



Gradient waveform design for variable density sampling in Magnetic Resonance Imaging

Nicolas Chauffert, Pierre Weiss, Jonas Kahn, Philippe Ciuciu

► To cite this version:

Nicolas Chauffert, Pierre Weiss, Jonas Kahn, Philippe Ciuciu. Gradient waveform design for variable density sampling in Magnetic Resonance Imaging. 2014. hal-01095320v1

HAL Id: hal-01095320

<https://inria.hal.science/hal-01095320v1>

Preprint submitted on 15 Dec 2014 (v1), last revised 30 Dec 2014 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Gradient waveform design for variable density sampling in Magnetic Resonance Imaging

Nicolas Chauffert, Pierre Weiss, Jonas Kahn and Philippe Ciuciu.

Abstract—Fast coverage of k -space is a major concern to speed up data acquisition in Magnetic Resonance Imaging (MRI) and limit image distortions due to long echo train durations. The hardware gradient constraints (magnitude, slew rate) must be taken into account to collect a sufficient amount of samples in a minimal amount of time. However, sampling strategies (e.g., Compressed Sensing) and optimal gradient waveform design have been developed separately so far. The major flaw of existing methods is that they do not take the sampling density into account, the latter being central in sampling theory. In particular, methods using optimal control tend to agglutinate samples in high curvature areas. In this paper, we develop an iterative algorithm to project any parameterization of k -space trajectories onto the set of feasible curves that fulfills the gradient constraints. We show that our projection algorithm provides a more efficient alternative than existing approaches and that it can be a way of reducing acquisition time while maintaining sampling density for piece-wise linear trajectories.

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 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 designs) should be addressed simultaneously but it is unclear how to design a feasible waveform corresponding to a k -space sampling with good space coverage properties. The gradient waveform design is thus generally performed sequentially: a first step aims at finding the trajectory support or at least control points, and a second step consists of designing the fastest gradient waveforms to traverse the trajectory.

Recent progresses in compressed sensing (CS) [1–4] give hints concerning the best strategies of k -space sampling. In particular, the k -space has to be sampled with a variable density, decaying from the lowest to the highest frequencies. To date, CS sampling schemes are limited to simple trajectories, such as 2D independent sampling and acquisition in the orthogonal readout direction. The design of physically plausible acquisition schemes with a prescribed sampling

density remains an open question. To better understand how the proposed approach goes beyond the state-of-the-art, we start with a short review of existing methods for designing gradient waveforms for k -space sampling. We show that existing methods do not encompass the case of curves with high curvature in terms of sampling density and scanning time, justifying a new strategy for designing gradient waveforms.

A. Related works

Some heuristic methods have been proposed to design simultaneously the sampling density and the gradient waveforms [5, 6] but these are not computationally efficient and it is not clear how to sample the k -space according to a prescribed distribution using these methods. Let us mention other algorithms that have been proposed to traverse specific curve shapes, such as spiral [7, 8] or concentric rings [9].

It is in general easier to consider a sampling trajectory in a first time and design the gradient waveform afterwards. To the best of our knowledge, the latter problem is currently solved using convex optimization [10, 11], optimal control [12], or optimal interpolation of k -space control points [13]. Given an input curve, both methods find a parameterization that minimizes the time required to traverse the curve subject to kinematic constraints. This principle suffers from two limitations in certain situations. First, reparameterizing the curve changes the density of samples along the curve. This density is now known to be a key factor in compressed sensing [1–4], since it directly impacts the number of required measurements to ensure accurate 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 an arbitrary curve may be time consuming, especially in the high curvature parts of the trajectory. In particular, the time to traverse piecewise linear sampling trajectories [2, 14–17] may become too long. Indeed, the magnetic field gradients have to be set to zero at each singular point of such trajectory. For this reason, new methods have to be pushed forward to fill this gap.

B. Contributions

In this paper, we propose an alternative method 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 usually shorter, ii) the distance between the input

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.

and output curves is the quantity to be minimized ensuring a low deviation to the original sampling distribution, iii) the maximal acquisition time may be fixed, enabling to find the closest curve in a given time and iv) it is flexible enough to handle additional hardware constraints (e.g., trajectory starting from the k -space center) in the same framework.

C. 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 + 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. Finally, we illustrate the behaviour of our algorithm on two particular cases: rosette and Travelling Salesman Problem-based trajectories (Section V).

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

In this section, we recall classical modelling of the acquisition constraints in MRI [11, 12]. We justify the lack of accuracy of current reparameterization methods in the context of variable density sampling, and motivate the introduction of a new *projection* algorithm.

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:

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

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 [11]. The gradient waveform is obtained by energizing gradient coils (arrangements of wire) with electric currents.

1) *kinematic constraints*: due to obvious physical constraints, these electric currents 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

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.

$\|f\|_{\infty, 2} := \sup_{t \in [0, T]} \left(\sum_{i=1}^d |f_i(t)|^2 \right)^{\frac{1}{2}}$. 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} :

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

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): $s(k \cdot TR) = 0, 0 \leq k \leq \left\lfloor \frac{T}{TR} \right\rfloor$.
- In addition to starting from the k -space center, one could impose starting at speed 0, i.e. $\dot{s}(0) = 0$.
- To avoid artifacts due to flow motion in the object of interest, gradient moment nulling (GMN) techniques have been introduced in [18]. In terms of constraints, nulling the i^{th} moment reads $\int_0^T t^i g(t) dt = 0$.

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

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

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).

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 [12] or convex optimization [11]. 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). 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

The idea we propose in this paper is to find the projection of the given input curve c onto the set of admissible curves S :

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

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 optimal control approach introduced here above: 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 . In particular, time to traverse a curve is in general shorter than using exact parameterization (see Fig. 1(f) where $T < T_{\text{Rep}}$).

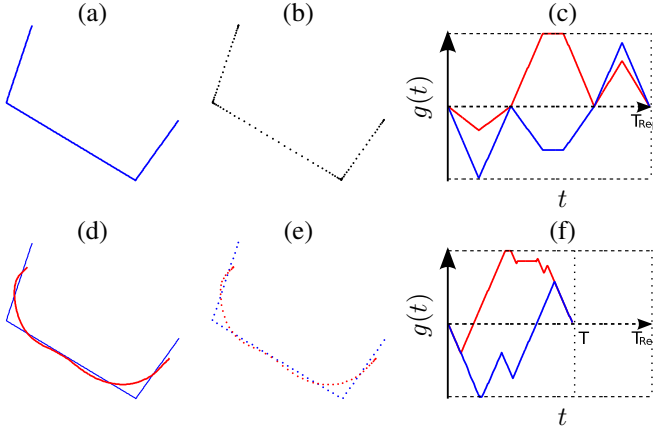


Fig. 1. Comparison of two methods to design gradient waveforms. **Top row:** Optimal control-based parameterization [12]. (a): input curve support. (b): discrete representation of the optimal reparameterization of the curve in S . (c): corresponding gradient waveforms (g_x, g_y). Dashed lines correspond to 0 and $\pm G_{\text{max}}$. **Bottom row:** Illustration of the projection algorithm. (d): the same input curve as in (a) parameterized at maximal speed, and the support of the projected curve onto the set 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): time to traverse the curve is reduced by 39%.

In the next section, we will explain why the empirical distribution of the curves obtained by this projection method is close to the one of the input curve, and we illustrate how the parameterization will distort the sampling distribution.

III. CONTROL OF THE SAMPLING DENSITY

Recent works have emphasized the importance of the sampling density [1–4] in an attempt to reduce the amount of

acquired data while preserving image quality at the reconstruction step. The choice of an accurate distribution p is crucial since it directly impacts the number of measurements required [19]. In this paper, we will denote by π a distribution defined over the k -space K . The profile of this distribution can be obtained by theoretical arguments [1–4] leading to distributions as the one depicted in Fig. 2(a). Some heuristical distributions are known to perform well in CS-MRI experiments (Fig. 2(b)). A comparison between these two approaches can be found in [20].

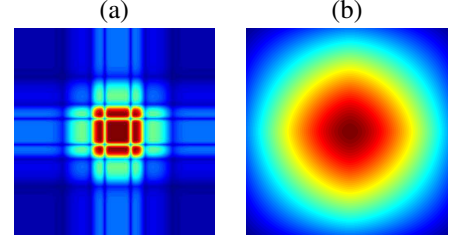


Fig. 2. Examples of 2D sampling distribution. (a): optimal distribution for a Symmlet transform [1, 2]. (b): radial distribution advocated in [21]: $p(k) \propto (1 - |k|/k_{\text{max}})^p$ with $p = 3$.

However, designing a trajectory that performs sampling according to a fixed distribution while satisfying gradient constraints, is really challenging and has never been addressed so far. Recent attempts did not manage gradient constraints to design continuous variable density sampling trajectories since the sampling curve cannot be traversed at constant speed, especially at angular points as shown in Fig. 1(b). This impacts the sampling density by concentrating probability mass in the portions of the curve associated with large curvature. In this section, we show that if we start from an inadmissible trajectory that covers the k -space with an empirical distribution close to the target one, then we can figure out how to design an admissible sampling trajectory with almost the same empirical distribution. Strategies to design such input curves are detailed in Appendix A. In what follows, we first derive theoretical guarantees, and then, we perform experiments to quantify the distribution distortion.

A. Theoretical guarantees

Let us show that the sampling distribution carried over a curve that our algorithm delivers can be close to a fixed target distribution. To this end, we start with the definition of the empirical distribution of a curve, and then we introduce a distance defined over distributions.

Definition 1 (Empirical measure of a curve). *Let γ denote the Lebesgue measure and $\gamma_T = \frac{\gamma}{T}$ 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:*

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

This means that the mass of a set ω is proportional to the time spent by the curve in ω . The natural distance arising in

our work is the between-distribution 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:*

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

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., [22]). 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 encodes the deformation map to turn one distribution into the other.

A global strategy to design feasible trajectories with an empirical distribution close to a target density π relies on:

- 1) finding an input curve c , such that P_c is close to π ;
- 2) estimating s^* the projection of c onto the set of constraints, by solving Eq. (4).

The objective is to show that $W_2(P_{s^*}, \pi)$ is small. Since W_2 is a norm, the triangle inequality holds:

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

The *initial guess* term $W_2(P_c, \pi)$ can be as small as possible if c is a Variable Density Sampler [2]. The proof is postponed to Appendix A for the ease of reading. Here, we are interested in the *distortion* term $W_2(P_{s^*}, P_c)$. The following proposition shows that the W_2 distance between the empirical distributions of the input curve and of the output curve (c and s^*) is controlled by the quantity $d(s^*, c)$ to be minimized when solving Eq. (4).

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

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

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

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

where $\gamma_{s,c}$ 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 $\gamma_{c,s}(\omega_1, \omega_2) = \frac{1}{T} \int_{t=0}^T 1_{\omega_1}(s(t)) 1_{\omega_2}(c(t)) dt$. The choice of this coupling is equivalent to choosing the transformation map as the association of locations of $c(t)$ and $s(t)$ for each 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 $\gamma_{s,c}$. \square

To sum up, solving the projection problem amounts to minimizing an upper bound of the W_2 distance between the target and the empirical distributions of the solution if we neglect the influence of the *initial guess* term.

B. Simulations

Next, we performed simulations to show that the sampling density is better preserved using our approach than relying on the optimal control approach. For doing so, we use travelling salesman-based (TSP) sampling trajectories [2, 15], which are an original way to design random trajectories which empirical distribution is any target distribution such as in Fig. 3(a). 10000 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 fixed rate Δt (as in Fig. 3(b)), to form an histogram (the empirical distribution shown in Fig. 3(c)), which was compared to π in Fig. 3(d). The error is actually not closed to zero, since the convergence result is asymptotic in the size of the curve that is bounded in this experiment.

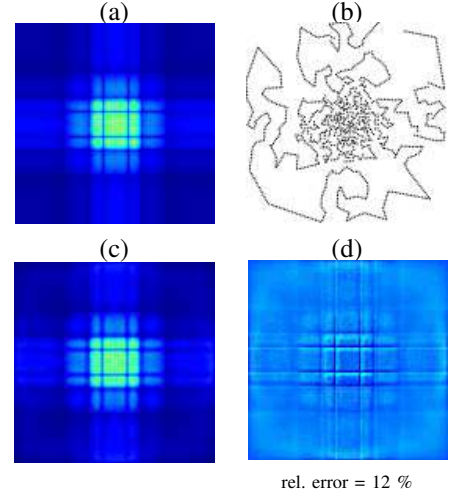


Fig. 3. Illustration of TSP VDS. (a) Target distribution. (b) Regular sampling of TSP curve parameterized with its arc-length. (c) Empirical distribution (histogram of samples). (d) Histogram difference (c)-(a).

In Fig. 4 (top row), we show that the classical reparameterization technique [12] lead to a major distortion of the sampling density, because of the behavior on the angular points illustrated in Fig. 1(b). Then, we considered three constant speed parameterizations and projected them onto the same set of constraints ($G_{\max} = 4 \text{ T.cm}^{-1}$ and $S_{\max} = 15 \text{ T.cm}^{-1}.\text{ms}^{-1}$). Among these 3 initial guesses, we first used an initial parameterization with low velocity (10 % of the maximal speed γG_{\max} with $\gamma = 42.576 \text{ MHz.T}^{-1}$), which projection fits the sampling density quite well. Then, we increased the velocity to reach first 50 % and then 100 % of the maximal speed. The distortion of the sampling density of the projected curve increased, but remained negligible in contrast to the exact reparameterization. Hence, this example illustrates that if we have a continuous trajectory with an empirical sampling distribution close to the target one, our projection algorithm enables to design feasible waveforms while sampling the k -space along a discretized trajectory with an empirical density close to the target one as well.

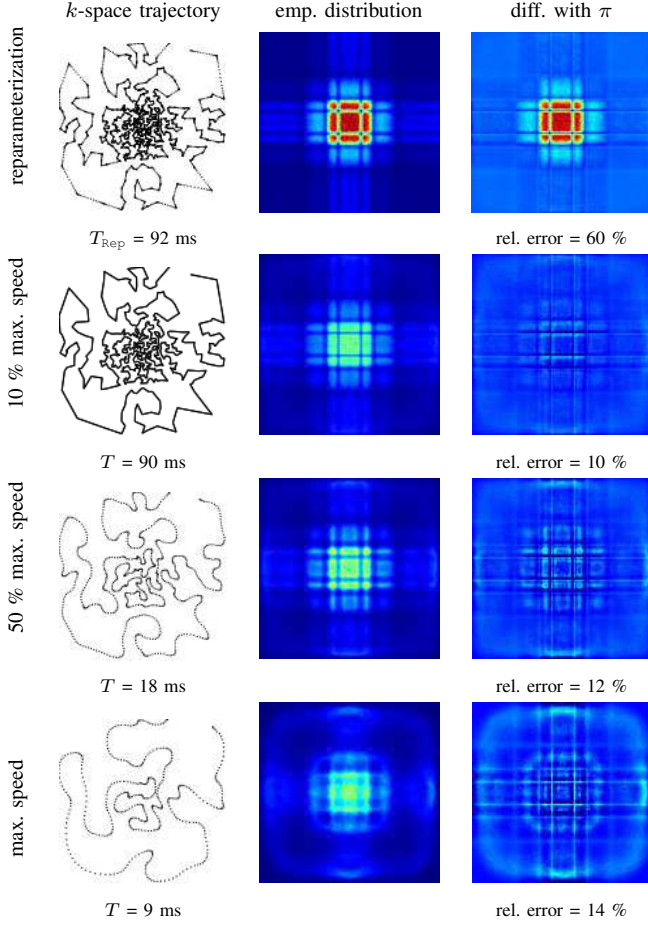


Fig. 4. Illustration of TSP trajectories traversed with optimal control (**top row**) and with our projection algorithm (**rows 2-4**). Columns represent the k -space trajectory, the empirical distribution and the difference with the target distribution π (Fig. 3(a)).

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 structure², it is unlikely that an explicit solution to Problem (4) can be found. In what follows, we thus propose a numerical algorithm to find the projection.

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

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

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

²it is just a polytope when the ℓ^∞ -norm is used.

An efficient projection algorithm: The discrete problem we consider is the same as problem (4) except that all objects are discretized. The set \mathcal{S} is now $\mathcal{S} := \{\mathbf{s} \in \mathbb{R}^{n \cdot d}, \|\dot{\mathbf{M}}\mathbf{s}\| \leq \alpha, \|\ddot{\mathbf{M}}\mathbf{s}\| \leq \beta\}$ with all norms discretized. Similarly, the discretized version of \mathcal{A} is denoted \mathcal{A} . The main idea of our algorithm is to take advantage of the structure of the dual problem to design an efficient projection algorithm. The following proposition specifies the dual problem and the primal-dual relationships.

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

$$\min_{\mathbf{s} \in \mathcal{S} \cap \mathcal{A}} \frac{1}{2} \|\mathbf{s} - \mathbf{c}\|_2^2 \quad (8)$$

$$= \sup_{\mathbf{q}_1, \mathbf{q}_2 \in \mathbb{R}^{n \cdot d}} F(\mathbf{q}_1, \mathbf{q}_2) - \alpha \|\mathbf{q}_1\|_* - \beta \|\mathbf{q}_2\|_*, \quad (9)$$

where

$$F(\mathbf{q}_1, \mathbf{q}_2) = \min_{\mathbf{s} \in \mathcal{A}} \langle \dot{\mathbf{M}}\mathbf{s}, \mathbf{q}_1 \rangle + \langle \ddot{\mathbf{M}}\mathbf{s}, \mathbf{q}_2 \rangle + \frac{1}{2} \|\mathbf{s} - \mathbf{c}\|_2^2. \quad (10)$$

Moreover, let $(\mathbf{q}_1^*, \mathbf{q}_2^*)$ denote any minimizer of the dual problem (9), \mathbf{s}^* denote the unique solution of the primal problem (8) and $\mathbf{s}^*(\mathbf{q}_1^*, \mathbf{q}_2^*)$ denote the solution of the minimization problem (10). Then $\mathbf{s}^* = \mathbf{s}^*(\mathbf{q}_1^*, \mathbf{q}_2^*)$.

The following proposition gives an explicit expression of $\mathbf{s}^*(\mathbf{q}_1^*, \mathbf{q}_2^*)$.

Proposition 3. The minimizer

$$\mathbf{s}^*(\mathbf{q}_1^*, \mathbf{q}_2^*) = \arg \min_{\mathbf{s} \in \mathcal{A}} \langle \dot{\mathbf{M}}\mathbf{s}, \mathbf{q}_1^* \rangle + \langle \ddot{\mathbf{M}}\mathbf{s}, \mathbf{q}_2^* \rangle + \frac{1}{2} \|\mathbf{s} - \mathbf{c}\|_2^2$$

is given by

$$\mathbf{s}^*(\mathbf{q}_1, \mathbf{q}_2) = \mathbf{z} + \mathbf{A}^+(\mathbf{v} - \mathbf{A}\mathbf{z}), \quad (11)$$

where $\mathbf{A}^+ = \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}$ denotes the pseudo-inverse of \mathbf{A} and $\mathbf{z} = \mathbf{c} - \dot{\mathbf{M}}^*\mathbf{q}_1 - \ddot{\mathbf{M}}^*\mathbf{q}_2$.

Let us now analyze the smoothness properties of F .

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

$$\nabla F(\mathbf{q}_1, \mathbf{q}_2) = - \begin{pmatrix} \dot{\mathbf{M}}\mathbf{s}^*(\mathbf{q}_1, \mathbf{q}_2) \\ \ddot{\mathbf{M}}\mathbf{s}^*(\mathbf{q}_1, \mathbf{q}_2) \end{pmatrix}. \quad (12)$$

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} .

The proofs are given in Appendix B. The dual problem (9) has a favorable structure for optimization: it is the sum of a differentiable convex function $F(\mathbf{q}_1, \mathbf{q}_2) = -F(\mathbf{q}_1, \mathbf{q}_2)$ and of a simple convex function $G(\mathbf{q}_1, \mathbf{q}_2) = \alpha \|\mathbf{q}_1\|_* + \beta \|\mathbf{q}_2\|_*$. The sum $F+G$ can thus be minimized efficiently using accelerated proximal gradient descents [23] (see Algorithm 1 below).

Moreover, by combining the convergence rate results of [23, 24] and some convex analysis (see Appendix D), we obtain the following convergence rate:

Algorithm 1: Projection algorithm in the dual space

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

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

Initialize $q^{(0)} = (q_1^{(0)}, q_2^{(0)})$ with $q_i^{(0)} = 0$ for $i = 1, 2$.

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

Set $t = 1/L$.

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

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

return $\tilde{s} = s^*(q_1^{(n_{it})}, q_2^{(n_{it})})$;

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

$$\|s^{(k)} - s^*\|_2^2 \leq \frac{2L\|q^{(0)} - q^*\|_2^2}{k^2}. \quad (13)$$

V. ILLUSTRATIONS

To compare our results with [12], we used the same gradient constraints. In particular, the maximal gradient norm G_{\max} is set to 4 T.cm⁻¹, and the slew-rate S_{\max} to 15 T.cm⁻¹.ms⁻¹. k -space size is 6 cm⁻¹ and sampling rate is $\Delta t = 4 \mu\text{s}$. For the ease of trajectory representation, we limit ourselves to 2D sampling curves, but the algorithm encompasses the 3D setting. The Matlab codes containing the projection algorithm as well as the scripts to reproduce the results depicted hereafter are available at <http://chauffertn.free.fr/codes.html>. First, we illustrate the output of Algorithm 1 for the case of rosette trajectories introduced in [25], showing a similar behaviour of our algorithm compared to optimal control. Then, we show an application to gradient waveform design to traverse TSP-based sampling trajectories. In this case, only our algorithm yields a fast k -space coverage.

A. Smooth trajectories

The first illustration in Fig. 5 shows that when the sampling curve is smooth enough, our algorithm has a similar behaviour than optimal control. Here, we considered the case of Rosette trajectory defined in [12, 25] as $k(t) = k_{\max} \sin(\omega_1 t) [\cos(\omega_2 t); \sin(\omega_2 t)]$ (with $\omega_1 = 1.419$ and $\omega_2 = 0.8233 \text{ rad.s}^{-1}$). In Fig. 5, we considered $c(t)$ the parameterization of $k(t)$ at constant speed (90% of γG_{\max}). The black curve represents the projection of c onto the set of constraints for Rotation Variant (RV) constraints (Fig. 5(a)) and for Rotation Invariant (RIV) constraints (Fig. 5(b)). Then, we show the corresponding gradient waveforms to traverse the trajectory (see Fig. 5(c,d) for optimal control solution and Fig. 5(e,f) for the output of our projection algorithm).

We observed that the time spent to traverse the curve is slightly shorter compared to the exact reparameterization, while the support of the curve is reduced. Of course, the initial parameterization of c has a strong impact, and there is a trade-off between the time spent to traverse the curve and the distance to the original support. We also noticed that RV constraints are less restrictive, enabling a faster traversal time

in the optimal control framework (Fig. 5(c)) or a lower support distortion in the proposed setting (Fig. 5(a)), as also noticed in [26].

In this example, there is no real difference between the optimal control and the projection approaches. The following example, based on piece-wise linear sampling curves, shows in contrast that optimal control yields really longer traversal time.

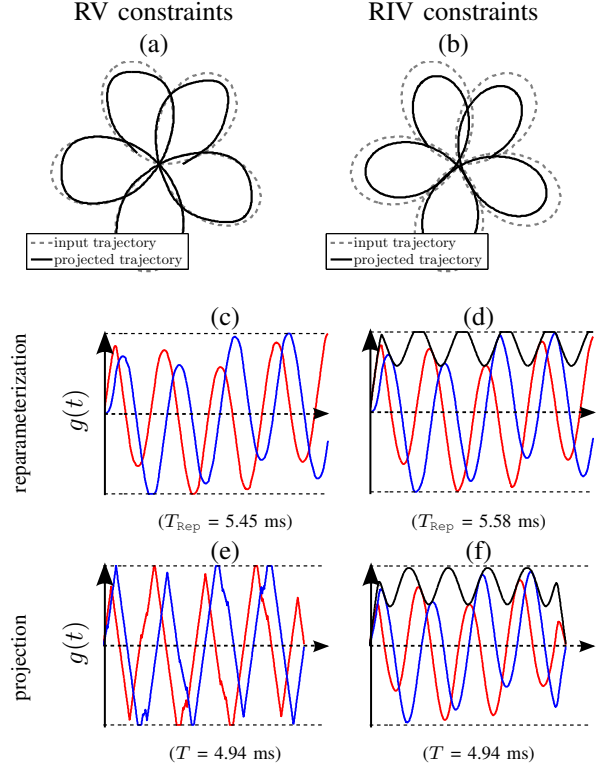


Fig. 5. Gradient waveforms to design rosette trajectories [25]. Left (resp. right), with Rotation-Variant (RV) (resp. Rotation-Invariant (RIV)) constraints ($\|\cdot\| = \|\cdot\|_{\infty}$ (resp. $\|\cdot\|_{\infty,2}$)). Corresponding gradient waveforms ($g_x(t)$, $g_y(t)$, $\|g(t)\|$). **Top row:** k -space trajectories between -6 and 6 cm⁻¹. **Middle row:** gradient waveforms corresponding to optimal reparameterization [12]. **Bottom row:** gradient waveforms corresponding to the proposed projection algorithm.

B. TSP sampling

Next, we performed similar experiments with a TSP trajectory, as introduced in [15]. We illustrate the results for RIV constraints, given an initial parameterization at 50 % of the maximal speed. In Fig. 6(a), we notice that the output trajectory is a smoothed version of the initial piece-wise linear trajectory. Although the output curve has a slightly different support, the traversal time is reduced from 58 ms (optimal reparameterization that can be computed exactly) to 16 ms (Fig. 6(b-c)). The main reason is that TSP trajectories embody singular points, requiring the gradients to be set to zero for each of these points. 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.

This second example shows that with existing methods, it is hopeless to implement TSP-based sequences in many MRI

modalities, since the time to collect data can be larger than any realistic repetition time. In contrast, our method enables to traverse such curves in a reasonable time which can be tuned according to the image weighted (T_1 , T_2 , proton density).

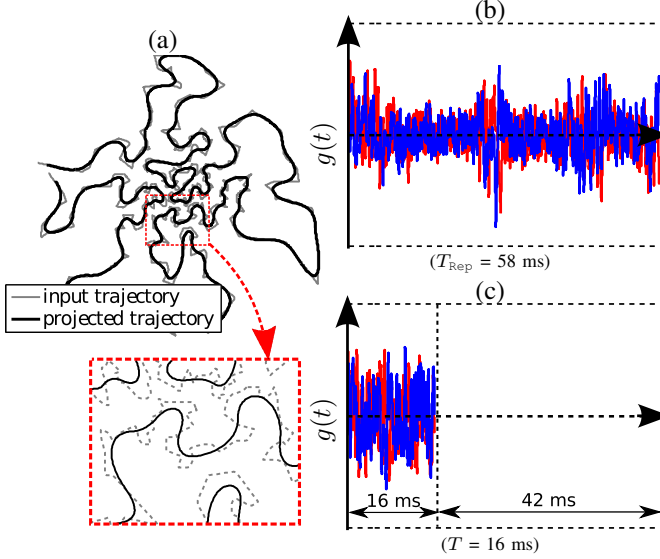


Fig. 6. Gradient waveforms to design TSP trajectories [25]. (a): k -space trajectories between -6 and 6 cm^{-1} . (b): gradient waveforms ($g = (g_x, g_y)$) corresponding to optimal reparameterization [12]. (c): gradient waveforms corresponding to the proposed projection algorithm.

VI. 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 main advantage of our approach is that one can reduce the time to traverse a curve if the latter has a lot of high curvature areas, like a solution of the TSP.

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.

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) [2]. 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 [2], 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 [27]. For example, let $n \in \mathbb{N}$ be the number of revolutions and $r : [0, 1] \mapsto \mathbb{R}^+$ a strictly increasing smooth function. Denote by r^{-1} its inverse function. Define the spiral for $t \in [0, n]$ by $c_n(t) = r\left(\frac{t}{n}\right) \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \end{pmatrix}$ and the target distribution π by:

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

then $P_{c_n} \rightharpoonup \pi$ when the number of revolutions 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 [2, 14]. Let us draw n k -space locations uniformly according to a density q define over the $d\mathbf{D}$ 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:

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

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. 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.

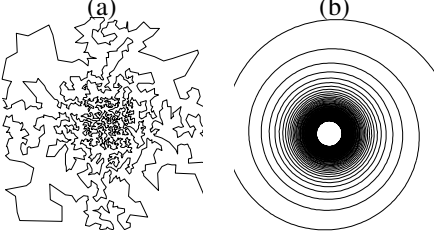


Fig. 7. Examples of Variable Density Sampling. (a): TSP trajectory with empirical distribution depicted in Fig. 2(a). (b): spiral trajectory with empirical radial distribution such as Fig. 2(b).

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., [22]):

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. [29, Theorem 31.3]).

APPENDIX C PROOF OF PROPOSITIONS 4 AND 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 rewritten 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 $\ker(\mathcal{A})^\perp = \text{im}(\mathcal{A}^*)$. Thus $s^*(q_1, q_2) = z + \mathcal{A}^*\lambda$ for some λ such that:

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

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

$$s^*(q_1, q_2) = z + \mathcal{A}^*(\mathcal{A}\mathcal{A}^*)^{-1}(v - \mathcal{A}z),$$

ending the proof.

Proposition 4 results from standard results of convex optimization. See for instance [30, Theorem 1] or the book [28].

APPENDIX D PROOF OF THEOREM 1.

The analysis proposed here follows closely ideas proposed in [24, 31–33]. We will need two results. The first one is a duality result from [32].

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. Assume that g is σ -strongly convex [28] and that $\text{Ari}(\text{dom}(f)) \cap \text{ri}(\text{dom}(g)) \neq \emptyset$. Let $\mathcal{A} \in \mathbb{R}^{m \times n}$ denote a matrix.*

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

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

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

The second ingredient is the standard convergence rate for accelerated proximal gradient descents given in [24].

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

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

To conclude, it suffices to set $g(s) = \frac{1}{2} \|s - c\|_2^2$, $f(q_1, q_2) = \iota_{B_\alpha}(q_1) + \iota_{B_\beta}(q_2)$ and $\mathcal{A} = \begin{pmatrix} \dot{M} \\ \ddot{M} \end{pmatrix}$. By doing so, the

Algorithm 2: Accelerated proximal gradient descent

Input: $q_0 \in \text{ri}(\text{dom}(f^*)) \cap \text{Ari}(\text{dom}(g^*))$.

n_{it}

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

Set $y_0 = q_0$. **for** $k = 1 \dots n_{it}$ **do**

$q^{(k)} = \text{prox}_{t f^*}(y^{(k-1)} + t A \nabla g^*(-A^* y^{(k-1)}))$
 $y^{(k)} = q^{(k)} + \frac{k-1}{k+2}(q^{(k)} - q^{(k-1)})$

projection problem rewrites $\min_{s \in \mathbb{R}^{nd}} p(s) = f(As) + g(s)$.

Its dual problem (9) can be rewritten more compactly as

$\min_{q=(q_1, q_2) \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}} d(q) = g^*(-A^* q) + f^*(q)$. Note that function g is 1-strongly convex. Therefore, Algorithm 2 can be used to minimize d , ensuring a convergence rate in $\mathcal{O}(\frac{1}{k^2})$ on the function values $d(y^{(k)})$. It then suffices to use Proposition 8 to obtain a convergence rate on the distance to the solution $\|s^{(k)} - s^*\|_2^2$.

REFERENCES

- [1] G. Puy, P. Vandergheynst, and Y. Wiaux, "On variable density compressive sampling," *IEEE Signal Processing Letters*, vol. 18, no. 10, pp. 595–598, 2011.
- [2] 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.
- [3] F. Krahmer and R. Ward, "Beyond incoherence: stable and robust sampling strategies for compressive imaging," preprint, 2012.
- [4] B. Adcock, A. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: asymptotic incoherence and asymptotic sparsity in compressed sensing," 2013.
- [5] R. Mir, A. Guesalaga, J. Spiniak, M. Guarini, and P. Irarrazaval, "Fast three-dimensional k-space trajectory design using missile guidance ideas," *Magnetic Resonance in Medicine*, vol. 52, no. 2, pp. 329–336, 2004.
- [6] B. M. Dale, J. S. Lewin, and J. L. Duerk, "Optimal design of k-space trajectories using a multi-objective genetic algorithm," *Magnetic Resonance in Medicine*, vol. 52, no. 4, pp. 831–841, 2004.
- [7] P. T. Gurney, B. A. Hargreaves, and D. G. Nishimura, "Design and analysis of a practical 3D cones trajectory," *Magnetic Resonance in Medicine*, 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," *Magnetic Resonance in Medicine*, vol. 71, no. 1, pp. 278–285, 2014.
- [9] H. H. Wu, J. H. Lee, and D. G. Nishimura, "MRI using a concentric rings trajectory," *Magnetic Resonance in Medicine*, vol. 59, no. 1, pp. 102–112, 2008.
- [10] 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.
- [11] B. A. Hargreaves, D. G. Nishimura, and S. M. Conolly, "Time-optimal multidimensional gradient waveform design for rapid imaging," *Magnetic Resonance in Medicine*, vol. 51, no. 1, pp. 81–92, 2004.
- [12] 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.
- [13] 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.*, 2014. in press.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] R. M. Willett, "Errata: Sampling trajectories for sparse image recovery," note, Duke University, 2011.
- [18] 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," *Medical Imaging, IEEE Transactions on*, vol. 29, no. 6, pp. 1252–1259, 2010.
- [19] 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.
- [20] 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.
- [21] M. Lustig, D. L. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, pp. 1182–1195, Dec. 2007.
- [22] C. Villani, *Optimal transport: old and new*, vol. 338. Springer, 2008.
- [23] 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.
- [24] A. Beck and M. Teboulle, "Gradient-based algorithms with applications to signal recovery," *Convex Optimization in Signal Processing and Communications*, 2009.
- [25] D. C. Noll, "Multishot rosette trajectories for spectrally selective MR imaging," *Medical Imaging, IEEE Transactions on*, vol. 16, no. 4, pp. 372–377, 1997.
- [26] S. Vaziri and M. Lustig, "The fastest gradient waveforms for arbitrary and optimized k-space trajectories," in *Proc. of the 10th IEEE ISBI conference*, (San Francisco, USA), pp. 708–711, Apr. 2013.
- [27] D. M. Spielman, J. M. Pauly, and C. H. Meyer, "Magnetic resonance fluoroscopy using spirals with variable sampling densities," *Magnetic Resonance in Medicine*, vol. 34, no. 3, pp. 388–394, 1995.
- [28] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I: Part I: Fundamentals*, vol. 305. Springer, 1996.
- [29] R. T. Rockafellar, *Convex analysis*. No. 28, Princeton university press, 1997.
- [30] Y. Nesterov, "Smooth minimization of non-smooth functions," *Math. Program.*, vol. 103, pp. 127–152, May 2005.
- [31] 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.
- [32] 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.
- [33] 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.