New Ways of Data Compression and Noiseless Recovery

Anahita Bayani
Registration Number: 613236
Supervisor: Professor Michele Pavon
University of Padua
Faculty of Engineering
Bachelor Degree in Information Engineering

23.09.13

Contents

Chapter 1

Introduction

A large variety of engineering disciplines have huge advantages in studying signals in their frequency domain, rather than the classic time domain.

The significant contribution in this direction has been provided by the Fourier transform and thus its decomposing signals into theirs costituent frequencies. Though being proposed by French matematician Joseph Fourier as early as 1811, the Fourier transform began to fulfil its real and vast potential only over a century later and especially thanks to the discrete Fourier transform, or DFT, which, alongside the plain old Fourier transform and the Fourier series, is part of the three varieties in which the Fourier transform comes. Particularly, what makes the DFT so powerful is the fast Fourier transform, the algorithm first described by James Cooley and John Tukey in 1965 and used to easily calculate DFTs on computers.

In this sense, a very new and exciting theory emerging only in the last few years has seen a flurry of researches developing tools towards a suggestive yet realistic reduction in time and power costs of many imaging and signal processing applications: Compressed Sensing, or CS. The innovative approach of CS seems to put the entire signal processing theory well beyond the renowned Shannon sampling theorem.

Moreover, considering that every data acquisition inevitably comes with noise, a neat recovery of the original data is of great interest in many practical applications.

In this regard, this thesis presents a brief introduction to the Fourier analysis and offers basic insight into CS theory, followed by an actual approach of denoising data applied to MRI acquisitions.

Chapter 2

Basic Concepts on Fourier Analysis

Let us begin providing useful definitions and notions on Fourier transforms and signal processing.

2.1 Continuous-Time Fourier Analysis

2.1.1 Periodic Signals

Recalling a periodic signal definition, that is

$$x(t) = x(t+T), \qquad \text{for all } t \tag{2.1}$$

for some positive, nonzero value of T, with T_0 the fundamental period of $x(t)$, that is the smallest positive value of T for which eq. (2.1) holds, and $2\pi/T_0$ the fundamental frequency, let us consider the periodic sinusoid signal

$$x(t) = \cos \omega_0 t \tag{2.2}$$

and the periodic complex exponential

$$x(t) = e^{j\omega_0 t}\,. \tag{2.3}$$

Relying on the above signals is particularly advantageous in the light of the immediate and useful properties that convolutional systems yield, expressly in terms of the frequency response, given the simple yet powerful way with which it affects the involving signals. Indeed, considering an input signal $x(t)$ and the output signal $y(t)$ resulting by the convolution of input and impulsive response $h(t)$, with $h(t) \in \ell^1$:

$$x(t) = e^{j\omega_0 t} \longrightarrow y(t) = H(j\omega_0)e^{j\omega_0 t}, \tag{2.4}$$

with $H(j\omega) = \int_{-\infty}^{\infty} h(\tau)e^{-j\omega\tau}\,d\tau$ being the frequency response of the system.

It is therefore interesting to analyse linear combinations of exponential signals, taking into account periodic signals firstly, in the form of the set of *harmonically related* complex exponentials

$$\phi_{\mathrm{k}}(t) = e^{jk\omega_0 t}, \qquad k = 0, \pm 1, \pm 2, \ldots \tag{2.5}$$

periodic with period T_0. Then, the following signal is also periodic with period T_0:

$$x(t) = \sum_{k=N_1}^{N_2} a_k \phi_k(t) = \sum_{k=N_1}^{N_2} a_k e^{jk\omega_0 t}. \tag{2.6}$$

In this regard, we refer to the components for $k = +n$ and $k = -n$ as the nth harmonic components.

We therefore determine the *Fourier series coefficients* through equation

$$a_k = \frac{1}{T_0} \int_{T_0} x(t) e^{-jk\omega_0 t}\, dt\,. \tag{2.7}$$

Remark 1. Energy over the period of the signal $x(t)$ highlights the preserving of geometrical properties:

$$E(T) = \int_{T_0} |x(t)|^2\, dt = T_0 \sum_{k=N_1}^{N_2} |a_k| < \infty \tag{2.8}$$

Remark 2. Eq.(2.7) holds for every periodic signal x of period T absolutely integrable over the period: $\int_0^T |x(t)|\, dt < \infty$.

In order to have a clear view of which signals are representable in the form of $x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}$, the following theorem is introduced:

Theorem 2.1.1 (Riesz-Fisher Theorem)**.** *Let $\{a_k\}_{k=-\infty}^{\infty}$ be a sequence of complex numbers for which*

$$\sum_{k=-\infty}^{\infty} |a_k|^2 < \infty \qquad (a = \{a_k\}_{k=-\infty}^{\infty} \in \ell^2) \tag{2.9}$$

holds, that is $\{a_k\}_{k=-\infty}^{\infty} \in \ell^2$, then there exist a signal x over $[0, T]$ of finite energy over $[0, T]$ such that eq. (2.7) *holds.*

Conversely, given a signal x of finite energy over $[0, T]$, its Fourier coefficients are well defined by (2.7) *where condition of* (2.9) *also holds.*

Furthermore, defining $s_N(t) = \sum_{k=-N}^{N} a_k e^{jk\omega_0 t}$, then s_N shows mean-square convergence *(in L^2)*

$$\lim_{N\to\infty} \int_0^T |x(t) - s_N(t)|^2\, dt \; = 0 \tag{2.10}$$

To reconcile absolute integrability of the signal as in Remark 2 to finite energy requested by Theorem 2.1.1, it is recalled that $L^2(0,T) \subseteq L^1(0,T)$.

Theorem 2.1.1 does not explicitly require periodic signals, however the following Remark applies:

Remark 3. Replacing x with its periodic repetition $rep_T x(t)$, Theorem 2.1.1 has validity on $\mathbb{R}$.

Therefore, under the Riesz-Fisher Theorem conditions, the Fourier series is defined as follows:

Definition 2.1 (Fourier Series)**.** The Fourier series of a periodic and absolutely convergent signal over a period, that is $\int_{T_0} |x(t)|\, dt < \infty$, is a representation of the signal in the form of

$$x(t) = \sum_{k=-\infty}^{+\infty} a_k e^{jk\omega_0 t}, \tag{2.11}$$

accordingly to Theorem 2.1.1, where the limit is in the mean-square sense.

Furthermore, a pointwise convergence to their Fourier series of continuous periodic signals is guaranteed by the so-called *Dirichlet Conditions*, that basically cover almost all the interesting signals. That is, under the following conditions, signals are representable by Fourier series:

1. $x(t)$ must be absolutely integrable over any period:

$$\int_{T_0} |x(t)|\, dt \, < \infty\,; \tag{2.12}$$

2. $x(t)$ has at most a finite number of maxima and minima during any single period;

3. in any finite interval of time, discontinuities, which must be finite, are at most in a finite number.

Remark 4. Dirichlet Conditions also imply the square-integrability over a period of the signal:

$$\int_{T_0} |x(t)|^2\, dt < \infty\,. \tag{2.13}$$

Thus, the Fourier series of continuous periodic signals converges and equals the original signal at every value of t, whereas periodic signals with discontinuities have their equal Fourier series representation everywhere except at the isolated points of discontinuity at which the convergence is to the average value of the signal left- and right-hand side limits.

2.1.2 Aperiodic Signals

As per aperiodic signals, we can always suit the above theories, constructing a periodic signal $\tilde{x}(t)$ from the aperiodic signal $x(t)$ of finite duration, $x(t)$ being one period for $\tilde{x}(t)$. This leads to rewriting and extending eqs. (2.11) and (2.7) as follows:

Definition 2.2 (Fourier Transform)**.** We define the Fourier Transform of an aperiodic signal $x(t)$ of finite energy ($\int_{-\infty}^{\infty} |x(t)|^2\, dt < \infty$) as

$$X(j\omega) = \int_{-\infty}^{+\infty} x(t) e^{-j\omega t}\, dt, \tag{2.14}$$

where the limit is in the mean-square sense.

and the

Definition 2.3 (Inverse Fourier Transform)**.**

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(j\omega) e^{j\omega t}\, d\omega\,, \tag{2.15}$$

again in a mean-square sense of the limit.

Pointwise convergence is also guaranteed as per the Dirichlet conditions.

2.2 Discrete-Time Fourier Analysis

Analogously to the above dealing with continuous-time signals, here we handle the same theory for discrete-time signals.

2.2.1 Periodic Signals

Let us consider a periodic discrete-time signal:

$$x[n] = x[n+N] \qquad \text{for some positive } N \tag{2.16}$$

and the set of all discrete-time complex exponential signals periodic with period N, that is

$$\phi_k(n) = e^{jk\frac{2\pi}{N}n}\,, \qquad k = 0, \pm 1, \pm 2, \ldots \tag{2.17}$$

that are all *harmonically related* having frequencies multiples of the same fundamental frequency $2\pi/N$.

Yet it is noteworthy to observe that, contrary to continuous-time signals, recalling that discrete-time complex exponentials differing in frequency by a multiple of 2π are identical, there are only N different signals in the set given in eq (2.17). Therefore, dealing with a more general representation

including linear combinations of the sequences $\phi_k[n]$ in eq (2.17), letting k vary over a range of N consecutive integers, we have:

$$x[n] = \sum_{k=\langle N \rangle} a_k \phi_k[n] = \sum_{k=\langle N \rangle} a_k e^{jk\frac{2\pi}{N}n} \,. \tag{2.18}$$

Definition 2.4 (Discrete-Time Fourier Series)**.** The discrete-time Fourier series of a periodic signal is a representation of the signal in the form of eq. (2.18).

We therefore determine the *Fourier series coefficients* through the equation:

$$a_k = \frac{1}{N} \sum_{n=\langle N \rangle} x[n] e^{-jk\frac{2\pi}{N}n} \tag{2.19}$$

2.2.2 Aperiodic Signals

As previously, an aperiodic sequence $x[n]$ of finite duration lets us to consider the periodic sequence $\tilde{x}[n]$ for which $x[n]$ is a single period. Then, analysing the Fourier representation of $\tilde{x}[n]$ and conveniently adapting it to $x[n]$, we are able to define the discrete-time counterparts of eqs. (2.14) and (2.15):

Definition 2.5 (Discrete-Time Fourier Transform)**.** We define the Discrete-Time Fourier Transform (DTFT) of an aperiodic signal $x[n]$ of finite energy ($\sum_{n=-\infty}^{+\infty} |x[n]|^2 < \infty$) as the mean-square limit

$$X(e^{j\theta}) = \sum_{n=-\infty}^{+\infty} x[n] e^{-j\theta n} \tag{2.20}$$

and the

Definition 2.6 (Inverse Discrete-Fourier Transform)**.**

$$x[n] = \frac{1}{2\pi} \int_{2\pi} X(e^{j\theta}) e^{j\theta n} \, d\theta \,. \tag{2.21}$$

2.2.3 Discrete Fourier Transform (DFT)

Dealing with discrete-time signals allows us to consider yet another analysis of the Fourier theory that leads to the so called Discrete Fourier Transform (DFT). This technique retains a particular importance in the light of the very efficient algorithm used for its calculation, that is the Fast Fourier Transform (FFT), which has helped the extensive development of discrete-time signals analysis.

Namely, as before, we consider a signal of finite length $x[n]$ and construct the periodic signal $\tilde{x}[n]$ for which $x[n]$ is one period. Then, from the Fourier series coefficients of $\tilde{x}[n]$ over the interval $\tilde{x}[n] = x[n]$ we have:

Definition 2.7 (Discrete Fourier Transform (DFT))**.** We denote with $\tilde{X}(k)$ the DFT of $x[n]$ calculated as

$$\tilde{X}(k) = a_k = \frac{1}{N}\sum_{n=0}^{N-1} x[n]e^{-jk\frac{2\pi}{N}n}\,, \qquad k = 0, 1, \ldots, N-1\,. \tag{2.22}$$

Remark 5. The Discrete Time Fourier Transform (DTFT) of the considered signal yields:

$$X(e^{j\theta}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\theta n} = \sum_{n=0}^{N-1} x[n]e^{-j\theta n}, \tag{2.23}$$

thus resulting in $Na_k = X(e^{jk\frac{2\pi}{N}})$ being samples of $X(e^{j\theta})$.

We can therefore observe that the finite duration signal can be seen both as specified by the finite set of nonzero values it assumes or by the finite set of values of $\tilde{X}(k)$ in its DFT.

2.3 Fourier Transform Properties

Several properties descend directly from the Fourier transforms, hence offering useful relationships between time and frequency domains of signals that are easily extended to corresponding properties for Fourier series too and facilitating computational complexity of Fourier transforms and inverse transforms.

Accordingly, follows a brief overview of the most common and important properties of signals ant their Fourier transforms through the Fourier series coefficients, here presented for continuous-time signals, but straightforwardly extending to discrete-time signals:

- Isometry: it preserves the ℓ^2 norm (meaning conservation of *distance* and *ortogonality*);
- Linearity:
$$\alpha x_1(t) + \beta x_2(t) \overset{\mathcal{F}}{\longleftrightarrow} \alpha a_k + \beta b_k;$$
- Time-Shifting:
$$x(t - t_0) \overset{\mathcal{F}}{\longleftrightarrow} a_k e^{-jk\omega_0 t_0};$$
- Frequency-Shifting:
$$x(t)e^{jM(2\pi/T)t} \overset{\mathcal{F}}{\longleftrightarrow} a_{k-M};$$
- Conjugation:
$$x^*(t) \overset{\mathcal{F}}{\longleftrightarrow} a^*_{-k};$$

- Time Reversal:
$$x(-t) \overset{\mathcal{F}}{\longleftrightarrow} a_{-k};$$
- Time Scaling:
$$x(\alpha t), (\alpha > 0, \text{periodic with period } T/\alpha) \overset{\mathcal{F}}{\longleftrightarrow} a_k;$$
- Periodic Convolution:
$$\int_T x(\tau) y(t-\tau)\, d\tau \overset{\mathcal{F}}{\longleftrightarrow} T a_k b_k;$$
- Multiplication:
$$x(t)y(t) \overset{\mathcal{F}}{\longleftrightarrow} \sum_{l=-\infty}^{+\infty} a_l b_{k-l};$$
- Symmetry:
  - $x(t)$ real-valued $\Leftrightarrow a_{-k} = a_k^*$;
  - $x(t)$ even $(x(-t) = x(t)) \Rightarrow \{a_k\}$ even $(a_{-k} = a_k)$;
  - $x(t)$ odd $(x(-t) = -x(t)) \Rightarrow \{a_k\}$ odd $(a_{-k} = -a_k)$;
  - $x(t)$ real-valued and even $\Leftrightarrow \{a_k\}$ real-valued and even;
  - $x(t)$ real-valued and odd $\Leftrightarrow \{a_k\}$ purely imaginary and odd.

Furthermore, confronting Fourier transform and inverse transform expressions, a neat *duality* emerges (though duality for discrete-time signals is to be traced in the Fourier series eqs. (2.18) and (2.19)).

Also of fundamental importance are *modulation* and *convolution* properties, alongside the powerful *Plancherel Theorem*, which offer an insight into the strong and effective means that Fourier analysis delivers.

Chapter 3

Compressed Sensing: Beyond the Sampling Theorem

The impressive and dramatic changes that Fourier analysis has brought since its introduction have recently further developed into new studies and applications based on **Compressed Sensing**, or CS. Particularly, CS offers an innovative approach to sampling signals, leading to results well beyond the ones of the Shannon Sampling Theorem.

To truly appreciate this new theory, let we firstly recall the Shannon Sampling Theory:

Theorem 3.0.1 (Shannon Sampling Theorem)**.** *Let $x(t)$ be a bandlimited signal, with its Fourier transform $X(j\omega) = 0$ for all $|\omega| > \omega_M$. Then, denoting with $\{x(nT), n \in \mathbb{Z}\}$ its samples, with*

$$\omega_s = \frac{2\pi}{T} \tag{3.1}$$

if

$$\omega_s > 2\omega_M \tag{3.2}$$

the signal is uniquely determined in a mean-squared sense by its samples.

In this regard, eqs. (3.1) and (3.2) are respectively defined as the *Nyquist frequency* and the *Nyquist rate*.

Therefore, according to the sampling theorem, in order to avoid aliasing and to reconstruct the exact signal from its samples, the sampling rate must be at least twice the Nyquist rate. Hence, to afford sampling at lower rates than the Nyquist rate and still recovering the exact signal, we cannot rely on the sampling theorem and it is in this sense that CS opens up to new strategies of sampling, allowing to reconstruct signals from even fewer samples.

3.1 CS Signal Reconstruction

The new sampling approach introduced by CS, is mainly based upon two principles: **Sparsity** of the signal and **Incoherence** of its sensing.

Namely, the sparsity of a signal refers to its concise representation in a suitable basis, whereas incoherence, on the other hand, expresses the dual spread of the signal in its frequency domain. The immediate convenience of this setting is that natural signals are actually sparse.

Therefore, in this frame, given a signal $f(t)$, we are interested in its *information* in the form of its correlation with waveforms $\phi_k(t)$, that is recording the values $f(t)$ is obtained by:

$$y_k = \langle f, \phi_k \rangle, \qquad k = 1, \dots, m. \tag{3.3}$$

Dealing with CS, specifically considering discrete signals $f \in \mathbb{R}^n$, the attention comes to cases of $m < n$, available measurements fewer than signal dimension. It is especially in this sense that the strength of CS emerges: we pose the problem of recovering $f \in \mathbb{R}^n$ from $y \in \mathbb{R}^m$, with $n > m$ through the *sensing matrix* $A_{m,n}$.

Accordingly, let us consider the vector $f \in \mathbb{R}^n$ and its representation in an orthonormal basis $\Psi = [\psi_1, \dots, \psi_n]$ through its coefficients $x_i = \langle f, \psi_i \rangle$, $i = 1, \dots, n$:

$$f(t) = \sum_{i=1}^{n} x_i \psi_i(t) \tag{3.4}$$

that is $f = \Psi x$, with Ψ the $n \times n$ matrix having ψ_i as columns.

Therefore, as per the above tenets, signal sparsity allows discarding small coefficients without appreciable loss.

Definition 3.1 (S-sparse Vector)**.** Let $f_S(t)$ be the result of (3.4) setting to zero all but the S largest values of x_i: $f_S := \Psi x_S$. We define S-sparse a vector with at most S nonzero entries.

Recalling the Φ basis used to sense f in (3.3), we can define a new parameter involving both the orthonormal bases Φ, Ψ of $\mathbb{R}^n$:

Definition 3.2 (Coherence)**.** The coherence between matrices Φ and Ψ is defined as the largest correlation among any two of their respective elements:

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max_{1<k,h<n} |\langle \phi_k, \psi_h \rangle|. \tag{3.5}$$

Remark 6. Definition (3.5) sets $\mu(\Phi, \Psi) \in [1, \sqrt{n}]$, meaning that the more correlated elements Φ and Ψ have, the larger their coherence is.

As per CS, the aim is to deal with pairs of orthonormal bases as little coherent as possible. In this sense, an interesting occurrence is the one of the canonical basis Φ to sense f, opposed to the Fourier basis Ψ to represent

f, respectively $\phi_k(t) = \delta(t-k)$ and $\psi_h(t) = \frac{1}{\sqrt{n}} e^{j2\pi ht/n}$, where $\mu(\Phi, \Psi) = 1$, underlying the maximal incoherence between spikes of the time domain and sinusoids of the frequency domain.

A significant feature of sparsity in CS comes from the large incoherence random matrices present: precisely we can state that $\mu(\Phi, \Psi) \simeq \sqrt{2 \lg n}$ with high probability. Relying on orthonormal bases to compute f through its coefficients x_i guarantees a negligible error in terms of the ℓ_2 norm:

$$\|f - f_S\|_{\ell_2} = \|x - x_S\|_{\ell_2}. \tag{3.6}$$

Furthermore, handling with values from (3.3) (only a subset $m < n$ of all the available), let us recover the signal f by ℓ_1 norm minimization (convex optimization):

$$y_k = \langle \phi_k, \Psi \tilde{x} \rangle, \quad k = 1, \ldots, m \qquad \text{selecting} \quad \min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_{\ell_1}. \tag{3.7}$$

In this regard, the following theorem sets a new CS-based sampling:

Theorem 3.1.1 (Incoherent Sampling Theorem)**.** *Let x be the S-sparse representation of $f \in \mathbb{R}^n$ in the basis Ψ, collecting a subset of m values in the Φ domain uniformly at random. Then if*

$$m \geq C \cdot \mu^2(\Phi, \Psi) \cdot S \cdot \lg n, \qquad C \in \mathbb{R}^+ \tag{3.8}$$

the convex optimization of (3.7) *is exact with overwhelming probability.*

Theorem 3.1.1 highlights the importance of low coherence, which directly determines how few samples are needed in order to perfectly recover signal f. Particularly, the lower endpoint of coherence, by its definition in eq. (3.5), conveys to only $S \lg n$ samples needed, instead of n. Moreover, only sparsity of the interested signal plays a key role, while all the information on the nonzero x_i coefficients could well be unknown a priori.

3.2 Incoherent counter to Linear Sampling

A straightforward comparison of the two sampling Theorems 3.0.1 and 3.1.1 exposes some of the advantages of CS:

- Shannon Sampling Theorem 3.0.1
  - the support of f in the frequency domain is a known connected set of size S,
  - exact signal reconstruction from S uniformly distributed samples of the time domain,
  - linear reconstruction by sinc interpolation;

- Incoherent Sampling Theorem 3.1.1
  - the support of f in the frequency domain is an *arbitrary and unknown* set of size S,
  - exact signal reconstruction from $\sim S \lg n$ almost arbitrarily distributed samples,
  - nonlinear reconstruction by convex programming.

Of appreciable remark for CS is the independence between measurement techniques and the signals involved. Further, CS proposes sampling at low rates in the sense that the relevant information is collected from just the few sensors computing f_S.

Finally, figure 3.1 recalls the CS-based methods to recover signals.

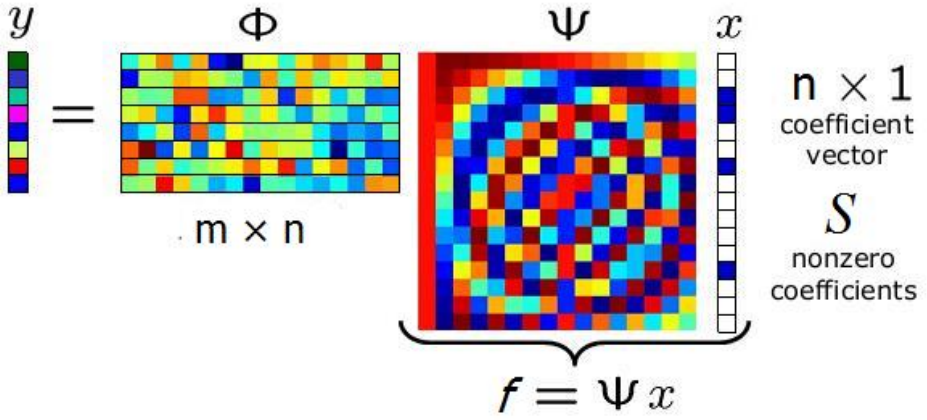

Figure 3.1: CS paradigm guarantees S-term quality recovery from measurements in the order of just $\sim S \lg(n)$.

Straightforwardly from the above theories, different and various applications emerge, each providing interesting and challenging views on new CS-based developments. Specifically, data acquisition and compression and channel coding are among the discussed and studied subjects. In this scenario, another possible unfolding attains to the so called **inverse problems**, involving fixed sensing environments but still recoverable signals by virtue of their sparse representation. Particularly, this is the case of Magnetic Resonance Imaging (MRI) which we aim to in the following chapter.

Chapter 4

CS and Denoising on MRI

CS and MRI match comes rather naturally, due to the enhanced accuracy provided by the former on the strict constraints of the latter. Especially, what renders MRI so suitable to CS applications, is sparsity, which, as seen in section 3.1, is a prominent feature in CS.

Magnetic Resonance as diagnostic tool is based on image processing through Fourier analysis and, specifically, recovering the requested signals by the use of the inverse Fourier Transform. On the other hand, the main purpose of MRI being medical demands of tight process performances, both in acquisition time and image quality terms. It is therefore in this sense that MRI fittingly sets into CS frame. Particularly, the non-linear recovery of highly incoherent and sparse samples as proposed by Theorem 3.1.1 applies to MRI, having sinusoids sensing waveforms recorded as their corresponding Fourier coefficients.

4.1 MRI Denoising

In MRI, as per every data acquisition, the collected information comes with noise and it is thus of fundamental importance to reduce it as much as possible. Though having to take into account noise effects onto the entire MRI process, acceptable estimation can be made in order to relate accurately to the original noise-free data.

Let us therefore consider an appropriate describing system: Y noisy observations of the interested data matrix $X^0_{m,n}$

$$Y_{i,j} = X^0_{i,j} + W_{i,j} \qquad W_{i,j} \overset{iid}{\sim} \mathcal{N}(0, \tau^2) \qquad \begin{matrix} i = 1, \dots, m \\ j = 1, \dots, n \end{matrix} \tag{4.1}$$

Hence, through eq. (4.1) we characterize disturbances as Gaussian noise, setting the goal of estimating X^0. A precise approximation is met under the quite mild assumption of X^0 having low rank, which is actually the reality

of MR images with columns of X^0 showing the high correlation between nearby wavelengths.

4.1.1 Thresholding Approximation

Assuming the Gaussian model of eq. (4.1) with X^0 having low rank, two different *unbiased estimators* can be considered. Primarily, let we consider the following

Definition 4.1 (Singular Value Decomposition)**.** The singular value decomposition of a matrix $Y_{m,n}$ is a factorization of the form

$$Y = U\Sigma V^* \tag{4.2}$$

with $U_{m,m}$ orthogonal (unitary) matrix, $\Sigma_{m,n}$ diagonal matrix and $V^*_{n,n}$ unitary conjugate transpose of V matrix.

Firstly, a Singular Value Hard Threshold (SVHT) approach is possible:

$$\mathrm{SVHT}_\lambda(Y) = \min_{X\in\mathbb{R}^{m\times n}} \frac{1}{2}\|Y - X\|^2 + \lambda \cdot \mathrm{rank}(X), \qquad \lambda \in \mathbb{R}^+. \tag{4.3}$$

Otherwise, SVHT being a discontinuous estimator in Y, a soft-thresholding method (SVT) is preferred instead:

$$\mathrm{SVT}_\lambda(Y) = \sum_{i=1}^{\min(m,n)} (\sigma_i - \lambda) + \mathbf{u}_i\mathbf{v}_i^*. \tag{4.4}$$

Both estimators propose a truncation of the singular value decomposition of Y, while, to completely appreciate the estimation quality, the mean-squared error is minimized through Stein's Unbiased Risk Estimate (SURE). Particularly, a SURE formulation for MR applications involving Gaussian iid noise as per 4.1 is

$$\mathrm{SURE}(\mathrm{SVT}_\lambda(Y)) = -2mn\tau^2 + \sum_{i=1}^{\min(m,n)} \min(\lambda^2, \sigma_i^2) + 2\tau^2\mathrm{div}(\mathrm{SVT}_\lambda(Y)), \tag{4.5}$$

with 'div' being the divergence of the nonlinear mapping SVT_λ.

Despite SVT being an efficient and direct approach to estimate the noise-free MR images, choosing an appropriate threshold parameter remains a challenging question.

4.2 MRI Applications

Applying denoising strategies to MRI applications is of primal importance, in order to relate to reliable morphological and clinical data. This can be

reflected onto an as high as possible signal to noise ratio (SNR). Particularly, thinking of cardiac MRI takes into account a very dynamic signal, due to natural breathing and cardiac cycled motion, thus further degrading SNR.

Therefore, denoising MRI data through SVT, implies a set of t separate $n \times n$ images referred to as in 4.1. Furthermore, MRI data demand multi-dimensional acquisitions, thus surmounting temporal dimensionality, that is $n^2 >> t$, resulting in a very thin $Y_{n^2,t}$ matrix. Under these circumstances, a direct SVT approach does not yield an accurate estimation, reflecting instead the low degrees of freedom onto flawed results of the temporal terms. In order to overcome this limitation, a different approach is possible: image sets can be analysed separately, based on a spatial block organization of the total space dimensionality. The proposed spatial block-wise solution is practically employed appropriately extracting k^2 rows from each $k \times k$ spatial block by the way of a *binary operator* R_b. Hence, to analyse the set Ω of blocks globally composing the image domain, a block-wise SVT (BSVT) is defined as

$$\mathrm{BSVT}_\lambda(Y) = c^{-1} \sum_{b \in \Omega} R_b^* \mathrm{SVT}_\lambda(R_b Y), \qquad c > 0. \tag{4.6}$$

Remark 7. The BSVT analysis on a singular block of $k = n$, $|\Omega| = 1$, yields to the standard SVT as defined in eq. (4.4)

Basically, BSVT estimates submatrices of Y by SVT, collecting a weighted sum of the results. In this context, recalling the unbiased risk estimator introduced by eq. 4.5, the divergence 'div' of the nonlinear mapping BSVT_λ is:

$$\mathrm{div}\ \mathrm{BSVT}_\lambda(Y) = c^{-1} \sum_{b \in \Omega} \mathrm{div} R_b^* \mathrm{SVT}_\lambda(R_b Y). \tag{4.7}$$

Finally, considering the mean-squared error (MSE) of BSVT_λ

$$\mathrm{MSE}(\lambda) = \mathbb{E}\|X_0 - \mathrm{BSVT}_\lambda(Y)\|^2, \tag{4.8}$$

the unbiased estimator of 4.8 is given by SURE as in 4.5. In this respect figure 4.1 shows a block-wise approach onto a series of MR images, whereas BSVT and global SVT performances are compared.

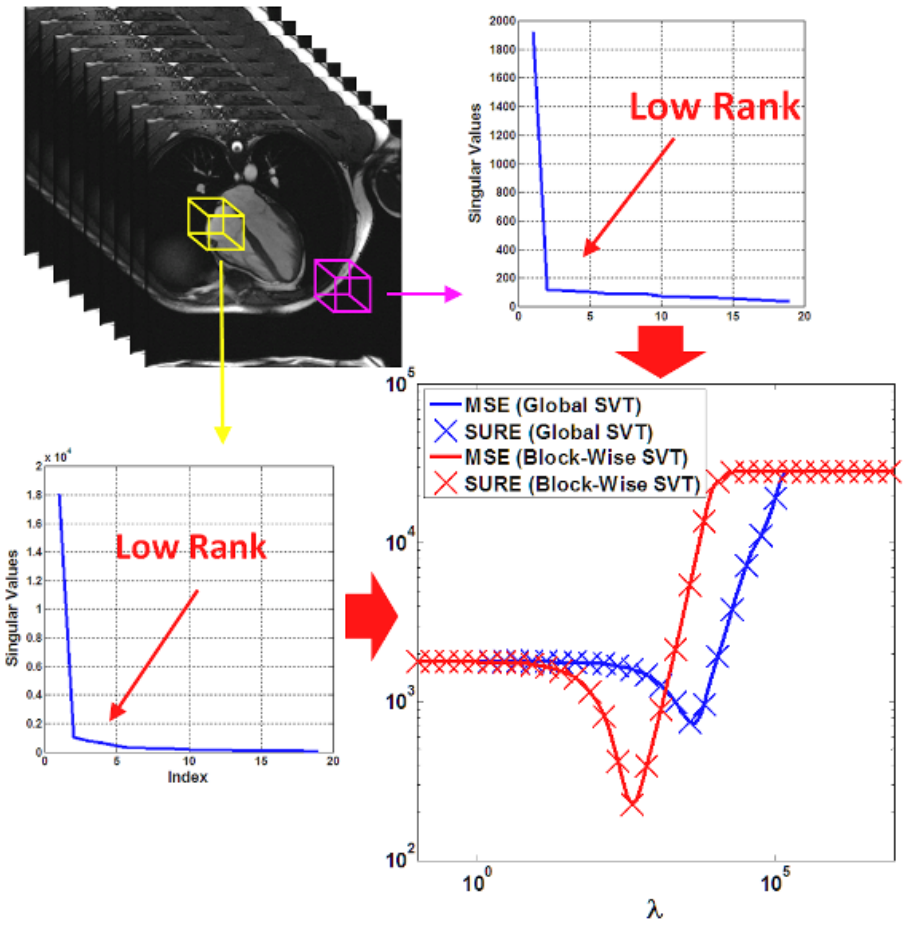

Figure 4.1: Singular values of Y formed from different extracted blocks, including a dynamic region (yellow) and background noise (magenta), while plots of MSE and SURE of both BSVT and global SVT as a function of threshold value λ result in a lower MSE for BSVT than global SVT.

4.2.1 SVT-denoised MRI Examples

Let us illustrate some practical examples of the above explained theories on actual SVT-based denoising of MRIs.

Firstly, a data simulation of a *myocardial perfusion* MRI is performed on a free-breathing model, having $n = 128$ and $t = 50$, adding iid Gaussian noise of $\tau = 30$. Results are collected implementing on 101 equispaced values in $[10^{-1}, 10^{7}]$ both global SVT and BSVT (the latter with $k = 7$ and $|\Omega| = n^2$). Particularly, figure 4.2 shows interesting results, comparing the original noise-free data to each of the noisy, SVT-denoised and BSVT-denoised data over three successive time frames (early, middle and late), while SVT threshold values are based on the MSE/SURE-minimizers as plotted in figure 4.1. In this sense, as figure 4.2 clearly suggests, both

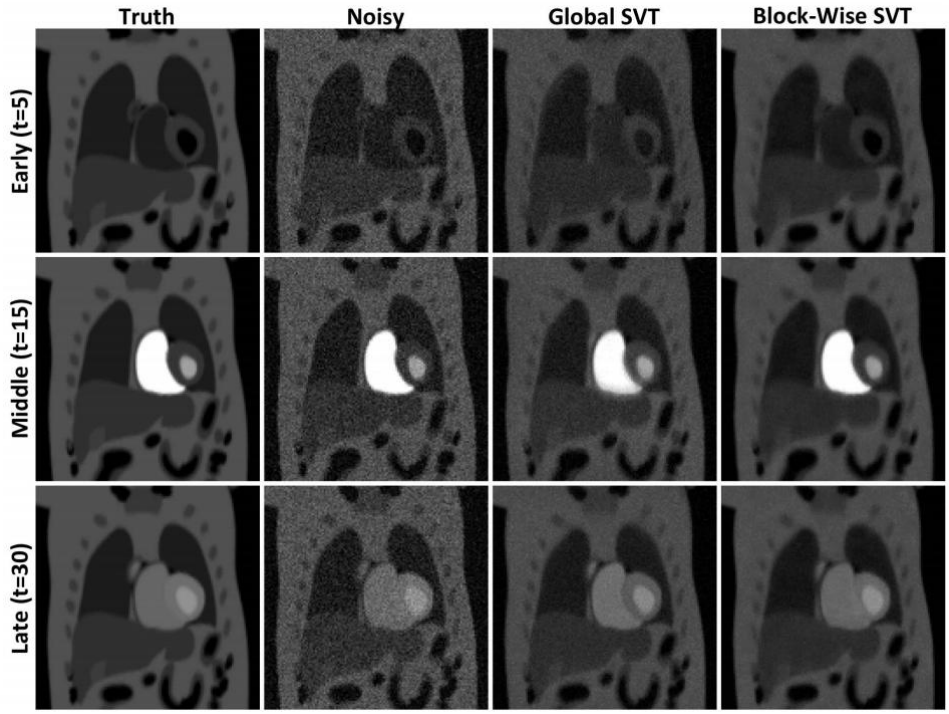

Figure 4.2: Real-time myocardial perfusion MRI simulation in original (truth), noisy and SVT/BSVT-denoised versions.

SVT and BSVT achieve an appreciable noise reduction, accurately preserving morphological and contrast aspects. On the other hand, in terms of smoothness and closeness to truth, is notably significant the greater precision BSVT yields. This is primarily due to the lower MSE of BSVT than global SVT.

A second interesting experiment is based on a *cine-cardiac* MRI series, setting $n = 192$, $t = 19$ and subsequently denoising by SVT. Background

noise is characterized with $\tau = 0.67$, while BSVT applies $k = 5$ on $|\Omega = n^2|$, choosing the parameter over 101 values equispaced over $[10^{-3}, 10^5]$. In this regard, figure 4.3 highlights the importance of threshold value parameter selection for both SVT and BSVT. Consequences of wrongly estimating λ

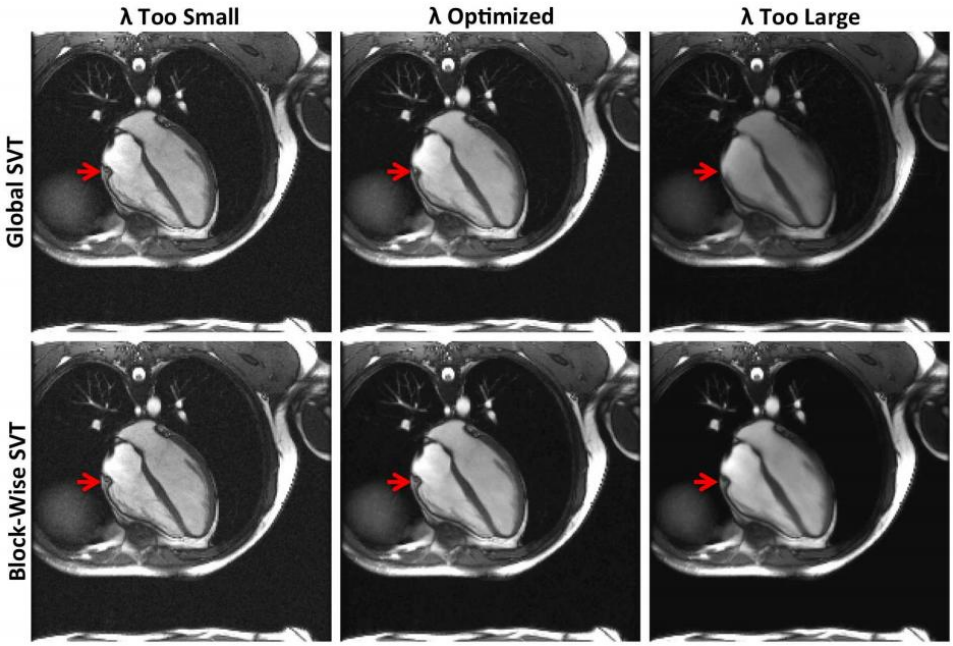

Figure 4.3: Cine-cardiac MRI series denoised by global SVT and BSVT in different threshold values.

are clearly visible from figure 4.3 for each denoising strategy. Although BSVT performing again slightly better in terms of noise reduction, critical results are achieved in cases of non-optimized thresholds. While a λ under-estimation does not remove noise acceptably, over-estimating λ causes a spatio-temporal blurring. In particular, blurring prominence appears differently, based on the denoising strategy: global SVT shows flawed results in areas of high motion (as for the tricuspid heart valve, indicated with the red arrow) and cleaner representations of more static areas, whereas BSVT has a quite reverse trend, with a neater denoising of dynamic areas and blurred less active areas.

The two previous examples underlie the strong efficacy of unbiased risk estimations to reduce noise via SVT. The proposed methods are clearly extended to every other than cardiac MRI. Furthermore, analogously to threshold values optimization, unbiased risk estimation can also be performed to determine BSVT optimal block-size.

Chapter 5

Conclusions

Through the previous chapters, an insight into new techniques of data processing were presented. Particularly, the given theories on both CS and SVT are especially interesting because of their suitability to natural and real cases of undersampled and noisy data. Of primal effectiveness is also the vast and straightforward generalization of the proposed methods, fitting to various and different practical applications in many fields, besides opening new challenges into research developments.

As per MRI applications, CS and SVT theories result in a complementary solution to the overwhelming demands of reconstructing undersampled matrices on both quality and fidelity constraints. While efficiently founding on recent research developments on matrix completion, the new theories are strongly establishing in the light of promising possibilities of SURE application also on undersampled problems. In this sense, parameter selection in undersampled contests enhances interest in studies and researches in this direction.

Besides the challenging and exciting future prospects CS and SVT offer, the surge of new CS and SVT-based applications is already visible and competently covers actual available technologies such as the previously explained MRIs, analogue-to-digital CS-based conversions and many others. On the other hand, the continuing interest in fastening and increasing quality of applications is of primal care and attraction for further researches and developments.

Chapter 6

Bibliography

[1] A New Type of Mathematics for Signal Processing. `http://www.stat.stanford.edu/~candes/A2IWeb/index.html`.

[2] SFFT: Sparse Fast Fourier Transform. `http://groups.csail.mit.edu/netmit/sFFT/`.

[3] Emmanuel J. Candès and Carlos Fernandez-Granda. Towards a Mathematical Theory of Super-Resolution. To appear in *Communication on Pure and Applied Mathematic*, 2012.

[4] Emmanuel J. Candès and Justin Romberg. Sparsity and Incoherence in Compressive Sampling. *Inverse Problems*, 2007.

[5] Emmanuel J. Candès, Carlos A. Sing-Long, and Joshua D. Trzasko. Unbiased Risk Estimates for Singular Value Thresholding and Spectral Estimators. To appear in *IEEE Transactions on Signal Processing*, 2012.

[6] Emmanuel J. Candès and Michael B. Wakin. People Hearing Without Listening. An Introduction to Compressive Sampling. *IEEE Signal Processing Magazine*, 2008.

[7] Larry Hardesty. Explained: The Discrete Fourier Transform. `http://web.mit.edu/newsoffice/2009/explained-fourier.html`, November 2009.

[8] Larry Hardesty. The Faster-than-Fast Fourier Transform. `http://web.mit.edu/newsoffice/2012/faster-fourier-transforms-0118.html`, January 2012.

[9] Larry Hardesty. Toward Practical Compressed Sensing. `http://web.mit.edu/newsoffice/2013/toward-practical-compressed-sensing-0201.html`, February 2013.

[10] Michael Lustig, David Donoho, and John M. Pauly. Sparse MRI: the Application of Compressed Sensing for Rapid MR Imaging. *Magnetic Resonance in Medicine*, 2007.

[11] Alan V. Oppenheim and Alan S. Willsky. *Signals and Systems*. Prentice Hall, 1997.

[12] Justin Romberg and Michael Wakin. Compressed Sensing: a Tutorial. *IEEE Statistical Signal Processing Workshop*, 2007.

[13] Yeyang Yu, Mingjian Hong, Feng Liu, Hua Wang, and Stuart Crozier. Compressed Sensing MRI with Singular Value Decomposition-based Sparsity Basis. *Physics in Medicine and Biology*, 2011.