum_{k=1}^N \sigma\left(\frac{k}{N}\right) \frac{(-1)^k - 1}{k} = \\ &= - \sum_{k \text{ odd} \geq 1}^N \frac{\sigma(\xi_k)}{\xi_k} \frac{2}{N} = - \int_0^1 \frac{\sigma(\xi)}{\xi} d\xi + \varepsilon_0(N), \quad \xi_k := \frac{k}{N}. \end{aligned}$$

with an error estimate, e.g., [6, 2],

$$(2.2) \quad \left| \int_0^\pi \mathcal{K}_N^\sigma(y) dy + 1 \right| \lesssim \varepsilon_0(N), \quad \varepsilon_0(N) := \frac{1}{N} \left\| \frac{\sigma_N(\xi)}{\xi} \right\|_{BV}$$

We set¹

$$K_N^\sigma f(x) := \mathcal{K}_N^\sigma * f(x) = \pi i \sum_{|k| \leq N} \operatorname{sgn}(k) \sigma_N\left(\frac{|k|}{N}\right) \widehat{f}(k) e^{ikx}.$$

Our purpose is to choose the *concentration factors*, $\sigma_N(|k|/N)$, such that $K_N^\sigma f(x) \approx [f](x)$. Thus, $K_N^\sigma f(x)$ will detect the edges, $[f](z_j)$, $j = 1, \dots, J$, by concentrating near these $\mathcal{O}(1)$ -edges which are to be separated from a much smaller scale of order $K_N^\sigma f(x) \approx 0$ in regions of smoothness. In the following theorem we present a rather general framework for edge detectors based on concentration factors. In particular, we track the precise dependence of the scale separation on the behavior of σ .

Theorem 2.1 (Concentration kernels). *Assume that $f(\cdot)$ is piecewise smooth such that the first-variation of f is of locally bounded variation,*

$$(2.3) \quad \omega_f(y) \equiv \omega_f(y; x) := \frac{f(x+y) - f(x-y) - [f](x)}{\sin(y/2)} \in BV[-\pi, \pi].$$

Let $\mathcal{K}_N^\sigma(x)$ be an admissible concentration kernel (2.1), such that

$$(2.4a) \quad \varepsilon_0(N) := \frac{1}{N} \left\| \frac{\sigma_N(\xi)}{\xi} \right\|_{BV} \ll 1,$$

$$(2.4b) \quad \varepsilon_1(N) := \frac{1}{N} |\sigma_N(1)| \ll 1,$$

$$(2.4c) \quad \varepsilon_2(N) := \frac{1}{N} \int_{\xi \sim \frac{1}{N}}^1 \frac{|\sigma'_N(\xi)|}{\xi} d\xi \ll 1$$

$$(2.4d) \quad \varepsilon_3(N) := \left| \sigma_N\left(\frac{1}{N}\right) \right| \ll 1.$$

Set

$$(2.5) \quad \varepsilon_N := \varepsilon_0(N) + \varepsilon_1(N) + \varepsilon_2(N) + \varepsilon_3(N).$$

¹We observe that $K_N^\sigma f$ is the *operator* associated with but otherwise different from the concentration kernel $\mathcal{K}_N^\sigma(x)$.

Then, the conjugate sum $K_N^\sigma f(x)$,

$$(2.6) \quad K_N^\sigma f(x) = \pi i \sum_{|k| \leq N} \operatorname{sgn}(k) \sigma_N \left(\frac{|k|}{N} \right) \widehat{f}(k) e^{ikx},$$

satisfies the concentration property,

$$(2.7) \quad K_N^\sigma f(x) \sim \begin{cases} [f](z_j) + \mathcal{O}(\varepsilon_N), & x \sim z_j, \quad j = 1, \dots, J \\ \mathcal{O}(\varepsilon_N), & \operatorname{dist}\{x, \{z_1, \dots, z_J\}\} \gg \varepsilon_N. \end{cases}$$

Proof. We simplify the proof in [5, 7]. The key to the proof is to observe that \mathcal{K}_N^σ is an appropriately normalized *derivative* of the delta function; in particular, since $\mathcal{K}_N^\sigma(\cdot)$ is odd

$$\begin{aligned} \mathcal{K}_N^\sigma * f(x) &= - \int_0^\pi K_N^\sigma(y) (f(x+y) - f(x-y)) dy \\ &= - \int_0^\pi \mathcal{K}_N^\sigma(y) (f(x+y) - f(x-y) - [f](x)) dy - [f](x) \times \int_0^\pi \mathcal{K}_N^\sigma(y) dy. \end{aligned}$$

Since σ is assumed normalized, the error estimate (2.2) tells us²

$$\int_0^\pi \mathcal{K}_N^\sigma(y) dy = -1 + \varepsilon_0(N), \quad \varepsilon_0(N) = \frac{1}{N} \left\| \frac{\sigma_N(\xi)}{\xi} \right\|_{BV},$$

and we end up with the error estimate

$$(2.8) \quad |K_N^\sigma f(x) - [f](x)| \lesssim \left| \int_0^\pi \sin(y/2) \mathcal{K}_N^\sigma(y) \omega_f(y; x) dy \right| + \mathcal{O}(\varepsilon_0(N)).$$

To upperbound the expression on the right, we sum by parts, deriving the identity

$$\begin{aligned} 2 \sin(y/2) \mathcal{K}_N^\sigma(y) &\equiv \overbrace{\sigma_N(1) \cos \left(N + \frac{1}{2} \right) y}^{I_1(y)} \\ &+ \sum_{k=1}^{N-1} \overbrace{\left(\sigma_N(\xi_k) - \sigma_N(\xi_{k+1}) \right) \cos \left(k + \frac{1}{2} \right) y}^{I_{2k}(y)} - \overbrace{\sigma_N(\xi_1) \cos \left(\frac{y}{2} \right)}^{I_3(y)}. \end{aligned}$$

²In our case, the errors for the trapezoidal and rectangular rules provided respectively in [2] and [6] coincide modulo the $\mathcal{O}(\varepsilon_1(N) + \varepsilon_3(N))$ terms.

The usual cancelation estimate, $|\int_0^\pi \cos((k + \frac{1}{2})y)w(y)dy| \lesssim \|w\|_{BV}/|k + \frac{1}{2}|$, implies

$$\begin{aligned} \left| \int_0^\pi I_1(y)\omega_f(y)dy \right| &\lesssim |\sigma_N(1)| \cdot \frac{1}{N} \|\omega_f(\cdot; x)\|_{BV} \leq \varepsilon_1(N) \|\omega_f(\cdot; x)\|_{BV}, \\ \left| \int_0^\pi I_{2k}(y)\omega_f(y)dy \right| &\lesssim |\sigma(\xi_{k+1}) - \sigma(\xi_k)| \cdot \frac{1}{k + \frac{1}{2}} \|\omega_f(\cdot; x)\|_{BV}, \\ \left| \int_0^\pi I_3(y)\omega_f(y)dy \right| &\lesssim \varepsilon_3(N) \cdot \|\omega_f(\cdot; x)\|_{L^\infty}. \end{aligned}$$

We conclude

$$\begin{aligned} &\left| \int_0^\pi yR_N(y)\omega_f(y; x)dy \right| \\ (2.9) \quad &\lesssim \left(\varepsilon_1(N) + \sum_{k=1}^{N-1} \frac{|\sigma(\xi_{k+1}) - \sigma(\xi_k)|}{\xi_{k+\frac{1}{2}}} \frac{1}{N} + \varepsilon_3(N) \right) \|\omega_f(\cdot)\|_{BV} \\ &\lesssim \left(\varepsilon_1(N) + \varepsilon_2(N) + \varepsilon_3(N) \right) \|\omega_f(\cdot)\|_{BV}. \end{aligned}$$

The result (2.7) follows from (2.8) and (2.9). \square

As an example, we consider the noise-free case with a concentration factor for which

$$(2.10) \quad \frac{\sigma_N(\xi)}{\xi} \in C^1.$$

Clearly $\varepsilon_0(N) \lesssim 1/N$, and since σ is bounded, $\varepsilon_1(N) \lesssim 1/N$. Moreover, since $\sigma'_N(\xi)$ is bounded,

$$\varepsilon_2(N) \lesssim \frac{1}{N} \int_{\xi \sim \frac{1}{N}}^1 \frac{1}{\xi} d\xi \lesssim \frac{\log(N)}{N}.$$

Finally, since $|\sigma_N(\xi)| \lesssim \xi$ then $\varepsilon_3(N) \lesssim 1/N$. Theorem 2.1 yields the following result of [5, 7] for edge detection in piecewise-smooth, noiseless data.

Corollary 2.2 (Concentration kernels). *Assume that $f(\cdot)$ is piecewise smooth such that (2.3) holds. Let $K_N^\sigma(x)$ be a normalized concentration kernel (2.1) with $\frac{\sigma(\xi)}{\xi} \in C^1$. Then $K_N^\sigma f(x)$ satisfies the concentration property with*

$\varepsilon_N = \log(N)/N$,

(2.11)

$$K_N^\sigma f(x) \sim \begin{cases} [f](z_j) + \mathcal{O}\left(\frac{\log(N)}{N}\right), & x \sim z_j, \quad j = 1, \dots, J \\ \mathcal{O}\left(\frac{\log(N)}{N}\right), & \text{dist}\{x, \{z_1, \dots, z_J\}\} \gg \frac{1}{N}. \end{cases}$$

3. NOISY DATA – A NEW PERSPECTIVE

3.1. Constrained minimization. Assume that f experiences a single jump discontinuity at location z of height $[f](z)$. This dictates a first order decay of the Fourier coefficients,

$$(3.1) \quad \widehat{f}(k) = [f](z) \frac{e^{-ikz}}{2\pi ik} + \widehat{s}(k) + \widehat{n}(k).$$

Here, $\widehat{s}(k)$ are associated with the regular part of f after extracting the jump $[f](z)$; their decay is of order $\sim |k|^{-2}$ or faster, depending on the smoothness of the regular part $s(\cdot)$. The new aspect of the problem enters through the $\widehat{n}(k)$'s, which are the Fourier coefficients of the noisy part corrupting the smooth part of the data. We assume $n(\cdot)$ to be white noise whose mean-square power at each frequency is $E(|\widehat{n}(k)|^2) = \eta$. With (3.1), the conjugate sum (2.6) becomes

$$\begin{aligned} K_N^\sigma f(x) &= [f](z) \times 2\pi i \sum_{k=1}^N \frac{\sigma_N\left(\frac{k}{N}\right)}{k} \cos k(x-z) \\ &\quad - 2\pi \sum_{k=1}^N \sigma\left(\frac{k}{N}\right) \widehat{s}(k) \sin kx - 2\pi \sum_{k=1}^N \sigma\left(\frac{k}{N}\right) \widehat{n}(k) \sin kx. \end{aligned}$$

We quantify the “energy” of each of the three sums on the right. E_J and E_R are associated with the discontinuous and regular parts of f ,

$$(3.2a) \quad \sum_{k=1}^N \left(\frac{\sigma\left(\frac{k}{N}\right)}{k}\right)^2 \approx \frac{1}{N} \int_0^1 \left(\frac{\sigma(\xi)}{\xi}\right)^2 d\xi =: E_J(\sigma),$$

$$(3.2b) \quad \sum_{k=1}^N \sigma^2\left(\frac{k}{N}\right) |\widehat{s}(k)|^2 \sim \frac{1}{N^3} \int_0^1 \frac{\sigma^2(\xi)}{\xi^4} d\xi =: E_R(\sigma),$$

and $E_\eta(\sigma)$ associated with the noisy part of f which was assumed to have variance η

$$(3.2c) \quad \sum_{k=1}^N \sigma^2\left(\frac{k}{N}\right) E(|\widehat{n}(k)|^2) \approx \eta N \int_0^1 \sigma^2(\xi) d\xi =: E_\eta(\sigma).$$

Our perspective for construction of edge detectors for such noisy data is to treat the problem as a *constrained minimization*. We seek a function, $\sigma(\xi)$, which minimizes the total energy, thus making the conjugate sum $K_N^\sigma f$ as localized as possible, subject to prescribed normalization constraint (2.1b),

$$(3.3) \quad \min \left\{ a_J E_J(\sigma) + a_\eta E_\eta(\sigma) + a_R E_R(\sigma) \mid \int_0^1 \frac{\sigma(\xi)}{\xi} d\xi = 1 \right\}.$$

This yields

$$\sigma(\xi) = C_\sigma \frac{\frac{1}{\xi}}{a_J \frac{1}{N\xi^2} + a_\eta \eta N + a_R \frac{1}{N^3\xi^4} +} = \frac{C_\sigma N^3 \xi^3}{a_J N^2 \xi^2 + a_\eta \eta N^4 \xi^4 + a_R}.$$

We ignore the relatively negligible contribution of the regular part which becomes even smaller as $s(\cdot)$ becomes smoother. Setting $a_R = 0$ we end up with concentration factors of the form

$$(3.4) \quad \sigma(\xi) = \frac{C_\sigma}{a_J} \cdot \frac{N\xi}{1 + \eta\beta^2 N^2 \xi^2}, \quad \beta := \sqrt{\frac{a_\eta}{a_J}}.$$

The corresponding concentration kernel depends only the *relative* size of the amplitudes $\beta^2 = a_\eta/a_J$: indeed, the normalization of $\sigma(\xi)/\xi$ (2.1b) causes the constant C_σ to satisfy

$$\int_0^1 \frac{\sigma(\xi)}{\xi} d\xi = \frac{C_\sigma}{a_J \sqrt{\eta}\beta} \tan^{-1}(\sqrt{\eta}\beta N) = 1,$$

and we end up with the *normalized* concentration factor

$$(3.5) \quad \sigma(\xi) \equiv \sigma_\eta(\xi) = \frac{1}{\tan^{-1}(\sqrt{\eta}\beta N)} \cdot \frac{\sqrt{\eta}\beta N \xi}{1 + \eta\beta^2 N^2 \xi^2}, \quad \beta = \sqrt{\frac{a_\eta}{a_J}}.$$

The corresponding edge detector then takes the form

$$(3.6) \quad K_N^{\sigma_\eta} f(x) = \frac{\pi \sqrt{\eta}\beta}{\tan^{-1}(\sqrt{\eta}\beta N)} \sum_{|k| \leq N} \frac{ik}{1 + \eta\beta^2 k^2} \widehat{f}(k) e^{ikx}.$$

The concentration factor $\sigma = \sigma_\eta$ now involves three factors: the ratio β , the noise variance η and the number of modes N . The concentration kernel (3.6) tends to de-emphasize both the low frequencies which are “corrupted” by the jump discontinuity(-ies) and the high frequencies which are corrupted by the noise. Different procedures yield different policies for the choice of $\beta = \beta(\eta)$; one will be discussed in the next subsection. It is worth noting the essential dependence of $\sigma_\eta(\xi)$ on the variance of the noise η . There are three scales involved — the small “smoothness” of order $\sim 1/N$, the noise scale of order $\sim \eta$ and the $\mathcal{O}(1)$ -scale of jump discontinuities. We distinguish between two cases. If $\eta \ll 1/N$ so that $\sqrt{\eta}\beta N \ll 1$, then the noise can be considered part of the smooth variation of f and $\sigma_\eta(\xi) \approx \xi$ recovers the usual concentration factor for noise-free data. Indeed, $\sigma_\eta(\xi) = \xi$ at the limit of $\eta \downarrow 0$. Otherwise, when the $\mathcal{O}(1/N)$ -smoothness scale is dominated by the $\mathcal{O}(\eta)$ -noise scale in the sense that $\sqrt{\eta}\beta \gtrsim 1/N$, in which we assume the noise to be still well-below the $\mathcal{O}(1)$ -scale of the jumps,

$$\frac{1}{N} \lesssim \sqrt{\eta}\beta \ll \mathcal{O}(1).$$

In this case, we can ignore the bounded factor $1/\tan^{-1}(\sqrt{\eta}\beta N)$, and we compute the small scale dictated by theorem 2.1. Setting $\zeta(\xi) := \sqrt{\eta}\beta N \xi$

we find,

$$\begin{aligned}\varepsilon_0(N) &= \frac{1}{N} \left\| \frac{\sigma_\eta(\xi)}{\xi} \right\|_{BV} \lesssim \frac{\sqrt{\eta}\beta N}{N} \int \frac{1}{(1+\zeta^2)^2} d\zeta \lesssim \sqrt{\eta}\beta, \\ \varepsilon_1(N) &= \frac{1}{N} |\sigma_\eta(1)| \lesssim \frac{1}{\sqrt{\eta}\beta N^2}, \\ \varepsilon_2(N) &= \frac{1}{N} \int_{1/N}^1 \frac{1-\zeta^2}{\xi(1+\zeta^2(\xi))^2} d\xi \lesssim \sqrt{\eta}\beta \log(\sqrt{\eta}\beta) \\ \varepsilon_3(N) &= \left| \sigma_\eta\left(\frac{1}{N}\right) \right| \lesssim \sqrt{\eta}\beta.\end{aligned}$$

and hence (2.4) holds with

$$\varepsilon_N \equiv \varepsilon_\eta := \sqrt{\eta}\beta \log(\sqrt{\eta}\beta).$$

It is remarkable to see how the small scale of smoothness in the noiseless case, $\mathcal{O}(\log(N)/N)$, is now replaced by the small scale of noise $\varepsilon_\eta = \mathcal{O}(\sqrt{\eta}\beta \log(\sqrt{\eta}\beta))$. We now appeal to (2.7): since $\eta \ll 1$, theorem 2.1 implies that $K_N^{\sigma_\eta} f$ separates the $\mathcal{O}(1)$ -scale of the edges from the noise scale of order $\varepsilon_\eta \ll 1$.

Theorem 3.1 (Edge detection in noisy data). *Assume that $f(\cdot)$ is piecewise smooth in the sense that (2.3) holds. assume that its spectral data contains white noise with variance $\eta \ll 1$. Let $K_N^{\sigma_\eta} f(x)$ be a normalized concentration kernel (3.6)*

$$K_N^{\sigma_\eta} f(x) = \frac{\pi\sqrt{\eta}\beta}{\tan^{-1}(\sqrt{\eta}\beta N)} \sum_{|k| \leq N} \frac{ik}{1 + \eta\beta^2 k^2} \hat{f}(k) e^{ikx}$$

associated with the concentration factor

$$\sigma_\eta(\xi) = \frac{1}{\tan^{-1}(\sqrt{\eta}\beta N)} \cdot \frac{\sqrt{\eta}\beta N \xi}{1 + \eta\beta^2 N^2 \xi^2}.$$

We distinguish between two cases:

- (i) if $\sqrt{\eta}\beta N \ll 1$ we set the small scale $\varepsilon = \varepsilon_N := \log(N)/N$;
 - (ii) if $1/N \lesssim \sqrt{\eta}\beta \ll 1$ we set the small scale $\varepsilon = \varepsilon_\eta := \sqrt{\eta}\beta \log(\sqrt{\eta}\beta)$.
- Then, $K_N^{\sigma_\eta} f(x)$ satisfies the following concentration property,

$$(3.7) \quad K_N^{\sigma_\eta} f(x) \sim \begin{cases} [f](z_j) + \mathcal{O}(\varepsilon), & x \sim z_j, \quad j = 1, \dots, J \\ \mathcal{O}(\varepsilon), & \text{dist}\{x, \{z_1, \dots, z_J\}\} \gg \varepsilon. \end{cases}$$

3.2. Balancing the different types of errors. In order to choose the free parameter β , it is important to know how β influences the error at the

output of our edge detector. Let us consider E_η . We have seen that:

$$\begin{aligned} E_\eta &= \eta N \int_0^1 \sigma^2(\xi) d\xi \\ &= \frac{\eta N}{(\tan^{-1}(\sqrt{\eta}\beta N))^2} \int_0^1 \frac{\eta\beta^2 N^2 \xi^2}{(1 + \eta\beta^2 N^2 \xi^2)^2} d\xi \\ &= \frac{\sqrt{\eta}}{\beta(\tan^{-1}(\sqrt{\eta}\beta N))^2} \int_0^{\sqrt{\eta}\beta N} \frac{\zeta^2}{(1 + \zeta^2)^2} d\zeta \stackrel{\sqrt{\eta}\beta N \gtrsim 1}{\sim} \frac{\sqrt{\eta}}{\beta}. \end{aligned}$$

As E_η is approximately the expected value of the square of the contribution of the noise to the output of the edge detector—it is approximately the variance of the contribution of the noise, if we want to consider the *size* of the noise, we should consider something related to the standard deviation of the contribution. It is customary to bound the noise by some number of standard deviations—we will use two standard deviations. We define the *effective size* of the noise's contribution to be

$$E_{\eta,eff} \equiv 2\sqrt{E_\eta} \sim \frac{\eta^{1/4}}{\sqrt{\beta}}.$$

From the results of Theorem 3.1 we find that far from jump the contribution from the jump is of order $\sqrt{\eta}\beta$. The *effective size* of this term is therefore

$$E_{J,eff} \sim \eta^{1/2}\beta.$$

Let us minimize $E_{\eta,eff} + E_{J,eff}$ with respect to β . We find that:

$$\beta \sim \eta^{-1/6}.$$

4. NUMERICAL RESULTS. I

To illustrate the results of the previous section, we present two sets of numerical results. We begin with the noiseless case, in Figure 4.1, where we set $\beta = 1$ corresponding to equal weights for the errors due to the noise and the discontinuous parts of the signal. We plot the output of the concentration factor when a periodic function with a single jump continuity is used as the input. A simple examination of the results shows that the output is what we predicted. The output is one at the (unit) jump and is zero away from it. As η gets smaller the value away from the jump tends to zero. Considering the figure, we find that the ratio of the size in the continuous region is $0.00365/0.00137 = 2.66 \approx \sqrt{10}$ which is the ratio of the square root of the η 's—as it should be.

Next, Figure 4.2 demonstrates the edge detected in noisy data using the concentration kernel (3.6) with the advocated $\beta = \pi\eta^{-1/6}$.

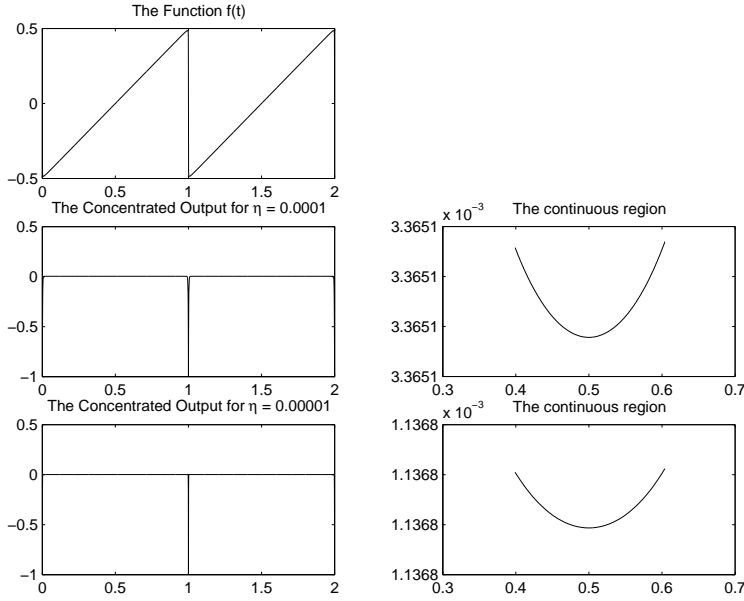


Figure 4.1: The output for various values of η when there is actually no noise at the input.

5. NOISY DATA AND SMOOTHNESS — CONCENTRATION KERNELS REVISITED

As an alternative approach to the L^2 -minimization offered in Section 3, we now replace the L^2 -“averaged” effect of the regular part taken in (3.2b), by the BV-like quantity

$$(5.1a) \quad E_R(\sigma) \approx \sum_{k=1}^N \left| \sigma\left(\frac{k}{N}\right) \right| \cdot |\hat{s}(k)|,$$

where the regular part is sufficiently smooth that

$$(5.1b) \quad |\hat{s}(k)| \sim \frac{1}{k^2}.$$

As in (3.3), we consider the constrained minimization

$$(5.2a) \quad \min \left\{ J(\sigma) \mid \int_0^1 \frac{\sigma(\xi)}{\xi} d\xi = 1 \right\},$$

with $J(\sigma) = a_J E_J(\sigma) + a_\eta E_\eta(\sigma) + a_R E_R(\sigma)$, where $E_J(\sigma)$ and $E_\eta(\sigma)$ given by (3.2a) and (3.2c) but with an alternative expression for the “energy” of the regular part motivated by (5.1): $E_R(\sigma) := \int_0^1 |\sigma(\xi)| \xi^{-2} d\xi / N$. We find that:

$$(5.2b) \quad J(\sigma) := \frac{a_J}{N} \int_0^1 \frac{\sigma^2(\xi)}{\xi^2} d\xi + a_\eta \eta N \int_0^1 \sigma^2(\xi) d\xi + \frac{a_R}{N} \int_0^1 \frac{|\sigma(\xi)|}{\xi^2} d\xi.$$

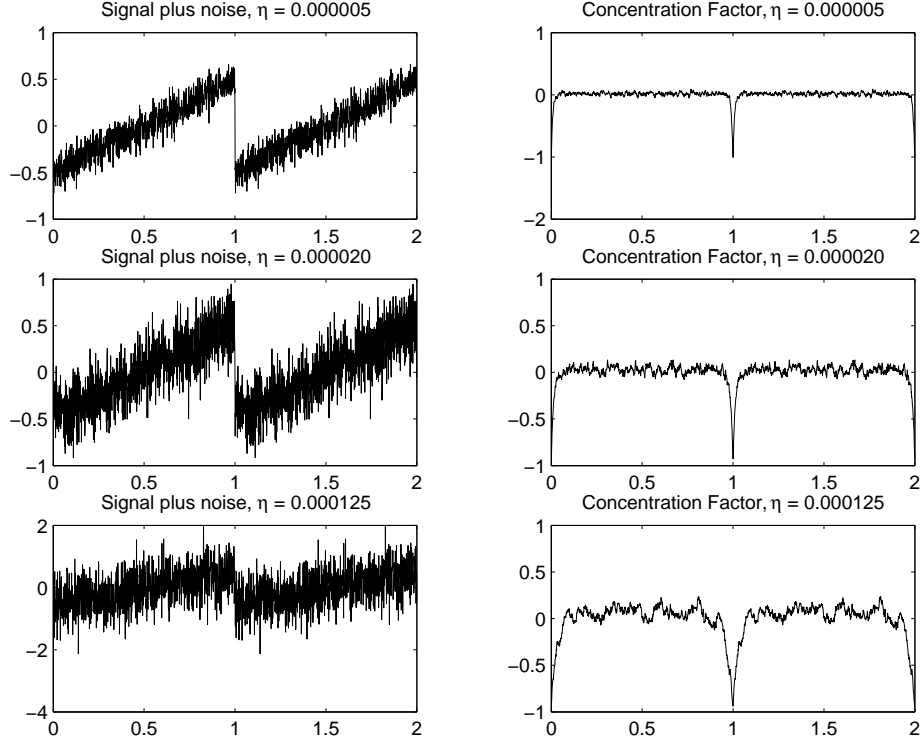


Figure 4.2: Detection of edges in noisy saw-tooth function corrupted with various values of η , using the concentration kernel (3.6) with $\beta = \pi\eta^{-1/6}$.

Proceedings formally, the solution for the first variation of (5.2) leads to

$$\frac{2a_J\sigma(\xi)}{N\xi^2} + 2\eta a_\eta N\sigma(\xi) + \frac{a_R \cdot \text{sgn}(\sigma(\xi))}{N\xi^2} = \frac{\lambda}{\xi}, \quad \text{sgn}(\sigma) = \begin{cases} 1, & \sigma > 0 \\ 0, & \sigma = 0 \\ -1, & \sigma < 0 \end{cases}.$$

We will show that the resulting optimal concentration factor is given by

$$(5.3) \quad \sigma(\xi) = C_\sigma \frac{(N\xi - k_0)_+}{1 + \eta\beta^2 N^2 \xi^2}, \quad k_0 := \frac{a_R}{\lambda}, \quad \beta = \sqrt{\frac{a_\eta}{a_J}}, \quad C_\sigma = \frac{\lambda}{2a_J}.$$

Indeed, to justify the passage to (5.3), one may consider a *regularized* version of the variational statement (5.2), $\min J_\epsilon(\sigma)$, where

$$J_\epsilon(\sigma) := \frac{a_J}{N} \int_0^1 \frac{\sigma^2(\xi)}{\xi^2} d\xi + a_\eta \eta N \int_0^1 \sigma^2(\xi) d\xi + \frac{a_R}{N} \int_0^1 \frac{|\sigma(\xi)|_\epsilon}{\xi^2} d\xi,$$

involves a *mollified* absolute value function:

$$|\sigma|_\epsilon := \begin{cases} |\sigma|, & |\sigma| \geq \epsilon, \\ \frac{\sigma^2}{2\epsilon} + \frac{\epsilon}{2}, & |\sigma| \leq \epsilon. \end{cases}$$

The solution of the corresponding regularized first variation yields the minimizer:

$$\sigma(\xi) = C_\sigma \frac{N\xi - k_0 \cdot \text{sgn}_\varepsilon(\sigma(\xi))}{\eta\beta^2 N^2 \xi^2 + 1}, \quad \text{sgn}_\varepsilon(\sigma) := \begin{cases} 1, & \sigma > \varepsilon, \\ \sigma/\varepsilon, & |\sigma| \leq \varepsilon, \\ -1, & \sigma < -\varepsilon. \end{cases}$$

Thus, we end up with the optimal concentration factor, $\sigma(\xi) = \sigma_\varepsilon(\xi)$,

$$\sigma_\varepsilon(\xi) = \begin{cases} C_\sigma \frac{N\xi - k_0}{1 + \eta\beta^2 N^2 \xi^2}, & N\xi - k_0 > \varepsilon, \\ C_\sigma \frac{\varepsilon N\xi}{\varepsilon(1 + \eta\beta^2 N^2 \xi^2) + C_\sigma k_0}, & N\xi - k_0 \leq \varepsilon, \end{cases}$$

and (5.3) is recovered by letting $\varepsilon \downarrow 0$. Clearly, the resulting optimal concentration factor is non-negative.

It remains to calculate the normalization factor, C_σ , for which

$$C_\sigma \int_{k_0/N}^1 \frac{1}{\xi} \frac{N\xi - k_0}{\eta\beta^2 N^2 \xi^2 + 1} d\xi = 1.$$

The integral on the left is found to be

$$\begin{aligned} \int_{k_0/N}^1 \left(\frac{-k_0}{\xi} + \frac{k_0 \eta \beta^2 N^2 \xi + N}{\eta \beta^2 N^2 \xi^2 + 1} \right) d\xi = \\ k_0 \ln \left(\frac{k_0}{N} \right) + \frac{k_0}{2} \ln \left(\frac{\eta \beta^2 N^2 + 1}{\eta \beta^2 k_0^2 + 1} \right) + \frac{1}{\sqrt{\eta} \beta} \left(\tan^{-1}(\sqrt{\eta} \beta N) - \tan^{-1}(\sqrt{\eta} \beta k_0) \right). \end{aligned}$$

We focus our attention on the “noisy” case when $1/N \lesssim \sqrt{\eta} \beta \ll 1$ so that the fourth term on the right is negligible while the second term on the right is approximated by

$$\frac{k_0}{2} \ln \left(\frac{\eta \beta^2 N^2 + 1}{\eta \beta^2 k_0^2 + 1} \right) \approx k_0 \ln(\sqrt{\eta} \beta N).$$

We end up with an approximated integral

$$\int_{k_0/N}^1 (\dots) d\xi \approx k_0 \ln(\sqrt{\eta} \beta N) + \frac{1}{\sqrt{\eta} \beta} \tan^{-1}(\sqrt{\eta} \beta N).$$

The balance between these two terms depends on the specific policy for β and the detailed balance between $\sqrt{\eta} \beta$ and N . Our normalized concentration factor takes the form

$$(5.4a) \quad \sigma_\eta(\xi) = \frac{1}{\sqrt{\eta} \beta k_0 \ln(\sqrt{\eta} \beta N) + \tan^{-1}(\sqrt{\eta} \beta N)} \cdot \frac{\sqrt{\eta} \beta (N\xi - k_0)_+}{1 + \eta \beta^2 N^2 \xi^2}$$

We can simplify this concentration factor in several ways; we mention two here.

(i) When N is large enough, we have $\tan^{-1}(\sqrt{\eta}\beta N) \approx \pi/2$ yielding

$$(5.4b) \quad \sigma_\eta(\xi) = \frac{1}{k_0 \ln(\sqrt{\eta}\beta N) + \pi/(2\sqrt{\eta}\beta)} \cdot \frac{(N\xi - k_0)_+}{1 + \eta\beta^2 N^2 \xi^2}$$

(ii) Observe that $\sigma_\eta(\xi)$ is rapidly decreasing at $\xi \approx 1$ with

$$\sigma_\eta(\xi) \sim \frac{1}{\eta\beta^2 N \log N}, \quad \xi \approx 1$$

so $\sigma_\eta(\xi)$ can be set to zero for $\xi \approx 1$ when N is large enough. In order to properly normalize the resulting concentration factor N must be replaced by N_0 . This leads us to:

$$(5.4c) \quad \sigma_\eta\left(\frac{k}{N}\right) = \begin{cases} 0, & k < k_0, \\ \frac{1}{k_0 \ln(\sqrt{\eta}\beta N_0) + \pi/(2\sqrt{\eta}\beta)} \cdot \frac{(k - k_0)_+}{1 + \eta\beta^2 k^2}, & k_0 < k < N_0, \\ 0, & N_0 < k < N. \end{cases}$$

6. NUMERICAL RESULTS. II

We consider two examples depicted in Figure 6.1. In the first case, we have a noise of variance $\eta = 2 \times 10^{-5}$ to be detected out of the first $N \gg 1000$ modes. With $\beta = \pi\eta^{-1/6} \sim 15$ and by tuning $k_0 = 8\pi$ and $N_0 = 1000$ we find

$$\sigma_\eta\left(\frac{k}{N}\right) = \begin{cases} 0, & k = 1, \dots, 24, 1001, 1002, \dots, \\ 0.3985 \cdot \frac{(k - 8\pi)_+}{1 + \eta(15k)^2}, & 25 \leq k \leq 1000. \end{cases}$$

In the second case, of noise variance $\eta = 4.5 \times 10^{-5}$ which led us to the choice of $\beta = \pi\eta^{-1/6} \sim 13$; setting $k_0 = 6\pi$ and $N_0 = 1000$ we have

$$\sigma_\eta\left(\frac{k}{N}\right) = \begin{cases} 0, & k = 1, \dots, 19, 1001, 1002, \dots, \\ 0.5070 \cdot \frac{(k - 6\pi)_+}{1 + \eta(13k)^2}, & 20 \leq k \leq 1000. \end{cases}$$

(Note that in calculating the constants we made use of the exact normalization factor C_σ . For our values of η and β the value $N_0 = 1000$ is not large enough to make the approximate value given in (5.4c) useful.)

Note that even with a large amount of white noise and of smooth signal, the location of the jump discontinuity is still clear. When considering jumps “corrupted” by low frequency data, we avoid low frequency signals by not using low frequency data. This helps keep the smooth signal from corrupting our results. On the other hand, because the jump discontinuity has most of its energy at low frequencies as well, our technique will increase the noise’s effect. Comparing Figures 4.2 and 6.1, we find that the latter is not as clean as the former in the sub-figure where the strength of the noise is the same.

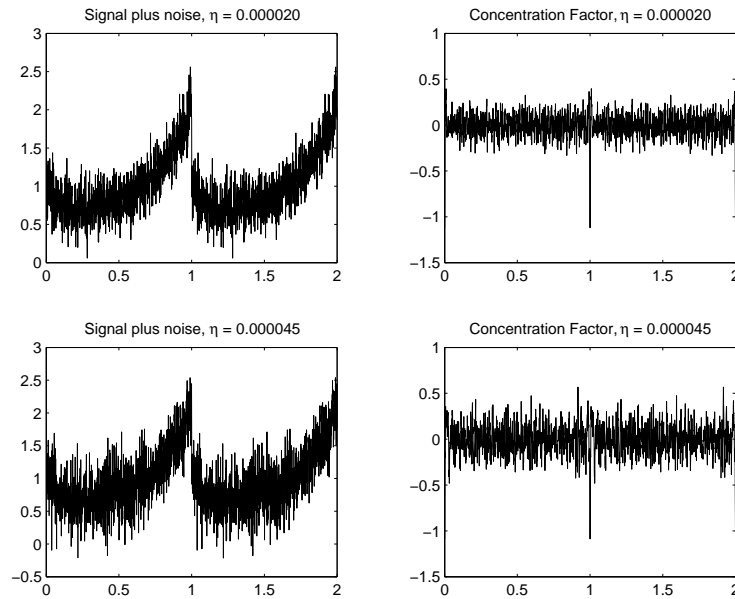


Figure 6.1: The output for various values of η when the input consists of a piecewise smooth function and noise (of the specified power spectral density).

REFERENCES

- [1] R. Archibald and A. Gelb, "Reducing the Effects of Noise in Image Reconstruction," *J. of Sci. Comp.*, 17 (2002), 167-180.
- [2] D. Cruz-Uribe and C. J. Neugebauer, "Sharp error bounds for the trapezoidal rule and Simpson's rule," *J. Ineq. Pure Appl. Math.* 3(4) (2002), article 49.
- [3] S. Engelberg, "Edge Detection Using Fourier Coefficients," *Amm. Math. Monthly*, to appear.
- [4] A. Gelb and E. Tadmor, "Detection of Edges in Spectral Data," *Appl. Comput. Harmonic Anal.* 7 (1999), 101-135.
- [5] A. Gelb and E. Tadmor, "Detection of Edges in Spectral Data II. Nonlinear Enhancement," *SIAM J. Numer. Anal.* 38 (2000), 1389-1408.
- [6] G. Polya and G. Szego, "Problems and Theorems in Analysis", Vol. I, Springer Verlag, New York, 1972,
- [7] E. Tadmor, "Filters, mollifiers and the computation of the Gibbs phenomenon", to appear.

(Shlomo Engelberg)

DEPARTMENT OF ELECTRONICS
JERUSALEM COLLEGE OF TECHNOLOGY
P.O.B. 16031, JERUSALEM, ISRAEL

E-mail address: shlomoe@jct.ac.il

URL: <http://www.cc.jct.ac.il/~shlomoe/>

(Eitan Tadmor)

DEPARTMENT OF MATHEMATICS, INSTITUTE FOR PHYSICAL SCIENCE AND TECHNOLOGY
AND CENTER OF SCIENTIFIC COMPUTATION AND MATHEMATICAL MODELING (CSCAMM)

UNIVERSITY OF MARYLAND
COLLEGE PARK, MD 20742 USA

E-mail address: tadmor@cscamm.umd.edu

URL: <http://www.cscamm.umd.edu/~tadmor>