From Compressed sensing to realistic sampling: the example of MRI

N. Chauffert

CEA/NeuroSpin Parietal Team

Apr. 10, 2014

Advisors: Philippe Ciuciu & Pierre Weiss
Joint work with: Jonas Kahn

Séminaire PGMO, École Polytechnique

Outline

Compressed Sensing MRI

Variable density sampling

Two examples of VDS

Outline

Compressed Sensing MRI

Variable density sampling

Two examples of VDS

Motivation - Magnetic Resonance Imaging

Objective: reduce acquisition times in MRI

- Patient comfort.
- Reducing geometrical distortions (patient moves).
- Reducing scanning costs.
- Improvement of spatio-temporal resolutions.

Our approach: devise *mathematically grounded* strategies to reduce acquisition times by changing *the sampling and reconstruction strategies*.

Basic facts - An MR scanner as a gigantic microwave

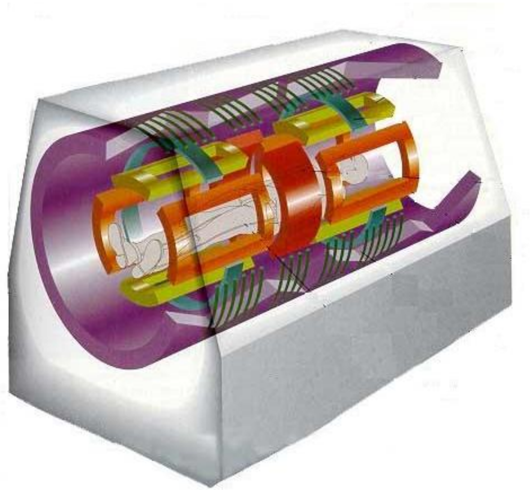
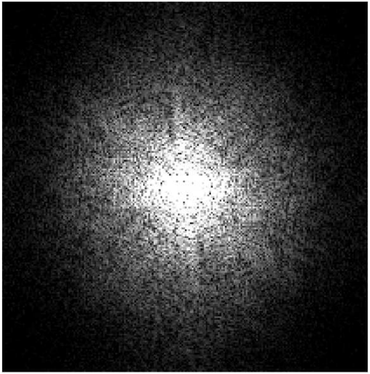


Figure: MRI acquisition.

Background

MRI sampling

MRI allows to measure a discrete set of **Fourier transform values**.



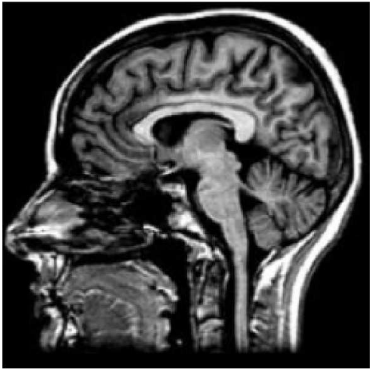


Figure: Left: Fourier transform of a 2D brain. Right: 2D MRI brain image.

Formalizing the sampling problem

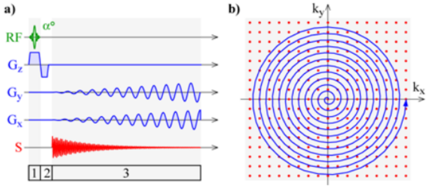


Figure: Pulse sequence and corresponding sampling trajectory.

Let $s : [0, T] \to \mathbb{R}^d$, $(d = 2, 3)$ denote the sampling curve. We have:

$$s(t) = s(0) + \gamma \int_0^t g(s)ds \text{ avec } g = (g_x, g_y).$$

The g field is called *gradient encoding*, it should satisfy:

- $\|g_x\|_\infty \leq \alpha$, $\|g_y\|_\infty \leq \alpha$ **(bounded speed)**.
- $\|g_x'\|_\infty \leq \beta$, $\|g_y'\|_\infty \leq \beta$ **(bounded "curvatures")**.

Similar to driving a car on the Fourier plane.

Formalizing the sampling problem

Let $u : [0,1]^d \to \mathbb{R}$ be an image and $\hat{u}$ denote its Fourier transform.

Our objective: reconstruct $\tilde{u}$ such that $\|u - \tilde{u}\| \leq \epsilon$

Minimize T_ϵ under the constraint that there exists $g : [0, T_\epsilon] \to \mathbb{R}^d$ s.t.

- g and g' are uniformly bounded.
- Sampling the curve $s(t) = s(0) + \gamma \int_0^t g(s)ds$ generates a set

$$E(s) = \{\hat{u}(s(k\Delta t))\}_{k \in \{0,\ldots,T_\epsilon/(\Delta t)\}}$$

  that allows reconstructing $\tilde{u}$ with precision ϵ.

Questions...

- How to choose the measurements?
- How to find s?
- How to reconstruct $\tilde{u}$ knowing $E(s)$?

A possible answer: Compressed Sensing

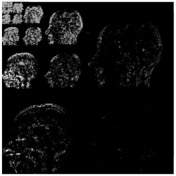
Notation

Let $x \in \mathbb{C}^n$ denote an s-sparse vector.
Let A denote the acquisition matrix. Let $\Omega \subseteq \{1, \cdots, n\}$ and $A_\Omega = (a_i^*)_{i \in \Omega}$.
We acquire a measurement vector:

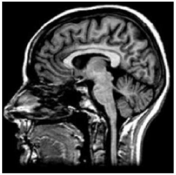
$$y = A_\Omega x.$$

Example

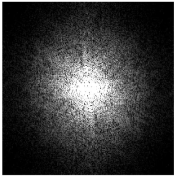
x



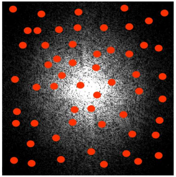
Ψx



$F^* \Psi x = Ax$



$A_\Omega x$



The foundations of variable density sampling

Theorem [Candès, Plan 2011]

Let x be an arbitrary s-sparse vector.
Let $(J_k)_{k\in\{1,\dots,m\}}$ denote a sequence of i.i.d. random variables taking value $i \in \{1,\dots,n\}$ with probability p_i.
Generate a random set $\Omega = \{J_1,\dots,J_m\}$ and measure $y = A_\Omega x$.
Take $\eta \in]0,1[$ and assume that:

$$m \geq C \max_{k\in\{1,\dots,n\}} \frac{\|a_k\|_\infty^2}{p_k} s \ln\left(\frac{n}{\eta}\right)$$

where C is a universal constant.
Then with probability $1-\eta$ vector x is the unique solution of the following problem:

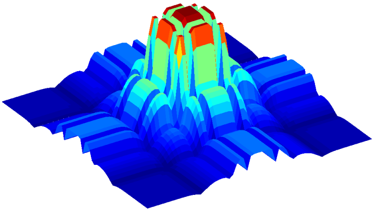
$$\min_{z\in\mathbb{C}^n, A_\Omega z = y} \|z\|_1.$$

The foundations of variable density sampling

Corollary

The "optimal" distribution reads $p_k = \frac{\|a_k\|_\infty^2}{\sum_{i=1}^n \|a_i\|_\infty^2}$.

Illustration of optimal sampling strategy for $A = F^*\Psi$ (MRI)



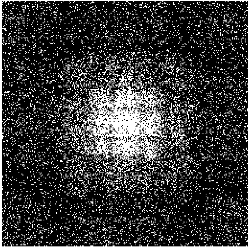


Figure: Left: Optimal drawing probability. Right: a realization of a set Ω with 1 coefficient kept out of 5.

The foundations of variable density sampling

Illustration of optimal sampling strategy for $A = F^*\Psi$

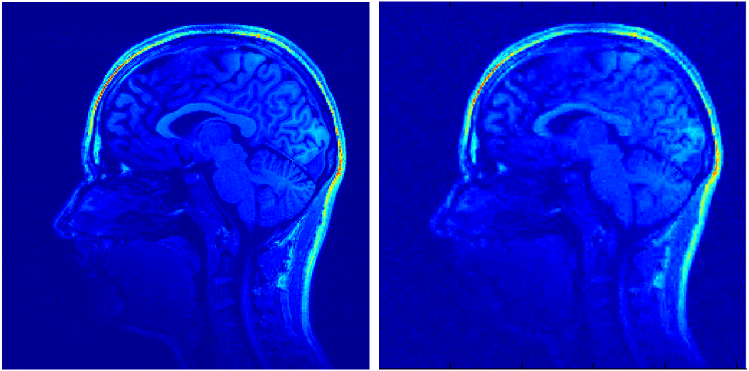


Figure: Left: original image. Right: reconstructed image (PSNR = 35.4dB).

Partial summary

- Existing strategies:

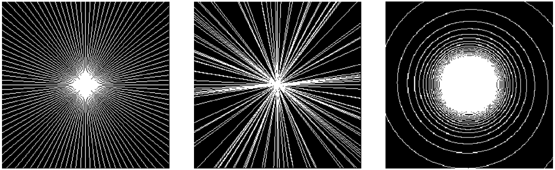


Figure: Traditional sampling schemes.

- Compressed Sensing:

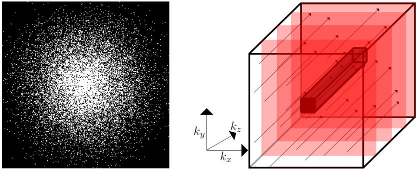


Figure: "Compressed Sensing" acquisitions.

Outline

Compressed Sensing MRI

Variable density sampling

Two examples of VDS

Sketch of the proof of the CS theorem

Proof outline - Part I: deterministic, convex analysis

Main ingredient

Let S denote the support of x.
A necessary condition for perfect recovery of x is:

$$\|(A_\Omega^S)^* A_\Omega^S - I_s\|_{2\to 2} \leq \delta \leqslant \frac{1}{2} \tag{1}$$

Equation (1) can be viewed as a near isometry property of A_Ω^S.

Sketch of the proof of the CS theorem

Proof outline - Part II: concentration inequalities

By assumption we have

$$I_s = A^{S*} A^S = \sum_{i=1}^{n} p_i \frac{a_i^S a_i^{S*}}{p_i}. \tag{2}$$

Let $X_i^S = \frac{a_{J_i}^S a_{J_i}^{S*}}{p_{J_i}}$ denote a rank 1 random matrix.
By Eq. (2), we get $\mathbb{E}(X_i^S) = I_s$ thus, by the C.L.T.:

$$\lim_{m \to +\infty} \frac{1}{m} \sum_{i=1}^{m} X_i^S = I_s \text{ a.s..}$$

Concentration inequalities (Bernstein) provide stronger results:

$$\mathbb{P}\left(\left\| \frac{1}{m} \sum_{i=1}^{m} X_i^S - I_s \right\|_{2 \to 2} \geq t \right) \leq 2s \exp\left(-\frac{mt^2}{C\mu s} \right).$$

A definition of VDS

Definition: variable density samplers [C. et al, submitted, 2014]

Let p denote a probability measure on a space Ξ (e.g. $\{1, \dots, n\}$ or $[0, 1]^d$).
A stochastic process $X = (X_i)_{i\in\mathbb{N}}$ or $X = (X_t)_{t\in\mathbb{R}_+}$ is called a p-variable density sampler if its empirical measure (or occupation measure) satisfies

$$\lim_{m\to+\infty} \frac{1}{m} \sum_{i=1}^{m} f(X_i) = p(f) \text{ a.s.}$$

or

$$\lim_{T\to+\infty} \frac{1}{T} \int_{t=0}^{T} f(X_t) = p(f) \text{ a.s.}$$

for all continuous bounded functions f.

Examples of variable density samplers

Example 1: point processes

Drawing independent random vectors in $\mathbb{R}^d$ with a distribution p is a p-variable density sampler.

Example 2: random walks

A random walk in $\mathbb{R}^d$ with a stationary distribution p is a p-variable density sampler.

Example 3: dynamical systems

The definition of variable density samplers is closely related to the ergodic hypothesis for dynamical systems.

What are the key properties of a *good* VDS ?

Outline

Sampling with random walks

Construction of a discrete Markov chain

Given a target probability distribution $p \in \mathbb{R}^n$.
Define a Markov chain $X = (X_i)_{i \in \mathbb{N}}$ on the set $\{1, \dots, n\}$.
Use Metropolis algorithm to construct a stochastic transition matrix $P \in \mathbb{R}^{n \times n}$ such that p is the stationary distribution of X i.e.

$$p = pP.$$

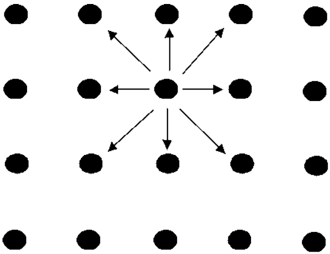


Figure: Authorized transitions to enforce continuity.

Sampling with random walks

Theorem [C. et al 2013]

Let $\Omega = (X_1, \ldots, X_m)$ denote the set of indices obtained at time m.
If

$$m \geq \frac{C}{\epsilon(P)} \max_{k \in \{1,\ldots,n\}} \frac{\|a_k\|_\infty^2}{p_k} s^2 \ln\left(\frac{n}{\eta}\right)$$

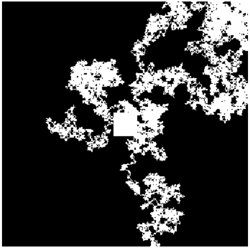
Then every s-sparse vector is recovered exactly by solving the ℓ^1 minimization problem with probability $1-\eta$.
The spectral gap $\epsilon(P)$ is the difference between the largest and the second largest eigenvalue of P.

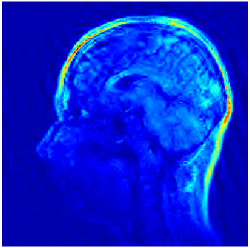
A doomed approach?

The spectral gap $\epsilon(P)$ usually depends on n and can be as small as $\frac{1}{n^{1/d}}$!

Sampling with random walks

A doomed approach?





PSNR=31.1 dB

Figure: Markov based sampling yields bad reconstruction results!

The Travelling Salesman sampler

An algorithmic alternative

- Throw a set of points according to a density q.
- Join them by finding the shortest path passing through all of them.

Problem: How to choose q ?

The Travelling Salesman sampler

An algorithmic alternative

- Throw a set of points according to a density q.
- Join them by finding the shortest path passing through all of them.

Problem: How to choose q ?

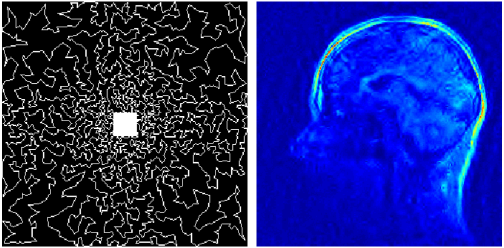
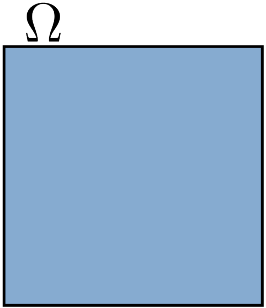


Figure: **The naive approach fails!** We can't just draw the initial points according to p and join them using the TSP.

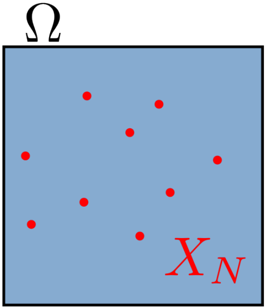
The Travelling Salesman sampler

- Let $\Xi = [0,1]^d$, with $d \geqslant 2$.
- $(x_i)_{i \in \mathbb{N}^*}$ a sequence of points in Ξ, *i.i.d.* drawn $\sim q$.
- $X_N = (x_i)_{i \leqslant N}$.
- Denote $T(X_N, \Xi)$ the length of the TSP amongst X_N.
- $\gamma_N : [0,1] \to \Xi$ denotes the parametrization of the curve at constant speed $T(X_N, \Xi)$.



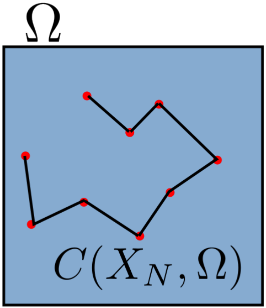
The Travelling Salesman sampler

- Let $\Xi = [0,1]^d$, with $d \geqslant 2$.
- $(x_i)_{i\in\mathbb{N}^*}$ a sequence of points in Ξ, *i.i.d.* drawn $\sim q$.
- $X_N = (x_i)_{i\leqslant N}$.
- Denote $T(X_N, \Xi)$ the length of the TSP amongst X_N.
- $\gamma_N : [0,1] \to \Xi$ denotes the parametrization of the curve at constant speed $T(X_N, \Xi)$.



The Travelling Salesman sampler

- Let $\Xi = [0,1]^d$, with $d \geqslant 2$.
- $(x_i)_{i\in\mathbb{N}^*}$ a sequence of points in Ξ, *i.i.d.* drawn $\sim q$.
- $X_N = (x_i)_{i\leqslant N}$.
- Denote $T(X_N, \Xi)$ the length of the TSP amongst X_N.
- $\gamma_N : [0,1] \to \Xi$ denotes the parametrization of the curve at constant speed $T(X_N, \Xi)$.



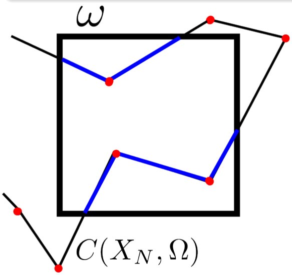
The Travelling Salesman sampler

The Lebesgue measure on an interval $[0,1]$ is denoted $\lambda_{[0,1]}$.

Distribution of the TSP solution

For any Borelian ω in Ξ:

$$P_N(\omega) = \lambda_{[0,1]}\left(\gamma_N^{-1}(\omega)\right).$$



Intuitive definition

$$P_N(\omega) = \frac{T_{|\omega}(X_N,\Xi)}{T(X_N,\Xi)}, \ \forall\omega.$$

The Travelling Salesman sampler

Theorem (TSP empirical measure (Chauffert, W. 2013))

Define the density:

$$p = \frac{q^{(d-1)/d}}{\int_{\Xi} q^{(d-1)/d}(x)dx}.$$

Almost surely w.r.t. the law $q^{\otimes \mathbb{N}}$ of the sequence $(x_i)_{i\in\mathbb{N}^*}$ of random points in the hypercube, the distribution P_N converges in distribution to p:

$$P_N \overset{(d)}{\to} p \qquad\qquad q^{\otimes \mathbb{N}}\text{-a.s.}$$

The Travelling Salesman sampler

Intuition

Consider a small hypercube:

- The number of point n is $\propto q$;

The Travelling Salesman sampler

Intuition

Consider a small hypercube:

- The number of point n is $\propto q$;
- The typical distance is proportional to $n^{-1/d}$ (or $q^{-1/d}$);

The Travelling Salesman sampler

Intuition

Consider a small hypercube:

- The number of point n is $\propto q$;
- The typical distance is proportional to $n^{-1/d}$ (or $q^{-1/d}$);
- $\Rightarrow$ The length of the TSP in the small cube is $\propto qq^{-1/d} = q^{(d-1)/d}$

The Travelling Salesman sampler

Intuition

Consider a small hypercube:

- The number of point n is $\propto q$;
- The typical distance is proportional to $n^{-1/d}$ (or $q^{-1/d}$);
- $\Rightarrow$ The length of the TSP in the small cube is $\propto qq^{-1/d} = q^{(d-1)/d}$

The Travelling Salesman sampler

Intuition

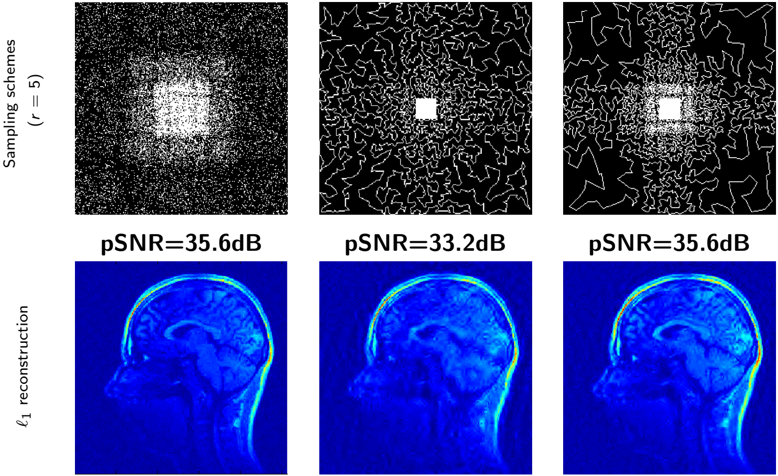
Consider a small hypercube:

- The number of point n is $\propto q$;
- The typical distance is proportional to $n^{-1/d}$ (or $q^{-1/d}$);
- $\Rightarrow$ The length of the TSP in the small cube is $\propto qq^{-1/d} = q^{(d-1)/d}$

Conclusion

To reach a target density p, one should choose $q \propto p^{d/(d-1)}$!

The Travelling Salesman sampler



The Travelling Salesman sampler

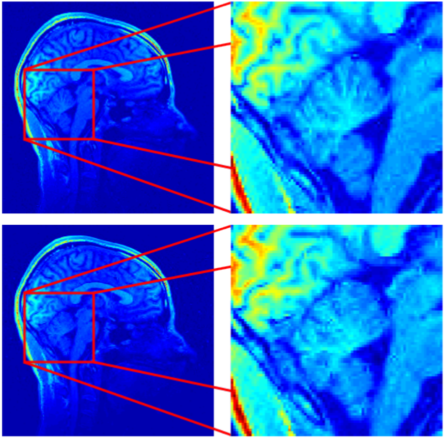


Figure: 3D reconstruction results for $r = 8.8$ for various sampling strategies. **Top row:** TSP-based sampling schemes (PSNR=42.1 dB). **Bottom row:** 2D random drawing and acquisitions along parallel lines (state-of-the-art) (PSNR=40.1 dB).

Conclusion 1/2

- We motivated VDS and introduced a definition.
- We proposed two continuous sampling strategies.
- We highlighted the key-properties of a *good* VDS.

Conclusion 2/2

	Markov approach	TSP approach
		
Mixing time	$\varepsilon(P)$	seems fast, how to quantify ?
Distribution	Metropolis	draw points w.r.t π^2
MRI constraints	can be improved	strong limitation

Objective

Find a trajectory which covers the space rapidly, and with the optimal distribution.

Questions?

Thank you for your attention !