Travelling salesman-based variable density sampling

Nicolas Chauffert, Philippe Ciuciu
CEA, NeuroSpin center,
INRIA Saclay, PARIETAL Team
145, F-91919 Gif-sur-Yvette, France
Email: firstname.lastname@cea.fr

Jonas Kahn
Laboratoire Painlevé, UMR 8524
Université de Lille 1, CNRS
Cité Scientifique Bât. M2
59655 Villeneuve d’Ascq Cedex, France
Email: jonas.kahn@math.univ-lille1.fr

Pierre Weiss
ITAV, USR 3505
PRIMO Team,
Université de Toulouse, CNRS
59655 Villeneuve d’Ascq Cedex, France
Email: pierre.weiss@itav-recherche.fr

Abstract—Compressed sensing theory indicates that selecting a few measurements independently at random is a near optimal strategy to sense sparse or compressible signals. This is infeasible in practice for many acquisition devices that acquire samples along continuous trajectories. Examples include magnetic resonance imaging (MRI), radio-interferometry, mobile-robot sampling, ... In this paper, we propose to generate continuous sampling trajectories by drawing a small set of measurements independently at random according to a distribution $\pi$ sampling scheme. It consists of picking a few samples independently at random according to a distribution $\pi$ and joining them using a travelling salesman problem solver. Our contribution lies in the theoretical derivation of the approximate probability density of the initial drawings. Preliminary simulation results show that this strategy is as efficient as independent drawings while being implementable on real acquisition systems.

I. INTRODUCTION

Compressed sensing theory provides guarantees on the reconstruction quality of sparse and compressible signals $x \in \mathbb{R}^n$ from a limited number of linear measurements $\langle a_k, x \rangle_{k \in K}$. In most applications, the measurement or acquisition basis $A = (a_k)_{k \in \{1, \ldots, n\}}$ is fixed (e.g. Fourier or Wavelet basis). In order to reduce the acquisition time, one then needs to find a set $K$ of minimal cardinality that provides satisfactory reconstruction results. It is proved in [1], [2] that a good way to proceed consists of drawing the indices of $K$ independently at random according to a distribution $\pi$ that depends on the sensing basis $A$. This result motivated a lot of authors to propose variable density random sampling strategies (see e.g. [3]–[7]). Fig. 1(a) illustrates a typical sampling pattern used in the MRI context. Simulations confirm that such schemes are efficient in practice. Unfortunately, they can hardly be implemented on real hardware where the physics of the acquisition processes imposes at least continuity of the sampling trajectory and sometimes a higher level of smoothness. Hence, actual CS-MRI solutions rely on adhoc solutions such as random radial or randomly perturbed spiral trajectories to impose gradient continuity. Nevertheless these strategies strongly deviate from the theoretical setting and experiments confirm their practical suboptimality.

In this work, we propose an alternative to the independent sampling scheme. It consists of picking a few samples independently at random according to a distribution $\pi$ and joining them using a travelling salesman problem (TSP) solver in order to design continuous trajectories. The main theoretical result of this paper states that $\pi$ should be proportional to $\pi^{d/(d-1)}$ where $d$ denotes the space dimension (e.g. $d = 2$ for 2D images, $d = 3$ for 3D images) in order to emulate an independent drawing from distribution $\tilde{\pi}$. Similar ideas were previously proposed in the literature [8], but it seems that no author made this central observation.

The rest of this paper is organized as follows. The notation and definitions are introduced in Section II. Section III contains the main result of the paper along with its proof. Section IV shows how the proposed theory can be implemented in practice. Finally, Section V presents simulation results in the MRI context.

II. NOTATION AND DEFINITIONS

We shall work on the hypercube $\Omega = [0, 1]^d$ with $d \geq 2$. Let $m \in \mathbb{N}$. The set $\Omega$ will be partitionned in $m^d$ congruent hypercubes $\{\omega_i\}_{i \in I}$ of edge length $1/m$. In what follows, $\{x_i\}_{i \in \mathbb{N}}$ denotes a sequence of points in the hypercube $\Omega$, independently drawn from a density $\pi : \Omega \mapsto \mathbb{R}_+$. The set of the first $N$ points is denoted $X_N = \{x_i\}_{i \leq N}$. For a set of points $F$, we consider the solution to the TSP, that is the shortest Hamiltonian path between those points. We denote $T(F)$ its length. For any set $R \subseteq \Omega$ we define $T(F, R) = T(F \cap R)$.

We also introduce $C(X_N, \Omega)$ for the optimal curve itself, and $\gamma_N : [0, 1] \mapsto \Omega$ the function that parameterizes $C(X_N, \Omega)$ by moving along it at constant speed $T(X_N, \Omega)$.

The Lebesgue measure on an interval $[0, 1]$ is denoted $\lambda_{[0,1]}$. We define the distribution of the TSP solution as follows.

Definition II.1 The distribution of the TSP solution is denoted $\tilde{\Pi}_N$ and defined, for any Borelian $B$ in $\Omega$ by:

$$\tilde{\Pi}_N(B) = \lambda_{[0,1]} \left( \gamma_N^{-1}(B) \right) .$$

Remark. The distribution $\tilde{\Pi}_N$ is defined for fixed $X_N$. It makes no reference to the stochastic component of $X_N$.

In order to prove the main result, we need to introduce other tools. For a subset $\omega_i \subseteq \Omega$, we denote the length of $C(X_N, \Omega) \cap \omega_i$ as $T_{\omega_i}(X_N, \Omega) = T(X_N, \Omega)\tilde{\Pi}_N(\omega_i)$. Using this definition, it follows that:

$$\tilde{\Pi}_N(B) = \frac{T_B(X_N, \Omega)}{T(X_N, \Omega)}, \quad \forall B.$$  

(1)
Let $T_B(F, R)$ be the length of the boundary TSP on the set $F \cap R$. The boundary TSP is defined as the shortest Hamiltonian tour on $F \cap R$ for the metric obtained from the Euclidean metric by the quotient of the boundary of $R$, that is $d(a, b) = 0$ if $a, b \in \partial R$. Informally, it matches the original TSP while being allowed to travel along the boundary for free. We refer to [9] for a complete description of this concept.

### III. MAIN THEOREM

Our main theoretical result reads as follows.

**Theorem III.1** Define the density $\tilde{\pi} = \frac{e^{(d-1)/d}}{\int_{\Omega} (\pi e^{(d-1)/d}(x))dx}$. Then almost surely with respect to the law $\pi \otimes N$ of the sequence $\{x_i\}_{i \in \mathbb{N}}$ of random points in the hypercube, the distribution $\tilde{\Pi}_N$ converges in distribution to $\tilde{\pi}$:

$$\tilde{\Pi}_N \xrightarrow{d} \tilde{\pi} \quad \pi \otimes N, \text{a.s.} \quad (2)$$

Intuition: Let us first provide a rough intuition of the result since the exact proof is technical. The distribution $\tilde{\Pi}_N$ in a small cube is the relative length of the TSP in this cube. The number of points $N_c$ in the cube is proportional to $\pi$. Approximately, the TSP connects the points with other points in the cube, typically their neighbours, since they are close. Now, the typical distance between two neighbours in the cube is proportional to $N_c^{-1/d}$ or $\pi^{-1/d}$. So that the total length of the TSP in the small cube is proportional to $\pi^{-1/d} = \pi^{1/d} \propto \tilde{\pi}$.

The remainder of this section is dedicated to proving this result. The following proposition is central to obtain the proof:

**Proposition III.2** Almost surely, for all $\omega_i \in \{\omega_i\}_{1 \leq i \leq n^d}$:

$$\lim_{N \to \infty} \tilde{\Pi}_N(\omega_i) = \tilde{\pi}(\omega_i) \quad (3)$$

$$\tilde{\Pi}_N = \frac{\int_{\Omega} \pi^{-1/d}(x)dx}{\int_{\Omega} (\pi^{-1/d}(x)dx)} \pi \otimes N, \text{a.s.} \quad (4)$$

The strategy consists in proving that $T_{\omega_i}(X_N, \Omega)$ tends asymptotically to $T(X_N, \omega_i)$. The estimation of each term can then be obtained by applying the asymptotic result of Beardwood, Halton and Hammersley [10]:

**Theorem III.3** If $R$ is a Lebesgue-measurable set in $\mathbb{R}^d$ such that the boundary $\partial R$ has zero measure, and $\{y_i\}_{i \in \mathbb{N}}$, with $Y_N = \{y_i\}_{i \leq N}$ is a sequence of i.i.d. points from a density $\rho$ supported on $R$, then, almost surely,

$$\lim_{N \to \infty} T(Y_N, R) = \rho(d) \int_R \rho^{-1/d}(x)dx, \quad (5)$$

where $\beta(d)$ depends on the dimension $d$ only.

We shall use a set of classical results on TSP and boundary TSP, that may be found in the survey books [9] and [11].

**Useful lemmas.** Let $F$ denote a set of $n$ points in $\Omega$.

1. The boundary TSP is superadditive, that is, if $R_1$ and $R_2$ have disjoint interiors:

$$T_B(F, R_1 \cup R_2) \geq T_B(F, R_1) + T_B(F, R_2) \quad (6)$$

2. The boundary TSP is a lower bound on the TSP, both globally and on subsets. If $R_2 \subset R_1$:

$$T(F, R) \geq T_B(F, R) \quad (7)$$

$$T_{R_2}(F, R_1) \geq T_B(F, R_2) \quad (8)$$

3. The boundary TSP approximates well the TSP [11, Lemma 3.7]:

$$|T(F, \Omega) - T_B(F, \Omega)| = O(n^{d-2}/(d-1)) \quad (9)$$

4. The TSP in $\Omega$ is well-approximated by the sum of TSPs in a grid of $n^d$ congruent hypercubes [9, Eq. (33)].

$$|T(F, \Omega) - \sum_{i=1}^{n^d} T(F, \omega_i)| = O(n^{d-2}/(d-1)) \quad (10)$$

We now have all the ingredients to prove the main results.

**Proof of Proposition III.2:**

$$\sum_{i \in I} T_B(X_N, \omega_i) \leq T_B(X_N, \Omega) \quad (6)$$

$$\leq T(X_N, \Omega) = \sum_{i \in I} T|_{\omega_i}(X_N, \Omega) \quad (7)$$

$$\leq \sum_{i \in I} T(X_N, \omega_i) + O(N^{d-1}/(d-2)) \quad (10)$$

Let $N_i$ be the number of points of $X_N$ in $\omega_i$. Since $N \leq N_i$, we may use the bound (9) to get:

$$\lim_{N \to \infty} \frac{T(X_N, \omega_i)}{N^{d-1}/d} = \lim_{N \to \infty} \frac{T_B(X_N, \omega_i)}{N^{d-1}/d}. \quad (11)$$

Using the fact that there are only finitely many $\omega_i$, the following equalities hold almost surely:

$$\lim_{N \to \infty} \frac{\sum_{i \in I} T_B(X_N, \omega_i)}{N^{d-1}/d} = \lim_{N \to \infty} \frac{\sum_{i \in I} T(X_N, \omega_i)}{N^{d-1}/d} \quad (11)$$

Since the boundary TSP is a lower bound (cf. Eqs. (8)-(7)) to both local and global TSPs, the above equality ensures that:

$$\lim_{N \to \infty} \frac{T_B(X_N, \omega_i)}{N^{d-1}/d} = \lim_{N \to \infty} \frac{T(X_N, \omega_i)}{N^{d-1}/d} \quad (12)$$

Finally, by the law of large numbers, almost surely $N_i/N \rightarrow \pi(\omega_i) = \int_{\omega_i} \rho(x)dx$. The law of any point $x_i$ conditioned on being in $\omega_i$ has density $\pi(\omega_i)$. By applying Theorem II.3 to the hypercubes $\omega_i$ and $\Omega$ we thus get:

$$\lim_{N \to \infty} \frac{T(X_N, \omega_i)}{N^{d-1}/d} = \beta(d) \int_{\omega_i} \rho(x)^{d-1}/d \, dx \quad \pi \otimes N, \text{a.s.}, \forall i.$$
Combining this result with Eqs. (12) and (1) yields Proposition III.2.

**Proof of Theorem III.1:** Let \( \varepsilon > 0 \) and \( m \) be an integer such that \( \sqrt{dmn^d} < \varepsilon \). Then any two points in \( \omega_i \) are at distance less than \( \varepsilon \).

Using Theorem III.2 and the fact that there is a finite number of \( \omega_i \), almost surely, we get:

\[
\lim_{N \to +\infty} \sum_{i \in I} |\bar{\Pi}_N(\omega_i) - \bar{\pi}(\omega_i)| = 0.
\]

Hence, for any \( N \) large enough, there is a coupling \( K \) of \( \bar{\Pi}_N \) and \( \bar{\pi} \) such that both corresponding random variables are in the same \( \omega_i \) with probability \( 1 - \varepsilon \).

A possible way of constructing \( K \) is to use a genetic algorithm [15]. In our experiments, we used a genetic algorithm [15]. The total number of points sampled is thus \( N_s \approx \left[ \frac{L(N)}{N^d} \right] \) if an arclength parameterization is used. A possible way of obtaining approximately \( \tilde{N} \) samples is thus to set:

\[
N = |\Delta t| L^{-1}(\tilde{N})).
\]

**Algorithm 1:** An algorithm to generate a continuous sampling pattern according to a target density.

**Input:** \( \bar{\pi} : \Omega \mapsto \mathbb{R}_+ \): a target sampling density.  
\( N \): an initial number of drawings.  

**Output:** A continuous sampling curve \( C \).

**begin**

Define \( \pi = \frac{\bar{\pi}^{d/(d-1)}(x)dx}{\int_{\Omega} \bar{\pi}^{d/(d-1)}(x)dx} \).

Draw \( N \) points independently at random according to density \( \pi \).

Link these points with a travelling salesman solver to generate the curve \( C \).

**end**

Applying this algorithm raises various questions: how to choose the target density \( \bar{\pi} \)? How to set the initial number of points \( N \)? Can the travelling salesman problem be solved for millions of points? We give various hints to the previous questions below.

a) **Choosing a density \( \bar{\pi} \):** We believe that this question is still treated superficially in the literature and deserves attention. Various strategies can be considered. A common empirical method consists in learning a density on image databases [4]. In the cases of Fourier measurements, this leads to the use of polynomially decreasing densities from low to high frequencies. The same strategy was proposed in [3] with no theoretical justification. The compressed sensing results allow to derive mathematically founded densities [2], [5]. However, as outlined in [7], an important ingredient is missing for these theories to provide good reconstruction results. The standard CS theory relies on the hypothesis that the signal is sparse, with no assumption on the sparsity structure. This makes the current theoretically founded sampling strategies highly suboptimal. Recent works partially address this problem (see e.g. the review paper [12]). However, to the best of our knowledge, the recent focus is on modifying the reconstruction algorithm, rather than deriving optimal sampling patterns.

b) **Choosing an initial number of points \( N \):** In applications, one usually wishes to sample \( N \) points out of the \( n \) possible ones. One should thus choose \( N \) so that the discretized TSP trajectory contains \( N \) points. This problem is well studied in the TSP literature [10], [13]. Theorem III.3 ensures that the length of the TSP trajectory obtained by drawing \( N \) points should be close to \( L(N) = N^{(d-1)/d} \beta(d) \int_R p^{(d-1)/d}(x)dx \) where \( \beta(d) \) can be evaluated numerically. Concentration results by Talagrand [13] show that this approximation is very accurate for moderate to large values of \( N \). In order to obtain a discrete set of measurements from the continuous trajectory generated by Algorithm 1, we may discretize it with a stepsize \( \Delta t \). The total number of points sampled is thus \( N_s \approx \left[ \frac{L(N)}{N^d} \right] \)

\[
N = |\Delta t| L^{-1}(\tilde{N})).
\]

**V. SIMULATION RESULTS IN MRI**

The proposed sampling algorithm was assessed in a 2D MRI acquisition setup where images are sampled in the 2D Fourier domain and compressible in the wavelet domain. Hence, \( A = F^* \Psi \) where \( F^* \) and \( \Psi \) denote the discrete Fourier and inverse discrete wavelet transform, respectively. Following [7], it can be shown that a near optimal sampling strategy consists of probing \( m \) independent samples of the 2D Fourier plane \((k_x, k_y)\) drawn independently from a target density \( \bar{\pi} \). The image is then reconstructed by solving the following \( l^1 \) problem using a Douglas-Rachford algorithm:

\[
x^* = \arg \min_{A_{mn}x = y} \| x \|_1
\]

where \( A_{mn} \in \mathbb{C}^{m \times n} \) is the sensing matrix, \( x^* \in \mathbb{C}^n \) is the reconstructed image and \( y \in \mathbb{C}^n \) is the collected data. A
compared to the initial drawing from the central distribution for practical applications such as MRI. In this paper, we propose and justify an original two-step approach based on a TSP solver to produce such continuous trajectories. It allows to emulate any variable density sampling strategy and could thus be used in a large variety of applications. In the above mentioned MRI example, this method improves the signal-to-noise ratio by 10 dB compared to more naive approaches and provides results similar to those obtained using unconstrained sampling schemes. From a theoretical point of view, we plan to assess the convergence rate of the empirical law of the travelling salesman trajectory to the target distribution $\pi^{(d-1)/d}$. From a practical point of view, we plan to develop algorithms that integrate stronger constraints into account such as the maximal curvature of the sampling trajectory, which plays a key role in many applications.

**ACKNOWLEDGMENT**

The authors would like to thank the mission pour l’interdisciplinarité from CNRS and the ANR SPHIM3D for partial support of Jonas Kahn’s visit to Toulouse and the CMI Excellence Laboratory for inviting Philippe Ciuciu on an excellence researcher position during winter 2013.

**REFERENCES**


Fig. 1: **Left:** different sampling patterns (with an acceleration factor $r = 5$). **Right:** reconstruction results. From top to bottom: independent drawing from distribution $\tilde{\pi}$ (a), the same followed by a TSP solver (c) and finally independent drawing from distribution $\pi^2$ followed by a TSP solver.

A typical realization is illustrated in Fig. 1(a) which in practice cannot be implemented since MRI requires probing samples along continuous curves. To circumvent such difficulties, a TSP solver was applied to such realization in order to join all samples through a continuous trajectory, as illustrated in Fig. 1(c). Finally, Fig. 1(e) shows a curve generated by a TSP solver after drawing the same amount of Fourier samples from the density $\pi^2$ as underlied by Theorem III.1. In all sampling schemes the number of probed Fourier coefficients was equal to one fifth of the total number (acceleration factor $r = 5$).

Figs. 1(b,d,f) show the corresponding reconstruction results. It is readily seen that an independent random drawing from $\pi^2$ followed by a TSP-based solver yields promising results. Moreover, a dramatic improvement of 10 dB was obtained compared to the initial drawing from $\tilde{\pi}$.

**VI. CONCLUSION**

Designing sampling patterns lying on continuous curves is central for practical applications such as MRI. In this paper, we