IMA Journal of Information and Inference (2019) Page 1 of 26
doi:10.1093/imamat/xxx000

Phase Harmonic Correlations and Convolutional Neural Networks

STÉPHANE MALLAT
Collège de France, Paris
DI, École Normale Supérieure, PSL Researh University, Paris
CCM, Flatiron Institute, New York

SIXIN ZHANG
DI, École Normale Supérieure, PSL Researh University, Paris
Center for Data Science, Peking University, Beijing, China

AND

GASPAR ROCHETTE
DI, École Normale Supérieure, PSL Researh University, Paris

[Received on 2 July 2019]

A major issue in harmonic analysis is to capture the phase dependence of frequency representations, which carries important signal properties. It seems that convolutional neural networks have found a way. Over time-series and images, convolutional networks often learn a first layer of filters which are well localized in the frequency domain, with different phases. We show that a rectifier then acts as a filter on the phase of the resulting coefficients. It computes signal descriptors which are local in space, frequency and phase. The non-linear phase filter becomes a multiplicative operator over phase harmonics computed with a Fourier transform along the phase. We prove that it defines a bi-Lipschitz and invertible representation. The correlations of phase harmonics coefficients characterise coherent structures from their phase dependence across frequencies. For wavelet filters, we show numerically that signals having sparse wavelet coefficients can be recovered from few phase harmonic correlations, which provide a compressive representation.

Keywords: Neural networks, harmonics, phase, wavelets

2000 Math Subject Classification: 42C40, 62M45

1. Introduction

Convolutional neural networks capture highly complex properties of one, two and three-dimensional signals, leading to remarkable applications for classification, regression and generation [LeCun et al., 2015]. For image and time-series applications, the learned filters in the first layer are usually well localized both in space or time, and along frequencies, with different phases [Krizhevsky et al., 2012, Luo & Mesgarani, 2018]. They resemble wavelets [Mallat, 2016]. These filters provide a representation of an input signal $x(u)$ which is local along u and along frequencies. However, rectifiers applied to first layer convolutional network coefficients also perform a filtering along phases, which creates multiple harmonics. We show that rectified coefficients define a local space-frequency-phase representation where local signal properties can be characterised by linear correlations across phases and frequencies.

The mathematical properties of rectifiers are more easily analyzed by extending first layer real network coefficients into complex values. This is done with a complex analytic extension of the network filters. A central result shows that rectifiers act as a filter along the analytic phase. By applying a Fourier transform along the phase, the phase filtering becomes a multiplicative operator over phase harmonics. Phase harmonics can be interpreted as non-linear frequency transpositions. For appropriate network filters, they define an invertible bi-Lipschitz signal representation. We shall study the particular case where network filters are wavelets, and thus define a complete and stable multiscale representation.

A major issue in harmonic analysis is to capture the dependence of phases across frequencies, which produce "coherent structures" such as isolated singularities. The phase is often considered as an annoying variable which yields decorrelated coefficients at different frequencies because of cancellation effects. Moreover, the phase is unstable at locations the modulus vanishes. Most recent harmonic analysis research has concentrated on the modulus of wavelet coefficients, which is sufficient to specify uniform or pointwise regularity of functions in Sobolev, Hölder or Besov spaces [Jaffard, 1991]. However, scrambling the phase by modifying the sign of wavelet coefficients in a basis destroys the geometric structures of signals and images, which shows that phase dependencies are crucial. The importance of the wavelet phase to capture coherent signal structures was emphasized by the pioneer work in [Grossmann et al., 1989]. It was also shown that correlating the phase of wavelet coefficients across frequencies improves the synthesis of image textures [Portilla & Simoncelli, 2000].

The main contribution of this paper is to define a stable local phase harmonic representation, where phase dependencies can be captured by linear autocorrelation matrices. Section 4 analyzes their mathematical properties. The use of autocorrelation matrices over deep convolutional network coefficients was introduced in [Gatys et al., 2015], to generate image textures which are perceptually similar to an original one. These results are simplified in [Ustyuzhaninov et al., 2017], by synthesizing textures from correlations of coefficients in a first convolutional layer, computed with local filters. We shall concentrate on recovery of signal with small $\mathbf{L^2}$ approximation errors as opposed to signals having similar perceptual properties.

Section 5 shows that fixed sets of wavelet harmonic correlations can be sufficient to recover close approximations of signals having sparse wavelet coefficients. These properties are similar to compressive sensing algorithms [Candes et al., 2006], where linear sensing measurements are replaced by non-linear harmonic autocorrelations. Numerical experiments indicate that approximation errors can have the same decay rate as sparse approximations in wavelet bases. However, similarly to compressive sensing, it requires that wavelet coefficients are sufficiently sparse.

Section 2 introduces the first main result which proves that a non-linear rectifier acts as a filter on the phase, and creates phase harmonics. Section 3 considers the case where network filters are wavelets. Section 4 studies phase harmonic means and autocorrelations to capture phase dependence. Section 5 shows that signals having sparse wavelet coefficients can be recovered from few wavelet harmonic correlations. Computations can be reproduced with a software in *https://github.com/kymatio/phaseharmonics*.

Notations: We write $\varphi(z)$ the complex phase of $z \in \mathbb{C}$. For any $x(u)$, $\|x\|^2 = \int |x(u)|^2 \, du$. The Fourier transform of $x(u)$ is written $\widehat{x}(\omega) = \int_{\mathbb{R}^d} x(u) \, e^{-i\omega.u} \, du$. The indicator function of a set S is 1_S.

2. Phase Harmonics

Convolutional network architectures typically learn filters of small support. For image and audio applications, filters in the first layer usually have a Fourier transform which is well localized [Mallat, 2016]. In this case, we show that rectified non-linearities act as a filter on the phase of the first layer coefficients. Section 2.1 defines the notion of phase by extending the real network filters into complex analytic

filters. By applying a Fourier transform along the phase variable, Section 2.3 proves that the phase filtering is a multiplicative filter over harmonic components. Lipschitz continuity properties of harmonic representations are proved in Section 2.4.

2.1 *Rectifiers as Analytic Phase Filters*

A one-layer convolutional network computes convolutions of a d-dimensional input signal $x(u)$ with a family of real valued filters which we write $\{\psi_m^r\}_m$. It then applies a pointwise non-linearity ρ. The resulting first layer convolutional coefficients are indexed by a position u and a channel m

$$Ux(u,m) = \rho(x \star \psi_m^r(u)).$$

The non-linearity is often chosen to be the rectifier $\rho(a) = \max(a,0)$.

If x is positive and if ψ_m^r is a low-pass filter which averages x then $\rho(x \star \psi_m^r(u)) = x \star \psi_m^r(u)$ because $x \star \psi_m^r(u)$ remains positive. This is not true if ψ_m^r is a band-pass filter: $\int \psi_m^r(u)\,du = 0$. In this case, we analyze the effect of the rectifier by computing an analytic extension of ψ_m^r.

ANALYTIC EXTENSION In one dimension $u \in \mathbb{R}$, the analytic part ψ_m of ψ_m^r is a complex filter whose Fourier transform $\widehat{\psi}_m$ is the restriction of $\widehat{\psi}_m^r$ over positive frequencies:

$$\widehat{\psi}_m(\omega) = 2\,\widehat{\psi}_m^r(\omega)\,1_{[0,\infty)}(\omega)\ . \tag{2.1}$$

Since ψ_m^r is real, $\widehat{\psi}_m^r(-\omega) = \widehat{\psi}_m^r(\omega)^*$, so ψ_m^r is the real part of ψ_m:

$$\psi_m^r(u) = \text{Real}(\psi_m(u)). \tag{2.2}$$

For example, if $\psi_m^r(u) = \cos(\lambda u + \alpha)$ with $\lambda > 0$ then its analytic part is $\psi_m(u) = e^{i(\lambda u+\alpha)}$.

If $u \in \mathbb{R}^d$ is multidimensional, the analytic part is defined from the restriction of the Fourier transform of $\widehat{\psi}_m^r$ over half of the frequency space

$$\widehat{\psi}_m(\omega) = 2\,\widehat{\psi}_m^r(\omega)\,1_{S_m}(\omega)\ ,$$

where S_m is a half space of $\mathbb{R}^d$ whose boundary is a hyperplane including the frequency 0. This analytic extension depends upon the choice of S_m, which is not unique. If possible, S_m is chosen so that $\widehat{\psi}_m^r(\omega)$ vanishes at the boundary of S_m, so that $\widehat{\psi}_\lambda(\omega)$ is not discontinuous at this boundary. Since $\widehat{\psi}_m^r(-\omega) = \widehat{\psi}_m^r(\omega)^*$, we verify that $\psi_m^r = \text{Real}(\psi_m)$.

Since $\psi_m^r = \text{Real}(\psi_m)$, convolutional network coefficients can be rewritten

$$Ux(u,m) = \rho(x \star \text{Real}(\psi_m)(u)) = \rho(\text{Real}(x \star \psi_m(u))). \tag{2.3}$$

PHASE FILTER Let us now show that a rectifier acts as a filter on the phase of these complex coefficients. This is valid for any homogeneous operator ρ, which means that

$$\forall (\beta, a) \in \mathbb{R}^+ \times \mathbb{R}\ ,\ \ \rho(\beta a) = \beta\,\rho(a).$$

A rectifier is homogeneous, but an absolute value $\rho(a) = |a|$ or the identity $\rho(a) = a$ are also homogeneous. If ρ is homogeneous then for any $z = |z|\,e^{i\varphi(z)} \in \mathbb{C}$

$$\rho(\text{Real}(z)) = |z|\,h(-\varphi(z))) \ \text{ with } \ h(\alpha) = \rho(\cos(\alpha)). \tag{2.4}$$

The function h is the 2π periodic phase filter. For a rectifier $h(\alpha) = \max(\cos\alpha, 0)$, for an absolute value $h(\alpha) = |\cos\alpha|$, and for the identity $h(\alpha) = \cos\alpha$.

In signal processing, the complex phase $\varphi(x \star \psi_m(u))$ is called the analytic phase of $x \star \psi_m^r$. Applying (2.4) to $z = x \star \psi_m(u)$ proves that a rectifier applied to output of band-pass filters acts as a filter on the analytic phase

$$Ux(u,m) = |x \star \psi_m(u)|\, h(-\varphi(x \star \psi_m(u))) \,. \tag{2.5}$$

The rectifier sets to zero all coefficients whose phases are in $[\pi/2, 3\pi/2]$.

2.2 *Phase Filtering and Selectivity*

The phase of $x \star \psi_m$ depends upon the phase of ψ_m. In the first layer of a convolutional network, there are usually several filters having similar frequency locations but different phases. We explain the role of the phase.

PHASE AND MEAN FREQUENCY We define the phase of real filters ψ_m^r from the complex phase of their analytic part ψ_m. This phase is evaluated at a mean frequency. The mean frequency $\lambda \in \mathbb{R}^d$ of $|\widehat{\psi}_m|^2$ is defined by

$$\lambda = \frac{\int \omega\, |\widehat{\psi}_m(\omega)|^2\, d\omega}{\int |\widehat{\psi}_m(\omega)|^2\, d\omega}. \tag{2.6}$$

We suppose that $|\widehat{\psi}_m(\lambda)| \neq 0$, so $\widehat{\psi}_m(\lambda) = e^{-i\alpha}\, |\widehat{\psi}_m(\lambda)|$. We introduce a filter $\psi_\lambda = e^{i\alpha}\psi_m$ having a 0 phase at the frequency λ.

If $\psi_\lambda(u)$ is localized both along u and in frequency then $x \star \psi_\lambda(u)$ provides an information about x which is local both around the spatial variable u and the frequency variable λ. These local "space-frequency" representations have been studied extensively [Mallat, 2001], depending upon the localization of ψ_λ and $\widehat{\psi}_\lambda$. Signal properties have been mostly analyzed through the complex modulus $|x \star \psi_\lambda(u)|$. It is called a spectrogram if ψ_λ is a windowed Fourier filter and a scalogram if ψ_λ is a wavelet.

Local space-frequency representations do not easily capture the dependencies of $x \star \psi_\lambda(u)$ and $x \star \psi_{\lambda'}(u)$ at different frequencies because they are not linearly correlated. Indeed,

$$\int \Big(x \star \psi_\lambda(u)\Big)\, \Big(x \star \psi_{\lambda'}(u)\Big)^*\, du = (2\pi)^{-d} \int |\widehat{x}(\omega)|^2\, \widehat{\psi}_\lambda(\omega)\, \widehat{\psi}_{\lambda'}(\omega)^*\, d\omega \approx 0$$

if the filters are separated in the sense that $\widehat{\psi}_\lambda\, \widehat{\psi}_{\lambda'} \approx 0$. This cancellation is due to phase oscillations. This is why the phase is often removed and signal properties are usually analyzed from the modulus $|x \star \psi_\lambda(u)|$. We shall see that the non-linear action of a rectifier introduces correlations while preserving phase information.

Since $\psi_\lambda = e^{i\alpha}\psi_m$ and $\psi_m^r = \mathrm{Real}(\psi_m)$, we get

$$\psi_m^r = \mathrm{Real}(e^{-i\alpha}\psi_\lambda). \tag{2.7}$$

We will use α as a free continuous parameter in $[0, 2\pi]$ to arbitrarily modify the phase of ψ_m^r. Let us replace the index m by (λ, α) which is more meaningful:

$$Ux(u, \lambda, \alpha) = \rho\Big(\mathrm{Real}(e^{-i\alpha} x \star \psi_\lambda(u))\Big). \tag{2.8}$$

Similarly to (2.5) we get

$$Ux(u,\lambda,\alpha) = |x\star\psi_\lambda(u)|\,h(\alpha-\varphi(x\star\psi_\lambda(u)))\,, \tag{2.9}$$

with $h(\alpha)=\rho(\cos\alpha)$. We study the properties of this local space-frequency-phase representation for arbitrary phase filters $h(\alpha)$.

Each coefficient $Ux(u,\lambda,\alpha)$ provides information which is local along u and λ but also along the phase α. If the phase filter h has a support $[-A,A]$ then $Ux(u,\lambda,\alpha)$ in (2.9) is non-zero only at points u where $\varphi(x\star\psi_\lambda(u))\in[\alpha-A,\alpha+A]$. For a rectifier, the support of $h(\alpha)=\max(\cos\alpha,0)$ is $[-\pi/2,\pi/2]$, which is not highly selective. However, this phase selection eliminates the phase cancellation effect by ensuring that $Ux(u,\lambda,\alpha)\geqslant 0$ so that the correlation $\int Ux(u,\lambda,\alpha)\,Ux(u,\lambda',\alpha')\,du$ becomes non-zero even for separated frequencies $\widehat{\psi}_\lambda\,\widehat{\psi}_{\lambda'}\approx 0$. The properties of these correlation coefficients are studied in Section 4. The definition of phase harmonic representations in (2.9) is extended to any phase filter $h(\alpha)$ to potentially improve selectivity along phases.

The phase α is similar to a local translation parameter. In one dimension, if $\psi_\lambda(u)=e^{i\lambda u}$ then $e^{-i\alpha}x\star\psi_\lambda(u)=\widehat{x}(\lambda)\,e^{i\lambda(u-\alpha/\lambda)}$ so α/λ is a global translation parameter along u. However, if $\psi_\lambda(u)$ is a localized sinusoidal wave of frequency λ then α/λ induces a local translation only within the support of $\psi_\lambda(u)$. For a two-dimensional filter ψ_λ having a localized support in u and whose Fourier transform is centered at $\lambda=(\lambda_1,\lambda_2)$, the phase α acts as a local translation in $\mathbb{R}^2$ by $\alpha/|\lambda|$ in the direction of λ.

PHASE SELECTIVITY The phase selectivity depends upon the support of h, which can be modified by a convolution of $Ux(u,\lambda,\alpha)$ along α. Such a convolution is linear along α and identical for each u and can thus be implemented in a deep convolutional neural networks when computing second layer coefficients. The circular convolution of 2π periodic functions is written

$$a\circledast b(\alpha)=\int_{[0,2\pi]} a(\beta)\,b(\alpha-\beta)\,d\beta.$$

Computing the convolution of $Ux(u,\lambda,\alpha)$ in (2.9) by $g(\alpha)$ along α gives

$$Ux(u,\lambda,.)\circledast g(\alpha)=|x\star\psi_\lambda(u)|\,h\circledast g(\alpha-\varphi(x\star\psi_\lambda(u)). \tag{2.10}$$

It changes the phase filter h into $h\circledast g$. The following theorem proves that $h\circledast g$ may have an arbitrarily narrow support and thus become highly selective in phase, for an appropriate choice of g.

THEOREM 2.1 If $h(\alpha)=\max(\cos\alpha,0)$ or $h(\alpha)=|\cos\alpha|$ then for any $\varepsilon>0$ there exists a bounded filter g such that $h\circledast g$ is differentiable, π periodic, supported in $[-\varepsilon,\varepsilon]$ modulo π, with $\int_0^{2\pi} h\circledast g(\alpha)\,d\alpha=2$, and

$$\lim_{\varepsilon\to 0} h\circledast g(\alpha)=\delta(\alpha \bmod \pi). \tag{2.11}$$

Proof: We show that we can define g so that $h\circledast g$ is a π periodic cubic box spline supported in $[-\varepsilon,\varepsilon]$ with $\widehat{h\circledast g}(0)=\int_0^{2\pi} h\circledast g(\alpha)\,d\alpha=2$. One can verify that the Fourier transform of this box spline is

$$\widehat{h\circledast g}(k)=\begin{cases} 2 & \text{if } k=0\\ 2\sin^4(k\varepsilon/4)\,(k\varepsilon/4)^{-4} & \text{if } k\neq 0 \text{ is even}\\ 0 & \text{if } k \text{ is odd.}\end{cases} \tag{2.12}$$

For a rectifier and an absolute value, $\widehat{h}(k)$ is non-zero when k is even. Since $\widehat{h\circledast g}=\widehat{h}\,\widehat{g}$ we obtain (2.12) by defining $\widehat{g}(k)=0$ if k is odd and $\widehat{g}(k)=\widehat{h}_g(k)/\widehat{h}(k)$ if k is even. We prove that g is bounded

by verifying that $\sum_k |\widehat{g}(k)| < \infty$. The analytical expressions of $\widehat{h}(k)$ for a rectifier and an absolute value are given by (2.17) and (2.18). They imply that $|\widehat{g}(k)| = O(|k|^{-2})$ and hence that g is bounded.

The restrictions of $h \circledast g$ to $[-\pi/2, \pi/2]$ and $[\pi/2, 3\pi/2]$ are bounded functions of integrals equal to 1. Their supports are respectively $[-\varepsilon, \varepsilon]$ and $[\pi - \varepsilon, \pi + \varepsilon]$. When ε goes to 0 the restrictions of $h \circledast g$ to $[-\pi/2, \pi/2]$ and $[\pi/2, 3\pi/2]$ thus converge to $\delta(\alpha)$ and $\delta(\alpha - \pi)$, which proves (2.11). □

This theorem proves that for a rectifier or an absolute value, one can improve arbitrarily the phase selectivity through convolutions along phases. The theorem proof gives a possible but non-unique choice for the filters g. For wavelet filters and images, Section 3 shows that lines of constant phase α define the geometry of multiscale edges.

2.3 *Phase Harmonics and Frequency Transpositions*

This section proves that the phase filtering becomes a multiplication over a representation of phase harmonics. These phase harmonics correspond to non-linear frequency transpositions.

PHASE HARMONICS We separate the linear part of U from the non-linear phase filtering by factorizing

$$U = HW. \tag{2.13}$$

The linear operator W computes convolutions with analytic filters of mean frequency λ having a zero phase at λ, where λ belongs to a set Λ:

$$\forall \lambda \in \Lambda \ , \ Wx(u, \lambda) = x \star \psi_\lambda(u) \ ,$$

The phase filter H is defined by

$$\forall z \in \mathbb{C} \ , \ Hz(\alpha) = |z| \, h(\alpha - \varphi(z)). \tag{2.14}$$

Since α is a phase translation over $[0, 2\pi]$, we prove that a Fourier transform relatively to α transforms H into a multiplication over a sequence of phase harmonics defined below.

DEFINITION 2.2 The phase harmonics of a complex number $z \in \mathbb{C}$ is a sequence defined for all $k \in \mathbb{Z}$ by

$$[z]^k = |z| \, e^{ik\varphi(z)} \ .$$

As opposed to z^k which applies the exponent k to the modulus and to the phase, the phase harmonic $[z]^k$ transforms the phase only. As a result Lemma 2.1 will prove that $[z]^k$ is Lipschitz for all $k \in \mathbb{Z}$ whereas z^k is not Lipschitz for $|k| \neq 1$. The Fourier coefficients of a 2π periodic function $h(\alpha)$ are written

$$\widehat{h}(k) = \frac{1}{2\pi} \int_0^{2\pi} h(\alpha) \, e^{-ik\alpha} \, d\alpha \ .$$

The following proposition proves that the Fourier transform of $Hz(\alpha)$ is proportional to the phase harmonics of z. It results that the Fourier transform of $Ux(u, \lambda, \alpha)$ relatively to α is proportional to the phase harmonics of $x \star \psi_\lambda(u)$.

PROPOSITION 2.3 For all $z \in \mathbb{C}$, the Fourier transform of $Hz(\alpha)$ is

$$\widehat{Hz}(k) = \widehat{h}(k) \, [z]^{-k} \ . \tag{2.15}$$

The Fourier transform of $Ux(u,\lambda,\alpha)$ along α is

$$\widehat{U}x(u,\lambda,k) = \widehat{h}(k)\,[x\star\psi_\lambda(u)]^{-k}\,. \tag{2.16}$$

Equation (2.15) and (2.16) are directly obtained by computing the Fourier transform of (2.14) and (2.9) along α. The Fourier transform $\widehat{U}x(u,\lambda,k)$ is a local space-frequency-harmonic representation where k is the harmonic exponent. The phase filter produces a harmonic multiplier $\widehat{h}(k)$, which typically attenuates high order harmonics.

If $Ux(u,\lambda,\alpha)$ is calculated with a homogeneous non-linearity ρ then $h(\alpha)=\rho(\cos\alpha)$. If $\rho(a)=a$ then $h(\alpha)=\cos\alpha$ so $\widehat{h}(k)=1/2$ if $k=\pm1$ and $\widehat{h}(k)=0$ otherwise. This is usually not an appropriate choice because it eliminates all high-order harmonics. If $\rho(a)=\max(a,0)$ then $h(\alpha)=\max(\cos\alpha,0)$ and a direct calculation of Fourier integrals gives

$$\widehat{h}(k) = \begin{cases} \frac{-(i)^k}{\pi(k-1)(k+1)} & \text{if } k \text{ is even} \\ \frac{1}{4} & \text{if } k=\pm1 \\ 0 & \text{if } |k|>1 \text{ is odd} \end{cases}\,. \tag{2.17}$$

One can verify that any real homogeneous operator can be written $\rho(a)=\gamma\max(a,0)+\beta a$ so $\widehat{h}(k)=0$ when $|k|>1$ is odd. We may however define other phase filters for which $\widehat{h}(k)\neq0$ for $|k|>1$ odd. If $\rho(a)=|a|$ then $h(\alpha)=|\cos\alpha|$ and hence

$$\widehat{h}(k) = \begin{cases} \frac{-2(i)^k}{\pi(k-1)(k+1)} & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases}\,. \tag{2.18}$$

Because the absolute value is even and thus loses sign information, we have $\widehat{h}(1)=\widehat{h}(-1)=0$.

FREQUENCY TRANSPOSITION We show that phase harmonics produce frequency transpositions. They multiply the frequencies of $x\star\psi_\lambda$ by a factor k and hence perform a non-linear dilation by k of the Fourier transform of $x\star\psi_\lambda$, without affecting its spatial localization.

The derivative of the phase $\varphi(x\star\psi_\lambda(u))$ of $x\star\psi_\lambda(u)$ at u is called the analytic instantaneous frequency in signal processing [Mallat, 2001]. It corresponds to the perceived frequency of a sound at a given time. Since $[x\star\psi_\lambda(u)]^k = |x\star\psi_\lambda(u)|e^{ik\varphi(x\star\psi_\lambda(u))}$, it has a phase derivative which is k times larger. However, its modulus is unchanged so $x\star\psi_\lambda$ and $[x\star\psi_\lambda]^k$ have the same localization along u. For time signals, a transformation which modifies the instantaneous frequency but does not modify the time distribution is called a frequency transposition, by analogy to music transpositions. For example, shifting a musical score by one octave multiplies by two the "frequencies" of all musical notes, but it does not change the tempo and the melody.

A frequency transposition performs a non-linear dilation in the Fourier domain. The Fourier support of $x\star\psi_\lambda$ is included in the support of $\widehat{\psi}_\lambda$, which is centered at λ. In a first approximation, the Fourier transform of $[x\star\psi_\lambda]^k$ is non negligible over a domain centered in $k\lambda$ whose width is approximately k times the support width of $\widehat{\psi}_\lambda$. This is proved by the following proposition, under very restrictive hypotheses. A ball in $\mathbb{R}^d$ is written $B(c,r)=\{u\in\mathbb{R}^d\ :\ |u-c|\leqslant r\}$.

PROPOSITION 2.4 If $|x\star\psi_\lambda(u)|$ and $e^{i\varphi(x\star\psi_\lambda(u))}$ have a Fourier transform respectively supported in $B(0,\Delta)$ and $B(\lambda,\Delta)$ then $[x\star\psi_\lambda]^k$ has a Fourier transform supported in $B(k\lambda,(|k|+1)\Delta)$.

Proof: Since $[x\star\psi_\lambda]^k = |x\star\psi_\lambda|\,(e^{i\varphi(x\star\psi_\lambda(u))})^k$, its Fourier transform is the convolution of the Fourier transform of $|x\star\psi_\lambda|$ and k successive convolutions of the Fourier transform of $e^{i\varphi(x\star\psi_\lambda(u))}$. The Fourier support of $[x\star\psi_\lambda]^k$ can be derived from the Fourier supports of $|x\star\psi_\lambda|$ and $e^{i\varphi(x\star\psi_\lambda(u))}$ because the convolutions of two functions included in $B(c,r)$ and in $B(c',r')$ is a function included in $B(c+c',r+r')$. □.

If $\widehat{\psi}_\lambda$ is supported in $B(\lambda,\Delta)$ then $x\star\psi_\lambda$ has a Fourier transform supported in $B(\lambda,\Delta)$ but in general $|x\star\psi_\lambda|$ and $e^{i\varphi(x\star\psi_\lambda(u))}$ do not have a Fourier transform of compact support. Indeed, $|x\star\psi_\lambda|$ may be singular if $x\star\psi_\lambda(u)$ has a zero-crossing, which produces a Fourier transform with a slow asymptotic decay. The hypotheses of this proposition are thus very restrictive. For wavelet filters, Section 3 verifies numerically that the Fourier transform of $[x\star\psi_\lambda]^k$ is essentially dilated by k for a white noise signal.

2.4 *Bi-Lipschitz Continuity and Inversion*

This section considers general phase filters $h(\alpha)$ and gives bi-Lipschitz bounds on H depending upon h. We derive that U is also bi-Lipschitz and invertible for appropriate filters ψ_λ.

For any 2π periodic $g(\alpha)$, we have

$$\|g\|^2 = \frac{1}{2\pi}\int_0^{2\pi} |g(\alpha)|^2\, d\alpha = \sum_{k\in\mathbb{Z}} |\widehat{g}(k)|^2.$$

We first prove that each phase harmonic $[z]^k$ is Lipschitz, even if $|k|\neq 1$, as opposed to usual integer power exponents z^k.

LEMMA 2.1 For any $(z,z')\in\mathbb{C}^2$ and any $k\in\mathbb{Z}$

$$|[z]^k - [z']^k| \leqslant \max(1,|k|)\,|z-z'|\,. \tag{2.19}$$

Proof: The inequality (2.19) is verified if $k=0$ because

$$|[z]^k - [z']^k| = ||z|-|z'|| \leqslant |z-z'|.$$

For $k\neq 0$ and $\varepsilon = \varphi(z')-\varphi(z)$, let us define

$$\begin{aligned} f(|z|,|z'|,\varepsilon) &= \frac{|[z]^k-[z']^k|^2}{|z-z'|^2} = \frac{||z|-|z'|e^{ik\varepsilon}|^2}{||z|-|z'|e^{i\varepsilon}|^2} \\ &= \frac{||z|^2+|z'|^2-2|z||z'|\cos(k\varepsilon)|}{||z|^2+|z'|^2-2|z||z'|\cos(\varepsilon)|} \end{aligned}$$

One can verify that

$$\sup_{|z|,|z'|,\varepsilon} f(|z|,|z'|,\varepsilon) = |k|^2,$$

which implies (2.19). This is done by proving that for $|z|$ and $|z'|$ fixed, the maximum of f is reached when ε tends to 0, and when ε tends to zero $f(|z|,|z'|,\varepsilon)$ is maximized when $|z|=|z'|$, with a supremum equal to $|k|^2$. □

The following theorem applies this lemma to prove that H is bi-Lipschitz.

THEOREM 2.5 For any $(z,z')\in\mathbb{C}^2$

$$\|Hz\| = \|h\|\,|z|\,, \tag{2.20}$$

and

$$\sqrt{2}\,|\widehat{h}(1)|\,|z-z'| \leqslant \|Hz-Hz'\| \leqslant \kappa\,|z-z'| \tag{2.21}$$

with

$$\kappa^2 = |\widehat{h}(0)|^2 + \sum_{k\in\mathbb{Z}} k^2\,|\widehat{h}(k)|^2 = \frac{1}{4\pi^2}\Big(\int_0^{2\pi} h(\alpha)\,d\alpha\Big)^2 + \frac{1}{2\pi}\int_0^{2\pi} |h'(\alpha)|^2\,d\alpha. \tag{2.22}$$

For a rectifier, $\|h\| = 1/2$, $\widehat{h}(1) = 1/4$ and $\kappa = \sqrt{1/4+1/\pi^2}$.

Proof: Since $\widehat{Hz}(k) = \widehat{h}(k)\,[z]^{-k}$, we first prove (2.20) by observing that

$$\|Hz\|^2 = \sum_{k=-\infty}^{+\infty} |\widehat{h}(k)|^2 |[z]^{-k}|^2 = |z|^2 \sum_{k=-\infty}^{+\infty} |\widehat{h}(k)|^2 = |z|^2\,\|h\|^2.$$

To verify (2.21) let us compute

$$\|Hz-Hz'\|^2 = \sum_{k=-\infty}^{+\infty} |\widehat{h}(k)|^2\,|[z]^{-k} - [z']^{-k}|^2. \tag{2.23}$$

Since $h(\alpha)$ is real, $|\widehat{h}(k)| = |\widehat{h}(-k)|$. Restricting the sum to $k = \pm 1$ gives the lower bound

$$2\,|\widehat{h}(1)|^2\,|z-z'|^2 \leqslant \|Hz-Hz'\|^2.$$

We obtain the upper bound inequality of (2.21) inserting the Lipschitz inequality (2.19) of Lemma 2.1 in (2.23) with

$$\kappa^2 = |\widehat{h}(0)|^2 + \sum_{k\in\mathbb{Z}} k^2\,|\widehat{h}(k)|^2, \tag{2.24}$$

Since $\widehat{h}(0) = 1/2\pi \int_0^{2\pi} h(\alpha)\,d\alpha$ and the Fourier transform of $h'(\alpha)$ is $ik\widehat{h}(k)$, applying the Plancherel formula on (2.24) proves (2.22). A direct calculation gives the values of $\|h\|$ and κ for a linear rectifier where $h(\alpha) = \max(\cos\alpha, 0)$. □

In the particular case where the phase filter can be written $h(\alpha) = \rho(\cos\alpha)$ then the Lipschitz constant κ in (2.22) is finite if ρ has bounded derivatives. For a rectifier the lower and upper Lipschitz constants of the theorem are respectively about 0.35 and 0.69 and thus within a factor 2. In this case, one can prove that the upper bound Lipschitz constant is not tight and can be reduced to $1/2$ [Zhang & Mallat, 2019].

The following proposition reviews a standard condition which guarantees that W is invertible and stable, by imposing that the Fourier domain is uniformly covered by the filters $\widehat{\psi}_\lambda$.

PROPOSITION 2.6 If there exists $0 \leqslant \eta < 1$ such that for almost all $\omega \in \mathbb{R}^d$

$$(1-\eta)^2 \leqslant \frac{1}{2}\sum_{\lambda\in\Lambda} \Big(|\widehat{\psi}_\lambda(\omega)|^2 + |\widehat{\psi}_\lambda(-\omega)|^2\Big) \leqslant (1+\eta)^2 \tag{2.25}$$

then for all $x \in \mathbf{L^2}(\mathbb{R}^d)^2$ real

$$x = \sum_{\lambda\in\Lambda} \mathrm{Real}(x \star \psi_\lambda \star \overline{\psi}_\lambda), \tag{2.26}$$

where the support of $\widehat{\overline{\psi}_\lambda}$ is the same as the support of $\widehat{\psi}_\lambda$ and

$$\widehat{\overline{\psi}_\lambda}(\omega) = \frac{\widehat{\psi_\lambda}(\omega)^*}{\sum_{\lambda\in\Lambda} |\widehat{\psi}_\lambda(\omega)|^2}.$$

Moreover, for all $(x,x') \in \mathbf{L^2}(\mathbb{R}^d)^2$ real

$$(1-\eta)^2 \|x-x'\|^2 \leqslant \|Wx - Wx'\|^2 \leqslant (1+\eta)^2 \|x-x'\|^2 . \tag{2.27}$$

Proof: Since x is real, $\widehat{x}(-\omega) = \widehat{x}(\omega)^*$. Equation (2.26) is proved by taking the Fourier transform on both side and observing that

$$\frac{1}{2}\sum_{\lambda\in\Lambda} \left(\widehat{\psi}_\lambda(\omega)\,\widehat{\overline{\psi}_\lambda}(\omega) + \widehat{\psi}_\lambda(-\omega)^*\,\widehat{\overline{\psi}_\lambda}(-\omega)^* \right) = 1.$$

Since W is linear, (2.27) can be proved by setting $x' = 0$. Observe that

$$\|Wx\|^2 = \sum_{\lambda\in\Lambda} \|x\star\psi_\lambda\|^2 = \frac{1}{(2\pi)^d} \sum_{\lambda\in\Lambda} \int |\widehat{x}(\omega)|^2\, |\widehat{\psi}_\lambda(\omega)|^2\, d\omega.$$

Inverting the sum and the integral and inserting (2.25) in this equation proves (2.27) for $x' = 0$. □

The following corollary derives that $U = HW$ is invertible and bi-Lipschitz. We write

$$\|Ux\|^2 = \frac{1}{2\pi}\sum_{\lambda\in\Lambda} \int_0^{2\pi} \int_{\mathbb{R}^d} |Ux(u,\lambda,\alpha)|^2\, du\, d\alpha .$$

COROLLARY 2.1 If condition (2.25) is satisfied for $0 \leqslant \eta < 1$ and $\widehat{h}(1) \neq 0$ then for all $x \in \mathbf{L^2}(\mathbb{R}^d)$ real

$$x(u) = \sum_{\lambda\in\Lambda} \text{Real}\Big(\frac{1}{\widehat{h}(1)^*}\, \widehat{Ux}(.,\lambda,1)^* \star \overline{\psi}_\lambda(u) \Big), \tag{2.28}$$

$$\|h\|\,(1-\eta)\,\|x\| \leqslant \|Ux\| \leqslant \|h\|\,(1+\eta)\,\|x\| \tag{2.29}$$

and for all $(x,x') \in \mathbf{L^2}(\mathbb{R}^d)^2$ real

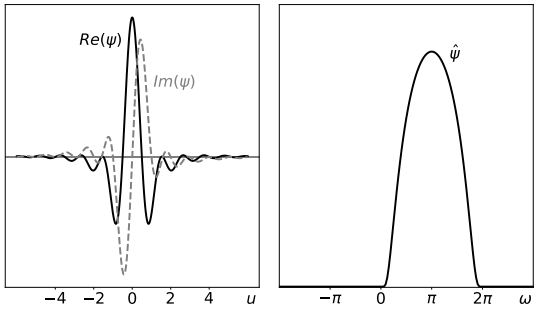
$$\sqrt{2}\,|\widehat{h}(1)|\,(1-\eta)\,\|x-x'\| \leqslant \|Ux - Ux'\| \leqslant \kappa\,(1+\eta)\,\|x-x'\| . \tag{2.30}$$

Proof: It results from (2.16) that $\widehat{Ux}(u,\lambda,1) = \widehat{h}(1)\, x\star\psi_\lambda(u)^*$, together with(2.26) it implies (2.28).

Since $U = HW$, (2.27) together with (2.20) proves (2.29), and (2.27) together with (2.21) proves (2.30). □

3. Wavelet Transforms

Wavelets are dilated filters ψ_λ which separate signal variations at multiple scales. They provide sparse representations of large classes of signals and images. Sections 3.1 and 3.2 review the properties of complex analytic wavelet transforms and introduce a new bump wavelet for numerical calculations.



(a) (b)

FIG. 1: (a): The real and imaginary parts of a bump wavelet $\psi(u)$, for $Q = 1$, are shown with a full and a dashed line. (b): Fourier transform $\widehat{\psi}(\omega)$.

3.1 *Analytic Wavelets for 1D Signals*

A one-dimensional wavelet transform is computed by convolutions with dilated wavelets. Analytic wavelets ψ have a Fourier transform $\widehat{\psi}(\omega)$ which is zero at negative frequencies. This section introduces a new analytic bump wavelet, which is used in numerical calculations. It provides sparse representations of piecewise regular signals.

We impose that $\widehat{\psi}$ is real valued which implies that the real part of ψ is even and its imaginary part is odd. Let ξ be the mean frequency of ψ according to (2.6). We also suppose that $\widehat{\psi}(\xi) > 0$. A wavelet transform with Q scales per octave is calculated by dilating ψ by $2^{j/Q}$, where j is an integer:

$$\psi_\lambda(u) = 2^{-j/Q}\,\psi(2^{-j/Q}u) \;\text{ and hence }\; \widehat{\psi}_\lambda(\omega) = \widehat{\psi}(2^{j/Q}\omega)\,.$$

The mean frequency of ψ_λ is

$$\lambda = 2^{-j/Q}\xi.$$

If the energy of $\widehat{\psi}(\omega)$ is concentrated in an interval centered at ξ of radius $\beta\xi$ then $\widehat{\psi}_\lambda$ is concentrated in an interval centered at λ of radius $\beta\lambda$.

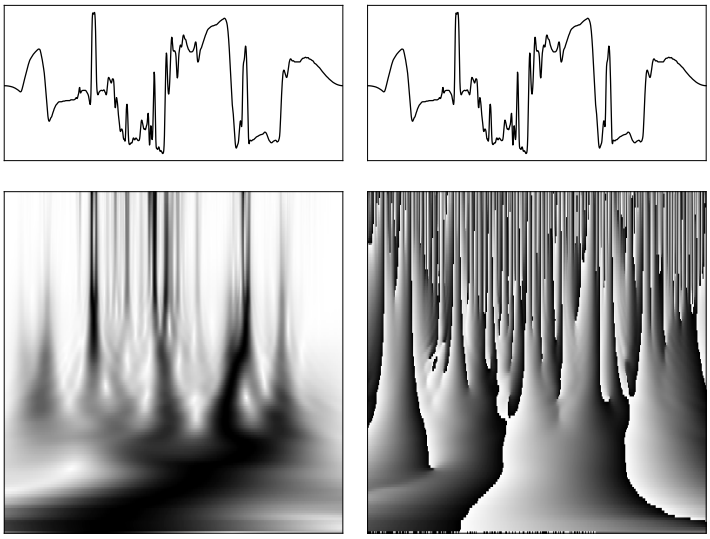
A real filter of phase α is defined by $\text{Real}(e^{-i\alpha}\psi_\lambda)$. The phase α is thus a symmetry parameter which makes the transition from even to odd filters, and which changes the filters sign when adding π.

Let 2^J be the maximum scale. Scales larger than 2^J are carried by a low-pass filter ψ_0 centered at the frequency $\lambda = 0$ and dilated by 2^J. We use a Gaussian filter whose Fourier transform is

$$\widehat{\psi}_0(\omega) = \exp\Big(-\frac{|\omega|^2}{2\sigma_J^2}\Big). \tag{3.1}$$

We consider an analytic wavelet whose Fourier transform $\widehat{\psi}(\omega)$ is a regular window centered at a frequency $\xi > 0$

$$\widehat{\psi}(\omega) = c\,g\Big(\frac{\omega-\xi}{\xi}\Big), \tag{3.2}$$



(a) (b)

FIG. 2: Top: original signal $x(u)$ as a function of u. (a): Wavelet transform modulus $|x * \psi_\lambda(u)|$ with $Q = 16$ scales per octave, as a function of $(u, \log_2 \lambda)$ along the horizontal and vertical axes. White and black points correspond respectively to small and large amplitudes. (b): Complex phase $\varphi(x * \psi_\lambda(u))$ as a function of $(u, \log_2 \lambda)$.

where $g(\omega)$ has a support in $[-1,1]$ with $g(1)=g(-1)=0$, so that $\widehat{\psi}(\omega)=0$ for $\omega \leqslant 0$. The window g is chosen so that ψ is well localized both in the spatial and Fourier domains, which has a tendency to produce wavelet coefficients which are more sparse. A bump window is an infinitely differentiable approximation of a Gaussian having a support equal to $[-1,1]$ [Jamshidi & Kirby, 2006]:

$$g(\omega) = \exp\Big(\frac{-|\omega|^2}{1-|\omega|^2}\Big)\, 1_{(-1,1)}(\omega)\ . \tag{3.3}$$

It defines a compact support $\widehat{\psi}(\omega)$ which is $\mathbf{C}^\infty$. The resulting *bump* wavelet ψ is a Schwartz class $\mathbf{C}^\infty$ analytic function, with a decay faster than any rational function. Since all derivatives of $\widehat{\psi}$ vanish at $\omega=0$, ψ has an infinite number of vanishing moments:

$$\forall k \in \mathbb{N}\ \ ,\ \ \int u^k\, \psi(u)\, du = 0\ .$$

Vanishing moments are important so that wavelet coefficients $x\star\psi_\lambda(u)$ are small in domains of u where x is regular [Mallat, 2001].

For wavelets, the Fourier transform condition (2.25) is called the Littlewood-Paley inequality. It guarantees that the wavelet transform is invertible and stable, with bounds which depend upon η. To minimize η, the constants in (3.1) and (3.2) are chosen to be

$$\sigma_J = 2^{\frac{-0.550}{Q}}\, 2^{-J+1}\,\xi \ \text{ and } \ c = (1.34\sqrt{Q}-0.05)^{-1}.$$

In numerical applications, we choose $\xi = 0.85\,\pi$. For these bump wavelets, $\eta = 0.091$ when $Q=1$ and $\eta \leqslant 0.035$ when $Q \geqslant 2$. Figure 1 shows the real and imaginary parts of ψ for $Q=1$, as well as its Fourier transform.

Figure 2 gives the modulus and the phase of the wavelet transform of a one-dimensional signal, calculated with a bump wavelet, with $Q=16$ scales per octave. Fine scales correspond to high frequencies λ. Large modulus coefficients $|x\star\psi_\lambda(u)|$ are sparse. They are located in the neighborhood of sharp signal transitions. The phase $\varphi(x\star\psi_\lambda(u))$ gives a local symmetry information on the transition of $x\star\psi_\lambda$ at u. Since the real and imaginary parts of ψ_λ are respectively symmetric and antisymmetric, the variations of $x\star\psi_\lambda$ is locally symmetric in the neighborhood of u if $\varphi(x\star\psi_\lambda(u))=0$ and antisymmetric if this phase is $\pi/2$. Changing the sign of $x\star\psi_\lambda$ adds π to the phase. When the phase is $\pi/2$ modulo π then the real part of $x\star\psi_\lambda(u)$ has a zero-crossing. Grossmann et. al. It was shown in [Grossmann et al., 1989] that lines of constant phase across scales capture properties of instantaneous frequencies. Zero-crossings of real wavelet transforms have been studied to reconstruct signals and detect the position of sharp transitions [Zhang & Mallat, 2019].

Section 2.3 explains that harmonics $[x\star\psi_\lambda]^k = |x\star\psi_\lambda|\, e^{ik\varphi(x\star\psi_\lambda)}$ perform a frequency transposition which approximately dilates the Fourier transform $x\star\psi_\lambda$ by a factor k. This is illustrated by Figure 3 for a Gaussian white noise x filtered by bump wavelet ψ_λ. For $k=0,1,2,3$, the modulus of the Fourier transform of $[x\star\psi_\lambda]^k$ has an energy centered at $k\lambda$ and concentrated over a domain dilated by $\max(k,1)$.

3.2 *Complex Steerable Wavelets for Images*

In two dimensions, we define wavelets by dilating and rotating a complex analytic wavelet $\psi(u)$ for $u \in \mathbb{R}^2$. Its Fourier transform $\widehat{\psi}(\omega)$ has a support included in the right half plane of $\mathbb{R}^2$. This wavelet is rotated in $\mathbb{R}^2$ by r_θ along several angles θ, and dilated at dyadic scales 2^j:

$$\psi_\lambda(u) = 2^{-2j}\psi(2^{-j}r_{-\theta}u) \ \text{ and hence } \ \widehat{\psi}_\lambda(\omega) = \widehat{\psi}(2^j r_\theta \omega)\ .$$

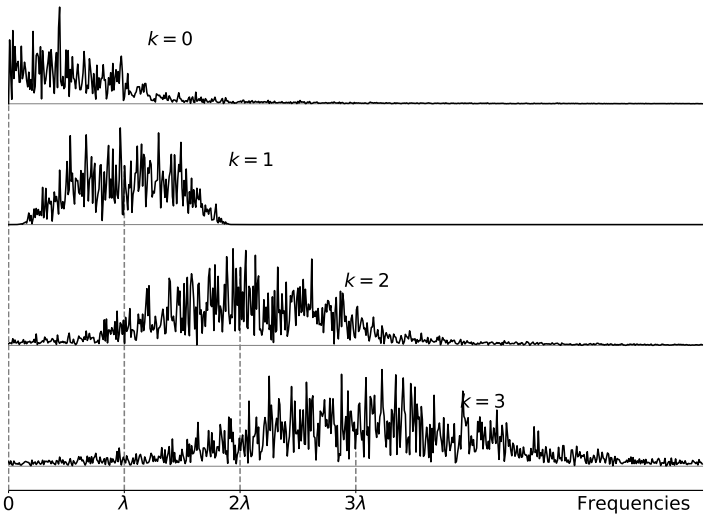


FIG. 3: Modulus of the Fourier transform of $[x \star \psi_\lambda]^k$, for a signal x which is a Gaussian white noise and a bump wavelet ψ_λ, as a function of frequencies. The center frequency λ is fixed and k increases from top to bottom. The Fourier supports are approximately dilated by k.

We use L angles $\theta = \pi\ell/L$ for $-L/2 < \ell \leqslant L/2$. If ξ is the center frequency of ψ then the center frequency of ψ_λ is

$$\lambda = 2^{-j} r_{-\theta}\, \xi \ .$$

Wavelets are computed up to a maximum scale 2^J. Scales larger than 2^J are carried by a low-pass filter ψ_0 centered at the frequency $\lambda = 0$ and dilated by 2^J. As in one dimension, we use a Gaussian filter whose Fourier transform is

$$\widehat{\psi}_0(\omega) = \exp\Big(-\frac{|\omega|^2}{2\sigma_J^2}\Big) \ . \tag{3.4}$$

Let us represent frequencies in polar coordinates $\omega = |\omega|\,(\cos\theta, \sin\theta)$. Steerable wavelets introduced in [Simoncelli & Freeman, 1995] have a center frequency $\xi = (\xi_0, 0)$ with L rotation angles, and a Fourier transform $\widehat{\psi}(\omega)$ which can be written:

$$\widehat{\psi}(\omega) = c\, g\Big(\frac{|\omega| - \xi_0}{\xi_0}\Big) \cos^{L-1}(\theta) 1_{|\theta|<\frac{\pi}{2}} . \tag{3.5}$$

Different wavelets are obtained by modifying the window g [Unser et al., 2011]. Steerable bump wavelets are computed with the one-dimensional bump window g in (3.3). Since $\widehat{\psi}$ has a compact support and $L-1$ bounded derivatives, ψ is a $\mathbf{C}^\infty$ function and $|\psi(u)| = O((1+|u|)^{-L+1})$. All partial derivatives of $\widehat{\psi}(\omega)$ are zero at $\omega = 0$ so ψ has an infinite number of vanishing moments

$$\forall (k_1, k_2) \in \mathbb{Z}^2 \ \ , \ \ \int_{\mathbb{R}^2} u_1^{k_1} u_2^{k_2}\, \psi(u_1, u_2)\, du_1 du_2 = 0 \ .$$

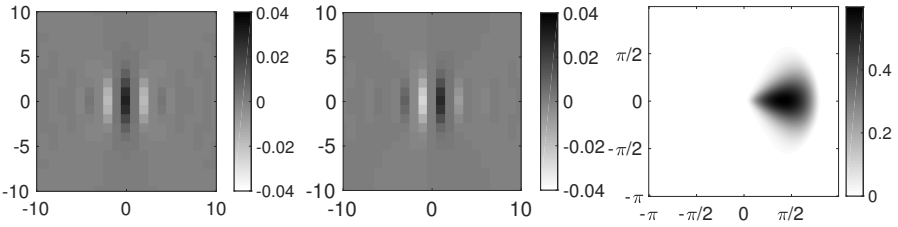


FIG. 4: (a): Real part of the two-dimensional bump wavelet $\psi(u)$ with $L = 8$. (b): Imaginary part of $\psi(u)$. (c): Fourier transform $\widehat{\psi}(\omega)$.

To minimize the Littlewood-Paley constant η in (2.25), the constants in (3.4) and (3.5) are chosen to be

$$\sigma_J = 2^{-0.550}\, 2^{-J+1}\, \xi \quad \text{and} \quad c = 1.29^{-1}\, 2^{L-1} \frac{(L-1)!}{\sqrt{L[2(L-1)]!}}.$$

In numerical applications, we choose $\xi = 0.85\,\pi$ and $L \geqslant 4$. One can verify that the Littlewood-Paley constant is $\eta = 0.091$ as for the one-dimensional wavelet (3.3) and does not depend upon L. Figure 4 shows the real and imaginary parts of ψ as well as its Fourier transform for $L = 8$ angles.

Figure 5 displays the phase and modulus of the wavelet transform of an image, calculated with bump steerable wavelets at angles $\theta = 0$ and $\theta = \pi/2$ and scales $2^j = 2, 4, 8$. Wavelet coefficients are sparse at fine scales. Large modulus coefficients $|x \star \psi_\lambda(u)|$ are along edges and sharp transitions. The decay of these modulus values when $|\lambda|$ increases depends upon the Lipschitz regularity of x at u [Jaffard, 1991]. The phase $\varphi(x \star \psi_\lambda(u))$ measures the local symmetry of the variations of $x \star \psi_\lambda(u)$ when u moves along the direction of θ. This symmetry is typically preserved along edges. This is why lines of constant phase in Figure 5 follow the geometry of edges when the modulus is non-zero. This observation is at the basis of many edge detection algorithms [Mallat, 2001]. The phase evolution across scales specifies the edge profile.

4. Phase Dependence Across Frequencies

This section characterises phase dependencies across frequencies, from the autocorrelation of $Ux(u, \lambda, \alpha)$. Section 4.1 introduces the autocorrelation and covariance matrices across λ and α integrated over u. Section 4.2 proves Lipschitz continuity properties of the autocorrelation.

4.1 *Phase Harmonic Autocorrelation*

Section 2.2 explains that $x \star \psi_\lambda$ is not correlated with $x \star \psi_{\lambda'}$ if the support of $\widehat{\psi}_\lambda$ and $\widehat{\psi}_{\lambda'}$ do not overlap. The phase filtering introduces correlations between such coefficients which allows us to characterise phase dependencies with autocorrelation matrices.

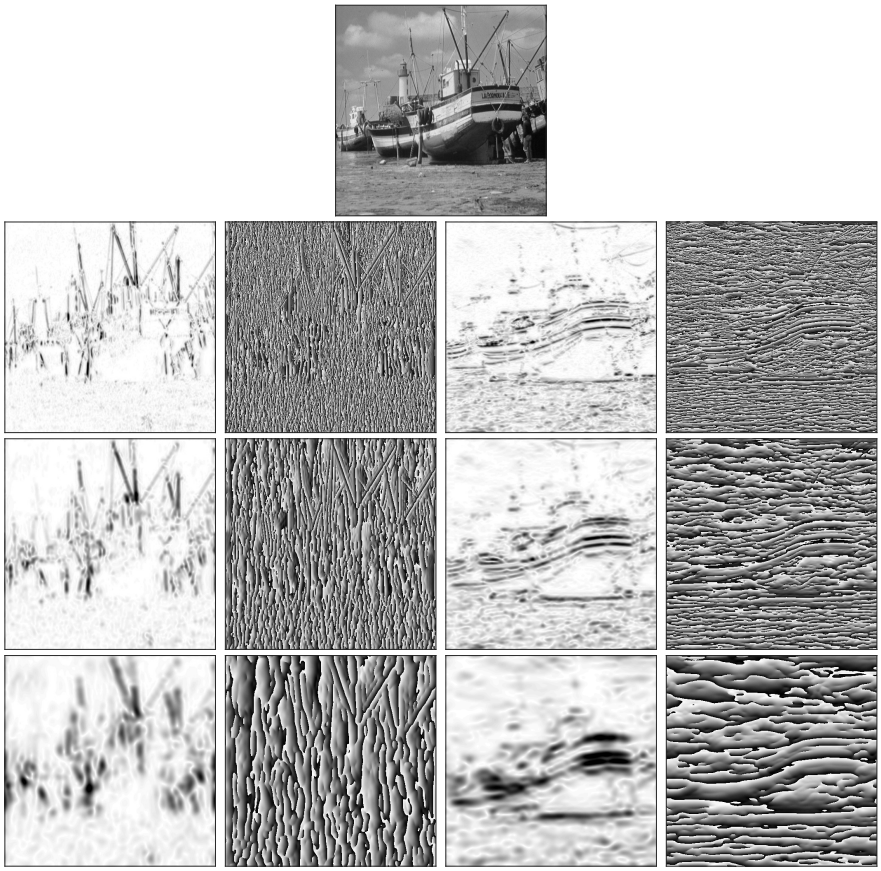


FIG. 5: Top: original image. Columns (a,c) display $|x \star \psi_\lambda(u)|$ at scales $2^j = 2,4,8$ from top to bottom. Small and large coefficients are displayed in white and black respectively. The wavelet orientation angle is $\theta = 0$ for (a) and $\theta = \pi/2$ for (c). Columns (b,d) display the phase images $\varphi(x \star \psi_\lambda)$, with $\theta = 0$ for (b) and $\theta = \pi/2$ for (d).

MEAN VECTOR We compute the mean of $Ux(u,\lambda,\alpha)$ along u for each (λ,α). The mean vector Mx is defined by

$$Mx(\lambda,\alpha) = \int Ux(u,\lambda,\alpha)\,du \tag{4.1}$$

$$= \int |x\star\psi_\lambda(u)|\,h(\alpha-\varphi(x\star\psi_\lambda(u)))\,du. \tag{4.2}$$

Since the Fourier transform of $Ux(u,\lambda,\alpha)$ along α is $\widehat{U}x(u,\lambda,k)=\widehat{h}(k)\,[x\star\psi_\lambda(u)]^{-k}$, the Fourier mean vector is:

$$\widehat{M}x(\lambda,k)=\int \widehat{U}x(u,\lambda,k)\,du=\widehat{h}(k)\int [x\star\psi_\lambda(u)]^{-k}\,du. \tag{4.3}$$

For $k=0$, $[x\star\psi_\lambda]^0=|x\star\psi_\lambda|$ so

$$\widehat{M}x(\lambda,0)=\widehat{h}(0)\,\|x\star\psi_\lambda\|_1.$$

Let us show that $\widehat{M}x(\lambda,k)\ll\widehat{M}x(\lambda,0)$ for $k\neq 0$. Since $\widehat{M}x(\lambda,k)$ is the Fourier transform of $M(\alpha,\lambda)$ along α, this implies that $M(\alpha,\lambda)$ remains nearly constant when α varies.

The integral of $[x\star\psi_\lambda]^k$ is the value of its Fourier transform at the zero frequency. If $k=1$ and $\lambda\neq 0$ then $\widehat{M}x(1,\lambda)=0$ because $\widehat{\psi}_\lambda(0)=0$. For $k>1$ the support of $[x\star\psi_\lambda]^k$ is approximately a dilation of the support of $x\star\psi_\lambda$ so its Fourier transform remains negligible at $\omega=0$. This is illustrated by Figure 3 for a one-dimensional wavelet filter ψ_λ. For one and two-dimensional bump wavelets, numerical computations show that for all $k\geqslant 1$

$$|\widehat{M}x(\lambda,k)|\leqslant\varepsilon\,\widehat{M}x(\lambda,0),$$

with $\varepsilon=0.017$ for the one-dimensional signal of Figure 2 and $\varepsilon=0.019$ for the image in Figure 5.

AUTOCORRELATION AND COVARIANCE The correlation of $Ux(u,\lambda,\alpha)$ and $Ux(u,\lambda',\alpha')$ is computed on average along u:

$$Cx(\lambda,\alpha,\lambda',\alpha') = \int Ux(u,\lambda,\alpha)\,Ux(u,\lambda',\alpha')^*\,du \tag{4.4}$$

$$= \int |x\star\psi_\lambda(u)|\,|x\star\psi_{\lambda'}(u)| \tag{4.5}$$

$$h(\alpha-\varphi(x\star\psi_\lambda(u)))\,h(\alpha'-\varphi(x\star\psi_{\lambda'}(u)))^*\,du.$$

It gives the correlation of $x\star\psi_\lambda(u)$ and $x\star\psi_{\lambda'}(u)$ at positions u where their phases are respectively in the neighborhoods of α and α'. For a fixed (λ,λ'), it is a full matrix which has a regular 2π periodic oscillation along (α,α'). This is shown in Figure 6 for a one-dimensional signal and an image.

The Fourier transform along α is an orthogonal change of basis. We show that it decorrelates a large portion of the coefficients and yields a sparse autocorrelation matrix. The resulting harmonics autocorrelation matrix is

$$\widehat{C}x(\lambda,k,\lambda',k') = \int \widehat{U}x(u,\lambda,k)\,\widehat{U}x(u,\lambda',k')^*\,du$$

$$= \widehat{h}(k)\,\widehat{h}(k')^*\int [x\star\psi_\lambda(u)]^{-k}\,[x\star\psi_{\lambda'}(u)]^{k'}\,du. \tag{4.6}$$

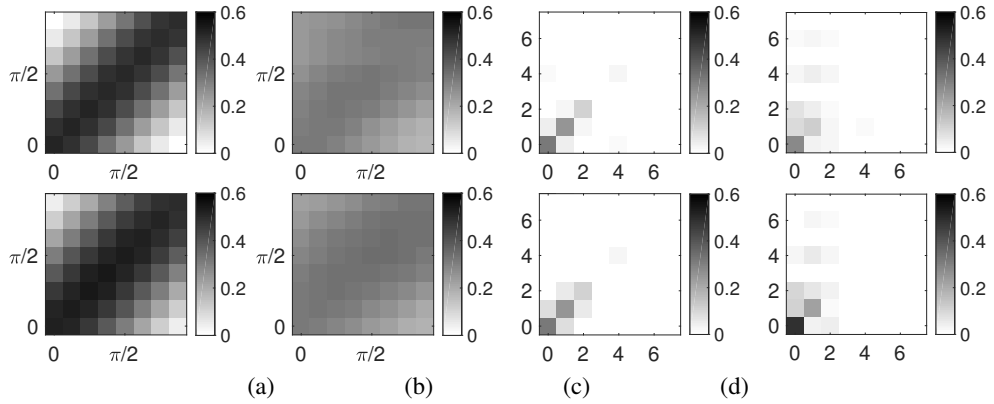


FIG. 6: Autocorrelation matrices Cx and $\widehat{C}x$ for the one-dimensional signal of Figure 2 (first row), and for the image of Figure 5 (second row). (a,b): Matrices of 8^2 coefficients $|Cx(\lambda,\alpha,\lambda',\alpha')|^{1/2}/(\|x\star\psi_\lambda\|\,\|x\star\psi_{\lambda'}\|)^{1/2}$ for $(\alpha,\alpha')\in[0,\pi]^2$, with $\lambda'=\lambda$ in (a) and $\lambda'=4\lambda$ in (b). (c,d): Matrices of 8^2 coefficients $|\widehat{C}x(\lambda,k,\lambda',k')|^{1/2}/(\|x\star\psi_\lambda\|\,\|x\star\psi_{\lambda'}\|)^{1/2}$ for $(k,k')\in[0,7]^2$ and $\lambda'=\lambda$ in (c) and $\lambda'=4\lambda$ in (d).

Diagonal coefficients are proportional to $\mathbf{L^2}$ norms

$$\widehat{C}x(\lambda,k,\lambda,k)=|\widehat{h}(k)|^2\,\|x\star\psi_\lambda\|^2. \tag{4.7}$$

The Plancherel formula applied to (4.6) proves that $\widehat{C}x(\lambda,k,\lambda',k')$ is nearly zero if the Fourier transforms of $[x\star\psi_\lambda(u)]^k$ and $[x\star\psi_{\lambda'}(u)]^{-k'}$ are concentrated over frequency domains which do not intersect. Since their center frequencies are approximately $k\lambda$ and $k'\lambda'$, the coefficients are negligible when $|k\lambda-k'\lambda'|$ is too large. The non-negligible coefficients for $k\lambda\approx k'\lambda'$ capture non-linear correlations across frequency bands.

For wavelet filters, numerical experiments show that the Fourier transform of $[x\star\psi_\lambda]^k$ has most of its energy concentrated in a ball centered in $k\lambda$ of radius $\beta\max(|k|,1)\,|\lambda|$, where β depends upon ψ. These are qualitative results, illustrated by Figure 3, but we have no mathematical characterisation of β. It results that $\widehat{C}x(\lambda,k,\lambda',k')$ is non-negligible if the distance between the center frequencies $k\lambda$ and $k'\lambda'$ satisfies

$$|k\lambda-k'\lambda'|\leqslant\beta\,(\max(|k|,1)\,|\lambda|+\max(|k'|,1)\,|\lambda'|). \tag{4.8}$$

When (k,k') varies for (λ,λ') fixed, (4.8) defines a band of non-negligible coefficients centered at $k\lambda=k'\lambda'$. Beyond this band, $\widehat{C}x(\lambda,k,\lambda',k')$ is nearly zero, which yields a sparse matrix. This property is illustrated by Figure 6 for the one-dimensional signal of Figure 2 and for the image of Figure 5. The first two columns of Figure 6 show that $Cx(\lambda,\alpha,\lambda',\alpha')$ is a full matrix for $\lambda=\lambda'$ and $\lambda'=4\lambda$ whereas the Fourier transform matrices $\widehat{C}x(\lambda,k,\lambda',k')$ are sparse.

The covariance matrix is

$$Kx=Cx-MxMx^*\,,$$

where V^* is the complex transpose of a vector V. This covariance measures the dependence of phases across frequencies at a same position u. In the Fourier phase basis, the first diagonal coefficients of

$\widehat{K}x = \widehat{C}x - \widehat{M}x\widehat{M}x^*$ are

$$\widehat{K}x(\lambda,0,\lambda,0) = |\widehat{h}(0)|^2 \Big(\|x \star \psi_\lambda\|^2 - \|x \star \psi_\lambda\|_1^2 \Big).$$

They measure the sparsity of $x \star \psi_\lambda(u)$, which is large if $\|x \star \psi_\lambda\| / \|x \star \psi_\lambda\|_1$ is small.

4.2 *Lipschitz Continuity for Spectral Norms*

We prove that the mean vector Mx is Lipschitz continuous and that the autocorrelation matrix Cx has a bounded trace and is Lipschitz continuous for the spectral norm. As a result, they define stable signal representations. Section 5 shows that these descriptors can provide accurate signal approximations.

To establish the result for $Mx(\lambda,\alpha) = \int Ux(u,\lambda,\alpha)\,du$, we suppose that x has a support in $[0,L]^d$ and that all convolutions are computed as periodic convolutions over this domain so that $x \star \psi_\lambda \in \mathbf{L^2}([0,L]^d)$. It introduces a normalization factor L^d.

For any $x \in \mathbf{L^2}(\mathbb{R}^d)$, Cx is a positive symmetric matrix whose spectral norm is:

$$\|Cx\|_{2,2} = \sup_{v \neq 0} \frac{\|Cx\,v\|}{\|v\|} \;\text{ with }\; \|v\|^2 = \frac{1}{2\pi} \sum_{\lambda \in \Lambda} \int_0^{2\pi} |v(\alpha,\lambda)|^2 \, d\alpha \;,$$

and

$$Cx\,v(\lambda,\alpha) = \int_0^{2\pi} \sum_{\lambda' \in \Lambda} C(\lambda,\alpha,\lambda',\alpha')\, v(\lambda',\alpha')\, d\alpha'.$$

The following theorem computes Lipschitz bounds for these norms.

THEOREM 4.1 If the filters ψ_λ satisfy (2.25) then for all $(x,x') \in \mathbf{L^2}([0,L]^d)^2$

$$\|Mx - Mx'\| \leqslant \kappa\,(1+\eta)\,L^{d/2}\,\|x - x'\| \;, \tag{4.9}$$

where κ is defined in (2.22). For all $x \in \mathbf{L^2}(\mathbb{R}^d)$

$$(1-\eta)^2\,\|h\|^2\,\|x\|^2 \leqslant \text{Trace}(Cx) \leqslant (1+\eta)^2\,\|h\|^2\,\|x\|^2 \;, \tag{4.10}$$

and all $(x,x') \in \mathbf{L^2}(\mathbb{R}^d)^2$

$$\|Cx - Cx'\|_{2,2} \leqslant \kappa^2\,(1+\eta)^2\,\|x - x'\|\,(\|x\| + \|x'\|). \tag{4.11}$$

Proof: We write the mean Mx in (4.1) in a vector form $Mx = \int Ux(u,.)\,du$, which gives

$$\|Mx - Mx'\|^2 = \Big\| \int (Ux(u,.) - Ux'(u,.))\,du \Big\|^2.$$

If $y \in \mathbf{L^2}([0,L]^d)$ then the Cauchy-Schwartz inequality implies

$$\Big| \int y(u) du \Big|^2 \leqslant L^d \int |y(u)|^2 du.$$

Applied to the vector $Ux(u,.) - Ux'(u,.)$ it gives

$$\|Mx - Mx'\|^2 \leqslant L^d \int \|Ux(u,.) - Ux'(u,.)\|^2\,du.$$

Applying the Lipschitz upper bound (2.30) proves (4.9).

The trace of Cx is computed from its diagonal coefficients in (4.7)

$$\widehat{Cx}(\lambda,k,\lambda,k) = |\widehat{h}(k)|^2 \, \|x \star \psi_\lambda\|^2,$$

and hence

$$\text{Trace}(Cx) = \sum_k |\widehat{h}(k)|^2 \sum_{\lambda \in \Lambda} \|x \star \psi_\lambda\|^2 = \|h\|^2 \, \|Wx\|^2 \, .$$

Applying (2.27) for $x' = 0$ proves (4.10).

The operator Cx is defined in (4.4) as the autocorrelation of Ux which we write in a vector form $Cx = \int Ux(u,.)\, Ux(u,.)^t du$, where $Ux(u,.)^t$ is the transpose of $Ux(u,.)$. We thus verify that

$$\begin{aligned} Cx' - Cx = \int \Big(& Ux'(u,.)(Ux'(u,.) - Ux(u,.))^t + \\ & (Ux'(u,.) - Ux(u,.))\, Ux(u,.)^t \Big) du \end{aligned} \tag{4.12}$$

Any matrix $C(a,b) = \int A(u,a)\, B(u,b)\, du$ has a norm which satisfies

$$\|C\|_{2,2} \leqslant \|A\| \, \|B\| \tag{4.13}$$

with $\|A\|^2 = \int \sum_a |A(u,a)|^2 du$ and $\|B\|^2 = \int \sum_b |B(u,b)|^2 du$. This is verified with the Cauchy Schwartz inequality by showing that for any vectors w and w':

$$|\langle w, Cw' \rangle| \leqslant \|w\| \, \|w'\| \, \|A\| \, \|B\| \, .$$

Applying (4.13) to (4.12) proves that

$$\|Cx' - Cx\|_{2,2} \leqslant \|Ux' - Ux\| \, (\|Ux'\| + \|Ux\|).$$

Since $\|Ux - Ux'\| \leqslant \kappa(1+\eta)\, \|x - x'\|$ and $\|Ux\| \leqslant \|h\|\,(1+\eta)\, \|x\|$ with $\|h\|^2 \leqslant \kappa^2$, we derive (4.11). □

The matrix Cx is a positive symmetric operator so $\|Cx\|_{2,2} \leqslant \text{Trace}(Cx)$. The theorem proves in (4.10) that Cx is a bounded matrix. It also proves in (4.11) that Cx is Lipschitz continuous when x varies in the neighborhood of x'. The theorem results also apply to $\widehat{M}x$ and $\widehat{C}x$ which are obtained from Mx and Cx through a Fourier orthogonal change of basis.

5. Compressive Approximations from Harmonic Correlations

Autocorrelations of one layer neural network coefficients have been studied in [Ustyuzhaninov et al., 2017] to generate stationary image textures having similar perceptual properties as an original texture x. In this section we study the reconstruction of x up to a global translation, as opposed to a perceptually similar signal.

To reconstruct $x(u)$ from its autocorrelation when shifting x along u is equivalent to recover x from its Fourier transform modulus $|\widehat{x}|^2$. This is only possible up to a global translation. It amounts to solve a phase retrieval problem which has been widely studied from mathematical [Akutowicz, 1956] and algorithmic point of views [Candes et al., 2013]. Several authors have shown that this recovery from autocorrelations can be solved with a reduced number of autocorrelation measurements if x is sparse

[Moravec et al., 2007, Shechtman et al., 2011]. The recovery is performed by minimizing an $\mathbf{l^1}$ norm to enforce sparsity.

Next sections study the recovery of x from the mean and autocorrelation of $Ux(u,\lambda,\alpha)$. The autocorrelation is not computed with a shift along u but along (λ,α). It relies on the phase correlation created by harmonics but it is mathematically more complicated because U is non-linear. We introduce a recovery algorithm based on a gradient descent, which also takes advantage of sparsity through $\mathbf{l^1}$ norm conditions. Section 5.2 shows that nearly optimal approximations can be obtained from wavelet harmonic correlations, if the signal has sparse wavelet coefficients. Computations are reproduced by a software in *https://github.com/kymatio/phaseharmonics*.

5.1 *Compressive Recovery*

We introduce a compressive recovery algorithm which computes a signal approximation from a limited set of P harmonic means and correlations. These means and correlations are invariant to translations so the recovery is up to a global translation.

Computations are carried over d-dimensional signals $x(u)$ uniformly sampled over N^d points, with $d=2$ for images. We set $\widehat{h}(k)=1$ over a limited range of k and $\hat{h}(k)=0$ beyond. The mean vector is computed with sums which are normalized by N^d

$$\widehat{M}x(\lambda,k) = N^{-d}\sum_u [x\star\psi_\lambda(u)]^k\,. \tag{5.1}$$

Section 4.1 shows that $\widehat{M}x(\lambda,k)$ is non-negligible only for $k=0$

$$\widehat{M}x(\lambda,0) = N^{-d}\|x\star\psi_\lambda\|_1. \tag{5.2}$$

This $\mathbf{l^1}$ norm specifies the sparsity of $x\star\psi_\lambda(u)$.

The autocorrelation of $\widehat{U}x$ is

$$\widehat{C}x(\lambda,k,\lambda',k') = N^{-d}\sum_u [x\star\psi_\lambda(u)]^{-k}\,[x\star\psi_{\lambda'}(u)]^{k'}. \tag{5.3}$$

Corollary 2.1 proves that $x(u)$ is recovered from $\widehat{U}x(u,\lambda,k)$ with a linear inverse operator. Recovering $\widehat{U}x$ from a limited number of autocorrelation coefficients $\widehat{C}x$ and means $\widehat{M}x$ can be interpreted as a quadratic recovery problem conditioned by $\mathbf{l^1}$ norm sparsity constraints. However, $\widehat{U}x(u,\lambda,k)$ is also correlated along u because the filters $\psi_\lambda(u)$ are regular. Instead of trying to recover $\widehat{U}x$ having such a regularity, we recover x directly.

LOSS MINIMIZATION We want to recover a signal $\tilde{x}$ such that $\widehat{M}\tilde{x}$ and $\widehat{C}\tilde{x}$ are equal to $\widehat{M}x$ and $\widehat{C}x$ over a predefined subset of P coefficients. It is equivalent to match $\widehat{M}x$ and the covariance $\widehat{K}x=\widehat{C}x-\widehat{M}x\widehat{M}x^*$. The loss $\mathscr{E}_P(x,y)$ of an approximation y of x is defined from a discrepancy between the means and covariances of $\widehat{U}y$ and $\widehat{U}x$. In a matrix form it is defined by

$$\mathscr{E}_P(x,y) = \|\widehat{K}y-\widehat{K}x+(\widehat{M}y-\widehat{M}x)\,(\widehat{M}y-\widehat{M}x)^*\|_F^2, \tag{5.4}$$

where $\|A\|_F$ is the Frobenius norm computed over a selected set of P indices (λ,k,λ',k').

We want to find $y=\tilde{x}_P$ which minimizes $\mathscr{E}_P(x,y)$. This is done with a gradient descent initialized with a Gaussian white noise y_0. The gradient descent computes y_{n+1} from y_n with a gradient step on

$\mathscr{E}_P(x,y)$ at y_n. We use an unconstrained gradient descent algorithm L-BFGS. The algorithm stops the minimization with the line-search Wolfe condition [Nocedal & Wright, 2006]. The loss is not convex so the gradient descent may be trapped in local minima. To improve local minima, we compute the gradient descent with 10 random initializations and we keep the solution $\tilde{x}_P = y_n$ having a minimum loss $\mathscr{E}_P(x,y_n)$.

5.2 *Recovery From Wavelet Harmonic Correlations*

We evaluate approximation errors of one and two-dimensional signals from P wavelet harmonic autocorrelations and means. We demonstrate numerically that this representation has compressive approximation properties. It recovers an accurate approximation of signals having sparse wavelet representations from fixed sets of wavelet harmonic correlations. The error decay rate is comparable to sparse non-linear approximations in a wavelet basis, which requires to adjust the choice of wavelet coefficients to each signal. However, wavelet harmonic correlations do not reconstruct signals having wavelet coefficients which are not sufficiently sparse.

CORRELATION SELECTION We first explain how to select P non-negligible wavelet harmonic correlation coefficients, independently from x.

Section 4.1 shows that $\widehat{C}x(\lambda,k,\lambda',k')$ is non-negligible only if $k\lambda$ is sufficiently close to $k'\lambda'$:

$$|k\lambda - k'\lambda'| \leqslant \beta\,(\max(|k|,1)\,|\lambda| + \max(|k'|,1)\,|\lambda'|). \tag{5.5}$$

for some constant β. Since $\widehat{C}x$ is symmetric, we can impose $|\lambda| \geqslant |\lambda'|$. We only keep low-order harmonics by setting $k = 0$ or $k = 1$, and for each λ' we restrict k' to the values which satisfy (5.5).

Let Q be the number of wavelet scales per octave. The wavelet frequencies λ and λ' correspond to scales $2^{j/Q} \sim |\lambda|^{-1}$ and $2^{j'/Q} \sim |\lambda'|^{-1}$. We limit the range Δ of scale interactions by imposing that

$$|j - j'| \leqslant \Delta\, Q. \tag{5.6}$$

In dimension d, a signal of N^d samples has at most $J \leqslant \log_2 N$ dyadic scales 2^j, so $\Delta \leqslant \log_2 N$. In all numerical experiments we set $J = \log_2 N$, and all convolutions are computed with periodic boundary conditionsf.

For one-dimensional signals of size N, the total number P of mean and correlation coefficients which satisfy (5.5) and (5.6) is

$$P \sim \Delta^2\, Q^2 \log_2 N\,. \tag{5.7}$$

For images of N^2 pixels, we set $Q = 1$ but the number of coefficients also depends upon the number L of wavelet angles, and hence

$$P \sim \Delta^2\, L^2 \log_2 N. \tag{5.8}$$

COMPRESSIVE RECOVERY Since $\widehat{C}x$ and $\widehat{M}x$ are invariant to translations, x can only be recovered up to a translation. The approximation error is calculated by translating $\tilde{x}_P$ so that it minimizes $\|x - \tilde{x}_P\|$. We evaluate numerical reconstruction errors from P correlation invariants with the Peak Signal to Noise Ratio (PSNR) in dB

$$\mathrm{PSNR}(x,\tilde{x}_P) = 20\log_{10}\frac{N^{d/2}\,\max_u |x(u)|}{\|x - \tilde{x}_P\|}\,,$$

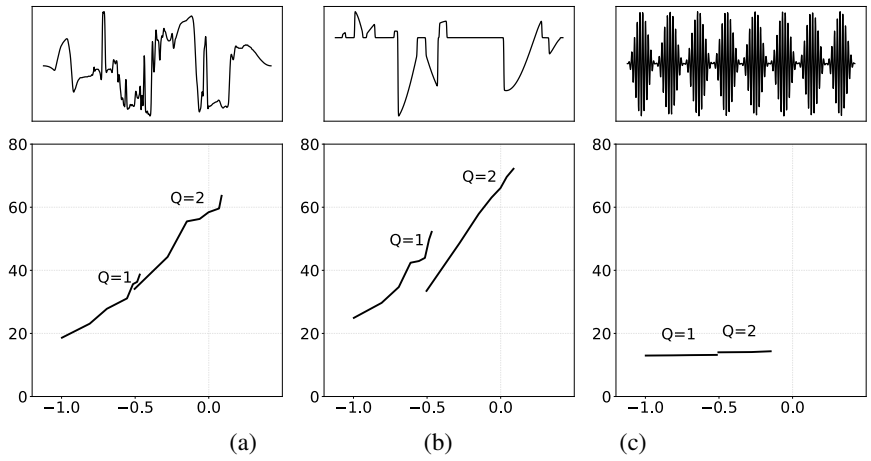


FIG. 7: For each figure, the top graph is the original signal x of $N = 1024$ points. The curve below gives the PSNR of signals $\tilde{x}$ reconstructed from P wavelet harmonic correlations and means, as a function of $\log_{10} P/N$. The first and second parts of each curve correspond to $Q = 1$ and $Q = 2$ respectively, for Δ varying.

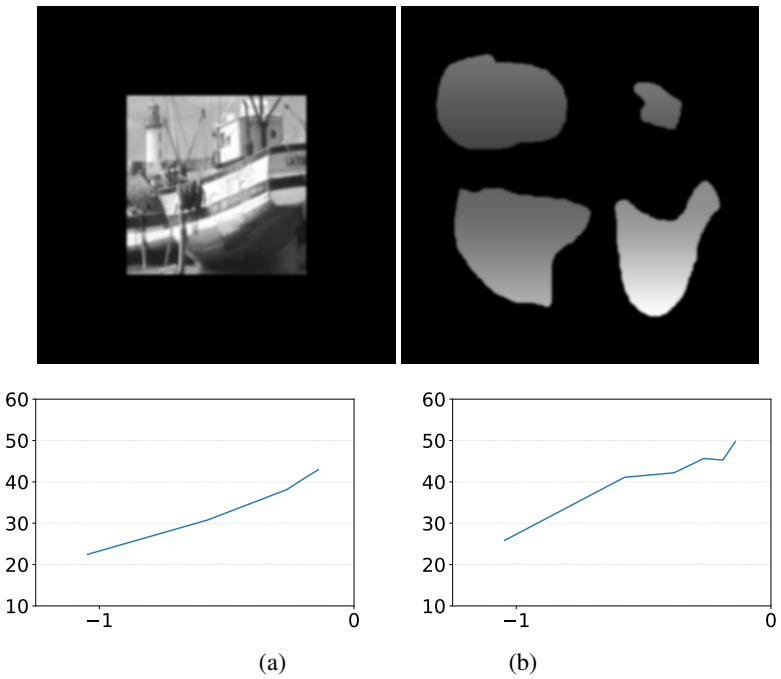
Above 35dB, reconstructed images are visually identical to the original ones, and signal plots superimpose so we do not display reconstructed signals. Figure 7 gives the PSNR error as a function of $\log_{10} P/N$ for one dimensional signals, and Figure 8 as a function of $\log_{10} P/N^2$ for images. The number of coefficients P varies by adjusting the range Δ of scale interactions in (5.7) and (5.8).

Figure 7(a,b) gives two examples of piecewise regular one-dimensional signals x having sparse wavelet coefficients. Large coefficients are located at sharp transitions. The PSNR curves show that when $P = N$, $\tilde{x}_P$ is above 60dB which corresponds to a relative error below 10^{-3}. The approximation error has a decay

$$\|x - \tilde{x}_P\| \leqslant C P^{-\chi}, \tag{5.9}$$

with $\chi \approx 2$. This is the same approximation rate as the one obtained by a non-linear adaptive approximation of such signals in a wavelet orthonormal basis. Non-linear approximations take advantage of sparsity by selecting the P largest wavelet coefficients of x, whose locations thus depend upon x [Mallat, 2001]. The decay rate $\chi = 2$ is the best one obtained over the class of signals which may be discontinuous but have a bounded total variation. In this case, the same error decay rate is obtained over fixed sets of P wavelet harmonic correlations, which do not depend upon x as in adaptive approximations. Correlations are averaged over all spatial positions which is why the choice of correlation coefficients is not adapted to the positions of non-zero wavelet coefficients. These non-linear approximation properties over fixed sets of measurements are similar to the ones obtained by compressive sensing algorithms [Candes et al., 2006]. However, the mathematical setting is more difficult because the linear sensing operator is replaced by non-linear harmonic autocorrelations.

Similarly to compressive sensing, if wavelet coefficients are not sufficiently sparse then the signal is



(a) (b)

FIG. 8: Each original image x shown at the top has $N^2 = 256^2$ pixels. The curve below gives the PSNR of signals $\tilde{x}$ reconstructed from P wavelet harmonic correlations and mean coefficients, as a function of $\log_{10} P/N^2$.

not reconstructed from wavelet harmonic correlations. The signal in Figure 7(c) is a dramatic example:

$$x(u) = (1 - \cos(\nu u)) \cos(\lambda u) \quad \text{with} \quad \nu \ll \lambda \ . \tag{5.10}$$

The reconstruction algorithm is unable to recover an approximation even when the number P of correlation coefficients reaches the signal size N. In this case the cosine of high frequency λ creates many non-zero wavelet coefficients $x \star \psi_\lambda(u)$. The cosine amplitude is modulated by another cosine of much lower frequency ν. The phase dependence of such frequencies need to be captured by harmonic coefficients providing the correlation of $x \star \psi_\lambda$ with $[x \star \psi_\nu]^k$ for $k = \lambda/\nu$. However we cannot recover the phase dependence of the two frequency components because $x \star \psi_\nu = 0$. When signals have localized sharp transitions, large wavelet coefficients propagate across scales as shown in Figure 2. At a location u, if $x \star \psi_{\lambda'}(u)$ and $x \star \psi_\lambda(u)$ are not zero then harmonic correlation coefficients capture the phase dependencies and can thus reconstruct the signal. This is the case for the two signals in Figure 7(a,b).

Figure 8 shows results similar to Figure 7, for images of N^2 pixels. These images have sparse wavelet coefficients located near edges. The reconstruction $\tilde{x}_P$ has a PSNR well above 40dB and hence a relative precision of 10^{-2} when P is close to N^2, for the natural image and the piecewise regular cartoon image. The error $\|x - \tilde{x}_P\|$ decays when increasing the number P of correlation terms, with an exponent $\chi \approx 1$ in (5.9). Again, this is the same approximation rate as the one obtained by a non-

linear adaptive approximation in a wavelet orthonormal basis, where the P largest wavelet coefficients are selected depending upon the image. This decay exponent is the optimal approximation rate for the class of images having a bounded total variation [Mallat, 2001]. However, the choice of the P harmonic wavelet correlations does not depend upon the image, as in compressive sensing approximation.

Similarly to the one-dimensional case, images whose wavelet coefficients are not sufficiently sparse cannot be reconstructed from wavelet phase harmonic correlations. One may construct such images with two-dimensional high frequency sinusoidal waves as in the one-dimensional example (5.10). Realizations of ergodic stationary processes are other counter examples.

ERGODICITY VERSUS COMPRESSIVE APPROXIMATIONS Realizations of ergodic stationary processes cannot be recovered from a limited number of wavelet harmonic correlations. This observation relates the signal recovery problem to approximations of stationary processes [Bruna & Mallat, 2019] and image texture synthesis [Simoncelli & Freeman, 1995, Ustyuzhaninov et al., 2017].

If x is a realization of an ergodic stationary process X and the domain size N^d is sufficiently large then the spatial empirical means (5.1) and correlations (5.3) provide accurate estimations of expected means $\mathbb{E}\left(\widehat{U}X(u,\lambda,k)\right)$ and correlations $\mathbb{E}\left(\widehat{U}X(u,\lambda,k)\,\widehat{U}X(u,\lambda,k)^*\right)$, computed relatively to the probability distribution of X. If x and x' are two realizations of X then $\|x-x'\|$ is typically large but they have nearly the same empirical means (5.1) and correlations (5.3). It results that if x is a realization of an ergodic process over a large domain size N^d then its recovery from $\widehat{M}x$ and $\widehat{C}x$ is not stable.

In this unstable ergodic regime, a gradient descent on the loss (5.4) will reconstruct different signals for different Gaussian white noise initializations. Such reconstructions are realizations of a random process whose probability measure was obtained by transporting the uniform Gaussian white noise measure with the gradient descent on the loss. One can prove that it defines a stationary microcanonical process, conditioned by the empirical means and autocorrelation coefficients $\widehat{M}x$ and $\widehat{C}x$ [Bruna & Mallat, 2019]. Texture synthesis algorithms from deep convolutional network coefficients follow this principle [Gatys et al., 2015, Ustyuzhaninov et al., 2017]. The properties of the stationary processes obtained with wavelet phase harmonic correlations are studied in [Zhang & Mallat, 2019].

This analysis shows that we must distinguish two cases. If x is a signal whose wavelet coefficients are sufficiently sparse compared to the total size N^d then it may be recovered from a relatively small number of wavelet harmonic correlations. On the contrary, we cannot recover x if we are in an ergodic regime where $\widehat{M}x$ and $\widehat{C}x$ are close approximations of empirical means and correlations of a stationary ergodic random process X. In this case, the reconstruction algorithm initialized over Gaussian white noise defines a stochastic model of X [Zhang & Mallat, 2019].

Acknowledgment

This work was supported by the ERC InvariantClass 320959.

REFERENCES

Akutowicz, E. J. (1956) On the determination of the phase of a Fourier integral. *Trans. of the American Mathematical Society*, **83**(1), 179–192.

Bruna, J. & Mallat, S. (2019) Multiscale Sparse Microcanonical Models. *arXiv:1801.02013*.

Candes, E., Romberg, J. & Tao, T. (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, **52**(2), 489–509.

Candes, E., Strohmer, T. & Voroninski (2013) Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, **66**, 1241–1271.

Gatys, L., Ecker, A. & Bethge, M. (2015) Texture Synthesis Using Convolutional Neural Networks. In Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M. & Garnett, R., editors, *Advances in Neural Information Processing Systems 28*, pages 262–270. Curran Associates, Inc.

Grossmann, A., Kronland-Martinet, R. & Morlet, J. (1989) Reading and understanding continuous wavelet transforms. In Combes, J., editor, *Wavelets, time-frequency representations and phase space*. Springer, Berlin.

Jaffard, S. (1991) Pointwise smoothness, two-microlocalisation and wavelet coefficients. *Publications Matematiques*, **35**, 155–168.

Jamshidi, A. & Kirby, M. J. (2006) Examples of Compactly Supported Functions for Radial Basis Approximations. In *International Conference on Machine Learning*.

Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012) ImageNet Classification with Deep Convolutional Neural Networks. In *Proc. of NIPS*, pages 1106–1114.

LeCun, Y., Bengio, Y. & Hinton, G. E. (2015) Deep learning. *Nature*, **521**(7553), 436–444.

Luo, Y. & Mesgarani, N. (2018) TaSNet: Time-Domain Audio Separation Network for Real-Time, Single-Channel Speech Separation. In *Proc. ICASSSP*, pages 696–700.

Mallat, S. (2001) *A Wavelet Tour of Signal Processing: The Sparse Way, 3rd Edition*. Academic Press.

Mallat, S. (2016) Understanding Deep Convolutional Networks. *Phil. Trans. of Royal Society A*, **374**(2065).

Moravec, M., Romberg, J. & Baraniuk, R. (2007) Compressive phase retrieval. In *SPIE International Symposium on Optical Science and Technology*.

Nocedal, J. & Wright, S. J. (2006) *Numerical Optimization*. Springer, New York, NY, USA, second edition.

Portilla, J. & Simoncelli, E. P. (2000) A Parametric Texture Model based on Joint Statistics of Complex Wavelet Coefficients. *International Journal of Computer Vision*, **40**, 49–71.

Shechtman, Y., Eldar, Y. C., Szameit, A. & Segev, M. (2011) Sparsity based sub-wavelength imaging with partially incoherent light via quadratic compressed sensing. *Opt. Express*, **19**, 14808–14822.

Simoncelli, E. P. & Freeman, W. T. (1995) The Steerable Pyramid: A Flexible Architecture for Multi-Scale Derivative Computation. In *IEEE Int. Conf. on Image Processing*, pages 444–447.

Unser, M., Chenouard, N. & Van De Ville, D. (2011) Steerable Pyramids and Tight Wavelet Frames in $L^2(R^d)$. *IEEE Transactions on Image Processing*, **20**(10), 2705–2721.

Ustyuzhaninov, I., Brendel, W., Gatys, L. & Bethge, M. (2017) What does it take to generate natural textures?. In *International Conference on Learning Representations*.

Zhang, S. & Mallat, S. (2019) Wavelet Phase Harmonic Covariance Models of Stationary Processes. *submitted to Jour. of Pure and Applied Harmonic Analysis*.