

A projection algorithm for gradient waveforms design in Magnetic Resonance Imaging

Nicolas Chauffert, Pierre Weiss, Jonas Kahn and Philippe Ciuciu *Senior Member*

Abstract—Collecting the maximal amount of useful information in a given scanning time is a major concern in Magnetic Resonance Imaging (MRI) to speed up image acquisition. The hardware constraints (gradient magnitude, slew rate, ...), physical distortions (e.g., off-resonance effects) and sampling theorems (Shannon, compressed sensing) must be taken into account simultaneously, which makes this problem extremely challenging. To date, the main approach to design gradient waveform has consisted of selecting an initial shape (e.g. spiral, radial lines, ...) and then traversing it as fast as possible. In this paper, we propose an alternative solution: instead of reparameterizing an initial trajectory, we propose to project it onto the convex set of admissible curves. This method has various advantages. First, it better preserves the density of the input curve which is critical in sampling theory. Second, it allows to smooth high curvature areas making the acquisition time shorter in some cases. We develop an efficient iterative algorithm based on convex programming and propose comparisons between the two approaches. For piecewise linear trajectories, our approach generates a gain of scanning time ranging from 20% (echo planar imaging) to 300% (travelling salesman problem) without degrading image quality in terms of signal-to-noise ratio (SNR). For smoother trajectories such as spirals, our method better preserves the sampling density of the input curve, making the sampling pattern relevant for compressed sensing, contrarily to the reparameterization based approaches.

Index Terms—gradient waveform design, k -space trajectories, variable density sampling, gradient hardware constraints, magnetic resonance imaging.

I. INTRODUCTION

THE advent of new hardware and sampling theories (e.g., Compressed Sensing or CS) provide unprecedented opportunities to reduce acquisition times in MRI. The design of gradient waveforms minimizing the acquisition time while providing enough information to reconstruct distortion-free images is however an important challenge. Ideally, these two concerns (sampling scheme and gradient waveform design) should be addressed simultaneously, but current theoretical results in sampling theories (either Shannon-based or CS-based) do not permit to incorporate complex physical constraints like the starting position or the traversal speed in k -space, despite recent progresses [1–4]

To date, the most widespread technique therefore consists of designing gradient waveforms sequentially: a first step aims to find the trajectory support or at least control points, and a second step essentially builds the gradient waveforms to

traverse this support or linking these control points. The first step either relies on Shannon sampling theorem [3, 4] or on the concept of variable density sampling (VDS) [1, 5, 6]. In Shannon theory, the samples located in the k -space should lie on a Cartesian grid with a sufficiently small grid step size. A typical instance of such schemes is the echo planar imaging (EPI) trajectory. The wealth of trajectories in VDS is constantly increasing and becomes more and more anchored in theory. It initially started with spirals [7, 8] and was progressively enriched with different patterns such as parallel or radial lines [9, 10], noisy spirals [11], Rosette trajectories [12], shell trajectories [13], ... The second step is currently solved by using reparameterization: the goal is to find a feasible waveform traversing the support in the minimum amount of time. This problem can be solved using optimal control [14], convex optimization [15, 16], or optimal interpolation of k -space control points [17]. These simple principles however suffer from potentially severe drawbacks. First, reparameterizing the curve changes the density of samples along the curve. This density is now known to be a key aspect in CS [1, 5, 6, 18], since it directly impacts the number of required measurements to ensure exact recovery (noiseless case) or accurate (noisy case) reconstruction. Second, the challenge of rapid acquisitions is to reduce the scanning time (echo train duration) and limit geometric distortions induced by inhomogeneities of the static magnetic field (B_0) by covering the k -space as fast as possible. The perfect fit to any arbitrary curve (support constraint) may be time consuming, especially in the high curvature parts of the trajectory. In particular, the time to traverse piecewise linear trajectories [1, 19–22] may become too long. Indeed, the magnetic field gradients have to be set to zero at each singular point of such trajectory. To overcome these two limitations, new gradient waveform design methods have to be pushed forward.

A. Contributions

In this paper, we propose an alternative to reparameterization based on a convex optimization formulation. Given any parameterized curve, our algorithm returns the closest curve that fulfills the gradient constraints. The main advantages of the proposed approach are the following: i) the time to traverse the k -space is fixed enabling to find the closest curve in a given time, ii) the distance between the input and output curves is the quantity to be minimized ensuring a low deviation to the original sampling distribution, iii) it is flexible enough to handle additional hardware constraints (e.g., trajectory starting from the k -space center, different kinematic constraints,...)

N. Chauffert and P. Ciuciu are with Inria Parietal Team / NeuroSpin center, CEA Saclay. Contact: nicolas.chauffert@gmail.com

P. Weiss is with PRIMO Team, ITAV, USR 3505, Université de Toulouse.

J. Kahn is with Laboratoire Painlevé, UMR 8524, Université de Lille 1, CNRS.

in the same framework. We propose an efficient first order dual algorithm to solve the resulting problem and provide theoretical guarantees in terms of convergence rate. We also demonstrate through theory and numerical experiments that the distortion to the initial density is minimized compared to the reparameterization approach. We eventually illustrate the performances of our approach on simulations.

B. Paper organization

In Section II, we review the formulation of MRI acquisition, by recalling the gradient constraints and introducing the projection problem. Then, in Section III, it is shown that curves generated by the proposed strategy (initial parameterization followed by the projection onto the set of physical constraints) may be used to design MRI sampling schemes with locally variable densities. In Section IV, we provide an optimization algorithm to solve the projection problem, and estimate its rate of convergence. Next, the behavior of our algorithm is illustrated in Section V on three complementary cases: one popular sampling scheme, namely EPI trajectory and two VDS strategies (travelling salesman problem or TSP-based curves and spirals), yet advertising the usefulness of the proposed approach for practical MRI applications. The pros and cons of our method are discussed in Section VI and concluding remarks are drawn in Section VII.

II. DESIGN OF k -SPACE TRAJECTORIES USING PHYSICAL GRADIENT WAVEFORMS.

In this section, we recall the standard modeling of the acquisition constraints in MRI [14, 16]. We justify the lack of accuracy of current reparameterization methods in the VDS context, and motivate the introduction of a new *projection* algorithm that preserves the sampling density.

A. Sampling in MRI

In MRI, images are sampled in the k -space domain along parameterized curves $s : [0, T] \mapsto \mathbb{R}^d$ where $d \in \{2, 3\}$ denotes the image dimensions. The i -th coordinate of s is denoted s_i . Let $u : \mathbb{R}^d \rightarrow \mathbb{C}$ denote a d dimensional image and $\hat{u}$ be its Fourier transform. Given an image u , a curve $s : [0, T] \rightarrow \mathbb{R}^d$ and a sampling step Δt , the image u shall be reconstructed using the set¹:

$$\mathcal{E}(u, s) = \left\{ \hat{u}(s(j\Delta t)), 0 \leq j \leq \left\lfloor \frac{T}{\Delta t} \right\rfloor \right\}. \quad (1)$$

B. Gradient constraints

The gradient waveform associated with a curve s is defined by $g(t) = \gamma^{-1} \dot{s}(t)$, where γ denotes the gyro-magnetic ratio [16]. The gradient waveforms being obtained by energizing orthogonal gradient coils with electric currents, they are submitted to hardware constraints.

¹For ease of presentation, we assume that the values of u in the k -space correspond to its Fourier transform and we neglect distortions occurring in MRI such as noise. We also neglect the energy decay due to signal relaxation.

1) *kinematic constraints*: Due to physical but also safety (i.e. avoid nerve stimulation) constraints, the electric currents passing through gradient coils have a bounded amplitude and cannot vary too rapidly (slew rate). Mathematically, these constraints read:

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

where $\|\cdot\|$ denotes either the ℓ^∞ -norm defined by $\|f\|_\infty := \max_{1 \leq i \leq d} \sup_{t \in [0, T]} |f_i(t)|$, or the $\ell^{\infty, 2}$ -norm defined by $\|f\|_{\infty, 2} := \sup_{t \in [0, T]} \left(\sum_{i=1}^d |f_i(t)|^2 \right)^{\frac{1}{2}}$. These constraints might be Rotation Invariant (RIV) if $\|\cdot\| = \|\cdot\|_{\infty, 2}$ or Rotation Variant (RV) if $\|\cdot\| = \|\cdot\|_\infty$, depending on whether each gradient coil is energized independently from others or not. The set of kinematic constraints is denoted $\mathcal{S}$:

$$\mathcal{S} := \left\{ s \in (\mathcal{C}^2([0, T]))^d, \|\dot{s}\| \leq \alpha, \|\ddot{s}\| \leq \beta \right\}, \quad (2)$$

where $\alpha = \gamma G_{\max}$ and $\beta = \gamma S_{\max}$.

2) *Additional affine constraints*: Specific MRI acquisitions may require additional constraints, such as:

- Imposing that the trajectory starts from the k -space center (i.e., $s(0) = 0$) to save time and avoid blips. The end-point can also be specified by $s(T) = s_T$, if s_T can be reached during travel time T .
- In the context of multi-shot MRI acquisition, several radio-frequency pulses are necessary to cover the whole k -space. Hence, it makes sense to enforce the trajectory to start from the k -space center at every TR (repetition time)²: $s(m \cdot TR) = 0, 0 \leq m \leq \lfloor \frac{T}{TR} \rfloor$.
- In addition to starting from the k -space center, one could impose the initial speed as for instance: $\dot{s}(0) = 0$.
- To avoid artifacts due to flow motion in the object of interest, gradient moment nulling (GMN) techniques have been introduced in [23] for spin or gradient echo sequences. In terms of constraints, nulling the i^{th} moment reads $\int_{t=0}^{TE} t^i g(t) dt = 0$, where TE denotes the echo time. For example, cancelling the first-order moment compensates the motion of spins moving with constant speed.

Each of these constraints can be modelled by an affine relationship. Hereafter, the set of affine constraints is denoted by $\mathcal{A}$:

$$\mathcal{A} := \left\{ s : [0, T] \rightarrow \mathbb{R}^d, A(s) = v \right\},$$

where v is a vector of parameters in $\mathbb{R}^p$ (p is the number of additional constraints) and A is a linear mapping from the curves space to $\mathbb{R}^p$.

A sampling trajectory $s : [0, T] \rightarrow \mathbb{R}^d$ will be said to be *admissible* if it belongs to the set $\mathcal{S} \cap \mathcal{A}$. In what follows, we assume that this set is non-empty, i.e. $\mathcal{S} \cap \mathcal{A} \neq \emptyset$. Moreover, we assume, without loss of generality, that the linear constraints are independent (otherwise some could be removed).

²corresponding to the delivery of every radio-frequency pulse.

C. Finding an optimal reparameterization

The traditional approach to design an admissible curve $s \in \mathcal{S}$ given an arbitrary curve $c : [0, T] \rightarrow \mathbb{R}^d$ consists of finding a reparameterization p such that $s = c \circ p$ satisfies the physical constraints while minimizing the acquisition time. This problem can be cast as follows:

$$T_{\text{Rep}} = \min T' \text{ such that } \exists p : [0, T'] \mapsto [0, T], c \circ p \in \mathcal{S}. \quad (3)$$

It can be solved efficiently using optimal control [14] or convex optimization [16]. The resulting solution $s = c \circ p$ has the same support as c . This method however suffers from an important drawback when used in the CS framework: it does not provide any control on the density of samples along the curve. For example, for a given curve support shown in Fig. 1(a), we illustrate the new parameterization (keeping the same support) and the corresponding magnetic field gradients (see Fig. 1(b) for a discretization of the curve and (c) for the gradient profile). We notice that the new parameterized curve has to stop at every angular point of the trajectory, yielding more time spent by the curve in the neighbourhood of these points (and more points in the discretization of the curve in Fig. 1(b)). This phenomenon is likely to modify the sampling distribution, as illustrated in Section III.

The next part is dedicated to introducing an alternative method relaxing the constraint of keeping the same support as c .

D. Projection onto the set of constraints

Here, we propose to find the projection of the given input curve c onto the set of admissible curves $\mathcal{S}$:

$$s^* := \operatorname{argmin}_{s \in \mathcal{S} \cap \mathcal{A}} \frac{1}{2} d^2(s, c) = \operatorname{argmin}_{s \in \mathcal{S} \cap \mathcal{A}} \frac{1}{2} \|s - c\|_2^2 \quad (4)$$

where $d^2(s, c) = \|s - c\|_2^2 := \int_{t=0}^T \|s(t) - c(t)\|_2^2 dt$. This method presents important differences compared to the above mentioned optimal control approach: i) the solution s^* and c have different support (see Fig. 1(d)) unless c is admissible; ii) the sets composed of the discretization of c and s^* at a given sampling rate are close to each other (Fig. 1(e)); iii) the acquisition time T is fixed and equal to that of the input curve c . Time to traverse a curve is generally different from optimal reparameterization. In particular for piecewise linear curves, it is generally lower (see Fig. 1(f) where $T < T_{\text{Rep}}$).

In the next section, we explain why the empirical distribution of the samples along the projected curve is closer to that of points lying on the input curve. Also, we illustrate how the parameterization can distort the sampling distribution.

III. CONTROL OF THE SAMPLING DENSITY

Recent works have emphasized the importance of the sampling density [1, 5, 6, 18] in the CS-MRI framework, i.e. in an attempt to reduce the amount of acquired data while preserving image quality at the reconstruction step. The choice of an accurate sampling distribution is crucial since it directly impacts the number of required measurements [24]. In this paper, we will denote by π a distribution defined over the

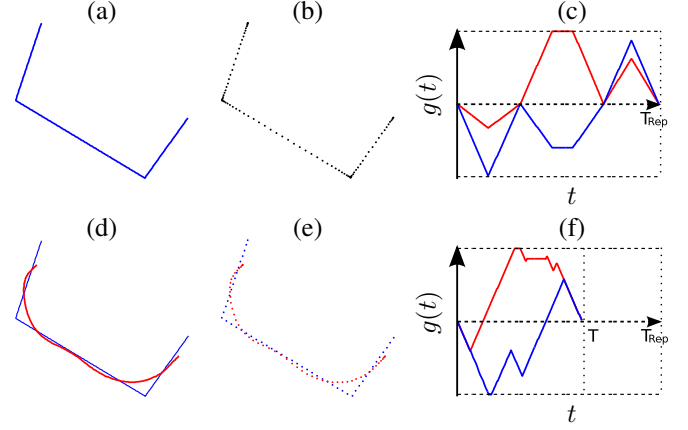


Fig. 1. Comparison of two methods to design gradient waveforms. **Top row:** Optimal control-based parameterization [14]. (a): input curve support. (b): discrete representation of the optimal reparameterization of the curve in $\mathcal{S}$. (c): corresponding gradient waveforms (g_x, g_y). Dashed lines correspond to 0 and $\pm G_{\text{max}}$. **Bottom row:** Illustration of the projection algorithm. (d): same input curve c as in (a) parameterized at maximal speed, and the support of the projected curve s^* onto $\mathcal{S}$. (e): discrete representation of the input and projected curves. (f): corresponding gradient waveforms (g_x, g_y) with the same time scale as in (c): the time to traverse the s^* is 39% shorter.

k -space K . The profile of this distribution can be obtained by theoretical arguments [1, 5, 6, 18] leading to distributions as the one depicted in Fig. 2(a). Some heuristic distributions (e.g., radial) are known to perform well in CS-MRI experiments (Fig. 2(b)). A comparison between these two approaches can be found in [25].

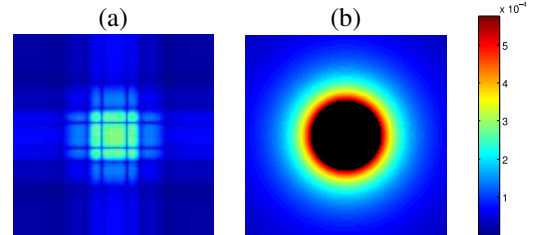


Fig. 2. Examples of 2D sampling distribution. (a): optimal distribution for a Symmlet transform [1, 5]. (b): radial distribution advocated in [1, 6]: $p(k) \propto 1/|k|^2$.

However, designing a trajectory that performs sampling according to a fixed distribution while satisfying gradient constraints is really challenging and has not been addressed so far. The classical approach consists of:

- 1) Finding an input curve (admissible or not) c with good distribution; We provide various strategies to achieve this step in Appendix A-B.
- 2) Estimating the fastest reparameterization of c that belongs to the set of constraints.

In this paper, we suggest to replace the second step by:

- 2') Estimating s^* the projection of c onto the set of constraints, by solving Eq. (4).

We show that step 2') is preferable to step 2) since it better preserves the sampling density (or empirical measure). We begin by showing it through a theoretical study in paragraph III-A and then validate it through numerical experiments

in paragraph III-B. The reader not interested by theoretical arguments can go directly to Subsection III-B.

A. Theoretical study of the density control

To formalize the notion of density, we need to introduce the definition of the empirical distribution of a curve.

Definition 1 (Empirical measure of a curve). *Let λ denote the Lebesgue measure and $\lambda_T = \frac{\lambda}{T}$ denote the Lebesgue measure normalized on the interval $[0, T]$. The empirical measure of a curve $s : [0, T] \mapsto K \subseteq \mathbb{R}^d$ is defined for any measurable set ω of K as:*

$$P_s(\omega) = \lambda_T(s^{-1}(\omega)).$$

This definition means that the mass of a set ω is proportional to the time spent by the curve in ω .

To measure the distortion between an input curve and the projected one, we need to design a distance between measures. In this work, we propose to use the Wasserstein distance W_2 defined hereafter:

Definition 2 (Wasserstein distance W_2). *Let M be a domain of $\mathbb{R}^d$ and $\mathcal{P}(M)$ be the set of measures over M . For $\mu, \nu \in \mathcal{P}(M)$, W_2 is defined as:*

$$W_2(\mu, \nu) = \left(\inf_{\sigma \in \Pi(\mu, \nu)} \int \|x - y\|_2^2 d\sigma(x, y) \right)^{\frac{1}{2}} \quad (5)$$

where $\Pi \subset \mathcal{P}(M \times M)$ denote the set of measures over $M \times M$ with marginals μ and ν on the first and second factors, respectively.

W_2 is a distance over $\mathcal{P}(M)$ (see e.g., [26]). Intuitively, if μ and ν are seen as mountains, the distance is the minimum cost of moving the mountains of μ into the mountains of ν , where the cost is the ℓ_2 -distance of transportation multiplied by the mass moved. Hence, the coupling σ encodes the deformation map to turn one distribution (μ) into the other (ν).

Let us now analyze the distortion between the empirical distribution of the projected curve P_{s^*} and the target distribution π . Since W_2 is a distance between measures, the triangle inequality holds:

$$W_2(P_{s^*}, \pi) \leq \underbrace{W_2(P_c, \pi)}_{\text{Initial distortion}} + \underbrace{W_2(P_{s^*}, P_c)}_{\text{Projection distortion}}. \quad (6)$$

The deviation is controlled by two terms: the *initial distortion* term $W_2(P_c, \pi)$ and the *projection distortion* term $W_2(P_{s^*}, P_c)$. The first term depends of the choice of the input curve c . This choice is crucial but is out the scope of this paper since it is not directly related to gradient waveform design. We still show in Appendix A that this term can be controlled precisely in a few cases of interest (spiral, TSP).

We are now interested in controlling the *Projection distortion* term $W_2(P_{s^*}, P_c)$. The following proposition shows that the W_2 distance between the empirical distributions of the input and output curves (c and s^* , respectively) is controlled by the quantity $d(s^*, c)$ to be minimized when solving Eq. (4).

Proposition 1. *For any two curves s and $c : [0, T] \rightarrow \mathbb{R}^d$:*

$$W_2(P_s, P_c) \leq d(s, c).$$

Proof. In terms of distributions, the quantity $d(s, c)$ reads:

$$d^2(s, c) = \int_{M \times M} \|x - y\|_2^2 d\sigma_{s,c}(x, y) \quad (7)$$

where $\sigma_{s,c}$ is the coupling between the empirical measures P_s and P_c defined for all couples of measure sets $(\omega_1, \omega_2) \in M^2$ by $\sigma_{s,c}(\omega_1, \omega_2) = \frac{1}{T} \int_{t=0}^T 1_{\omega_1}(s(t)) 1_{\omega_2}(c(t)) dt$, where 1_ω denote the indicator function of ω . The choice of this coupling is equivalent to choosing the transformation map as the association of locations of $c(t)$ and $s(t)$ for every t . We notice that the quantity to be minimized in Eq. (7) is an upper bound of $W_2(P_s, P_c)^2$, with the specific coupling $\sigma_{s,c}$. $\square$

To sum up, solving the projection problem (4) and finding s^* amounts to minimizing an upper-bound of $W_2(P_{s^*}, \pi)$, the Wasserstein distance between the target density π and the empirical distribution P_{s^*} , if we neglect the influence of the *initial parameterization* c . In some sense, our projection algorithm is therefore the best way to obtain a feasible curve and to preserve the input curve empirical measure. As will be seen in the next paragraph, densities are indeed much better preserved using projections than reparameterizations.

B. Numerical study of the density control

Next, we performed simulations to show that the sampling density is better preserved using our algorithm compared to the optimal control approach. For doing so, we use travelling salesman-based (TSP) sampling trajectories [1, 20], which are an original way to design random trajectories which empirical distribution is any target density π such as the one represented in Fig. 2(a). 10,000 such independent TSP were drawn and parameterized with arc-length: note that these parameterizations are not admissible in general. Then, we sampled each trajectory at constant rate Δt (as in Fig. 3 (top-row, left)), to form an histogram depicting the empirical distribution shown in Fig. 3 (top-row, center). The latter was eventually compared to π in Fig. 3 (top-row, right). It is worth noting that the error was actually not close to zero, since the convergence result enounced in [1] is asymptotic, i.e. when the length of the TSP curve tends to infinity whereas the latter remains bounded in this experiment.

In Fig. 3 (second row), we show that the classical reparameterization technique [14] leads to a major distortion of the sampling density, because of its behavior on the angular points already illustrated in Fig. 1(b). Then, we considered three constant speed parameterizations and projected them onto the same set of constraints ($G_{\max} = 40 \text{ mT.m}^{-1}$ and $S_{\max} = 150 \text{ mT.m}^{-1}.\text{ms}^{-1}$). Among these three initial candidates, we started by using an initial parameterization with low velocity (10 % of the maximal speed $\gamma G_{\max}$ with $\gamma = 42.576 \text{ MHz.T}^{-1}$ for proton imaging), which projection fits the sampling density quite well. Then, we increased the velocity to progressively reach 50 % and even 100 % of the maximal speed. The distortion of the sampling density

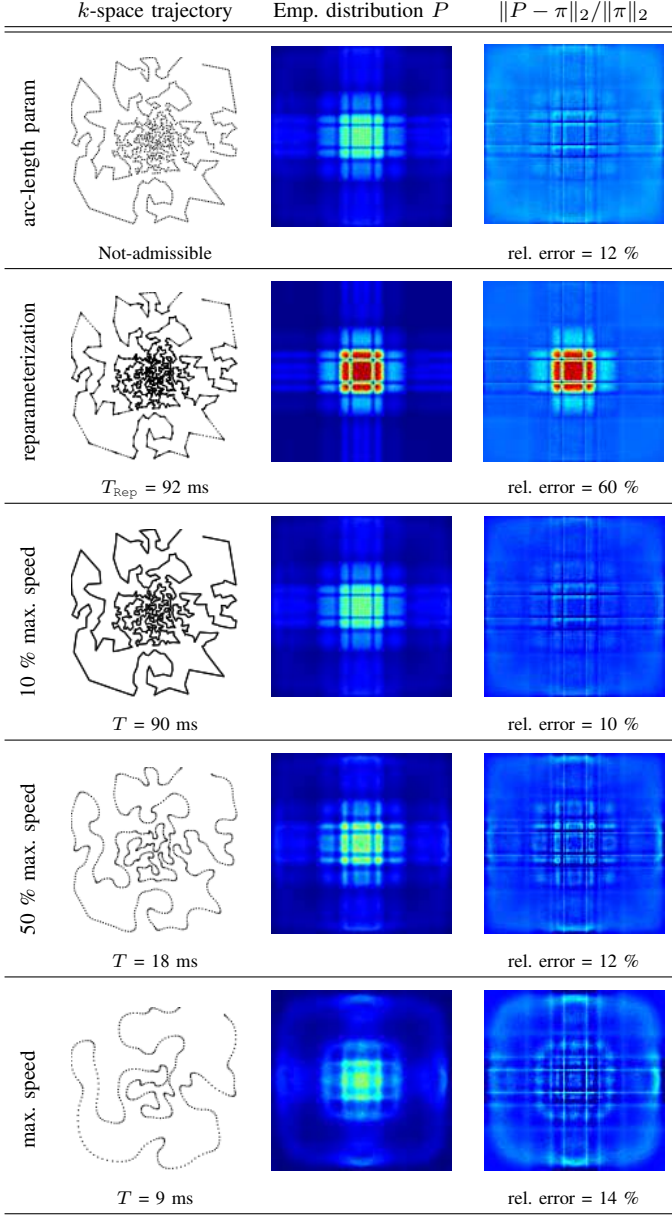


Fig. 3. Illustration of TSP trajectories traversed with arc-length parameterization (top row), optimal control (second row) and with our projection algorithm (rows 3-5). Columns represent the k -space trajectory (left), the empirical distribution P (center) and the difference with the target distribution π shown in Fig. 2(a) (right). At the bottom, the relative error $\|P - \pi\|_2 / \|\pi\|_2$ between the two is reported.

of the projected curve increased, but remained negligible in contrast to what we observed for the exact reparameterization. Hence, this example illustrates that starting from a continuous trajectory whose an empirical sampling distribution is close to the target π , our projection algorithm yields feasible gradient waveforms while sampling the k -space along a discretized trajectory whose empirical density is close to π too.

IV. FINDING FEASIBLE WAVEFORMS USING CONVEX OPTIMIZATION

Since the set of constraints $\mathcal{S} \cap \mathcal{A}$ is convex, closed and non-empty, Problem (4) always admits a unique solution. Even

though $\mathcal{S}$ has a rather simple structure³, it is unlikely that an explicit solution to Problem (4) can be found. In what follows, we thus propose a numerical algorithm to compute the projection.

Problem discretization: A discrete-time curve s is defined as a vector in $\mathbb{R}^{n \cdot d}$ where n is the number of time points. Let $s(i) \in \mathbb{R}^d$ denote the curve location at time $(i-1)\delta t$ with $\delta t = \frac{T}{n-1}$. The discrete-time derivative $\dot{s} \in \mathbb{R}^{n \cdot d}$ is defined using first-order differences:

$$\dot{s}(i) = \begin{cases} 0 & \text{if } i = 1, \\ (s(i) - s(i-1))/\delta t & \text{if } i \in \{2, \dots, n\}. \end{cases}$$

In the discrete setting, the first-order differential operator can be represented by a matrix $\dot{\mathbf{M}} \in \mathbb{R}^{n \cdot d \times n \cdot d}$, i.e. $\dot{s} = \dot{\mathbf{M}}s$. We define the discrete second-order differential operator by $\ddot{\mathbf{M}} = -\dot{\mathbf{M}}^* \dot{\mathbf{M}} \in \mathbb{R}^{n \cdot d \times n \cdot d}$.

An efficient projection algorithm: The discrete primal problem we consider is the same as (4) except that all objects are discretized. It reads:

$$\min_{s \in \mathcal{S} \cap \mathcal{A}} \frac{1}{2} \|s - c\|_2^2, \quad (\mathcal{P})$$

where $\mathcal{S} := \{s \in \mathbb{R}^{n \cdot d}, \|\dot{\mathbf{M}}s\| \leq \alpha, \|\ddot{\mathbf{M}}s\| \leq \beta\}$ with all norms discretized, and $\mathcal{A}$ are the discrete counterparts of $\mathcal{S}$ and $\mathcal{A}$, respectively. Next, the main idea is to take advantage of the structure of the dual problem of $\mathcal{P}$ to design an efficient projection algorithm. The following proposition specifies this dual problem and the primal-dual relationships.

Proposition 2. Let $\|q\|_* := \sup_{\|s\| \leq 1} \langle s, q \rangle$ denote the dual norm of $\|\cdot\|$. The following equality holds:

$$\min_{s \in \mathcal{S} \cap \mathcal{A}} \frac{1}{2} \|s - c\|_2^2 = \sup_{q_1, q_2 \in \mathbb{R}^{n \cdot d}} F(q_1, q_2) - \alpha \|q_1\|_* - \beta \|q_2\|_*, \quad (8)$$

where

$$F(q_1, q_2) = \min_{s \in \mathcal{A}} \langle \dot{\mathbf{M}}s, q_1 \rangle + \langle \ddot{\mathbf{M}}s, q_2 \rangle + \frac{1}{2} \|s - c\|_2^2. \quad (9)$$

Moreover, let (q_1^*, q_2^*) denote any minimizer of the dual problem (8), s^* denote the unique solution of the primal problem (??) and $s^*(q_1^*, q_2^*)$ denote the solution of the minimization problem (9). Then $s^* = s^*(q_1^*, q_2^*)$.

Proof. The proof is given in Appendix B. $\square$

The following proposition gives an explicit expression of $s^*(q_1^*, q_2^*)$.

Proposition 3. The minimizer

$$s^*(q_1^*, q_2^*) = \arg \min_{s \in \mathcal{A}} \langle \dot{\mathbf{M}}s, q_1^* \rangle + \langle \ddot{\mathbf{M}}s, q_2^* \rangle + \frac{1}{2} \|s - c\|_2^2$$

is given by

$$s^*(q_1, q_2) = z + A^+(v - Az), \quad (10)$$

³it is just a polytope when the ℓ^∞ -norm is used.

where $\mathbf{A} \in \mathbb{R}^{p \times n-d}$ is a matrix encoding the affine constraints, and $\mathbf{A}^+ = \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}$ denotes its pseudo-inverse⁴. In addition, $\mathbf{z} = \mathbf{c} - \dot{\mathbf{M}}^*\mathbf{q}_1 - \ddot{\mathbf{M}}^*\mathbf{q}_2$.

Proof. The proof is given in Appendix C. $\square$

Let us now analyse the smoothness properties of F .

Proposition 4. *Function $F(\mathbf{q}_1, \mathbf{q}_2)$ is concave differentiable with gradient given by*

$$\nabla F(\mathbf{q}_1, \mathbf{q}_2) = - \begin{pmatrix} \dot{\mathbf{M}}\mathbf{s}^*(\mathbf{q}_1, \mathbf{q}_2) \\ \ddot{\mathbf{M}}\mathbf{s}^*(\mathbf{q}_1, \mathbf{q}_2) \end{pmatrix}. \quad (11)$$

Moreover, the gradient mapping ∇F is Lipschitz continuous with constant $L = |||\dot{\mathbf{M}}^*\dot{\mathbf{M}} + \ddot{\mathbf{M}}^*\ddot{\mathbf{M}}|||$, where $|||\mathbf{M}|||$ denotes the spectral norm of $\mathbf{M}$.

Proposition 4 is a direct application of [27, Theorem 1] (see also [28]). The dual problem (8) has a favorable structure for its optimization: it is the sum of a differentiable convex function $\tilde{F}(\mathbf{q}_1, \mathbf{q}_2) = -F(\mathbf{q}_1, \mathbf{q}_2)$ and of a simple convex function $G(\mathbf{q}_1, \mathbf{q}_2) = \alpha\|\mathbf{q}_1\|_* + \beta\|\mathbf{q}_2\|_*$. The sum $\tilde{F} + G$ can thus be minimized efficiently using accelerated proximal gradient descents [29] (see Algorithm 1 below).

Algorithm 1: Projection algorithm in the dual space

Input: $\mathbf{c} \in \mathbb{R}^{n-d}$, $\alpha, \beta > 0$, n_{it} .

Output: $\tilde{\mathbf{s}} \in \mathbb{R}^{n-d}$ an approximation of the solution $\mathbf{s}^*$.

Initialize $\mathbf{q}^{(0)} = (\mathbf{q}_1^{(0)}, \mathbf{q}_2^{(0)})$ with $\mathbf{q}_i^{(0)} = 0$ for $i = 1, 2$.

Set $\mathbf{y}^{(0)} = \mathbf{q}^{(0)}$.

Set $\ell = 1/L$.

for $k = 1 \dots n_{it}$ **do**

$$\begin{cases} \mathbf{q}^{(k)} = \text{prox}_{\ell G}(\mathbf{y}^{(k-1)} - \ell \nabla \tilde{F}(\mathbf{y}^{(k-1)})) \\ \mathbf{y}^{(k)} = \mathbf{q}^{(k)} + \frac{k-1}{k+2}(\mathbf{q}^{(k)} - \mathbf{q}^{(k-1)}) \end{cases}$$

return $\tilde{\mathbf{s}} = \mathbf{s}^*(\mathbf{q}_1^{(n_{it})}, \mathbf{q}_2^{(n_{it})})$.

Moreover, by combining the convergence rate results of [29, 30] and some convex analysis (see Appendix D), we obtain the following convergence rate:

Theorem 1. *Algorithm 1 ensures that the distance to the minimizer decreases as $\mathcal{O}(\frac{1}{k^2})$:*

$$\|\mathbf{s}^{(k)} - \mathbf{s}^*\|_2^2 \leq \frac{2L\|\mathbf{q}^{(0)} - \mathbf{q}^*\|_2^2}{k^2}. \quad (12)$$

V. NUMERICAL EXPERIMENTS

To compare our results with [14], we used the same gradient constraints. In particular, the maximal gradient norm $G_{\max}$ was set to 40 mT.m⁻¹, and the slew-rate $S_{\max}$ to 150 mT.m⁻¹.ms⁻¹. We assume that the constraints are Rotation Invariant (RIV). The image field of view (FOV) is assumed to be 20 cm and $K_{\max} = N/(2 \cdot \text{FOV})$ where N is the target spatial grid size for image reconstruction. The sampling rate was fixed to $\Delta t = 4 \mu\text{s}$ except for spiral imaging. For the ease of trajectory representation, we limit

⁴Since the constraints are supposed to be linearly independent, $\mathbf{A}^+$ is well-defined.

ourselves to 2D sampling curves, although our algorithm encompasses the 3D setting.

The Matlab codes embedding the projection algorithm as well as the scripts to reproduce the results depicted hereafter are available at <http://chauffertn.free.fr/codes.html>. Hereafter, the supplementary affine constraints (e.g., nulling moments) are not taken into account. However, they have been implemented in the code so that every end-user can play with. Simulations were performed on a Linux Ubuntu (64 bits) workstation with an Intel Xeon(R) CPU E5-2630 v2 @2.60GHz processor and 64 GB of RAM. The computation time required to run the experiments range from 2 min. (EPI with 17,225 points) to 4 min. (TSP trajectory with 45,000 points) and the number of iterations of Algorithm 1 to achieve convergence was 15,000, to satisfy $\|(s^{(k+1)} - s^{(k)})/s^{(k)}\| < 10^{-3}$.

To measure the impact of the proposed projection algorithm and compare it with the optimal reparameterization, we also performed image reconstruction and computed image quality in terms of Signal-to-Noise-Ratio (SNR). To this end, we performed simulations by starting from a high-resolution $N \times N$ MRI phantom ($N = 1024$) depicted in Fig. 4. Next, we massively undersampled its Fourier transform by the two competing sampling strategies and analyzed image quality after non-Cartesian reconstruction. For the sake of self-containedness, all investigated trajectories are depicted in Fig. 7 and quantitative results corresponding traversal times and SNR of reconstructed images are reported in Tab. I. In

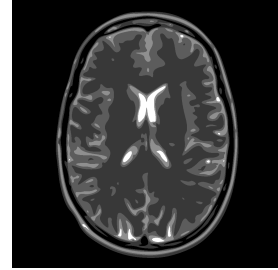


Fig. 4. MRI phantom of size $N \times N$ ($N = 1024$) used for the experiments.

what follows, we first discuss the results of our method in the context of classical (piecewise linear) EPI trajectory. Then, we illustrate the behavior of our algorithm on two VDS: classical spiral (smooth) trajectories and TSP-based (piecewise linear) trajectories.

A. EPI trajectories

EPI trajectories are a classical way of probing the k -space on a 2D-Cartesian grid. We compared a standard EPI with ramp-sampling (a sample was measured every Δt from $t = 0$ to T_{rep}) on $N = 128$ lines, parameterized with optimal control and a trajectory that traverses the k -space at constant speed (70% of the maximal gradient intensity), projected onto $\mathcal{S}$ using our algorithm.

As shown in Fig. 7 (third row), the projected trajectory has a smaller support than standard EPI. In particular, the resolution in the readout direction is slightly decreased. However, the time to traverse k -space is shorter ($T = 68.9$ ms) using our

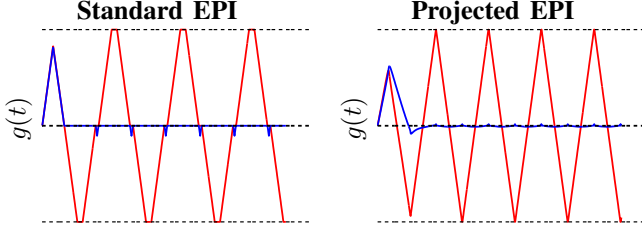


Fig. 5. Comparison between magnetic field gradients $g = (g_x, g_y)$ during the first 5 ms for standard EPI trajectory (left) and projected EPI (right).

algorithm as compared to the EPI trajectory ($T_{\text{Rep}} = 89.6$ ms). To provide a better insight on this acceleration factor, we depict in Fig. 5 the first 5 ms of the gradient waveforms for the two approaches. The corresponding acquired lines are colored in red in Fig. 7. While standard EPI is able to acquire 6.5 lines (Fig. 5-left) in this amount of time, the projected trajectory achieves the extended coverage of 8.5 lines (Fig. 5-right). Indeed, gradient blips are smoothed providing a substantial time reduction. In terms of image quality, we observed that the degradation of resolution along one direction has no significant impact since the SNR of reconstructed image is higher for the projected trajectory compared to standard EPI (Tab. I). Hence, in the EPI context, the projection algorithm allows us to traverse the k -space faster without degrading the image quality.

B. Spiral trajectories

The case of spiral trajectories is more tricky as explained below. For any radial density π , there exists a spiral that performs k -space sampling according to π . This trajectory is parameterized by $c(t) = r(t/T) \exp(i2\pi n\theta(t/T))$ and thus controlled by its time-varying modulus $r(t)$ and phase $\theta(t)$ and by the number of revolutions $n \in \mathbb{R}_+$ over the fixed traversal time T . The relation between π and $r(\cdot)$ is given in Appendix A, Eq. (14), hence the choice of $r(\cdot)$ determines π , whereas $\theta(\cdot)$ and n control the shape of the spiral. For fixed T and $r(\cdot)$, finding $\theta(\cdot)$ and n such that the spiral is optimal in the sense that the kinematics constraints $\mathcal{S}$ are saturated, is an open issue. Indeed, in the literature [31], it has been shown that different types of gradient parameterizations may yield different sampling patterns, hence various π . However, to the best of our knowledge, the inverse problem which consists of inferring the parameterization from the target density π , has never been solved.

Here, we provide a partial solution that relies on two ingredients: first, setting the function $r(\cdot)$ according to (14) and second choosing a constant angular speed ω such that $\theta(t) = \omega t$. This approach actually remains suboptimal since considering a constant ω imposes too low gradient magnitudes at the beginning of the trajectory (i.e. for the k -space center). The pair (ω, n) must satisfy the constraints in $\mathcal{S}$. For instance, to saturate the magnitude gradient constraint one may choose (ω, n) such that: $2\pi n\omega K_{\text{max}} = \gamma G_{\text{max}}$.

In our experiment, we adopted this strategy for the above defined G_{max} and the selected K_{max} (see Fig. 7). We also set $T = 200$ ms and $\Delta t = 24$ μ s, in order to meet an

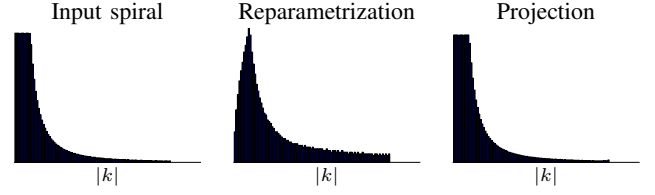


Fig. 6. Decay of the spiral for an input spiral with density $\pi(k) \propto 1/|k|^2$. Histogram of the values $r(t)$ for input spiral (not admissible), optimal reparametrization, and projection.

additional memory size constraint⁵. In Fig. 6, we illustrate how the sampling density $\pi(k) \propto 1/|k|^2$ is impacted after optimal control reparametrization whereas it is preserved when applying our projection algorithm. The histogram peak associated with the reparameterization is shifted to the right i.e. towards high frequencies meaning that the low frequencies are undersampled. This is the direct consequence of using a too fast traversal speed (see Fig. 7: the samples of the spiral are more spaced in the reparameterization scenario). The traversal time of the spiral is indeed $T_{\text{Rep}} = 42$ ms with reparametrization and T after projection (Tab. I). This also explains the significant difference of image quality by almost 5 dB in favor of the projection approach.

On the contrary, if the initial parameterization is not admissible (ω too large), we observed that the output trajectory of the projection algorithm concentrates on concentric circles corresponding to the maximal speed allowed by the gradient magnitude constraint (results not shown). In contrast, the optimal reparameterization is not impacted since it only depends on the support of the spiral. Hence, the choice of the initial parametrization is crucial for spiral imaging, and it seems that neither our algorithm nor reparameterization technique provides a universal answer to the issue of spiral sampling in MRI.

C. TSP sampling

In the same spirit of Fig. 3, we performed numerical experiments using a TSP trajectory [1, 20]. To perform a comparison at constant traversal time, we draw two sets of 4,500 and 45,000 “cities” in order to design a short and a long trajectory (Fig. 7 top row-right). The short curve is traversed with optimal reparameterization in a given time $T_{\text{Rep}} = 160$ ms (Fig. 7 middle row-right). The longer curve is parameterized at constant speed such that $T = T_{\text{Rep}}$, that corresponds to 25 % of the maximal speed γG_{max} . Then, this parameterization is projected onto $\mathcal{S}$ (Fig. 7 last row).

We notice that for a fixed time, the curve obtained with our algorithm provides a larger k -space coverage compared to optimal reparameterization. The main reason is that TSP trajectories embody singular points that require the gradients to be set to zero for each of them. Therefore, a sampling trajectory with singular points is time consuming. The main advantage of our algorithm is that the trajectory can be smoothed around these points, which saves a lot of acquisition time. In terms of image quality, the main consequence is that

⁵the buffer size of the analog-to-digital converter is 8912 in standard MRI scanners.

TABLE I
COMPARISON BETWEEN TRAVERSAL TIME AND RECONSTRUCTION SNR
FOR OPTIMAL REPARAMETERIZATION AND PROJECTION

Resolution		EPI	Spiral	TSP-based
		128	512	512
Optimal reparam.	T_{rep} (ms)	89.6	42.4	180
	SNR (dB)	20.1	7.9	11.3
Projection	T (ms)	68.9	200	180
	SNR (dB)	21.4	12.7	14.5

our projection algorithm outperforms the reparameterization approach by 3.2 dB.

This example demonstrates that existing methods do not permit to implement TSP-based sequences in many MRI modalities (e.g., short TE for a small number of “cities”), since the time to collect data can be larger than any realistic repetition time (here, the traversal time of the longer trajectory based on optimal reparameterization would require 1.1 s). In contrast, our method enables traversal of such curves in a reasonable time which can be tuned according to the image weighting (T_1 , T_2 or proton density).

D. Nonlinear image reconstruction

To demonstrate the effectiveness of the proposed approach not only for gradient waveform design but also for imaging, we performed nonlinear image reconstruction as prescribed in the CS context [9, 32]. Additionally, to fully take advantage of the projection algorithm, our reconstruction scheme was non-Cartesian. Hence, we used non-uniform Fourier transforms [33] to compute the k -space values out of the grid (on locations $s(i), i = 1, \dots, n$). For comparison purposes, we started from a high resolution phantom u (see Fig. 4) that was used to compute the sets $\mathcal{E}(u, s_{\text{rep}})$ and $\mathcal{E}(u, s_{\text{proj}})$. The latter are given by Eq. (1) where s_{rep} and s_{proj} denote the optimal reparameterization and projected trajectory, respectively. Next, the images were reconstructed using non-linear ℓ_1 penalization, i.e.:

$$u^* = \arg \min_{\tilde{u}} \left\| \sum_{i=1}^n (\widehat{u - \tilde{u}})(s(i)) \right\|_2^2 + \lambda \|\Phi \tilde{u}\|_1 \quad (13)$$

where Φ is a sparsifying transform (here Daubechies wavelets), λ is a hyper-parameter, and s is either s_{rep} or s_{proj} . The minimizer of (13) was computed using accelerated proximal gradient descent ([29], FISTA [34]). The image solutions (u_{rep}^* and u_{proj}^*) were then compared to a low resolution version of the $N \times N$ phantom where N ranged from 128 to 512 to compute SNR values in Tab. I. On top of this, it is worth noting that we could still improve the SNR of reconstructed images by resorting either to more redundant decompositions such as tight frames [35] or even by learning dictionaries over which the image can be sparsely decomposed [36]. However, this aspect is beyond the scope of our current proof of concept.

VI. DISCUSSION

In this paper, it has been shown that our projection algorithm has potential interests for smoothing sampling curves such as

EPI or TSP-based trajectories. In this context, our algorithm delivers physically plausible trajectories while drastically reducing the traversal time and improving image quality. This is a direct consequence of its ability to project any piecewise linear initial parameterization onto admissible trajectories with different support. In applications such as functional MRI, this offers the opportunity to shorten the echo train length and then to optimally select the effective echo time so as to maximize the blood oxygenated level-dependent contrast (e.g., $TE = 30$ ms at 3T). Finally, our method can be used in addition with other acceleration methods such as parallel imaging [37, 38] or simultaneous multi-slice imaging technique [39].

Beyond this context, our projection method provides a more accurate control of the sampling density as shown for variable density sampling on spirals. This has a positive impact on image reconstruction quality at the expense of longer traversal times. Setting a fair trade-off between image quality and acquisition time is a usual concern in MRI that may depend on the application at hand (e.g., static vs dynamic imaging). Interestingly, our algorithm prescribes the acquisition time a priori what actually provides the practitioner with an effective control on such trade-off. As we illustrated on TSP sampling, this acquisition time is tightly linked to sampling accuracy with respect to the target density. Hence, our approach clearly compensates a major drawback of reparameterization methods that do not offer such control: the traversal time can be too fast hence an insufficient number of data are collected (spiral case), or too slow and not implementable (TSP-based sampling case).

Usually in MRI acquisition, a number of trajectories are interleaved to provide enough k -space samples. So far, we have not demonstrate the optimization of a set of interleaves except that the segmentation of the trajectory can directly enter in our global optimization problem through affine constraints if the interleaving sequence is thought of as a way of crossing the k -space center at evenly spaced time intervals. More generally, we can prove theoretically and practically that if the combination of two input trajectories provides a good k -space coverage, the combination of the two projected curves admits the same property. The theoretical argument comes from the following observation: if we consider two interleaves c_1 and c_2 and apply our projection method by searching for s_1 and s_2 from initial candidates c_1 and c_2 respectively, we actually control an upper bound of $W_2(P_{s_1}, P_{c_1}) + W_2(P_{s_2}, P_{c_2})$. Practical illustration of this property is available in our Matlab toolbox.

On the other hand, our projection method has also limitations. In particular, the projected trajectory strongly depends on the initial parameterization. As we illustrated, parameterizing a given initial curve at different speeds provides very different projected trajectories. This clearly calls for extensions that might iterate until convergence between the two key steps, namely approximating the target density and finding an admissible trajectory from this approximation [40]. In such generalizations, the first step can be seen as a density-consistency stage where the sampled k -space locations might change from one iteration to the next to fit a target density. We believe that this idea might become the most important

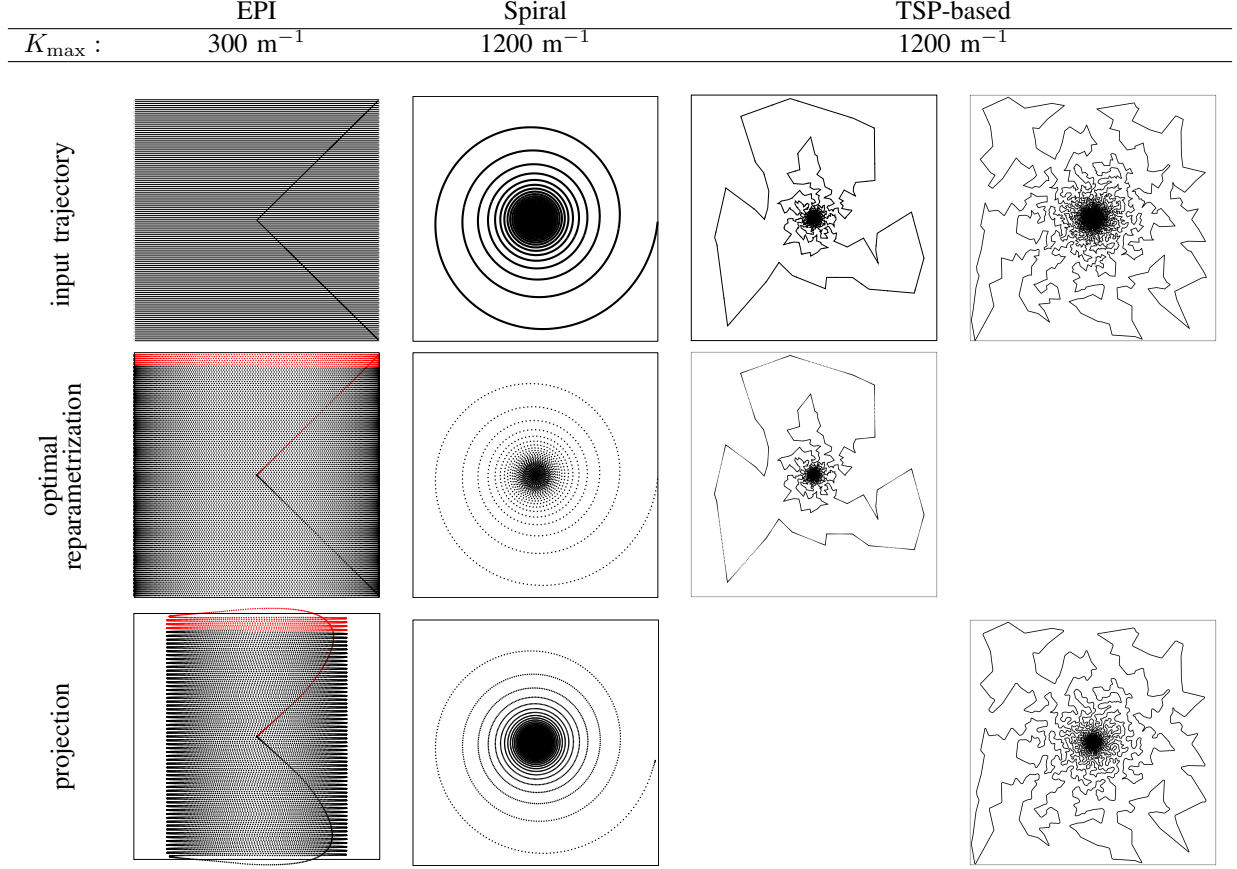


Fig. 7. Representation of input trajectory, optimal reparametrization and projection for EPI, spiral and TSP-based trajectories. The frame $[-K_{\max}, K_{\max}]^2$ is depicted with various values of $K_{\max}$ that depend on the reconstruction resolution.

aspect of our contribution in the future: projections are one of the most basic tools from optimization and might serve in many different contexts.

VII. CONCLUSION

We developed an algorithm to project any parameterized curve onto the set of curves which can be implemented on actual MRI scanners. Our method is an alternative to the existing gradient waveform design based on optimal control. The major advantages are that: i) the sampling time is fixed which is crucial to adapt the proposed scheme to any MR imaging modality; ii) the sampling density is close to the target one, as required by compressed sensing theory; iii) the behavior of our algorithm is similar to the state-of-the art for smooth trajectories, but it provides shorter k -space coverage when the trajectory comprises numerous high curvature points, as illustrated in the TSP and EPI cases.

ACKNOWLEDGMENTS

This work benefited from the support of the FMJH Program Gaspard Monge in optimization and operation research, and from the support to this program from EDF. We are also grateful to our colleagues Alexandre Vignaud who carefully reread the manuscript to better target the MRI readership, and Elvis Dohmatob for his remarks.

APPENDIX A

DENSITY DEVIATION, CONTROL OF W_2 -DISTANCE.

In Section III-A, we aim at controlling the Wasserstein distance $W_2(P_{s^*}, \pi)$, where π is a target fixed sampling distribution, and P_{s^*} is the empirical distribution of the projected curve. We used the triangle inequality (6) to bound this quantity by $W_2(P_{s^*}, P_c) + W_2(P_c, \pi)$. Here, we show that the quantity $W_2(P_c, \pi)$ can be as small as possible if c is Variable Density Sampler (VDS) [1]. First, we define the concept of VDS, and then we provide two examples. Next, we show that if c is a VDS, $W_2(P_c, \pi)$ tends to 0 as the length of c tends to infinity.

A. Definition of a VDS

First, we need to introduce the definition of weak convergence for measure:

Definition 3. A sequence of measures $\mu_n \in \mathcal{P}(K)$, the set of distributions defined over K , is said to weakly converge to μ if for any bounded continuous function ϕ

$$\int_K \phi(x) d\mu_n(x) \rightarrow \int_K \phi(x) d\mu(x).$$

We use the notation $\mu_n \rightharpoonup \mu$.

According to [1], a (generalized) π -VDS is a set of times T_n , such that $T_n \rightarrow \infty$ when $n \rightarrow \infty$, and a sequence of

curves $c_{T_n} : [0, T_n] \rightarrow \mathbb{R}^d$ such that $P_{c_{T_n}} \rightharpoonup \pi$ when n tends to infinity. A consequence of the definition is that the relative time spent by the curve in a part of the k -space is proportional to its density. Before showing that this implies that $W_2(P_{c_{T_n}}, \pi)$ tends to 0, we give two examples of VDS.

B. VDS examples

We give two examples to design continuous sampling trajectories that match a given distribution. The two examples we propose provide a sequence of curves, hence a sequence of empirical measures that weakly converge to the target density.

1) *Spiral sampling*: The spiral-based variable density sampling is now classical in MRI [31, 41]. For example, let $n \in \mathbb{R}_+$ be the number of revolutions, $r : [0, 1] \mapsto \mathbb{R}^+$ a strictly increasing smooth function, and $\theta : [0, 1] \rightarrow [0, 2\pi]$. Denote by r^{-1} the inverse function of r . Define the spiral for $t \in [0, n]$ by $c_n(t) = r\left(\frac{t}{n}\right) \exp\left(i \cdot n \cdot \theta\left(\frac{t}{n}\right)\right)$ and the target distribution π by:

$$\pi(x, y) = \begin{cases} \frac{r^{-1}(\sqrt{x^2+y^2})}{2\pi \int_{r(0)}^{r(1)} r^{-1}(\rho) \rho d\rho} & \text{if } r(0) \leq \sqrt{x^2+y^2} \leq r(1) \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

then $P_{c_n} \rightharpoonup \pi$ when n tends to infinity.

2) *Travelling Salesman-based sampling*: The idea of using the shortest path amongst a set of points (the “cities”) to design continuous trajectories with variable densities has been justified in [1, 19]. Let us draw n k -space locations uniformly according to a density q define over the d D k -space ($d = 2$ or 3), and join them by the shortest path (the Travelling Salesman solution). Then, denote by c_n a constant-speed parameterization of this curve. Define the density:

$$\pi = \frac{q^{(d-1)/d}}{\int q^{(d-1)/d}(x) dx}$$

Then $P_{c_n} \rightharpoonup \pi$ when the number of cities n tends to infinity.

These two sampling strategies are efficient to cover the k -space according to target distributions, as depicted in Fig. 7(top row) where TSP (resp. spiral) is a VDS for distribution depicted in Fig. 2(a) (resp. (b)). For spiral sampling, the target distribution may be any 2D radial distribution, whereas the Travelling salesman-based sampling enable us to consider any 2D or 3D density.

C. Control of W_2 distance

Let us now assume without loss of generality that $K = [-k_{\max}, k_{\max}]^d$.

Let us recall a central result about W_2 (see e.g., [26]):

Proposition 5. *Let $M \subset \mathbb{R}^d$, $\mu \in \mathcal{P}(M)$ and μ_n be a sequence of $\mathcal{P}(M)$. Then, if M is compact*

$$\mu_n \rightharpoonup \mu \Leftrightarrow W_2(\mu_n, \mu) \rightarrow 0$$

An immediate consequence of this proposition and of the compactness of K is the following proposition:

Proposition 6. *Let $(c_{T_n})_{n \geq 1}$ be a π -VDS, and $\varepsilon > 0$. Then, there exists $n \geq 1$ such that $c_{T_n} : [0, T_n] \rightarrow K$ fulfills:*

$$W_2(P_{c_{T_n}}, \pi) \leq \varepsilon.$$

To sum up, Proposition 6 ensures that we can find an input curve which empirical distribution is as close to the target distribution π as we want.

APPENDIX B

PROOF OF PROPOSITION 2

Definition 4 (indicator function). *Let $B \subseteq \mathbb{R}^n$. The indicator of B is denoted ι_B and defined by:*

$$\iota_B(x) = \begin{cases} 0 & \text{if } x \in B \\ +\infty & \text{otherwise} \end{cases}$$

Let us now recall a classical result of convex optimization [28, P. 195]:

Proposition 7. *Let $B_\alpha = \{x \in \mathbb{R}^n, \|x\| \leq \alpha\}$. Then the following identity holds:*

$$\iota_{B_\alpha}(x) = \sup_{y \in \mathbb{R}^n} \langle x, y \rangle - \alpha \|y\|_*.$$

Now, we can prove Proposition 2.

$$\begin{aligned} & \min_{s \in \mathcal{S} \cap \mathcal{A}} \frac{1}{2} \|s - c\|_2^2 \\ &= \min_{s \in \mathcal{A}} \frac{1}{2} \|s - c\|_2^2 + \iota_{B_\alpha}(\dot{M}s) + \iota_{B_\beta}(\ddot{M}s) \\ &= \min_{s \in \mathcal{A}} \frac{1}{2} \|s - c\|_2^2 + \sup_{q_1, q_2 \in \mathbb{R}^{n \cdot d}} \langle \dot{M}s, q_1 \rangle - \alpha \|q_1\|_* \\ & \quad + \langle \ddot{M}s, q_2 \rangle - \beta \|q_2\|_* \\ &= \sup_{q_1, q_2 \in \mathbb{R}^{n \cdot d}} \min_{s \in \mathcal{A}} \frac{1}{2} \|s - c\|_2^2 + \langle s, \dot{M}^* q_1 \rangle + \langle s, \ddot{M}^* q_2 \rangle \\ & \quad - \alpha \|q_1\|_* - \beta \|q_2\|_* \end{aligned}$$

The relationship between the primal and dual solutions reads $s^* = \arg \min_{s \in \mathcal{A}} \frac{1}{2} \|s - c\|_2^2 + \langle s, \dot{M}^* q_1^* \rangle + \langle s, \ddot{M}^* q_2^* \rangle$. The sup and the min can be interverted at the third line, due to standard theorems in convex analysis (see e.g. [42, Theorem 31.3]).

APPENDIX C

PROOF OF PROPOSITIONS 3

To show Proposition 3, first remark that

$$\begin{aligned} & \arg \min_{s \in \mathcal{A}} \langle \dot{M}s, q_1 \rangle + \langle \ddot{M}s, q_2 \rangle + \frac{1}{2} \|s - c\|_2^2 \\ &= \arg \min_{s \in \mathcal{A}} \frac{1}{2} \|s - (c - \dot{M}q_1 - \ddot{M}^* q_2)\|_2^2. \end{aligned}$$

Therefore, $s^*(q_1, q_2)$ is the orthogonal projection of $z = c - \dot{M}q_1 - \ddot{M}^* q_2$ onto $\mathcal{A}$. Since $\mathcal{A}$ is not empty, $\mathcal{A}\mathcal{A}^+v = v$, and the set $\mathcal{A} = \{s \in \mathbb{R}^{n \cdot d}, \mathcal{A}s = v\}$ can be decomposed as

$$\mathcal{A} = \mathcal{A}^+v + \ker(\mathcal{A}).$$

The vector $z - s^*(q_1, q_2)$ is orthogonal to $\mathcal{A}$, it therefore belongs to $\ker(\mathcal{A})^\perp = \text{im}(\mathcal{A}^*)$. Hence $s^*(q_1, q_2) = z + \mathcal{A}^*\lambda$ for some λ such that:

$$\mathcal{A}(z + \mathcal{A}^*\lambda) = v.$$

This leads to $\lambda = (\mathbf{A}\mathbf{A}^*)^{-1}(\mathbf{v} - \mathbf{A}\mathbf{z})$. We finally get

$$\mathbf{s}^*(\mathbf{q}_1, \mathbf{q}_2) = \mathbf{z} + \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}(\mathbf{v} - \mathbf{A}\mathbf{z}),$$

ending the proof.

APPENDIX D PROOF OF THEOREM 1.

Let us first recall that the relative interior of a convex set C $\text{ri}(C)$ is the interior of C relative to the affine hull of C [28]. The analysis proposed to prove Theorem 1 closely follows ideas proposed in [30, 43–45]. We will need two results. The first one is a duality result from [44].

Proposition 8. *Let $f : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{\infty\}$ and $g : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ denote two closed convex functions, and $\mathbf{A} \in \mathbb{R}^{m \times n}$ denote a matrix. Assume that g is σ -strongly convex [28] and that $\text{Ari}(\text{dom}(f)) \cap \text{ri}(\text{dom}(g)) \neq \emptyset$.*

Let $p(\mathbf{x}) = f(\mathbf{A}\mathbf{x}) + g(\mathbf{x})$ and $d(\mathbf{y}) = -g^(\mathbf{A}^*\mathbf{y}) - f^*(\mathbf{y})$. Let $\mathbf{x}^*$ be the unique minimizer of p and $\mathbf{y}^*$ be any minimizer of d .*

Then g^ is differentiable with $\frac{1}{\sigma}$ Lipschitz-continuous gradient. Moreover, by letting $\mathbf{x}(\mathbf{y}) = \nabla g^*(-\mathbf{A}^*\mathbf{y})$:*

$$\|\mathbf{x}(\mathbf{y}) - \mathbf{x}^*\|_2^2 \leq \frac{2}{\sigma}(d(\mathbf{y}) - d(\mathbf{y}^*)).$$

The second ingredient is the standard convergence rate for accelerated proximal gradient descents given in [30, Theorem. 4.4].

Proposition 9. *Under the same assumptions as Proposition 8, consider Algorithm 2.*

Algorithm 2: Accelerated proximal gradient descent

Input: $\mathbf{q}_0 \in \text{ri}(\text{dom}(f^*)) \cap \text{Ari}(\text{dom}(g^*))$ and n_{it}

Initialize Set $\ell = 1/L$, with $L = \frac{\|\mathbf{A}\|^2}{\sigma}$.

Set $\mathbf{y}_0 = \mathbf{q}_0$. **for** $k = 1 \dots n_{it}$ **do**
 $\mathbf{q}^{(k)} = \text{prox}_{\ell f^*}(\mathbf{y}^{(k-1)} + \ell \mathbf{A} \nabla g^*(-\mathbf{A}^*\mathbf{y}^{(k-1)}))$
 $\mathbf{y}^{(k)} = \mathbf{q}^{(k)} + \frac{k-1}{k+2}(\mathbf{q}^{(k)} - \mathbf{q}^{(k-1)})$

Then $\|\mathbf{y}^{(n_{it})} - \mathbf{y}^*\|_2^2 = \mathcal{O}\left(\frac{\|\mathbf{A}\|^2}{\sigma \cdot n_{it}^2}\right)$.

To conclude, it suffices to set $g(\mathbf{s}) = \frac{1}{2}\|\mathbf{s} - \mathbf{c}\|_2^2$, $f(\mathbf{q}_1, \mathbf{q}_2) = \nu_{B_\alpha}(\mathbf{q}_1) + \nu_{B_\alpha}(\mathbf{q}_2)$ and $\mathbf{A} = \begin{pmatrix} \mathbf{M} \\ \mathbf{0} \end{pmatrix}$. By doing so, the projection problem rewrites $\min_{\mathbf{s} \in \mathbb{R}^{nd}} p(\mathbf{s}) = f(\mathbf{A}\mathbf{s}) + g(\mathbf{s})$. Its dual problem (8) can be rewritten more compactly as $\min_{\mathbf{q}=(\mathbf{q}_1, \mathbf{q}_2) \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}} d(\mathbf{q}) = g^*(-\mathbf{A}^*\mathbf{q}) + f^*(\mathbf{q})$. Note that function g is 1-strongly convex. Therefore, Algorithm 2 can be used to minimize d , ensuring a convergence rate in $\mathcal{O}\left(\frac{L}{k^2}\right)$ on the function values $d(\mathbf{y}^{(k)})$, where $L = \|\mathbf{A}\|^2$. It then suffices to use Proposition 8 to obtain a convergence rate on the distance to the solution $\|\mathbf{s}^{(k)} - \mathbf{s}^*\|_2^2$. This ends the proof of Theorem 1.

REFERENCES

- [1] N. Chauffert, P. Ciuciu, J. Kahn, and P. Weiss, "Variable density sampling with continuous trajectories. Application to MRI," *SIAM Journal on Imaging Science*, vol. 7, no. 4, pp. 1962–1992, 2014.
- [2] C. Boyer, J. Bigot, and P. Weiss, "Compressed sensing with structured sparsity and structured acquisition," *arXiv preprint arXiv:1505.01619*, 2015.
- [3] J. Unnikrishnan and M. Vetterli, "Sampling high-dimensional bandlimited fields on low-dimensional manifolds," *Information Theory, IEEE Transactions on*, vol. 59, no. 4, pp. 2103–2127, 2013.
- [4] K. Gröchenig, J. L. Romero, J. Unnikrishnan, and M. Vetterli, "On minimal trajectories for mobile sampling of bandlimited fields," *Applied and Computational Harmonic Analysis*, 2014.
- [5] G. Puy, P. Vandergheynst, and Y. Wiaux, "On variable density compressive sampling," *IEEE Signal Processing Letters*, vol. 18, no. 10, pp. 595–598, 2011.
- [6] F. Krahmer and R. Ward, "Beyond incoherence: stable and robust sampling strategies for compressive imaging," preprint, 2012.
- [7] P. T. Gurney, B. A. Hargreaves, and D. G. Nishimura, "Design and analysis of a practical 3D cones trajectory," *Magn. Reson. Med.*, vol. 55, no. 3, pp. 575–582, 2006.
- [8] J. G. Pipe and N. R. Zwart, "Spiral trajectory design: A flexible numerical algorithm and base analytical equations," *Magn. Reson. Med.*, vol. 71, no. 1, pp. 278–285, 2014.
- [9] M. Lustig, D. L. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MRI imaging," *Magn. Reson. Med.*, vol. 58, pp. 1182–1195, Dec. 2007.
- [10] L. Feng, R. Grimm, K. T. Block, H. Chandarana, S. Kim, J. Xu, L. Axel, D. K. Sodickson, and R. Otazo, "Golden-angle radial sparse parallel mri: Combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric mri," *Magn. Reson. Med.*, vol. 72, no. 3, pp. 707–717, 2014.
- [11] M. Lustig, J. H. Lee, D. L. Donoho, and J. M. Pauly, "Faster imaging with randomly perturbed, under-sampled spirals and 11 reconstruction," in *Proceedings of the 13th annual meeting of ISMRM, Miami Beach*, p. 685, 2005.
- [12] D. C. Noll, "Multishot rosette trajectories for spectrally selective MR imaging," *IEEE Trans. Med. Imag.*, vol. 16, no. 4, pp. 372–377, 1997.
- [13] Y. Shu, S. J. Riederer, and M. A. Bernstein, "Three-dimensional mri with an undersampled spherical shells trajectory," *Magn. Reson. Med.*, vol. 56, no. 3, pp. 553–562, 2006.
- [14] M. Lustig, S. J. Kim, and J. M. Pauly, "A fast method for designing time-optimal gradient waveforms for arbitrary k-space trajectories," *IEEE Trans. Med. Imag.*, vol. 27, no. 6, pp. 866–873, 2008.
- [15] O. P. Simonetti, J. L. Duerk, and V. Chankong, "An optimal design method for magnetic resonance imaging gradient waveforms," *IEEE Trans. Med. Imag.*, vol. 12, no. 2, pp. 350–360, 1993.
- [16] B. A. Hargreaves, D. G. Nishimura, and S. M. Conolly, "Time-optimal multidimensional gradient waveform design for rapid imaging," *Magn. Reson. Med.*, vol. 51, no. 1, pp. 81–92, 2004.
- [17] M. Davids, M. Ruttorf, F. Zoellner, and L. Schad, "Fast and robust design of time-optimal k-space trajectories in MRI," *IEEE Trans. Med. Imag.*, vol. 34, no. 2, pp. 564–577, 2015.
- [18] B. Adcock, A. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: asymptotic incoherence and asymptotic sparsity in compressed sensing," 2013.
- [19] N. Chauffert, P. Ciuciu, J. Kahn, and P. Weiss, "Travelling salesman-based compressive sampling," in *Proc. of 10th SampTA conference*, (Bremen, Germany), pp. 509–511, July 2013.
- [20] N. Chauffert, P. Ciuciu, P. Weiss, and F. Gamboa, "From variable density sampling to continuous sampling using Markov chains," in *Proc. of 10th SampTA conference*, (Bremen, Germany), pp. 200–203, July 2013.
- [21] H. Wang, X. Wang, Y. Zhou, Y. Chang, and Y. Wang, "Smoothed random-like trajectory for compressed sensing MRI," in *Proc. of the 34th annual IEEE EMBC*, pp. 404–407, 2012.
- [22] R. M. Willett, "Errata: Sampling trajectories for sparse image recovery," note, Duke University, 2011.
- [23] K. Majewski, O. Heid, and T. Kluge, "MRI pulse sequence design with first-order gradient moment nulling in arbitrary directions by solving a polynomial program," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1252–1259, 2010.
- [24] H. Rauhut, "Compressive Sensing and Structured Random Matrices," in *Theoretical Foundations and Numerical Methods for Sparse Recovery* (M. Fornasier, ed.), vol. 9 of *Radon Series Comp. Appl. Math.*, pp. 1–92, deGruyter, 2010.

- [25] N. Chauffert, P. Ciuciu, and P. Weiss, "Variable density compressed sensing in MRI. Theoretical vs. heuristic sampling strategies," in *Proc. of 10th IEEE ISBI conference*, (San Francisco, USA), pp. 298–301, Apr. 2013.
- [26] C. Villani, *Optimal transport: old and new*, vol. 338. Springer, 2008.
- [27] Y. Nesterov, "Smooth minimization of non-smooth functions," *Math. Program.*, vol. 103, pp. 127–152, May 2005.
- [28] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I: Part I: Fundamentals*, vol. 305. Springer, 1996.
- [29] Y. Nesterov, "A method of solving a convex programming problem with convergence rate $\mathcal{O}(1/k^2)$," in *Soviet Mathematics Doklady*, vol. 27, pp. 372–376, 1983.
- [30] A. Beck and M. Teboulle, "Gradient-based algorithms with applications to signal recovery," *Convex Optimization in Signal Processing and Communications*, 2009.
- [31] D. H. Kim, E. Adalsteinsson, and D. M. Spielman, "Simple analytic variable density spiral design," *Magn. Reson. Med.*, vol. 50, no. 1, pp. 214–219, 2003.
- [32] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [33] J. Keiner, S. Kunis, and D. Potts, "Using nfft 3—a software library for various nonequispaced fast fourier transforms," *ACM Transactions on Mathematical Software (TOMS)*, vol. 36, no. 4, p. 19, 2009.
- [34] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.
- [35] A. Florescu, E. Chouzenoux, J.-C. Pesquet, P. Ciuciu, and S. Ciochina, "A majorize-minimize memory gradient method for complex-valued inverse problems," *Signal Processing*, vol. 103, pp. 285–295, 2014.
- [36] Y. Huang, J. Paisley, Q. Lin, X. Ding, X. Fu, and X.-P. Zhang, "Bayesian nonparametric dictionary learning for compressed sensing mri," *Image Processing, IEEE Transactions on*, vol. 23, no. 12, pp. 5007–5019, 2014.
- [37] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, "SENSE: sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, pp. 952–962, Jul. 1999.
- [38] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase, "Generalized autocalibrating partially parallel acquisitions GRAPPA," *Magn. Reson. Med.*, vol. 47, pp. 1202–1210, Jun. 2002.
- [39] D. A. Feinberg, S. Moeller, S. M. Smith, E. Auerbach, S. Ramanna, M. Gunther, M. F. Glasser, K. L. Miller, K. Ugurbil, and E. Yacoub, "Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging," *PloS one*, vol. 5, no. 12, p. e15710, 2010.
- [40] N. Chauffert, P. Ciuciu, J. Kahn, and P. Weiss, "A projection algorithm on measures sets," submitted to *Constructive Approximation*, May 2015.
- [41] D. M. Spielman, J. M. Pauly, and C. H. Meyer, "Magnetic resonance fluoroscopy using spirals with variable sampling densities," *Magn. Reson. Med.*, vol. 34, no. 3, pp. 388–394, 1995.
- [42] R. T. Rockafellar, *Convex analysis*. No. 28, Princeton university press, 1997.
- [43] P. Weiss, L. Blanc-Féraud, and G. Aubert, "Efficient schemes for total variation minimization under constraints in image processing," *SIAM journal on Scientific Computing*, vol. 31, no. 3, pp. 2047–2080, 2009.
- [44] C. Boyer, P. Weiss, and J. Bigot, "An algorithm for variable density sampling with block-constrained acquisition," *SIAM Journal on Imaging Sciences*, vol. 7, no. 2, pp. 1080–1107, 2014.
- [45] A. Beck and M. Teboulle, "A fast dual proximal gradient algorithm for convex minimization and applications," *Operations Research Letters*, vol. 42, no. 1, pp. 1–6, 2014.