Compressed Sensing Along Physically Plausible Trajectories In Magnetic Resonance Imaging

N. Chauffert

Thesis defense

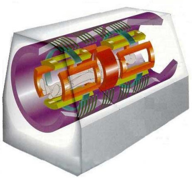
September 28, 2015

Advisors: Philippe Ciuciu & Pierre Weiss.

Background on Magnetic Resonance Imaging (1/5)

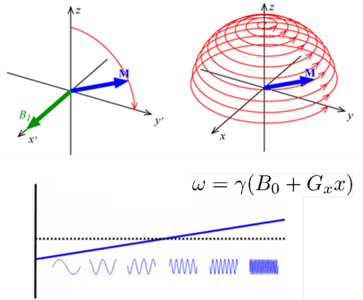
MRI is a non-invasive imaging modality to probe water molecules.

- Strong, static, homogeneous magnet (1.5 to 3T in hospitals, 7T and soon 11.7T at NeuroSpin).
- A Radio-frequency (RF) pulse to excite the spins.
- Receiving coils.



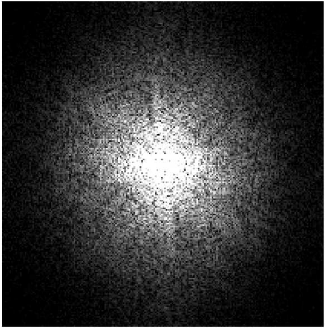
Background on Magnetic Resonance Imaging (2/5)

- Primary magnetic field (B_0). Align the spins in the z-direction
- Tip the global magnetization into the transverse (x,y) plane using a RF pulse at Larmor frequency $\omega_0 = \gamma B_0$.
- Release the RF pulse and measure transverse relaxation.
- Gradient magnets. Localize the MR signal.



Background on Magnetic Resonance Imaging (3/5)

Aquisitions are performed in the Fourier domain (k-space):



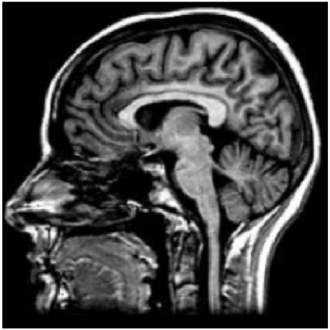


Figure: Left: 2D slice of MRI brain in the Fourier domain (k-space). Right: 2D slice of MRI brain in the image domain.

Background on Magnetic Resonance Imaging (4/5)

Mathematical modelling

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

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

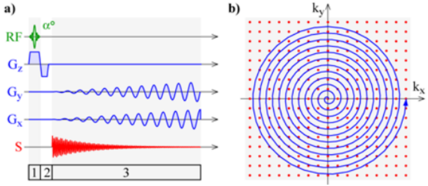


Figure: Pulse sequence and corresponding sampling trajectory.

Background on Magnetic Resonance Imaging (5/5)

The *gradient encoding* g should satisfy:

$$\|g\|_\infty \leq G_{\max} \quad \text{and} \quad \|\dot{g}\|_\infty \leq S_{\max}.$$

Admissible sampling curves

An *admissible sampling curve in MRI* is a curve belonging to the set:

$$\mathcal{S}_{\mathrm{MRI}} = \left\{ s : [0, T] \mapsto \mathbb{R}^3, \|\dot{s}\|_\infty \leqslant \alpha, \|\ddot{s}\|_\infty \leqslant \beta \right\}$$

Similar to driving a car on the Fourier plane.

Goals of this thesis (1/2)

Reducing scanning time

- Improve patient comfort.
- Reduce distortions due to patient moves.
- Reduce geometric distortions by decreasing readout times.
- Reducing scanning costs.
- Improve either spatial, temporal or angular resolution (MRI/fMRI/dw-MRI).

Goals of this thesis (2/2)

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

Our objective: reconstruct $\tilde{\rho}$ such that $\|\rho - \tilde{\rho}\|_2 \leq \epsilon$

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 $s(t) = s(0) + \int_0^t g(t)dt$ generates a set

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

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

Goals of this thesis (2/2)

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

Our objective: reconstruct $\tilde{\rho}$ such that $\|\rho - \tilde{\rho}\|_2 \leq \epsilon$

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 $s(t) = s(0) + \int_0^t g(t)dt$ generates a set

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

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

Questions...

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

Outline

Outline

Reducing the number of measurements using CS (1/3)

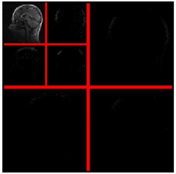
Compressed sensing theory:

- ρ is sparse in a given basis (e.g. wavelets), $\rho = \Psi x$, where $x \in \mathbb{C}^n$ is s-sparse.
- Acquisition matrix: $A = F^* \Psi$.

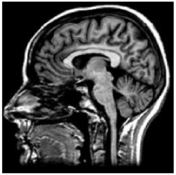
Let $x \in \mathbb{C}^n$ denote an s-sparse representation of the image.
Let $\Gamma \subseteq \{1, \cdots, n\}$ and $A_\Gamma = (a_i^*)_{i \in \Gamma}$. We acquire a measurement vector:

$$y = A_\Gamma x.$$

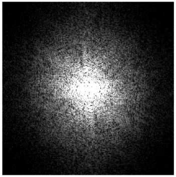
x



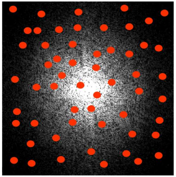
$\rho = \Psi x$



$F^* \Psi x = Ax$



$A_\Gamma x$



Reducing the number of measurements using CS (1/3)

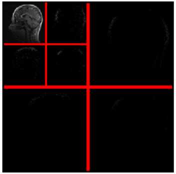
Compressed sensing theory:

- ρ is sparse in a given basis (e.g. wavelets), $\rho = \Psi x$, where $x \in \mathbb{C}^n$ is s-sparse.
- Acquisition matrix: $A = F^* \Psi$.

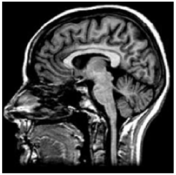
Let $x \in \mathbb{C}^n$ denote an s-sparse representation of the image.
Let $\Gamma \subseteq \{1, \cdots, n\}$ and $A_\Gamma = (a_i^*)_{i \in \Gamma}$. We acquire a measurement vector:

$$y = A_\Gamma x.$$

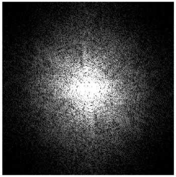
x



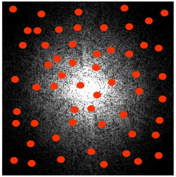
$\rho = \Psi x$



$F^* \Psi x = Ax$



$A_\Gamma x$



ℓ_1 reconstruction (promoting sparsity)

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

Reducing the number of measurements using CS (2/3)

A first CS theorem [Candès and Plan, 2011]

Theorem

Construct Γ by **uniform** *and i.i.d. drawing the lines of A.*
Let x be a sparse vector, containing s non-zero entries. Assume that:

$$m \geqslant C \cdot s \cdot \left(n \cdot \max_{1 \leqslant k \leqslant n} \|a_k\|_\infty^2 \right) \cdot \log\left(\frac{n}{\eta}\right) \quad (1)$$

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

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

Reducing the number of measurements using CS (2/3)

A first CS theorem [Candès and Plan, 2011]

Theorem

Construct Γ by **uniform** *and i.i.d. drawing the lines of A.*
Let x be a sparse vector, containing s non-zero entries. Assume that:

$$m \geqslant C \cdot s \cdot \left(n \cdot \max_{1 \leqslant k \leqslant n} \|a_k\|_\infty^2 \right) \cdot \log\left(\frac{n}{\eta}\right) \tag{1}$$

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

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

In MRI, $\max_{1 \leqslant k \leqslant n} \|a_k\|_\infty^2 = O(1)$, hence $m \gg n$.

This is called the coherence barrier

Reducing the number of measurements using CS (3/3)

Breaking the coherence barrier

- Change the acquisition model using tailored RF pulse:
  - Compressed Sensing with random encoding [Haldar et al., 2011].
  - Spread Spectrum MRI [Puy et al., 2012].
- Variable density sampling: draw with higher probability the measurements corresponding to coherent vectors.

Variable Density Sampling - Theoretical Foundations (1/3)

Theorem [Chauffert et al., 2013]

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 $\Gamma = \{J_1,\dots,J_m\}$ and measure $y = A_\Gamma x$. Take $\eta \in]0,1[$ and assume that:

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

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

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

Variable Density Sampling - Theoretical Foundations (1/3)

Theorem [Chauffert et al., 2013]

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 $\Gamma = \{J_1,\dots,J_m\}$ and measure $y = A_\Gamma x$. Take $\eta \in]0,1[$ and assume that:

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

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

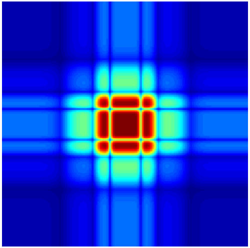
$$\min_{z\in\mathbb{C}^n, A_\Gamma z = y} \|z\|_1.$$

Optimal distribution $\pi_k \propto \|a_k\|_\infty^2$.

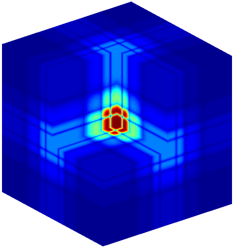
Coherence is now $\max_{k\in\{1,\dots,n\}} \frac{\|a_k\|_\infty^2}{p_k} = \sum_k \|a_k\|_\infty^2 = O(\log(n))$ in MRI.

Variable Density Sampling - Theoretical Foundations (2/3)

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



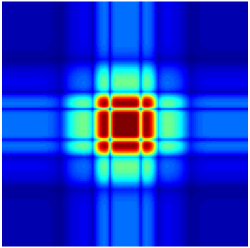
π in 2D



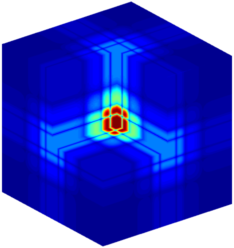
π in 3D

Variable Density Sampling - Theoretical Foundations (2/3)

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

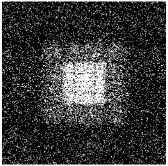


π in 2D



π in 3D

Example of sampling pattern obtained in 2D :



Variable Density Sampling - Theoretical Foundations (3/3)

- Recent results take the signal structure into account [Adcock et al., 2013, Boyer et al., 2015].
- To date, the best sampling distributions are heuristics [Chauffert et al., 2014a].
- From now on, π designs a target sampling distribution.

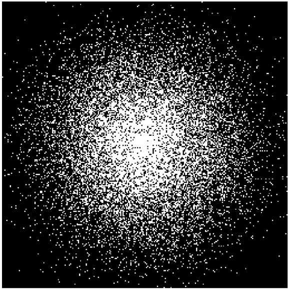
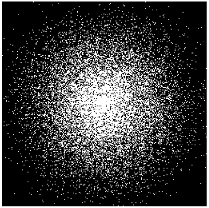


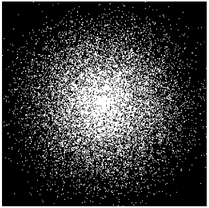
Figure: Example of sampling pattern obtained with CS theory

CS-MRI today

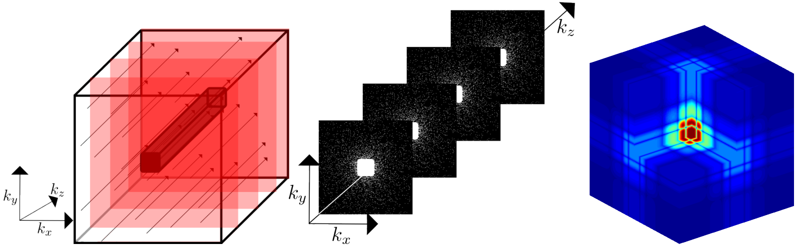


This is NOT feasible ! ($s \notin \mathcal{S}_{\mathrm{MRI}}$)

CS-MRI today



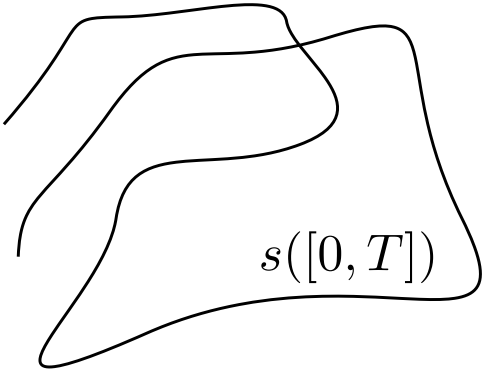
This is NOT feasible ! ($s \notin \mathcal{S}_{\mathrm{MRI}}$)



CS-MRI is sub-optimal ! [Lustig et al., 2007]

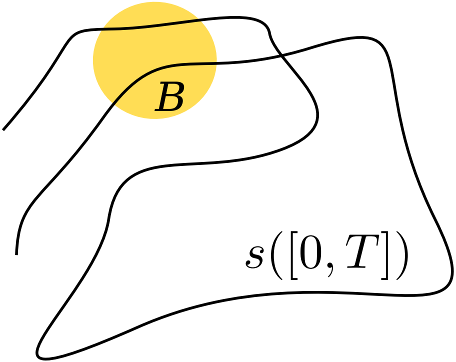
Variable Density Sampling - Definitions

Pushforward measure - illustration



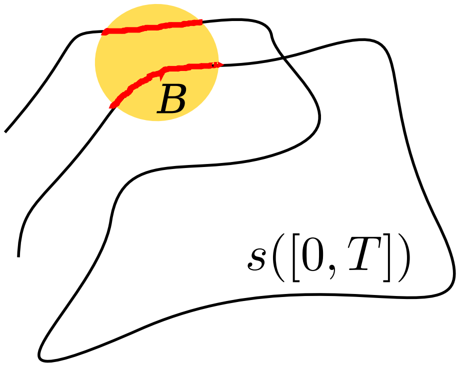
Variable Density Sampling - Definitions

Pushforward measure - illustration



Variable Density Sampling - Definitions

Pushforward measure - illustration



$$\nu(B) = s_* \lambda_T(B) = \lambda_T(s^{-1}(T))$$

λ_T is the (normalized) Lebesgue measure.

Variable Density Sampling - Definitions

Pushforward measure

Let $\Omega = [0,1]^d$, where $d = 2$ or 3 denote the space dimension. We equip Ω with the Borel algebra $\mathcal{B}$. Let (X, Σ) be a measurable space and $s : X \to \Omega$ be a measurable mapping. $\mu : X \to [0; +\infty[$ denote a measure. The *pushforward measure* ν of μ is defined by:

$$\nu(B) = s_*\mu(B) = \mu(s^{-1}(B)), \quad \forall B \in \mathcal{B}$$

Variable Density Sampling - Definitions

Pushforward measure

Let $\Omega = [0,1]^d$, where $d = 2$ or 3 denote the space dimension. We equip Ω with the Borel algebra $\mathcal{B}$. Let (X, Σ) be a measurable space and $s : X \to \Omega$ be a measurable mapping. $\mu : X \to [0; +\infty[$ denote a measure. The *pushforward measure* ν of μ is defined by:

$$\nu(B) = s_* \mu(B) = \mu(s^{-1}(B)), \quad \forall B \in \mathcal{B}$$

Ex. 1: Measures supported by curves

Ex. 2: Atomic measures

$s : \{1, \ldots, m\} \to \Omega$, where $s(i) = p_i$ denotes the i-th point. Set μ as the *counting measure* defined for any set $I \subseteq \{1, \ldots, m\}$ by $\mu(I) = \frac{|I|}{m}$. Then ν is defined by

$$\nu = \frac{1}{m} \sum_{i=1}^{m} \delta_{p_i}.$$

Variable Density Sampling - Definitions

Weak convergence

A sequence of measures (μ_n) is said to weakly converge to μ, if for any bounded continuous function Φ,

$$\int_\Omega \Phi(x)\mathrm{d}\mu_n(x) \to \int_\Omega \Phi(x)\mathrm{d}\mu(x)$$

Shorthand notation: $\mu_n \rightharpoonup \mu$.

Variable Density Sampling - Definitions

Weak convergence

A sequence of measures (μ_n) is said to weakly converge to μ, if for any bounded continuous function Φ,

$$\int_\Omega \Phi(x)\mathrm{d}\mu_n(x) \to \int_\Omega \Phi(x)\mathrm{d}\mu(x)$$

Shorthand notation: $\mu_n \rightharpoonup \mu$.

Variable density sampler

A sequence of (random) trajectories $s_n : X_n \to \Omega$ is said to be a π-Variable Density Sampler if

$$s_{n*}\mu \rightharpoonup \pi \quad \text{almost surely}$$

Variable Density Sampling - Definitions

Weak convergence

A sequence of measures (μ_n) is said to weakly converge to μ, if for any bounded continuous function Φ,

$$\int_\Omega \Phi(x)\mathrm{d}\mu_n(x) \to \int_\Omega \Phi(x)\mathrm{d}\mu(x)$$

Shorthand notation: $\mu_n \rightharpoonup \mu$.

Variable density sampler

A sequence of (random) trajectories $s_n : X_n \to \Omega$ is said to be a π-Variable Density Sampler if

$$s_{n*}\mu \rightharpoonup \pi \quad \text{almost surely}$$

Examples

i.i.d. drawing, random walks ...

Outline

Construction of a discrete Markov chain

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

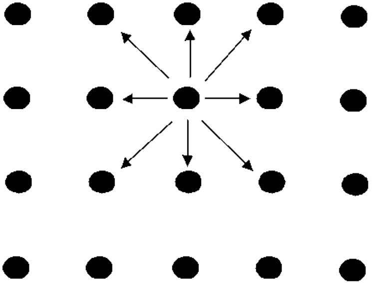


Figure: Authorized transitions to enforce continuity.

CS results

Theorem [Chauffert et al., 2015]

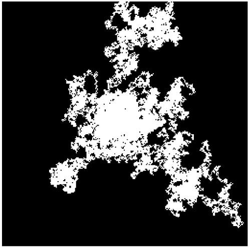
Let x be an s-sparse random vector.
Let $\Gamma = X_1, \ldots, X_m$ denote a set of m indexes selected using a Markov chain. Assume that $X_1 \sim \pi$. Then, if

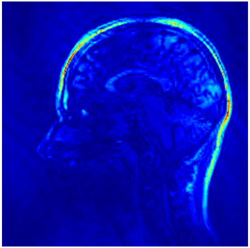
$$m \geq \frac{C}{\varepsilon(P)} \cdot s \cdot \left(\sum_k \|a_k\|_\infty^2 \right) \log^2 \left(\frac{6n}{\eta} \right),$$

every s-sparse vectors are recovered exactly by solving the ℓ^1 minimization problem for matrix A_Γ with probability $1 - \eta$.
$\varepsilon(P)$: spectral gap of the chain (difference between the largest and the second largest eigenvalues of P).

Sampling with random walks

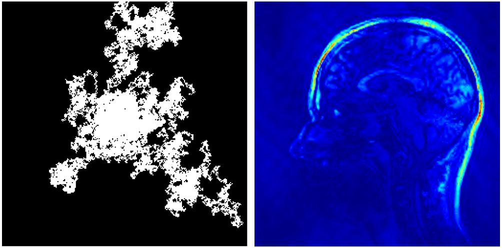
A practical example (20% measurements, PSNR=30dB)





Sampling with random walks

A practical example (20% measurements, PSNR=30dB)



- Time to cover the k-space is slow (controlled by $\epsilon(P)$)
- Local approach

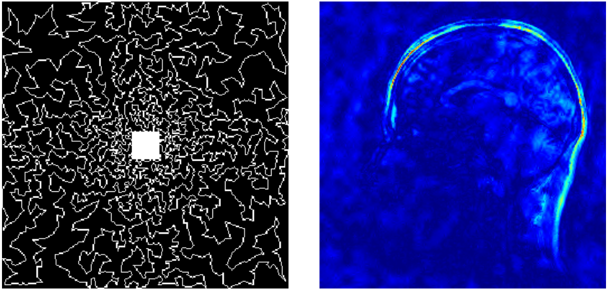
Travelling Salesman Problem (TSP) sampling

Idea : cover the k-space more quickly with a global approach.

Travelling Salesman Problem (TSP) sampling

Idea : cover the k-space more quickly with a global approach.

(a) (b) PSNR=24.1dB



- Pushforward measure far from π. From which distribution should we sample the initial points to reach a given target distribution?

The Travelling Salesman sampler

- Let

$$q = \frac{\pi^{d/(d-1)}}{\int_\Omega \pi^{d/(d-1)}}$$

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

The Travelling Salesman sampler

- Let

$$q = \frac{\pi^{d/(d-1)}}{\int_\Omega \pi^{d/(d-1)}}$$

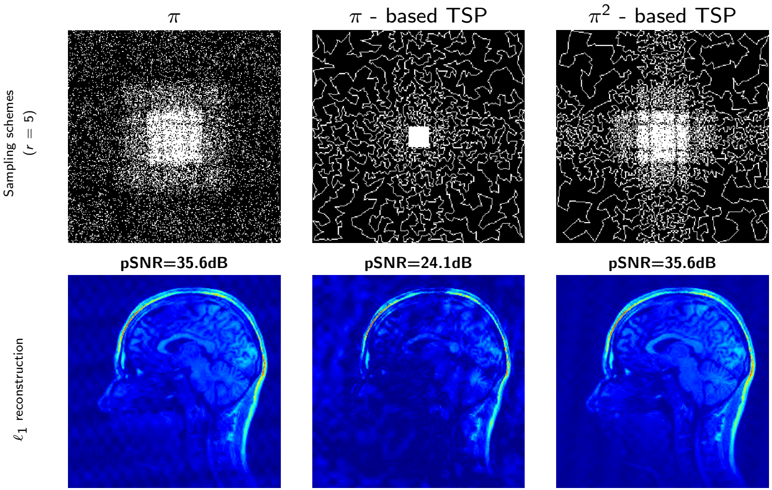
- $(x_i)_{i\in\mathbb{N}^*}$ a sequence of points in Ω, *i.i.d.* drawn $\sim q$.
- $X_N = (x_i)_{i\leqslant N}$.
- Denote $T(X_N)$ the length of the TSP amongst X_N.
- $\gamma_N : [0, T(X_N)] \to \Omega$ denotes the parametrization of the curve at speed 1.

Theorem (TSP is a VDS [Chauffert et al., 2014a])

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, $(\gamma_N)_{N\in\mathbb{N}}$ is a π-variable density sampler, i.e.,

$$\gamma_{N*}\lambda_{T(X_N)} \rightharpoonup \pi$$

The Travelling Salesman sampler - illustration



The Travelling Salesman Sampler - Illustration

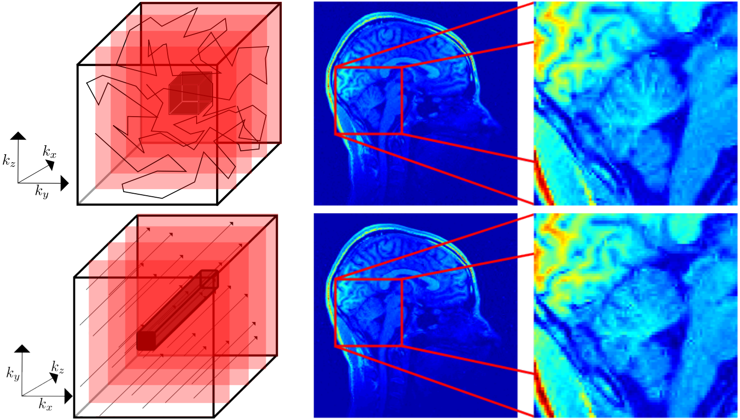


Figure: 3D reconstruction results for $r = 8.8$ for various sampling strategies. **Top row:** TSP-based sampling schemes (PSNR=42.1 dB). **Bottom row:** 2D random drawing and acquisitions along parallel lines [Lustig et al., 2007] (PSNR=40.1 dB).

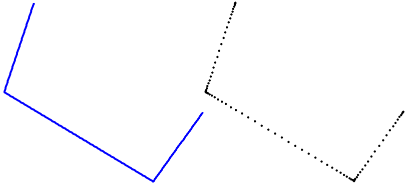
The Parameterization Problem

Finding a parameterization in $\mathcal{S}_{\mathrm{MRI}}$ corresponding to a curve support is not easy !

The Parameterization Problem

Finding a parameterization in $\mathcal{S}_{\text{MRI}}$ corresponding to a curve support is not easy !

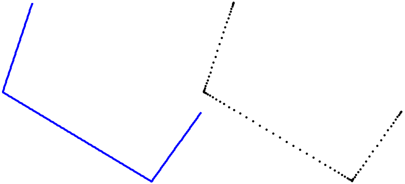
- Classical approach, find an admissible parameterization [Hargreaves et al., 2004, Lustig et al., 2008]:



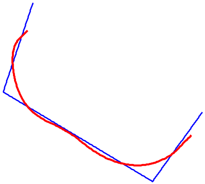
The Parameterization Problem

Finding a parameterization in $\mathcal{S}_{\mathrm{MRI}}$ corresponding to a curve support is not easy !

- Classical approach, find an admissible parameterization [Hargreaves et al., 2004, Lustig et al., 2008]:



- Projection onto $\mathcal{S}_{\mathrm{MRI}}$ [Chauffert et al., 2014b]



The projection operator

For an input parameterized curve $c : [0; T] \rightarrow \Omega$, define:

$$P_{\mathcal{S}_{\mathrm{MRI}}}(c) = \operatorname*{Arg\,min}_{s \in \mathcal{S}_{\mathrm{MRI}}} \int_{t \in [0;T]} (s(t) - c(t))^2 \mathrm{d}t$$

Main properties [Chauffert et al., 2014b]

- Fast resolution using accelerated proximal gradient descent on the dual.
- The sampling time is fixed (equal to T).
- The sampling distribution is well preserved (approximation of Wasserstein distance W_2).

The projection operator

For an input parameterized curve $c : [0; T] \rightarrow \Omega$, define:

$$P_{\mathcal{S}_{\text{MRI}}}(c) = \operatorname*{Arg\,min}_{s \in \mathcal{S}_{\text{MRI}}} \int_{t \in [0;T]} (s(t) - c(t))^2 \mathrm{d}t$$

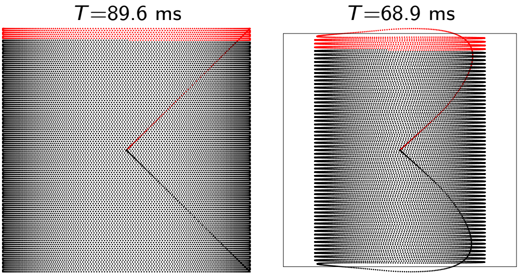
Main properties [Chauffert et al., 2014b]

- Fast resolution using accelerated proximal gradient descent on the dual.
- The sampling time is fixed (equal to T).
- The sampling distribution is well preserved (approximation of Wasserstein distance W_2).

$\Rightarrow$ More importantly, $P_{\mathcal{S}_{\text{MRI}}}$ is the cornerstone of a global approach, described in part 3.

Application to classical MRI trajectories

EPI trajectories



- Resolution is 128×128 (2 mm istropic).
- Very high k_y frequencies are not acquired after projection.
- Reconstruction results are similar.

Interim summary

- 2 key properties for a VDS:
  - sampling distribution;
  - fast k-space coverage.
- Sub-optimal 2-step approaches (random walks/TSP + projection).

Interim summary

- 2 key properties for a VDS:
  - sampling distribution;
  - fast k-space coverage.
- Sub-optimal 2-step approaches (random walks/TSP + projection).

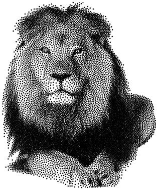
How to design feasible sampling trajectories with good coverage speed and good sampling distribution?

Outline

Introduction of a new metric

Useful to compare parameterizations and (probability) distributions.
Here : $s : \{1, \dots, m\} \to \Omega$ and $\pi : \Omega \to \mathbb{R}$ a distribution.

“s”



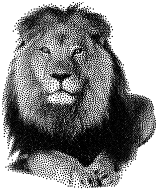
π

Related to dithering problem [Teuber et al., 2011].

Introduction of a new metric

Useful to compare parameterizations and (probability) distributions.
Here : $s : \{1, \dots, m\} \to \Omega$ and $\pi : \Omega \to \mathbb{R}$ a distribution.

"s"



π

$h \star s$

$h \star \pi$

h: a Gaussian kernel.

Related to dithering problem [Teuber et al., 2011].

A projection problem

Working with measures

Let $\mathcal{P}$ denote a set of admissible parameterizations and $\mathcal{M}(\mathcal{P})$ the set of pushforward measures associated with elements of $\mathcal{P}$: Sampling trajectories $s \in \mathcal{P} \to \Omega$ are seen through $s_*\mu \in \mathcal{M}(\mathcal{P})$.

$$\mathcal{M}(\mathcal{P}) = \{\nu = s_*\mu,\ s \in \mathcal{P}\}.$$

m-point measures:

Set of sums of m Dirac delta functions: $\mathcal{M}(\Omega^m) = \left\{\nu = \frac{1}{m}\sum_{i=1}^m \delta_{p_i},\ p_i \in \Omega\right\}$.

Admissible curves for MRI:

$$\mathcal{M}(\mathcal{S}_{\mathrm{MRI}}) = \{\nu = s_*\mu, s \in \mathcal{S}_{\mathrm{MRI}}\}.$$

We want $\nu \in \mathcal{M}(\mathcal{P})$ to be "as close as possible to" π, the target distribution.

Measuring distances between measures

Constructing a metric

Let π denote the *target density*.
Let ν denote the *pushforward measure*.

Let $h : \Omega \to \mathbb{R}$ denote a continuous function with a Fourier series that does not vanish. The following mapping:

$$dist(\pi, \nu) = \|h \star (\pi - \nu)\|_2^2$$

defines a distance (or metric) on $\mathcal{M}_\Delta$, the space of probability measures on Ω.

Properties of the projection problem

Goal: solve numerically, for arbitrary $\mathcal{M}(\mathcal{P})$:

$$\inf_{\nu \in \mathcal{M}(\mathcal{P})} dist(\pi, \nu)$$

Theorem

- If $\mathcal{P} = \Omega^m$, the sequence of solutions $\nu_m \rightharpoonup \pi$.
- If $\mathcal{P} = \mathcal{S}_{\text{MRI}}$, the sequence of solutions $\nu_T \rightharpoonup \pi$.

Numerical implementation

The general construction (similar to finite elements)

- Approximate $\mathcal{M}(\mathcal{P})$ by a subset $\mathcal{N}_p \subset \Omega^p$ of n-point measures:

$$\mathcal{N}_p = \mathcal{M}(\mathcal{Q}_p) = \left\{ \nu = \frac{1}{p} \sum_{i=1}^{p} \delta_{q_i}, \quad \text{for } q = (q_i)_{1 \leq i \leq p} \in \mathcal{Q}_p \right\},$$

  where $\mathcal{Q}_p$ is the discretized version of $\mathcal{P}$.
- Use a projected gradient descent to obtain an approximate projection ν_p^* on $\mathcal{N}_p$:

$$\nu_p^* \in \operatorname*{Arg\,min}_{\nu \in \mathcal{N}_p} \frac{1}{2} \|h \star (\nu - \pi)\|_2^2 ,$$

- Reconstruct an approximation $\nu \in \mathcal{M}(\mathcal{P})$ from ν_p^*.

Numerical Resolution

Variational formulation:

$$\min_{\nu \in \mathcal{N}_p} \frac{1}{2} \|h \star (\nu - \pi)\|_2^2 =$$

$$\min_{q \in \mathcal{Q}_p} J(q) = \underbrace{\frac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{p} H(q_i - q_j)}_{\text{Repulsion potential}} - \underbrace{\sum_{i=1}^{p} \int_\Omega H(x - q_i) d\pi(x)}_{\text{Attraction potential}},$$

where H is defined by $\hat{H}(\xi) = |\hat{h}|^2(\xi)$.

- Repulsion potential: fast k-space coverage
- Attraction potential: right target density π
- Generalization of Poisson disk sampling strategy [Bridson, 2007, Vasanawala et al., 2011]

Numerical Resolution

Projected gradient descents in the non-convex case

Assume that H is differentiable with L-Lipschitz continuous gradient. Consider the following algorithm:

$$q^{(k+1)} \in P_{\mathcal{Q}_p}\left(q^{(k)} - \tau \nabla J(q^{(k)})\right).$$

The sequence $(q^{(k)})_k$ converges to a critical point of the functional J.[Attouch et al., 2013].

Numerical Resolution

Projected gradient descents in the non-convex case

Assume that H is differentiable with L-Lipschitz continuous gradient. Consider the following algorithm:

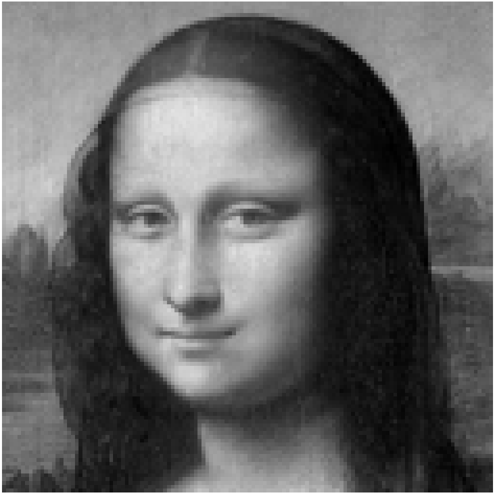
$$q^{(k+1)} \in P_{\mathcal{Q}_p}\left(q^{(k)} - \tau \nabla J(q^{(k)})\right).$$

The sequence $(q^{(k)})_k$ converges to a critical point of the functional J.[Attouch et al., 2013].

Remark

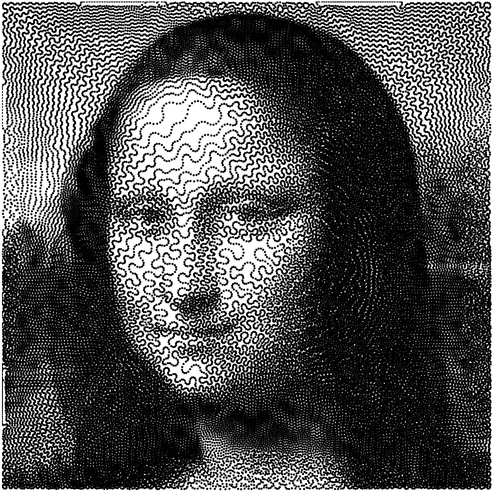
In MRI, $P_{\mathcal{Q}_p} = P_{\mathcal{S}_{MRI}}$!

Example



π = Mona Lisa.

Example



Representation of Mona Lisa by an element of $\mathcal{S}_{\mathrm{MRI}}$.

Application to MRI

Parameters:

- Image size: $n = 256 \times 256$ (resolution: 1 mm isotropic).
- $m = n/4$ decomposed in two segments of 8,192 samples each to make each trajectory shorter than 200 ms (164 ms).
- If sampling time is too large, multi-shot or segmented trajectories might be necessary.

Standard resolution imaging: sampling patterns (1/2)

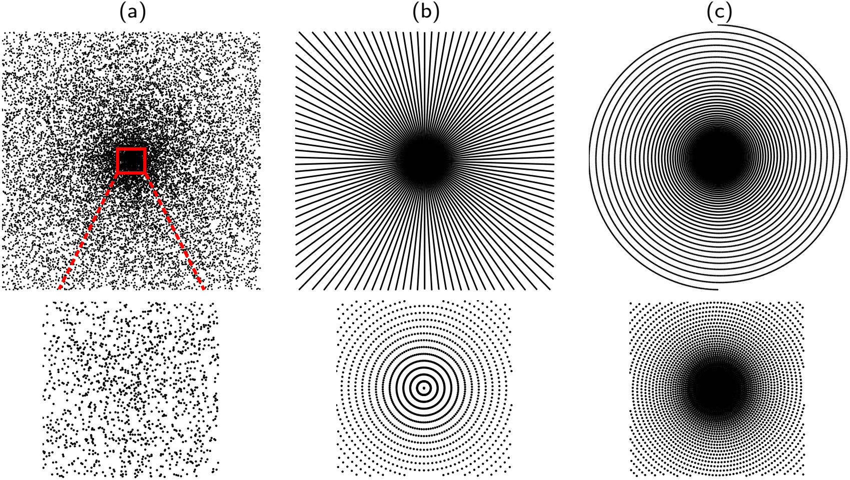


Figure: Classical sampling schemes (a-c). Top row: (a): independent drawing; (b): radial lines ; (c): spiral trajectory. Second row: zooms on the k-space centers.

Standard resolution imaging: sampling patterns (1/2)

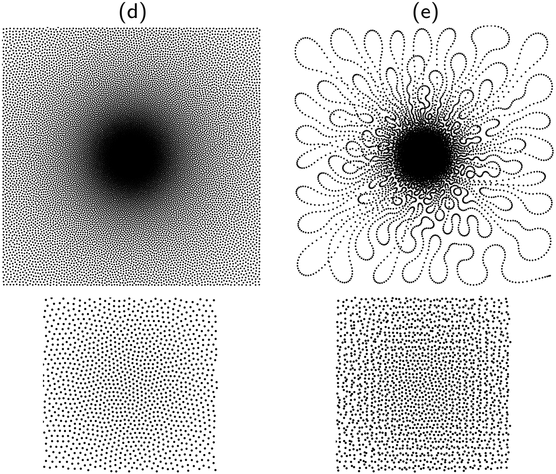
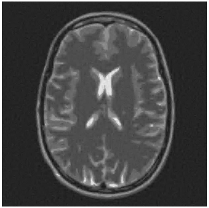


Figure: Sampling schemes obtained with the proposed projection algorithm (d-f). Top row: (d): isolated points; (e): admissible curves for MRI. Bottom row: zooms on the k-space center.

Standard resolution imaging: Reconstructed images

(a) SNR=17.7 dB

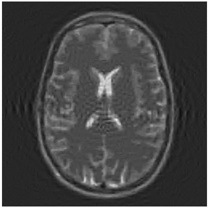


(i.i.d.)

(b) SNR=15.4 dB

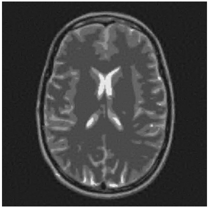
(radial)

(c) SNR=13.2 dB



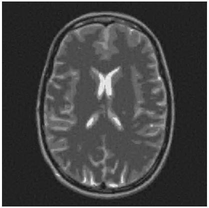
(spiral)

(d) SNR=18.3 dB



(m-points measure)

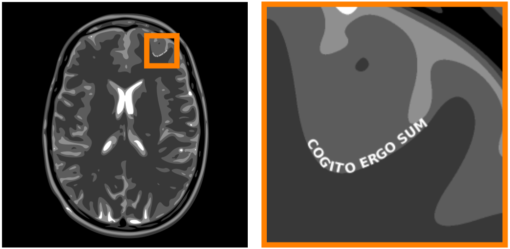
(e) SNR=18.0 dB



(admissible curve for MRI)

Figure: Reconstruction results for the sampling patterns presented.

Very high-resolution imaging



Parameters:

Image size: $n = 2048 \times 2048$ (resolution: 100 μm isotropic). $m = 0.048n$ decomposed in:

- 196 radial lines of 1,024 equispaced samples;
- 8 rotated versions of the same spiral made up by 25,000 samples.
- 8 curves of 25,000 samples each.

Very-high resolution imaging: Competing trajectories (1/2)

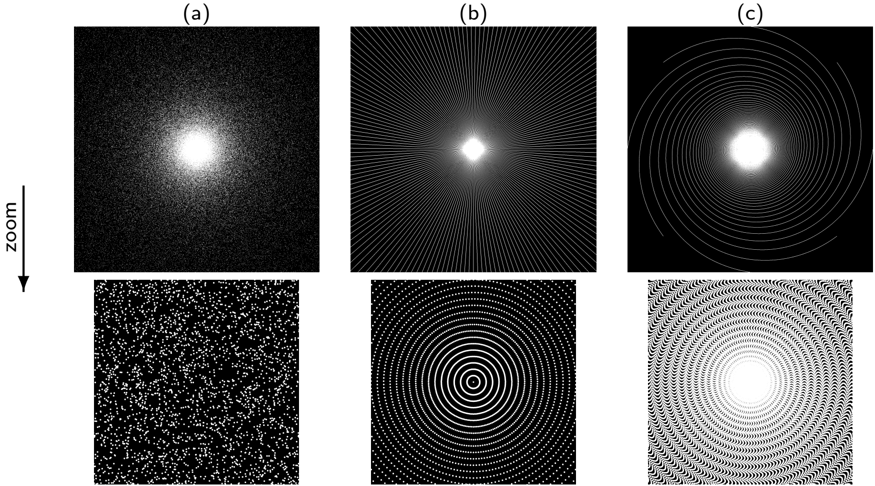


Figure: Standard sampling schemes composed of 200,000 samples. (a): i.i.d. drawings. (b): Radial lines. (c): 4 interleaved spirals.

Very-high resolution imaging: Competing trajectories (2/2)

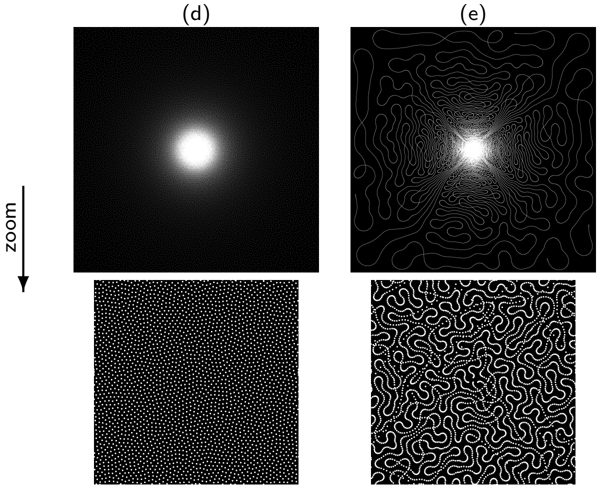
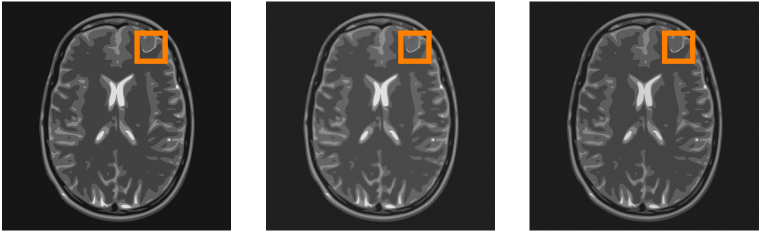


Figure: Sampling schemes yielded by our algorithm and composed of 200,000 samples. (d): Isolated measurements. (e): 4 feasible curves in MRI.

Very-high resolution imaging: Reconstructed images (1/2)

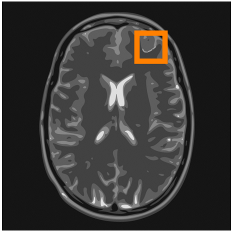
(a) SNR=26.7 dB (b) SNR=20.6 dB (c) SNR=21.0 dB



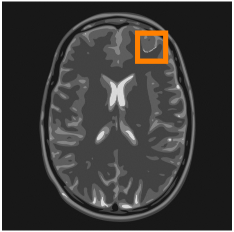
(i.i.d.) (radial) (spiral)

Very-high resolution imaging: Reconstructed images (2/2)

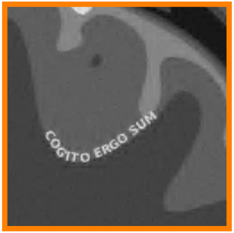
(d) SNR=27.0 dB



(e) SNR=23.5 dB



(m-points measure)



(admissible curve for MRI)

Interim summary

- A global approach by projection on measure sets.
- A convergent projection algorithm for computing local minimizer.
- The method is generic enough to include additional constraints (e.g., multishot).

Conclusion & Outlook (1/2)

Theoretical contributions

- Identification of CS-MRI questions and mathematical formalism of VDS.
- Demonstration of key properties of a VDS with 3 independent contributions:
  - closed form of "optimal sampling distribution" for MRI.
  - CS result for Random walks.
  - TSP sampling with guarantees on the distribution.
- A projection algorithm onto the set of MRI kinematics constraints.
- A measure projection algorithm with several potential applications.

Theoretical outlook

- Fill the gap between heuristic and optimal sampling distributions.
- Obtain theoretical guarantees in Compressed Sensing for trajectories obtained by measure projection...

Conclusion & Outlook (2/2)

Promising results on simulations

- 3D continuous CS-MRI outperforms classical 3D CS-MRI.
- On 2D simulations, curves obtained by projection provide better results compared to spiral or radial by at least 3 dB.

Outlook

- Design 3D trajectories by projection.
- Better manage MRI constraints such as off-resonance effects. Adapt the parameters to different imaging modalities.
- Manage discrepency between prescribed and actual trajectory.
- Implement MR sequences on a 7T scanner at NeuroSpin (PhD thesis of C. Lazarus beginning in 2015).

Codes

- Matlab codes for Cartesian CS-MRI in 2D and 3D, including TSP sampling.
- Toolbox for curve projection.

Journal publications (+ 6 conference papers)

- **Variable density sampling with continuous trajectories.** N. C., P. Ciuciu, J. Kahn and P. Weiss, SIAM Journal on Imaging Science, Vol. 7, Issue 4, pp. 1962–1992 (2014).
- **A projection algorithm for gradient waveforms design in Magnetic Resonance Imaging.** N. C, P. Weiss, J. Kahn and P. Ciuciu. In revision in IEEE Transactions on Medical Imaging
- **A projection method on measures sets.** N. C., P. Ciuciu, J. Kahn and P. Weiss (2015). Submitted à Constructive Approximation (2015)
- **On the generation of sampling schemes for Magnetic Resonance Imaging.** C. Boyer, N. C., P. Ciuciu, J. Kahn, P. Weiss (2015). Submitted soon.
- **A concentration inequality for matrix-valued Markov chains.** In preparation

Thanks

To my advisors :

Philippe Ciuciu

Pierre Weiss

To :

Claire Boyer

Jonas Kahn

And to Benoit Larrat, Sebastien Mériaux, Alexandre Vignaud...

Projection of videos - "$\pi(t)$" distribution

Projection of videos - projection on the set of 1000-point measures

Projection of videos - projection on a set of admissible curves

From Meisje met de Parel (Vermeer, 1665) to ...

Thank you for your attention !

Questions ?

Bibliography I

[Adcock et al., 2013] Adcock, B., Hansen, A., Poon, C., and Roman, B. (2013).
Breaking the coherence barrier: asymptotic incoherence and asymptotic sparsity in compressed sensing.
arXiv preprint arXiv:1302.0561.

[Attouch et al., 2013] Attouch, H., Bolte, J., and Svaiter, B. F. (2013).
Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods.
Mathematical Programming, 137(1-2):91–129.

[Boyer et al., 2015] Boyer, C., Bigot, J., and Weiss, P. (2015).
Compressed sensing with structured sparsity and structured acquisition.
arXiv preprint arXiv:1505.01619.

[Bridson, 2007] Bridson, R. (2007).
Fast Poisson disk sampling in arbitrary dimensions.
In *ACM SIGGRAPH*, volume 2007, page 5.

[Candès and Plan, 2011] Candès, E. J. and Plan, Y. (2011).
A probabilistic and RIPless theory of compressed sensing.
Information Theory, IEEE Transactions on, 57(11):7235–7254.

[Chauffert et al., 2014a] Chauffert, N., Ciuciu, P., Kahn, J., and Weiss, P. (2014a).
Variable density sampling with continuous trajectories. Application to MRI.
SIAM Journal on Imaging Science, 7(4):1962–1992.

[Chauffert et al., 2015] Chauffert, N., Ciuciu, P., Kahn, J., and Weiss, P. (2015).
A concentration inequality for matrix-valued Markov chain.
Preprint.

[Chauffert et al., 2013] Chauffert, N., Ciuciu, P., and Weiss, P. (2013).
Variable density compressed sensing in MRI. Theoretical vs. heuristic sampling strategies.
In *Proc. of 10th IEEE ISBI conference*, pages 298–301, San Francisco, USA.

[Chauffert et al., 2014b] Chauffert, N., Weiss, P., Kahn, J., and Ciuciu, P. (2014b).
A projection algorithm for gradient waveforms design in magnetic resonance imaging.
Technical report. http://chauffertn.free.fr/Publis/2014/GradientWaveformDesign.pdf.

Bibliography II

[Haldar et al., 2011] Haldar, J. P., Hernando, D., and Liang, Z.-P. (2011).
Compressed-sensing MRI with random encoding.
Medical Imaging, IEEE Transactions on, 30(4):893–903.

[Hargreaves et al., 2004] Hargreaves, B. A., Nishimura, D. G., and Conolly, S. M. (2004).
Time-optimal multidimensional gradient waveform design for rapid imaging.
Magnetic Resonance in Medicine, 51(1):81–92.

[Lustig et al., 2007] Lustig, M., Donoho, D. L., and Pauly, J. M. (2007).
Sparse MRI: The application of compressed sensing for rapid MR imaging.
Magnetic Resonance in Medicine, 58(6):1182–1195.

[Lustig et al., 2008] Lustig, M., Kim, S. J., and Pauly, J. M. (2008).
A fast method for designing time-optimal gradient waveforms for arbitrary k-space trajectories.
Medical Imaging, IEEE Transactions on, 27(6):866–873.

[Puy et al., 2012] Puy, G., Marques, J. P., Gruetter, R., Thiran, J., Van De Ville, D., Vandergheynst, P., and Wiaux, Y. (2012).
Spread spectrum magnetic resonance imaging.
Medical Imaging, IEEE Transactions on, 31(3):586–598.

[Teuber et al., 2011] Teuber, T., Steidl, G., Gwosdek, P., Schmaltz, C., and Weickert, J. (2011).
Dithering by differences of convex functions.
SIAM Journal on Imaging Sciences, 4(1):79–108.

[Vasanawala et al., 2011] Vasanawala, S., Murphy, M., Alley, M. T., Lai, P., Keutzer, K., Pauly, J. M., and Lustig, M. (2011).
Practical parallel imaging compressed sensing MRI: Summary of two years of experience in accelerating body MRI of pediatric patients.
In *Proc. ISBI*, pages 1039–1043.

Supplementary material

The key ingredient of the proof

$\mathcal{G}$ a finite graph with N vertices and (X_n) an irreducible and reversible Markov chain (X_n) on $\mathcal{G}$. P its transition matrix with stationary distribution π. $f : \mathcal{G} \to \mathbb{H}^d$, the set of Hermitian matrices of size $d \times d$. Assume that $X_1 \sim q$ and that:

$$\sum_{y \in \mathcal{G}} \pi(y) f(y) = 0 \quad \text{and} \quad \lambda_{\max}(f(y)) \leqslant R, \forall y \in \mathcal{G}.$$

Define :

$$\sigma_n^2 := n \cdot \lambda_{\max}\big(\sum_{y \in \mathcal{G}} \pi(y) f(y)^2\big)$$

Then, for all $t > 0$,

$$\mathbb{P}\left(\lambda_{\max}\left(\sum_{i=1}^{n} f(X_i)\right) \geqslant t\right) \leqslant d \cdot \sup\left(\frac{q_i}{\pi_i}\right) \cdot \exp\left(-\frac{\varepsilon(P) t^2}{4\sigma_n^2 + 2Rt\varepsilon(P)/3}\right).$$

The Travelling Salesman sampler intuition

Let q be the distribution of the "cities".

Intuition

Consider a small hypercube:

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

The Travelling Salesman sampler intuition

Let q be the distribution of the “cities”.

Intuition

Consider a small hypercube:

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

The Travelling Salesman sampler intuition

Let q be the distribution of the "cities".

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

Let q be the distribution of the "cities".

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

Let q be the distribution of the "cities".

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 $q \propto p^{d/(d-1)}$!