Global Diffusion Tractography
by Simulated Annealing

Marc C. Robini*, *Member, IEEE*, Matthew Ozon, Carole Frindel, Feng Yang, and Yuemin Zhu

Abstract—*Objective:* Our goal is to develop a robust global tractography method for cardiac diffusion imaging. *Methods:* A graph is stretched over the whole myocardium to represent the fiber structure, and the solutions are minima of a graph energy measuring the fidelity to the data along with the fiber density and curvature. The optimization is performed by a variant of simulated annealing that offers increased design freedom without sacrificing theoretical convergence guarantees. *Results:* Numerical experiments on synthetic and real data demonstrate the capability of our tractography algorithm to deal with low angular resolution, highly noisy data. In particular, our algorithm outperforms the Bayesian model-based algorithm of Reisert *et al.* (*NeuroImage*, vol. 54, no. 2, 2011) and the graph-based algorithm of Frindel *et al.* (*Magn. Reson. Med.*, vol. 64, no. 4, 2010) at the noise levels typical of in vivo imaging. *Conclusion:* The proposed algorithm avoids the drawbacks of local techniques and is very robust to noise, which makes it a promising tool for *in vivo* diffusion imaging of moving organs. *Significance:* Our approach is global in terms of both the fiber structure representation and the minimization problem. It also allows us to adjust the trajectory density by simply changing the vertex-lattice spacing in the graph model, a desirable feature for multiresolution tractography analysis.

Index Terms—Cardiac imaging, diffusion tensor imaging, global tractography, simulated annealing.

I. INTRODUCTION

A. Motivation and Purpose

Diffusion-tensor MRI (DT-MRI) is currently the only available technique to study water diffusion in living tissues without alteration, and hence the modality of choice for assessing the cardiac architecture *in vivo*. Indeed, the anisotropic form and organization of cardiomyocytes confer anisotropic characteristics to the diffusion of water molecules in the myocardium, so that DT-MRI is suitable for studying the structure of the heart wall [1], [2]. However, since cardiomyocytes are about

100 μm long, cardiac diffusion tractography merely depicts a macroscopic directional tendency; in other words, individual fiber tracts have no meaning per se in cardiac imaging. Yet, because cardiomyocyte arrangements form elongated structures with preferential local orientations, cardiac tractograms can be viewed as sets of coarse-scale fiber bundles. For simplicity, we will refer to such fiber bundles as "fibers", but always keeping in mind the true meaning of this shortcut term.

Our purpose is to design a method to extract global cardiac architectures from low resolution, highly noisy DT-MRI data such as those obtained under *in vivo* imaging constraints. As opposed to local tractography methods, which treat the fibers as independent streamlines and are quite sensitive to measurement errors, global strategies use the accumulation and interaction of information over the whole dataset and are thus more robust to noise. We propose a new formulation to global tractography in which the solution takes the form of a boolean-weighted undirected graph representing short-range spatial interactions. Our model leads to the minimization of an energy function that includes prior knowledge on fiber organization in addition to fidelity to the data. The optimization is performed by a variant of the simulated annealing (SA) algorithm with optimal convergence properties; so our approach is "doubly global" in the sense that both the fiber structure representation and the minimization problem are global.

B. Background and Connection With Previous Work

The first proposed (and most widely used) tractography approach is rooted in fluid mechanics, where streamlines are collections of fluid particles with instantaneous displacements collinear with the velocity field. In DT-MRI, streamlines are identified to fibers and estimated from the diffusion-tensor field either deterministically [3], [4] or probabilistically [5]–[7]. Streamlining has been used in human connectome studies [8]–[10] and in the description of the cardiac fiber architecture [11]. Due to its local nature, streamlining is sensitive to noise and is best suited to high-angular resolution data.

Global tractography techniques trace back to the early 2000s [12], but their development was boosted by the Bayesian approach of Jbabdi *et al.* [13] in which tractography is constrained by global connections between brain regions. In [14], O'Donnell *et al.* propose a probabilistic fiber atlas that allows multiple-subject registration of whole brain tractograms. In a similar vein, Li *et al.* [15] use a graph model whose vertices are predefined cortex regions of interest. Atlas-free methods take

advantage of graph representations in a different way: a graph is attached to the tensor field and the fiber-structure solutions are the edge configurations that minimize a global energy [16]–[18]. The method proposed in this paper belongs to this category and is related to a previous work of ours [16]; in both cases, a graph is stretched over the whole myocardium to represent the fiber structure, and the optimal tractograms minimize an energy function whose variables are binary edge weights. However, our approach here differs from that in [16] in two ways. First, we design a new energy function to increase robustness to noise; in particular, the data-fidelity penalty is no longer discrete and the prior penalties introduce more valuable information on fiber shape. Second, we perform optimization using a recent generalization of SA, namely, stochastic continuation (SC) [19]–[21], which offers more design freedom without sacrificing theoretical convergence guarantees.

C. Organization

This paper is organized as follows. Section II is devoted to the proposed model for global tractography and Section III describes the associated stochastic optimization algorithm. We present our numerical experiments in Section IV, followed by concluding remarks.

II. GLOBAL GRAPH-BASED TRACTOGRAPHY

A. Assumptions

When dealing with low angular resolution and high noise, the poor quality of the data can be compensated by the fact that, at the macroscopic scale, the myocardium is a network of fiber bundles that neither cross nor split apart. In this case, we can use the single-tensor model: the noise-free diffusion signal along the ith encoding direction is of the form

$$S_i = S_0 \, e^{-b \boldsymbol{g}_i^T \boldsymbol{D} \boldsymbol{g}_i} \tag{1}$$

where S_0 is the signal with no diffusion gradient applied, b is a positive constant, $\boldsymbol{g}_i \in \mathbb{R}^3$ is the ith gradient direction, and $\boldsymbol{D} \in \mathbb{R}^{3 \times 3}$ is the tensor characterizing the diffusion (in practice, $\boldsymbol{D}$ is estimated from the reference signal S_0 and the diffusion signals S_i by singular value decomposition [22]). This model allows us to use a standard acquisition scheme such as spin-echo with a small number of encoding directions; it is therefore compatible with *in vivo* imaging constraints, as opposed to multitensor representations obtained by high angular resolution diffusion imaging [23], diffusion spectrum imaging [24], or q-ball imaging [25].

The information used by our tractography algorithm is a vector field $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3$ carrying information on the principal diffusion direction and its reliability. For every $\boldsymbol{x} \in \mathbb{R}^3$,

$$\boldsymbol{f}(\boldsymbol{x}) = \text{FA}(\boldsymbol{x}) \boldsymbol{v}_1(\boldsymbol{x}) \tag{2}$$

where $\text{FA}(\boldsymbol{x})$ and $\boldsymbol{v}_1(\boldsymbol{x})$ are the fractional anisotropy and the normalized principal eigenvector of the tensor at position $\boldsymbol{x}$, respectively. This compression of the tensor information reduces computational costs and is fully justified for noisy diffusion data; indeed, while high fractional anisotropy translates to

high confidence that the fiber orientation coincides with the principal eigenvector, small fractional anisotropy indicates that no preferential orientation can be reliably inferred from the tensor.

B. Graph Model

We represent the fiber architecture by a boolean-weighted undirected graph

$$\mathcal{G} = (\mathcal{V}, \mathcal{E}, \boldsymbol{\omega}) \tag{3}$$

where the set of vertices $\mathcal{V}$ is supported by a cubic lattice covering the myocardium, the set of edges $\mathcal{E}$ connects the vertex pairs defined by a nearest-neighbor system on $\mathcal{V}$, and $\boldsymbol{\omega} \in \{0,1\}^{|\mathcal{E}|}$ is a set of boolean weights indicating the fiber locations (the notation $|\mathcal{E}|$ stands for the number of edges in $\mathcal{G}$). More precisely, the set of edges is defined by

$$\mathcal{E} = \{\{v, v'\} \subset \mathcal{V} \mid v' \in \mathcal{N}(v)\} \tag{4}$$

$$\mathcal{N}(v) = \{v' \in \mathcal{V} \mid 0 < \|\boldsymbol{x}(v') - \boldsymbol{x}(v)\| \leqslant \rho\} \tag{5}$$

where $\boldsymbol{x}(v) \in \mathbb{R}^3$ denotes the spatial position of v and the radius $\rho > 0$ is fixed ($\|\cdot\|$ denotes the usual Euclidean norm). The weight $\boldsymbol{\omega}(e)$ of an edge $e \in \mathcal{E}$ has the following interpretation:

$$\boldsymbol{\omega}(e) = \begin{cases} 1 & \text{if } e \text{ is part of a fiber} \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

We call e active if $\boldsymbol{\omega}(e) = 1$ and inactive if $\boldsymbol{\omega}(e) = 0$.

C. Energy Function

The graph model facilitates the representation of the interactions between the lattice points, and hence the introduction of prior knowledge to design a suitable energy function for global tractography. We first describe the data-fidelity penalty, which is similar to the fiber score of Aganj *et al.* [26] in that we use line integrals along the fiber paths and we measure the confidence in the data via the fractional anisotropy. Then, we introduce prior information in the form of fiber density and curvature penalties to obtain the final formulation of the tractography problem. Our choice encourages the active edges of the graph model to be consistent with the data in anisotropic regions (that is, where there is no ambiguity in the interpretation of the data), whereas in isotropic regions, the data-fidelity constraint is relaxed and the diffusion direction is inferred from neighboring anisotropic locations via the fiber density and curvature penalties.
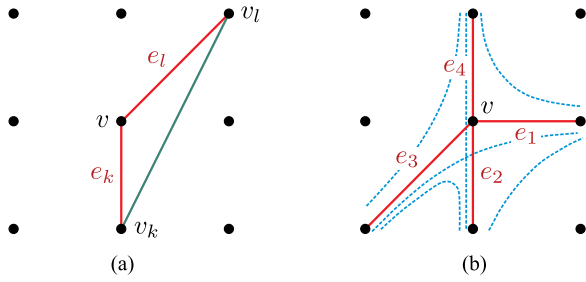
1) Data Fidelity: Let $F : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}_+$ be the segment penalty function defined by

$$F(\boldsymbol{a}, \boldsymbol{b}) = \overline{\text{FA}}([\boldsymbol{a}, \boldsymbol{b}]) - \frac{1}{\|\boldsymbol{b} - \boldsymbol{a}\|} \int_{[\boldsymbol{a}, \boldsymbol{b}]} |\boldsymbol{f} \cdot d\boldsymbol{x}| \tag{7}$$

where $\overline{\text{FA}}([\boldsymbol{a}, \boldsymbol{b}])$ denotes the average fractional anisotropy over the segment $[\boldsymbol{a}, \boldsymbol{b}]$ and $\int_{[\boldsymbol{a}, \boldsymbol{b}]} |\boldsymbol{f} \cdot d\boldsymbol{x}|$ is the nonoriented line integral of $\boldsymbol{f}$ along $[\boldsymbol{a}, \boldsymbol{b}]$. Equivalently,

$$F(\boldsymbol{a}, \boldsymbol{b}) = \int_0^1 \text{FA}(\boldsymbol{s}(t)) \left(1 - \frac{|\langle \boldsymbol{v}_1(\boldsymbol{s}(t)), \boldsymbol{b} - \boldsymbol{a} \rangle|}{\|\boldsymbol{b} - \boldsymbol{a}\|} \right) dt \tag{8}$$

where $\boldsymbol{s}(t) = \boldsymbol{a} + t(\boldsymbol{b} - \boldsymbol{a})$ and $\langle \cdot, \cdot \rangle$ denotes the usual inner product. The segment penalty $F(\boldsymbol{a}, \boldsymbol{b})$ is zero when the principal

Fig. 1. (a) Neighborhood of a vertex v supporting two active edges e_k and e_l. The penalty $F_{k,l}$ is obtained by integrating over the green segment. (b) Four active edges containing v and possibly corresponding continuous fibers (blue dotted curves) — the data-fidelity penalty at v is given by (11).

diffusion direction v_1 is parallel to $b - a$ at each point of $[a, b]$. Conversely, $F(a, b)$ reaches its maximum value $\overline{\mathrm{FA}}([a, b])$ when v_1 is orthogonal to $b - a$ everywhere on $[a, b]$.

The adequacy between the graph model and the diffusion data is measured using segment penalties associated with adjacent pairs of edges, that is, penalties

$$F_{k,l} = F(x(v_k), x(v_l)) \tag{9}$$

with v_k and v_l being such that $e_k = \{v_k, v\}$ and $e_l = \{v, v_l\}$ are active edges for some common vertex v, as schematized in Fig. 1(a). Let $Z_v(\omega)$ denote the number of pairs of active edges containing a given vertex v. The data-fidelity penalty at v is defined by

$$U_0(v, \omega) = \frac{1}{Z_v(\omega)} \sum_{\{v_k, v_l\} \subset \mathcal{N}(v)} \omega(e_k)\omega(e_l)F_{k,l} \tag{10}$$

if $Z_v(\omega) \geqslant 1$ and $U_0(v, \omega) = 0$ if $Z_v(\omega) = 0$. For example, in the case shown in Fig. 1(b), the data-fidelity penalty at v is the average of the six segment penalties obtained by pairing the four active edges, that is,

$$U_0(v, \omega) = \frac{1}{6} \sum_{\{k,l\} \in \mathcal{C}(4,2)} F_{k,l} \tag{11}$$

where $\mathcal{C}(4, 2)$ is the set of 2-combinations of $[1 \,.\, .\, 4]$.

If $Z_v(\omega)$ is nonzero, $U_0(v, \omega)$ is the average of the segment penalties $F(x(v_k), x(v_l))$ over the pairs of distinct active edges $\{v_k, v\}$ and $\{v, v_l\}$. Therefore, as long as v is adjacent to two or more active edges, $U_0(v, \omega)$ does not favor dense over sparse active-edge configurations or vice versa. Besides, $U_0(v, \omega)$ is independent of the resolution of the vertex lattice, since the line integral $\int_{[a,b]} |f \cdot dx|$ in (7) is normalized by the integration length.

2) Fiber Density: At the macroscopic scale, the diffusion trajectories in the myocardium can be viewed as curvilinear objects without intersections [2]. This observation is also supported by the helical ventricular myocardial band model [27], [28], which allows us to describe fiber bundles as 1-D curves lying on the 2-D myocardial band embedded in $\mathbb{R}^3$. In the graph model, this translates to a homogeneous cross-sectional fiber density fixed by the resolution of the vertex lattice.

Let $d_v(\omega) \in \mathbb{N}$ be the active degree of a vertex v, that is,

$$d_v(\omega) = \sum_{v' \in \mathcal{N}(v)} \omega(\{v, v'\}). \tag{12}$$

To discourage fiber splitting and fiber crossing, we attach to each vertex v the density penalty

$$U_1(v, \omega) = \begin{cases} 1/2 & \text{if } d_v(\omega) \in \{0, 1\} \\ 0 & \text{if } d_v(\omega) = 2 \\ d_v(\omega) & \text{if } d_v(\omega) \geqslant 3. \end{cases} \tag{13}$$

This term favors vertices with two adjacent active edges, and hence fiber structures in which each vertex belongs to a single fiber. The cases $d_v(\omega) = 0$ and $d_v(\omega) = 1$ correspond, respectively, to a nonconnected vertice and a fiber ending. Since these events are much more likely than fiber splitting and fiber crossing, they are assigned a penalty significantly smaller than in the case where $d_v(\omega) \geqslant 3$.

3) Fiber Curvature: Histological and X-ray computed tomography observations in healthy hearts show that the cardiomyocytes are well-aligned at the microscopic scale and that they form fiber bundles with small or moderate curvature at the macroscopic scale [2], [29]. In the graph model, defining the fiber curvature at a vertex position $x(v)$ makes sense only if v is incident with exactly two active edges, say $\{v_k, v\}$ and $\{v, v_l\}$, in which case we denote by $\theta(v)$ the angle between the vectors $x(v_k) - x(v)$ and $x(v_l) - x(v)$. We propose to limit the formation of high-curvature trajectories via the penalty

$$U_2(v, \omega) = \begin{cases} 1 - f_\sigma(\theta(v) - \pi/2) & \text{if } d_v(\omega) = 2 \\ 1 & \text{otherwise} \end{cases} \tag{14}$$

where f_σ denotes the logistic function with steepness parameter $\sigma > 0$:

$$f_\sigma(t) = \frac{1}{1 + \exp(-\sigma t)}. \tag{15}$$

If $d_v(\omega) = 2$, the term $1 - f_\sigma(\theta(v) - \pi/2)$ penalizes angles smaller than $\pi/2$ (that is, high local curvature) and promotes angles close to π (that is, aligned active edges). If $d_v(\omega) \neq 2$, the curvature penalty is set to its maximum value 1 to avoid an effect opposite to that of the density penalty $U_1(v, \omega)$. The definition of $U_2(v, \omega)$ amounts to thinking of $\theta(v)$ as a random variable following a logistic distribution with mean $\pi/2$ and variance $\pi^2/(3\sigma^2)$. Following this interpretation, $f_\sigma(t)$ is the probability that $\theta(v) \geqslant t + \pi/2$, and we set σ so that $f_\sigma(\pi/6) = 0.9$, or, equivalently, $P(\theta(v) \leqslant 2\pi/3) = 0.1$.

4) Graph Energy and Fiber-Structure Solutions: The raw solutions to the graph-based tractography problem are the edge-weight configurations that minimize a function combining the data-fidelity and prior penalties. This function is called the graph energy and is defined on the set $\Omega = \{0, 1\}^{|\mathcal{E}|}$ by

$$U(\omega) = \sum_{v \in \mathcal{V}} \left(U_0(v, \omega) + \alpha U_1(v, \omega) + \beta U_2(v, \omega) \right) \tag{16}$$

where α and β are positive parameters respectively controlling the strengths of the fiber density and curvature penalties.

In this way, the priors $U_1(\boldsymbol{\omega}) = \sum_v U_1(v, \boldsymbol{\omega})$ and $U_2(\boldsymbol{\omega}) = \sum_v U_2(v, \boldsymbol{\omega})$ balance the action of the data-fidelity penalty $U_0(\boldsymbol{\omega}) = \sum_v U_0(v, \boldsymbol{\omega})$ by favoring solutions with sparse splittings and crossings and moderate average curvature. Put another way, the minima of U are compromises between two extreme cases: on the one hand, minimizing U_0 alone produces totally disconnected active-edge configurations (because $U_0(v, \boldsymbol{\omega}) = 0$ if $Z_v(\boldsymbol{\omega}) = 0$), and on the other hand, minimizing a positive linear combination of U_1 and U_2 yields straight-line fibers.

A discrete fiber-structure solution is a set $\{\phi_i\}_{i \in \mathcal{I}}$ of noncyclic paths $\phi_i = (v_k^i)_{k \in [1..K_i]}$ extracted from a raw solution $\boldsymbol{\omega}^*$. More precisely, for every $i \in \mathcal{I}$ and every $k \in [1..K_i - 1]$, the vertex v_{k+1}^i is in the neighborhood of v_k^i, the edge $\{v_k^i, v_{k+1}^i\}$ is active in $\boldsymbol{\omega}^*$, and v_k^i appears only once in ϕ_i, that is,

$$v_{k+1}^i \in \mathcal{N}(v_k^i), \quad \boldsymbol{\omega}^*(\{v_k^i, v_{k+1}^i\}) = 1$$
$$\text{and} \quad v_l^i \neq v_k^i \text{ if } l \neq k. \tag{17}$$

Since the positions of the vertices are fixed, we view each discrete fiber ϕ_i as a set of control points of a Bézier curve $\varphi_i : [0, 1] \to \mathbb{R}^3$ representing a real fiber bundle:

$$\varphi_i(t) = \sum_{k=0}^{K_i} B_k^{K_i}(t) \, \boldsymbol{x}(v_k^i) \tag{18}$$

where B_k^K denotes the kth Bernstein basis polynomial of degree K. We call $\{\varphi_i\}_{i \in \mathcal{I}}$ a smooth fiber-structure solution.

III. STOCHASTIC OPTIMIZATION

Our global tractography problem is a large-scale labeling problem which consists in finding binary edge-weights so that the graph energy U in (16) is the lowest possible. The main difficulty lies in the huge computational cost, as U has many local minima and the state space Ω has large dimensionality (for example, the dimension $|\mathcal{E}| = \frac{1}{2} \sum_v |\mathcal{N}(v)|$ of Ω is close to 10^5 for the 26 nearest-neighbor system and a small $20 \times 20 \times 20$ vertex lattice).

The most popular deterministic methods for solving large-scale discrete labeling problems are iterated conditional modes (ICM) [30], loopy belief propagation [31], tree-reweighted message passing [32], and graph cuts [33]. However, except for ICM, these methods are restricted to specific classes of energy functions that do not contain U. As regards probabilistic approaches, evolutionary algorithms are not adapted to large-scale optimization problems (although some have interesting global convergence properties [34]), which leaves us with SA. Roughly speaking, SA can be viewed as a stochastic version of ICM whose key feature is to allow uphill moves to escape local minima. These moves are accepted with a probability controlled by a temperature parameter that decreases slowly enough to ensure probabilistic convergence to a global minimum.

This section is divided into five sections: Sections III-A and III-B provide a short description of SA along with finite-time convergence results (we refer to [20] and [21] for a comprehensive introduction to annealing-based algorithms); Section III-C is dedicated to the design of a candidate-solution generation mechanism adapted to the proposed graph energy; Section III-D

gives implementation details; and Section III-E discusses the tuning of the temperature schedule.

A. Simulated Annealing Algorithm

Let U be a real-valued function to be minimized on a finite state space Ω. An SA algorithm with energy U is a discrete-time, nonhomogeneous Markov chain $(X_n)_{n \geqslant 0}$ with transitions controlled by a *cooling schedule* and a *communication mechanism*. The cooling schedule is a sequence of temperatures $(T_n)_{n \geqslant 1}$ decreasing to zero, and the communication mechanism is a Markov matrix $q : \Omega^2 \to [0, 1]$ giving the probabilities of the possible moves in the state space. This matrix must have *symmetric support* and be *irreducible*, that is, $q(\boldsymbol{\omega}, \boldsymbol{\omega}') > 0$ if and only if $q(\boldsymbol{\omega}', \boldsymbol{\omega}) > 0$, and any state can be reached from any other state with a positive probability and in finitely many moves. The transitions probabilities of the annealing chain $(X_n)_n$ are given by

$$\mathsf{P}(X_n = \boldsymbol{\omega}' \,|\, X_{n-1} = \boldsymbol{\omega})$$
$$= q(\boldsymbol{\omega}, \boldsymbol{\omega}') \exp\left(-\max\left\{\frac{U(\boldsymbol{\omega}') - U(\boldsymbol{\omega})}{T_n}, 0\right\}\right) \tag{19}$$

and the probability to stay at the current state $\boldsymbol{\omega}$ is $1 - \sum_{\boldsymbol{\omega}' \neq \boldsymbol{\omega}} \mathsf{P}(X_n = \boldsymbol{\omega}' \,|\, X_{n-1} = \boldsymbol{\omega})$. In other words, downhill moves are unconditionally accepted, whereas an uphill move from $\boldsymbol{\omega}$ to $\boldsymbol{\omega}'$ at iteration n is accepted with probability $\exp(-(U(\boldsymbol{\omega}') - U(\boldsymbol{\omega}))/T_n)$. In practice, a finite-time realization $(\boldsymbol{\omega}_n)_{n \in [0..N]}$ of $(X_n)_n$ is generated as follows:

---

pick an initial state $\boldsymbol{\omega}_0 \in \Omega$
for $n = 1$ **to** N
    draw a state $\boldsymbol{\omega}'$ from the
         probability distribution $q(\boldsymbol{\omega}_{n-1}, \cdot)$
    set $\Delta \longleftarrow U(\boldsymbol{\omega}') - U(\boldsymbol{\omega}_{n-1})$
    if $\Delta \leqslant 0$ **then** set $\boldsymbol{\omega}_n \longleftarrow \boldsymbol{\omega}'$
    else
        choose $\delta \in [0, 1]$ uniformly at random
        if $\delta \leqslant \exp(-\Delta/T_n)$ **then** set $\boldsymbol{\omega}_n \longleftarrow \boldsymbol{\omega}'$
        else set $\boldsymbol{\omega}_n \longleftarrow \boldsymbol{\omega}_{n-1}$
        end(if)
    end(if)
end(for)

B. Main Convergence Results

The intuition behind SA is that the distribution of X_n concentrates on the global minima of U as T_n decreases to zero. Although early theory advocates logarithmic cooling [35], [36], successful applications of SA use exponential cooling for practical convergence reasons. The justification for exponential cooling is given in [37], where it is proved that the convergence-speed exponent of SA has an upper bound $\nu^\star$ and that this bound can be reached in the following sense: there exist $T_0 > 0$ and a positive sequence $(\zeta_N)_{N \geqslant 1}$ such that the final state X_N^N of the annealing chain $(X_n^N)_{n \in [1..N]}$ controlled by the cooling

schedule $(T_0 \exp(-n\zeta_N))_{n\in[1..N]}$ satisfies

$$\ln \mathsf{P}\big(U\big(X_N^N\big) > \min\nolimits_\Omega U\big) \sim \ln N^{-\nu^\star}. \tag{20}$$

It follows that for every $\nu \in (0, \nu^\star)$,

$$\mathsf{P}\big(U\big(X_N^N\big) > \min\nolimits_\Omega U\big) \leqslant N^{-\nu} \tag{21}$$

when N is large enough, which shows that SA with exponential cooling can have a convergence-speed exponent arbitrarily close to the maximum possible value $\nu^\star$.

In practice, the efficiency of SA depends strongly on the choice of the communication mechanism and on the tuning of the cooling schedule. Before discussing our choices in the next sections, it should be noted that the constraints on the design of SA algorithms can be relaxed by letting the communication mechanism vary with temperature. This variant of SA belongs to the class of SC algorithms [20], [21] and shares the optimal convergence properties of standard SA under the following conditions.

1) As T decreases to zero, the temperature-dependent communication mechanism q_T converges pointwise to a Markov matrix q;
2) The limit communication mechanism q has symmetric support, is irreducible, and can rest anywhere in the sense that $q(\boldsymbol{\omega}, \boldsymbol{\omega}) > 0$ for all $\boldsymbol{\omega}$;
3) The set of possible moves eventually freezes, that is, $q_T(\boldsymbol{\omega}, \boldsymbol{\omega}') > 0$ if and only if $q(\boldsymbol{\omega}, \boldsymbol{\omega}') > 0$ when T is small enough.

C. Communication Strategy

Our communication strategy uses three basic mechanisms $q^{(j)}, j \in \{1, 2, 4\}$, defined on the set $\Omega = \{0, 1\}^{|\mathcal{E}|}$ of all edge-weight configurations. These mechanisms are illustrated in Fig. 2 and allow the following types of moves.

1) *Single-edge moves* ($j = 1$): a single-edge move is implemented by picking and edge e uniformly at random and then inverting its weight $\boldsymbol{\omega}(e)$ (from active to inactive or vice versa).
2) *Edge-pair moves* ($j = 2$): an edge-pair move is a combination of two single-edge moves sharing a common vertex; it is implemented by picking a vertex v and two distinct edges $e_k = \{v_k, v\}$ and $e_l = \{v, v_l\}$ uniformly at random and then inverting the weights $\boldsymbol{\omega}(e_k)$ and $\boldsymbol{\omega}(e_l)$.
3) *Edge-chain moves* ($j = 4$): an edge-chain move is a combination of four single-edge moves; it consists in inverting the weights of the edges in a walk $e_{k_1} e_{k_2} e_{k_3} e_{k_4}$ selected by picking uniformly at random $e_{k_1} = \{v_{l_1}, v_{l_2}\}$, $e_{k_2} = \{v_{l_2}, v_{l_3}\} \in \mathcal{E} \setminus \{e_{k_1}\}$ and so on.

Let $\Omega_j(\boldsymbol{\omega})$ denote the set of states $\boldsymbol{\omega}'$ such that $q^{(j)}(\boldsymbol{\omega}, \boldsymbol{\omega}')$ is nonzero, that is,

1) $\Omega_1(\boldsymbol{\omega})$ is the set of states that differ from $\boldsymbol{\omega}$ in exactly one edge;
2) $\Omega_2(\boldsymbol{\omega})$ is the set of states that differ from $\boldsymbol{\omega}$ in exactly two edges e_k and e_l such that $|e_k \cap e_l| = 1$; and
3) $\Omega_4(\boldsymbol{\omega})$ is the set of states that differ from $\boldsymbol{\omega}$ in exactly four edges $e_{k_1}, \ldots, e_{k_4}$ such that $|e_{k_i} \cap e_{k_{i+1}}| = 1$ for each $i \in \{1, 2, 3\}$.
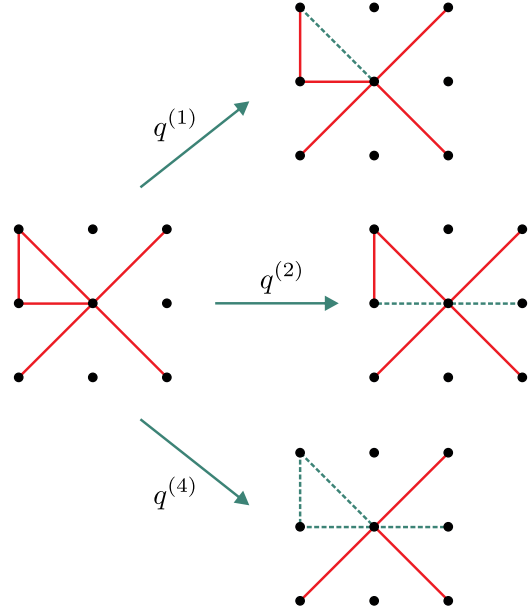

Fig. 2. Examples of moves starting from the edge configuration on the left using the communication mechanisms $q^{(j)}, j \in \{1, 2, 4\}$. The red solid lines represent the active edges, and the green dotted lines indicate the edges whose weights are to be inverted if the move is accepted.

In all three cases, $\boldsymbol{\omega}' \in \Omega_j(\boldsymbol{\omega}) \Longleftrightarrow \boldsymbol{\omega} \in \Omega_j(\boldsymbol{\omega}')$, and thus $q^{(1)}, q^{(2)}$, and $q^{(4)}$ have symmetric support. Besides, since any state can be reached from any other state by inverting a finite number of weights, $q^{(1)}$ is irreducible. However, $q^{(2)}$ and $q^{(4)}$ are not, as edge-pair and edge-chain moves do not change the parity of the number of active edges. A simple way to enforce the irreducibility of $q^{(2)}$ and $q^{(4)}$ is to use these mechanisms in conjunction with single-edge moves. This leads to the compound mechanism

$$q = \kappa_1 q^{(1)} + \kappa_2 q^{(2)} + \kappa_4 q^{(4)} \tag{22}$$

where the coefficients $\kappa_j \in [0, 1]$ are such that $\kappa_1 > 0$ and $\kappa_1 + \kappa_2 + \kappa_4 \leqslant 1$. Each κ_j is the probability of choosing $q^{(j)}$ for generating a move, and the quantity $1 - (\kappa_1 + \kappa_2 + \kappa_4)$ is the probability of letting the algorithm rest in the current state. The pseudocode implementation of q is the following.

---

choose $\kappa \in [0, 1]$ uniformly at random
if $\kappa < \kappa_1$ **then** draw $\boldsymbol{\omega}'$ from $q^{(1)}(\boldsymbol{\omega}, \cdot)$
else if $\kappa < \kappa_1 + \kappa_2$ **then** draw $\boldsymbol{\omega}'$ from $q^{(2)}(\boldsymbol{\omega}, \cdot)$
else if $\kappa < \kappa_1 + \kappa_2 + \kappa_4$ **then**
    draw $\boldsymbol{\omega}'$ from $q^{(4)}(\boldsymbol{\omega}, \cdot)$
else set $\boldsymbol{\omega}' \longleftarrow \boldsymbol{\omega}$
end(if)

Ideally, the communication mechanism should ensure efficient exploration of the state space at high temperatures as well as a "reasonable" acceptance rate at low-temperatures. A simple way to balance these conflicting demands is to use a

temperature-dependent mechanism of the form

$$q_T = \sum_{j \in \{1,2,4\}} \kappa_j(T) q^{(j)}. \qquad (23)$$

The resulting SC algorithm shares the convergence properties of standard SA if each function $\kappa_j : (0, +\infty) \to [0, 1]$ has a limit $\kappa_j(0^+)$ as T decreases to zero and if $\kappa_1(0^+) > 0$ and $\kappa_1(0^+) + \kappa_2(0^+) + \kappa_4(0^+) < 1$. The functions κ_1 and κ_4 should be respectively decreasing and increasing to favor smaller moves at low temperatures and larger moves at high temperatures. The choice of κ_2 depends on whether edge-chain moves are used or not. If $\kappa_4 = 0$, then κ_2 should be increasing so that edge-pair moves are more likely at high temperatures. Otherwise, we can favor edge-pair moves at mid-temperatures and edge-chain moves at high temperatures by choosing κ_2 to be increasing on $(0, T^*]$ and decreasing on $[T^*, +\infty)$ for some $T^* > 0$.

D. Implementation Details

The computation of the graph energy (16) is expensive, but it does not need to be performed at each SA iteration. Rather, we only have to compute the energy difference $U(\boldsymbol{\omega}') - U(\boldsymbol{\omega})$ between the candidate and current solutions.

By definition, $U(\boldsymbol{\omega}') - U(\boldsymbol{\omega})$ breaks down into the differences in data fidelity, fiber density, and fiber curvature. The computation of the difference in data fidelity, namely,

$$U_0(\boldsymbol{\omega}') - U_0(\boldsymbol{\omega}) = \sum_{v \in \mathcal{V}} \left(U_0(v, \boldsymbol{\omega}') - U_0(v, \boldsymbol{\omega}) \right) \qquad (24)$$

is by far the most expensive and therefore deserves particular attention. For every vertex v and every edge e containing v, we define

$$Y_{v,e}(\boldsymbol{\omega}) = \frac{2}{\left(d_v(\boldsymbol{\omega}) - \boldsymbol{\omega}(e) \right) \left(d_v(\boldsymbol{\omega}) - \boldsymbol{\omega}(e) + 1 \right)} \qquad (25)$$

with the convention that $Y_{v,e}(\boldsymbol{\omega}) = 0$ if $d_v(\boldsymbol{\omega}) = \boldsymbol{\omega}(e)$. If $\boldsymbol{\omega}'$ differs from $\boldsymbol{\omega}$ in exactly one edge $e = \{v, v'\}$, it can be shown that

$$U_0(\boldsymbol{\omega}') - U_0(\boldsymbol{\omega}) = (1 - 2\boldsymbol{\omega}(e))$$
$$\times \left(Y_{v,e}(\boldsymbol{\omega}) \sum_{u \in \mathcal{N}_+(v) \setminus \{v'\}} F(\boldsymbol{x}(u), \boldsymbol{x}(v)) \right.$$
$$\left. + Y_{v',e}(\boldsymbol{\omega}) \sum_{u \in \mathcal{N}_+(v') \setminus \{v\}} F(\boldsymbol{x}(u), \boldsymbol{x}(v')) \right) \qquad (26)$$

where $\mathcal{N}_+(v) = \{u \in \mathcal{N}(v) \mid \boldsymbol{\omega}(\{u, v\}) = 1\}$ (note that the segment penalties in the first sum are all different from those in the second sum). In the case of multiple edge moves such as edge-pair and edge-chain moves, the difference in data fidelity is simply obtained by chaining single-edge moves; that is, if $\boldsymbol{\omega}'$ and $\boldsymbol{\omega}$ differ in j edges $e_{k_1}, \ldots, e_{k_j}$, then

$$U_0(\boldsymbol{\omega}') - U_0(\boldsymbol{\omega}) = \sum_{i \in [1..j]} \left(U_0(\boldsymbol{\omega}^{(i)}) - U_0(\boldsymbol{\omega}^{(i-1)}) \right) \qquad (27)$$

where $\boldsymbol{\omega}^{(0)} = \boldsymbol{\omega}$ and for every $i \in [1..j]$, $\boldsymbol{\omega}^{(i)}$ is the configuration generated from $\boldsymbol{\omega}^{(i-1)}$ by inverting the weight $\omega(e_{k_i})$ so that $U_0(\boldsymbol{\omega}^{(i)}) - U_0(\boldsymbol{\omega}^{(i-1)})$ has the form (26).

To speed-up the optimization process, the segment penalties $F(\boldsymbol{x}(u), \boldsymbol{x}(v))$ are precomputed and stored in a sparse matrix with rows and columns indexed by $\mathcal{V}$. The numerical evaluation of (26) then requires at most $d_v(\boldsymbol{\omega}) + d_{v'}(\boldsymbol{\omega}) + 2\boldsymbol{\omega}(e) + 3$ floating-point operations and $2\boldsymbol{\omega}(e) + 4$ fixed-point operations. On average, this computational cost is multiplied by 2 for edge-pair moves and by 4 for edge-chain moves.

E. Tuning of the Cooling Schedule

The theoretical results in [21] suggest to use piecewise-constant exponential cooling schedules of the form

$$T_n = T_{\max} \left(\frac{T_{\min}}{T_{\max}} \right)^{\frac{1}{\gamma - 1} (\lceil n/L \rceil - 1)} \qquad (28)$$

where γ is the number of constant-temperature stages, each of length L (we use $L = |\mathcal{E}|$ in our experiments), and $\lceil \cdot \rceil$ is the ceiling function. Usually, the total number of iterations $N = \gamma L$ is fixed by the available computing resources, and thus the efficiency of annealing is governed by the start and end temperatures $T_1 = T_{\max}$ and $T_N = T_{\min}$.

It is common practice to select $T_{\max}$ and $T_{\min}$ so that the uphill acceptance rates $\chi_{\max}$ and $\chi_{\min}$ at the beginning and end of the annealing chain satisfy $1 > \chi_{\max} > 0.5 \gg \chi_{\min} > 0$ [38]. Here, we do so by using finite-time realizations of the homogeneous Markov chains with transition matrices $q_\infty = \lim_{T \to +\infty} q_T$ and $q_0 = \lim_{T \to 0^+} q_T$. The first step is to simulate these chains independently until their realizations each contain a given number M of uphill moves, that is, until we have a set $\mathcal{M}_\infty$ of M pairs $(\boldsymbol{\omega}, \boldsymbol{\omega}')$ such that $U(\boldsymbol{\omega}') > U(\boldsymbol{\omega})$ and $q_\infty(\boldsymbol{\omega}, \boldsymbol{\omega}') > 0$, and a set $\mathcal{M}_0$ defined similarly for q_0. Then, $T_{\max}$ and $T_{\min}$ are chosen to be the temperatures such that the average acceptance probabilities of the moves in $\mathcal{M}_\infty$ and $\mathcal{M}_0$ are equal to given values of $\chi_{\max}$ and $\chi_{\min}$. More precisely, $T_{\max}$ is the solution of

$$\sum_{(\boldsymbol{\omega}, \boldsymbol{\omega}') \in \mathcal{M}_\infty} \exp\left(-\frac{U(\boldsymbol{\omega}') - U(\boldsymbol{\omega})}{T} \right) = M \chi_{\max} \qquad (29)$$

and $T_{\min}$ is the solution of the similar equation with $\mathcal{M}_0$ and $\chi_{\min}$ replacing $\mathcal{M}_\infty$ and $\chi_{\max}$.

Because exponential cooling is not much affected by an overestimation of $T_{\max}$ or an underestimation of $T_{\min}$, we have some latitude in choosing the parameters M, $\chi_{\max}$, and $\chi_{\min}$. From our experience, efficient schedules are obtained by taking M of the order of 10 to 100 times the number of variables of U, $\chi_{\max} \in [0.6, 0.9]$, and $\chi_{\min} \in [10^{-4}, 10^{-3}]$ (in our experiments, we use $M = 50|\mathcal{E}|$, $\chi_{\max} = 0.8$, and $\chi_{\min} = 10^{-3}$).
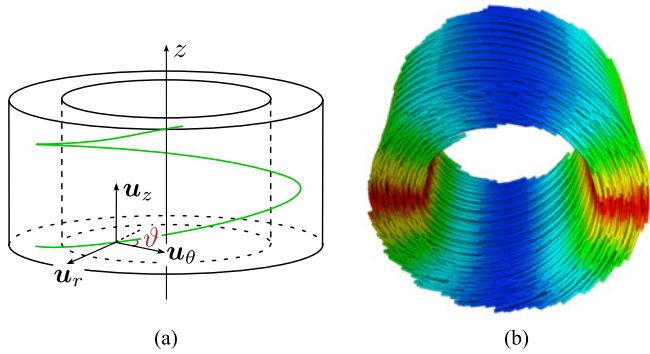
Fig. 3. Numerical phantom: (a) the green curve represents the shape of the tensor data, a helical vector field with fixed angle ϑ; (b) synthetic fibers obtained for $\vartheta = \pi/8$ (the colors reflect the local fiber orientations).

IV. EXPERIMENTAL RESULTS

A. Numerical Phantom Experiments

We begin with numerical phantom experiments so that a ground truth is available to assess performance. The phantom is depicted in Fig. 3 and consists of a set of helical fibers located in the cylindrical region

$$\mathcal{R} = \left\{ (x,y,z) \in \mathbb{R}^3 \mid 8 \leqslant \|(x,y)\| \leqslant 14, \ |z| \leqslant 9 \right\} \quad (30)$$

where the dimensions are in millimeter. The corresponding noise-free DT field $\boldsymbol{D}^*$ has eigenvalues $(\lambda_1, \lambda_2, \lambda_3) = (1, 0.2, 0.2)$ m$^2 \cdot$ s^{-1} everywhere and eigenvectors

$$\begin{cases} \boldsymbol{v}_1 = \cos\vartheta \, \boldsymbol{u}_\theta + \sin\vartheta \, \boldsymbol{u}_z \\ \boldsymbol{v}_2 = -\boldsymbol{u}_r \\ \boldsymbol{v}_3 = \boldsymbol{v}_1 \wedge \boldsymbol{v}_2 \end{cases} \quad (31)$$

where $\boldsymbol{u}_r = (\cos\theta, \sin\theta, 0)$, $\boldsymbol{u}_\theta = (-\sin\theta, \cos\theta, 0)$, and $\boldsymbol{u}_z = (0, 0, 1)$ are the cylindrical basis vectors.

A noisy tensor field is obtained by first adding Rician noise to the diffusion-weighted (DW) volumes generated from $\boldsymbol{D}^*$ (we use the single-tensor model with six gradient directions) and then computing back the tensors from the corrupted DW volumes using singular value decomposition. The standard deviation σ_R of the Rician noise is defined via the decibel level of the signal-to-noise ratio:

$$\mathrm{SNR} \, [\mathrm{dB}] = 20 \log \left(\sigma^* / \sigma_\mathrm{R} \right) \quad (32)$$

where σ^* is the standard deviation of the noise-free DW volumes in the region $\mathcal{R}$. Fig. 4 shows slices of the DW volumes and DT fields for SNRs ranging from $+\infty$ (no noise added) to 0 dB. To fix ideas, the SNR is of the order of 10 dB for *in vivo* cardiac data.

1) Error Measures: We assess the quality of the results using a measure of angular deflection with respect to the expected helical curves and a measure of fiber dissimilarity inspired from [39].

Let $\Phi = \{\varphi_i\}_{i \in \mathcal{I}}$ be a smooth fiber-structure solution obtained as explained in Section II-C4, and let Ψ denote the set of helical fibers of the phantom. We define the distance between
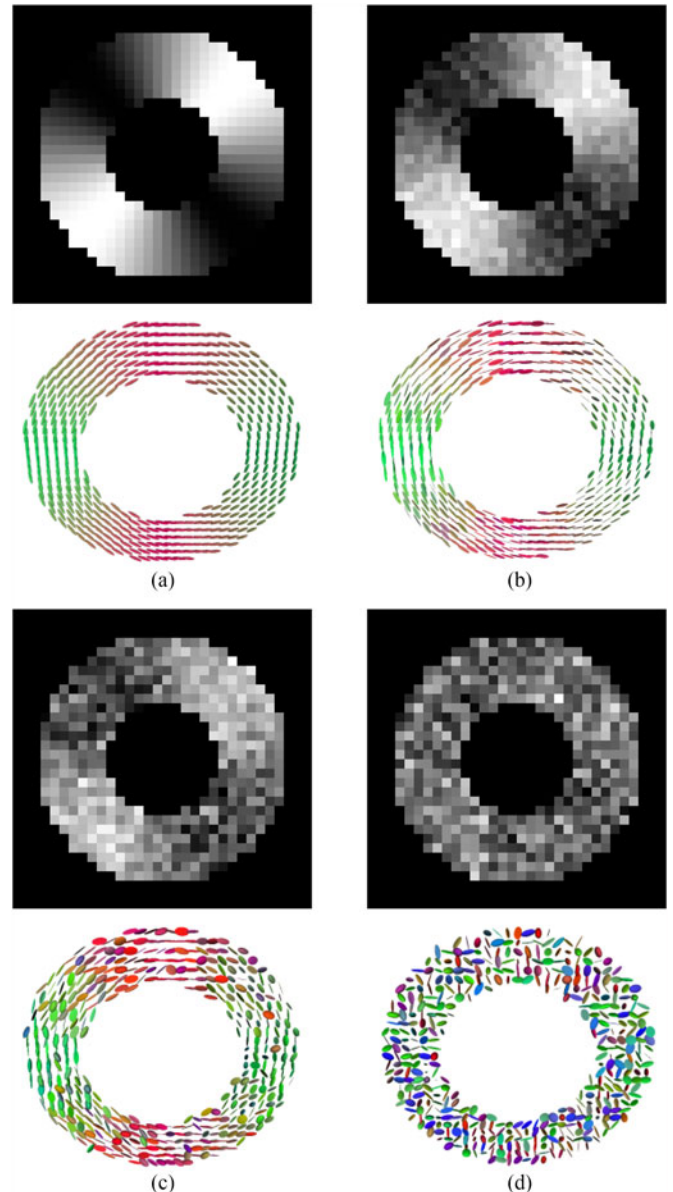

Fig. 4. Slices of the DW volumes and DT fields for different SNRs: (a) $+\infty$ (no noise); (b) 25 dB; (c) 15 dB; (d) 0 dB. (In the ellipsoidal representations of the tensors, the colors encode the orientations and the shape reflects the fractional anisotropy.)

two fibers $\varphi \in \Phi$ and $\psi \in \Psi$ by

$$d(\varphi, \psi) = \left(\frac{1}{\ell(\varphi)} \int_0^1 \|\varphi(t) - \psi(t)\|^2 \|\varphi'(t)\| \mathrm{d}t \right)^{1/2} \quad (33)$$

where $\ell(\varphi)$ is the arc length of φ. For every $i \in \mathcal{I}$, we denote by ψ_i the closest synthetic fiber to φ_i, that is,

$$\psi_i = \underset{\psi \in \Psi}{\arg\min} \, d(\varphi_i, \psi) \quad (34)$$

and we define the average angle between φ_i and ψ_i by

$$\bar{\theta}(\varphi_i) = \frac{1}{\ell(\varphi_i)} \int_0^1 \theta(\varphi'_i(t), \psi'_i(t)) \|\varphi'_i(t)\| \mathrm{d}t \quad (35)$$

where $\theta(\boldsymbol{u}, \boldsymbol{v}) = \arccos(\langle \boldsymbol{u}, \boldsymbol{v} \rangle / (\|\boldsymbol{u}\| \cdot \|\boldsymbol{v}\|))$.

The angular deflection μ_{ang} and the fiber dissimilarity μ_{sim} are defined as follows:

$$\mu_{\mathrm{ang}} = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \sin \overline{\theta}(\varphi_i) \qquad (36)$$

$$\mu_{\mathrm{sim}} = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} d(\varphi_i, \psi_i) \sin \overline{\theta}(\varphi_i). \qquad (37)$$

The former reflects the adequacy of Φ to the principal diffusion direction field, and the latter characterizes the adequacy between Φ and Ψ in terms of distance and angular accuracy. The smaller their values, the better the quality of the estimated fibers.

2) Deterministic Versus Stochastic Optimization:

We first compare the fiber structures estimated from the 25 dB SNR data using three optimization strategies: ICM, standard SA, and SA with temperature-dependent communication (referred to as SC). Our choice of the ICM method to illustrate the benefits of SA compared to a deterministic approach is motivated by two reasons: first, as already mentioned in Section III, other popular deterministic methods for solving large-scale discrete labeling problems do not apply to the graph energy (16), and second, ICM can be likened to SA at zero temperature.

The algorithms considered are ICM with single-edge moves (ICM_1), ICM with edge-pair moves (ICM_2), SA with single-edge moves (SA_1), SC without edge-chain moves ($\mathrm{SC}_{1,2}$), and SC with edge-chain moves ($\mathrm{SC}_{1,2,4}$). The communication mechanism of $\mathrm{SC}_{1,2}$ explores the state space by changing one or two edges at a time, starting with a high probability of selecting edge-pair moves and ending with predominant single-edge moves; it has the form (23) with

$$\kappa_1(T) = \frac{\ln(T_{\mathrm{max}}/T)}{\ln(T_{\mathrm{max}}/T_{\mathrm{min}})}$$

$$\kappa_2(T) = 1 - \kappa_1(T), \quad \text{and} \quad \kappa_4(T) = 0. \qquad (38)$$

(For the above choice of κ_1 and a cooling schedule of the form (28), $(\kappa_1(T_n))_{n \in [1..N]}$ is a piecewise-constant sequence increasing linearly from 0 to 1.) The communication mechanism of $\mathrm{SC}_{1,2,4}$ combines edge-pair and edge-chain moves during the first half of the iterations and edge-pair and single-edge moves during the second half: the function κ_1 is such that $\kappa_1(T_n)$ is zero for $n < \lceil N/2 \rceil$ and then increases from 0 to 1, and κ_2 and κ_4 are defined by

$$\kappa_2(T) = 1 - \max(\kappa_1(T), \kappa_4(T))$$

$$\text{and} \quad \kappa_4(T) = \kappa_1(T_{\mathrm{min}} + T_{\mathrm{max}} - T). \qquad (39)$$

The graph model has vertices on a $29 \times 29 \times 19$ cubic lattice with a 1 mm step size and edges defined by the 26-nearest neighbor system. The prior strength parameters α and β are, respectively, set to 0.2 and 0.5 — these values yield consistently good solutions in terms of fiber similarity and are kept the same in all the experiments. The cooling schedules of SA and SC are tuned as described in Section III-E and the number of iterations N is set to 10^5 times the number of edges.

The fiber structures produced by the different algorithms are shown in Fig. 5. The associated energy and fiber dissimilarity

(a) ICM_1 **(b) ICM_2**

(c) SA_1 **(d) $\mathrm{SC}_{1,2}$**
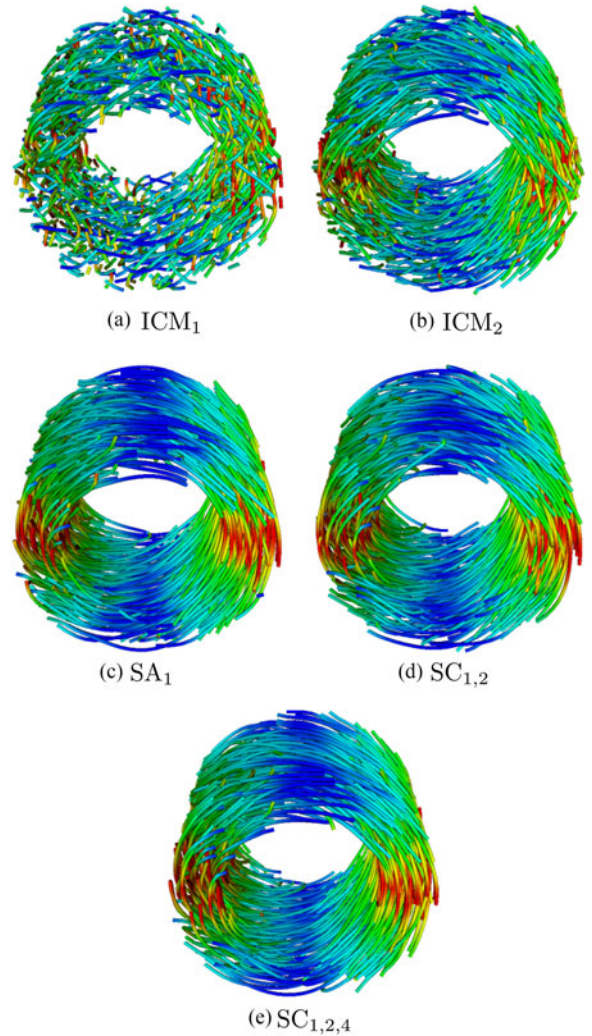
(e) $\mathrm{SC}_{1,2,4}$

Fig. 5. Fiber structures obtained by minimizing the graph energy (16) using different optimization algorithms (the colors reflect the local fiber orientations).

TABLE I
ENERGY (RELATIVE TO ICM_1) AND FIBER DISSIMILARITY OF THE FIBER STRUCTURES SHOWN IN FIG. 5

Algorithm	ICM_1	ICM_2	SA_1	$\mathrm{SC}_{1,2}$	$\mathrm{SC}_{1,2,4}$
Energy	0	-1541	-2480	-2478	-2482
μ_{sim}	0.0839	0.0337	0.0135	0.0128	0.0122

values are reported in Table I (the energy values are given relative to that achieved by ICM with single-edge moves). We observe that ICM gets stuck in poor local minima, especially when using only single-edge moves, which emphasizes the complexity of the energy landscape and suggests that reliable fiber structures are located in deep basins of attractions. (The performance of ICM may be improved by introducing larger moves, but then the computational load would exceed that of SA without even guaranteeing probabilistic convergence to a global minimum.) In contrast, the results obtained using SA and SC are quantita-

	SNR	0 dB	5 dB	10 dB	15 dB
μ_{ang}	$\mathcal{A}_S$	0.6312	0.4656	0.2498	0.1210
	$\mathcal{A}_B$	0.3253	0.2608	0.2306	0.2196
	$\mathcal{A}_G$	0.7186	0.6987	0.4983	0.1770
	$\mathcal{A}^\star$	0.3237	0.2454	0.2230	0.2159
μ_{sim}	$\mathcal{A}_S$	0.0774	0.0365	0.0090	0.0017
	$\mathcal{A}_B$	0.0745	0.0311	0.0274	0.0207
	$\mathcal{A}_G$	0.1975	0.1674	0.1133	0.0069
	$\mathcal{A}^\star$	0.0671	0.0244	0.0104	0.0099

tively close to the true fibers and visually similar. This indicates two things: first, the solutions obtained by stochastic optimization are close to a global minimum (or at least near the bottom of a deep basin of attraction), and second, the proposed graph energy is well defined in the sense that its deepest minima coincide with fiber structures close to the true fibers.

3) Robustness to Noise: We assess the robustness to noise by comparing the SC algorithm $\mathrm{SC}_{1,2,4}$ with the streamline algorithm $\mathcal{A}_S$ of Basser *et al.* [4], the global Bayesian model-based algorithm $\mathcal{A}_B$ of Reisert *et al.* [17] and the global graph-based algorithm $\mathcal{A}_G$ of Frindel *et al.* [16]. The streamline algorithm $\mathcal{A}_S$ is implemented using a fourth-order Runge–Kutta integration scheme with a 0.3 mm-step size (the tracking is stopped when either the fractional anisotropy is smaller than $1/4$ or the angle between two successive directions is smaller than $2\pi/5$). For the Bayesian model-based algorithm $\mathcal{A}_B$, we use the implementation in the DTI&FiberTools package[1] with parameters manually tuned for best performance for heart tractography ($\mathcal{A}_B$ was originally designed for brain tractography). The parameter setting of the graph-based algorithm $\mathcal{A}_G$ is as in [16].

For simplicity, we denote $\mathrm{SC}_{1,2,4}$ by $\mathcal{A}^\star$ from now on. Table II gives the error measures of the fiber structures produced by the four competing algorithms for SNR values of 0, 5, 10, and 15 dB. (As an illustration, Fig. 6 displays the solutions obtained for a 10 dB SNR.) We observe that the angular deflection achieved by $\mathcal{A}_B$ and $\mathcal{A}^\star$ is little affected by noise (contrary to $\mathcal{A}_G$ and $\mathcal{A}_S$) and that $\mathcal{A}^\star$ yields the smallest angular deflection in the $[0, 10]$ dB SNR range. In terms of fiber similarity, $\mathcal{A}^\star$ respectively outperforms $\mathcal{A}_S$, $\mathcal{A}_B$, and $\mathcal{A}_G$ in the $[0, 5]$, $[0, 15]$, and $[0, 10]$ dB SNR ranges. Concerning the results of the streamline algorithm, it should be stressed that shorter fibers are closer to the true fibers in terms of the distance (33), resulting in that μ_{sim} is biased to the disadvantage of long fibers, and hence to the advantage of $\mathcal{A}_S$. Consider for example the fiber structures displayed in Fig. 6(a) and (d). While the fibers estimated by $\mathcal{A}_S$ wave around the true fibers, $\mathcal{A}^\star$ produces a smoother tractogram which is more faithful to the expected helical shape, but μ_{sim} is smaller for $\mathcal{A}_S$ than for $\mathcal{A}^\star$ (0.0090 against 0.0104). The reason
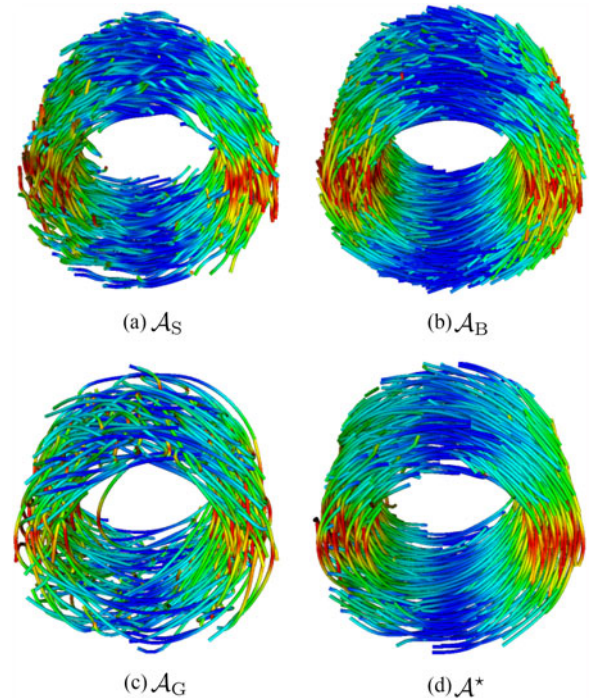

Fig. 6. Fiber structures estimated from the DW volumes corrupted by 10 dB SNR Rician noise: (a) streamline algorithm [4]; (b) Bayesian model-based algorithm [17]; (c) graph-based algorithm [16]; and (d) proposed algorithm.

is that the average fiber length is shorter for $\mathcal{A}_S$ than for $\mathcal{A}^\star$, and thus so is the distance to the set of true fibers, which compensates for the fact that $\mathcal{A}_S$ yields a larger angular deflection than $\mathcal{A}^\star$.

In summary, our algorithm is the most stable and outperforms $\mathcal{A}_S$, $\mathcal{A}_B$, and $\mathcal{A}_G$ for SNR values below 10, 15, and 12 dB, respectively. The increase in performance compared to $\mathcal{A}_G$ is due to (i) the new formulation of the data-fidelity penalty (which performs path integration rather than vertex-wise operations), (ii) the introduction of the fiber curvature penalty, and (iii) the improvement of the optimization strategy. Our results show in particular that a graph-based representation alone does not solve the issue of low SNR data and that the definition of the graph energy is equally important.

B. Experiments on Real Data

The aim of this section is to demonstrate the capability of the proposed algorithm $\mathcal{A}^\star$ to deal with real cardiac data with low angular resolution and low SNR. Our data were acquired *ex vivo* from the heart of a healthy patient who died accidentally, using a Siemens Avanto 1.5T scanner with an echo-planar sequence. The sequence parameters were as follows: 128×128 image size, 54 axial contiguous slices, $2 \times 2 \times 2$ mm^3 voxel size, $N_d = 12$ gradient directions, $N_e = 6$ excitations, 8600 ms repetition time, and 1000 s $\cdot$ mm^{-2} b-value.

We ran the algorithms $\mathcal{A}_S$, $\mathcal{A}_B$, $\mathcal{A}_G$, and $\mathcal{A}^\star$ using the full data (that is, $N_d = 12$ and $N_e = 6$) and a partial data set obtained by keeping six directions uniformly distributed on the

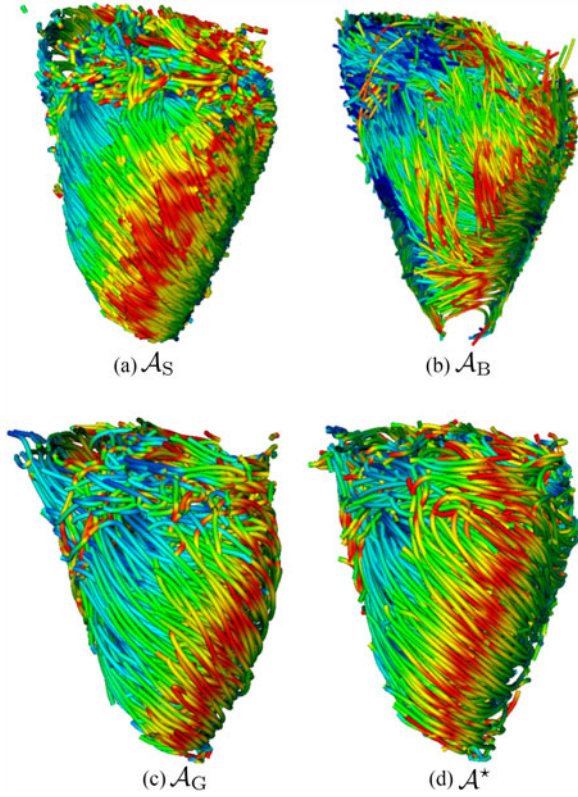[1]www.uniklinik-freiburg.de/mr-en/research-groups/diffperf/fibertools.html

Fig. 7. Fiber structures estimated from human cardiac data acquired using 12 gradient directions and 6 excitations: (a) streamline algorithm [4]; (b) Bayesian model-based algorithm [17]; (c) graph-based algorithm [16]; and (d) proposed algorithm.

half-sphere and one excitation per direction (that is, $N_d = 6$ and $N_e = 1$). The fiber structures estimated from the full data are shown in Fig. 7, where we observe a left-handed helical orientation coherent with the architecture of a healthy human heart. The algorithm $\mathcal{A}^\star$ produces a better organized structure than the other three algorithms, particularly in the basal ring. Fig. 8 shows mid-cavity slices of the fiber structures estimated from the full and partial data sets (first and second column, respectively). The streamline algorithm $\mathcal{A}_S$ (first row) is more sensitive to data loss than the graph-based algorithms $\mathcal{A}_G$ and $\mathcal{A}^\star$ (third and bottom row), and the Bayesian algorithm $\mathcal{A}_B$ (second row) fails to fill the full thickness of the left ventricular wall. Also, the deterioration in fiber quality resulting from data loss is much stronger for $\mathcal{A}_G$ than for $\mathcal{A}^\star$, which agrees with our findings in the numerical phantom experiments.

The spatial sampling effect observed in Fig. 8 for the graph-based algorithms is due to the low resolution of the data and the projective view rather than to the graph model. Indeed, seemingly straight-line fiber pieces are also visible for streamlining (and in the same regions as for the graph-based algorithms), but less clearly because of the higher cross-sectional fiber density. This higher density produces a good visual effect in the full-data case, but the problem with streamlining is that a low SNR is accompanied by a high fiber-splitting rate and a low fiber density, resulting in missing myocardial segments at the noise levels typical of *in vivo* cardiac imaging.
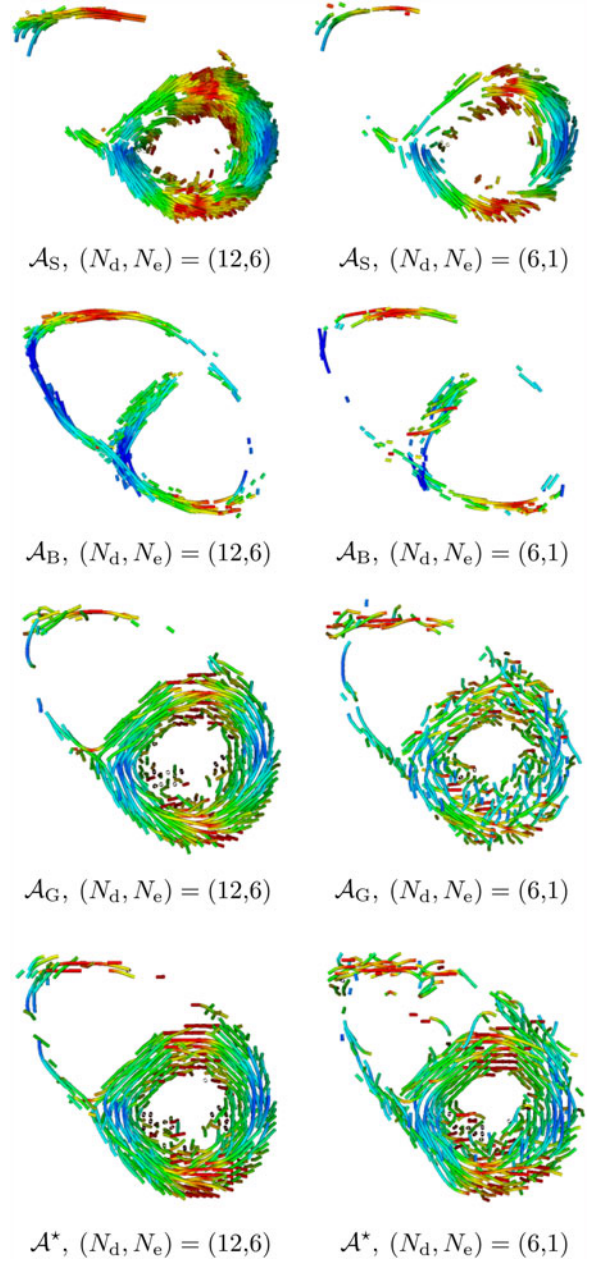

Fig. 8. Mid-cavity slices of the solutions produced by the tractography algorithms (the displayed fiber segments are located in a 4 mm thick slice parallel to the short axis plan). Left column: fiber structures estimated from the full data (12 directions, 6 excitations). Right column: fiber structures estimated from the partial data (6 directions, 1 excitation).

The ground truth being unknown, we cannot confirm the observations made from Figs. 7 and 8 using the error measures employed in the phantom experiments. Instead, we use the average fiber length

$$\mu_L = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \ell(\varphi_i) \tag{40}$$

and the average fiber curvature

$$\mu_\kappa = \frac{1}{|\mathcal{I}|} \sum_{i \in \mathcal{I}} \frac{1}{\ell(\varphi_i)} \int_0^1 \frac{\|\varphi_i'(t) \wedge \varphi_i''(t)\|}{\|\varphi_i'(t)\|^2} \, dt \tag{41}$$

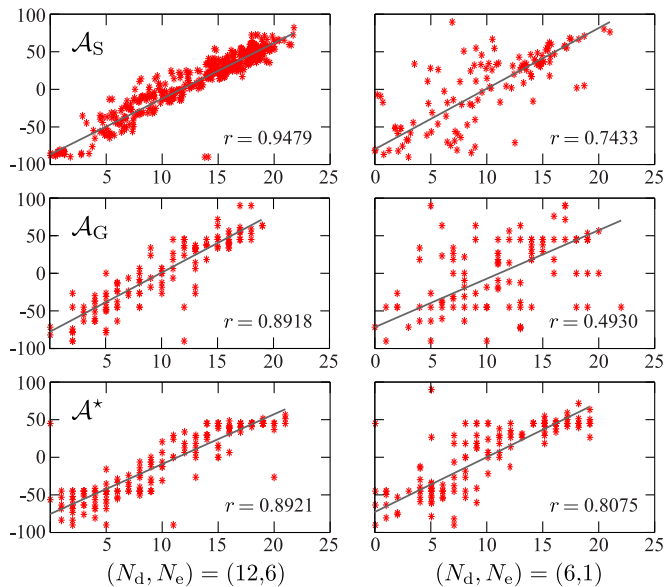|		Algorithm			
(N_d, N_e)	Measure	$\mathcal{A}_S$	$\mathcal{A}_B$	$\mathcal{A}_G$	$\mathcal{A}^\star$
$(12, 6)$	μ_L(mm)	21.1	44.30	153.3	126.1
	μ_κ (mm^{-1})	0.063	0.020	0.017	0.008
$(6, 1)$	μ_L(mm)	20.9	37.07	126.3	90.6
	μ_κ (mm^{-1})	0.059	0.031	0.020	0.012

Fig. 9. Distribution of the helix angle (degrees) in the mid inferolateral segment as a function of the transmural depth (mm). Left column: full data (12 directions, 6 excitations). Right column: partial data (6 directions, 1 excitation). Each plot contains a linear regression line and r is Pearson's correlation coefficient.

where $\wedge$ denotes the vector product operator. The reason for this choice is that long and smooth fibers are in agreement with popular cardiac architecture models such as the helical ventricular myocardial band model [27], [28], the nested "pretzel" geodesic model [40], the three-layer ventricle model [41] and the laminar structure model [42] (a review of these models can be found in [43]). Besides, it is now recognized that cardiac fibers are locally aligned with helical curves wrapping around the ventricles (see, for example, [44], [45] and the references therein), which in our case translates to long and smooth fiber paths relative to the size of the human heart. The average fiber length and curvature of the solutions obtained by the four algorithms are listed in Table III. The streamline and Bayesian algorithms $\mathcal{A}_S$ and $\mathcal{A}_B$ produce short fibers with high curvature, whereas the graph-based algorithms $\mathcal{A}_G$ and $\mathcal{A}^\star$ yield substantially longer and smoother fibers. Furthermore, $\mathcal{A}^\star$ outperforms $\mathcal{A}_G$ in terms of smoothness.

We conclude by examining the helix angle in the mid inferolateral segment (zone 11 of the 17-segment model [46]), which is linearly correlated with the transmural depth [11]. Fig. 9 shows the distributions of the helix angle versus the transmural depth for the algorithms $\mathcal{A}_S$, $\mathcal{A}_G$, and $\mathcal{A}^\star$. These two quantities are well correlated for all three algorithms when using the full data (left column), and $\mathcal{A}^\star$ yields the highest correlation in the case of the partial data (right column). We also observe that the graph-based algorithms maintain the fiber density constant (contrary to streamlining) and that $\mathcal{A}^\star$ is the most stable in terms of variation of the correlation coefficient.

V. CONCLUSION

We proposed a new graph-based method for cardiac tractography which is global in terms of both modeling and optimization. The use of a graph model facilitates the introduction of prior knowledge about healthy cardiac architectures; this information is injected via density and curvature penalties that reduce uncertainties in low angular resolution, highly noisy data. In order to find deep minima of the graph energy, we constructed an SC algorithm with a communication mechanism carefully designed to inherit the theoretical properties of SA.

The conducted experiments show that our global tractography algorithm is robust to noise and outperforms the Bayesian model-based algorithm of Reisert *et al.* [17] and the graph-based algorithm of Frindel *et al.* [16] at the noise levels typical of *in vivo* cardiac imaging. Besides, the graph model is not constrained by the resolution of the data, as the vertices do not need to be attached to the voxels of the DW volumes. Therefore, our approach can be used to extract fiber structures with different fiber densities by simply adjusting the step size of the cubic lattice supporting the vertices. This can be useful, for example, for multiresolution tractography analysis of the cardiac architecture.

REFERENCES

[1] E. Hsu *et al.*, "Magnetic resonance myocardial fiber-orientation mapping with direct histological correlation," *Amer. J. Physiol. Heart Circ. Physiol.*, vol. 274, no. 5, pp. H1627–H1634, 1998.

[2] D. Scollan *et al.*, "Histological validation of myocardial microstructure obtained from diffusion tensor magnetic resonance imaging," *Amer. J. Physiol. Heart Circ. Physiol.*, vol. 275, no. 6, pp. H2308–H2318, 1998.

[3] S. Mori *et al.*, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann. Neurol.*, vol. 45, no. 2, pp. 265–269, 1999.

[4] P. Basser *et al.*, "*In vivo* fiber tractography using DT-MRI data," *Magn. Reson. Med.*, vol. 44, no. 4, pp. 625–632, 2000.

[5] G. Parker and D. Alexander, "Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue," *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, vol. 360, no. 1457, pp. 893–902, 2005.

[6] T. Behrens *et al.*, "Probabilistic diffusion tractography with multiple fibre orientations: what can we gain?" *NeuroImage*, vol. 34, no. 1, pp. 144–155, 2007.

[7] J. Berman *et al.*, "Probabilistic streamline *q*-ball tractography using the residual bootstrap," *NeuroImage*, vol. 39, no. 1, pp. 215–222, 2008.

[8] T. Conturo *et al.*, "Tracking neuronal fiber pathways in the living human brain," *Proc. Nat. Acad. Sci. USA*, vol. 96, no. 18, pp. 10 422–10 427, 1999.

[9] R. Xue *et al.*, "In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging," *Magn. Reson. Med.*, vol. 42, no. 6, pp. 1123–1127, 1999.

[10] S. Wakana *et al.*, "Reproducibility of quantitative tractography methods applied to cerebral white matter," *NeuroImage*, vol. 36, no. 3, pp. 630–644, 2007.

[11] H. Lombaert *et al.*, "Human atlas of the cardiac fiber architecture: study on a healthy population," *IEEE Trans. Med. Imag.*, vol. 31, no. 7, pp. 1436–1447, Apr. 2012.

[12] L. O'Donnell *et al.*, "New approaches to estimation of white matter connectivity in diffusion tensor MRI: Elliptic PDEs and geodesics in a tensor-warped space," in *Proc. 5th Int. Conf. Med. Image Comput. Comput.-Assisted Intervention* 2002, pp. 459–466.

[13] S. Jbabdi *et al.*, "A Bayesian framework for global tractography," *NeuroImage*, vol. 37, no. 1, pp. 116–129, 2007.

[14] L. O'Donnell *et al.*, "Unbiased groupwise registration of white matter tractography," in *Proc. 15th Int. Conf. Med. Image Comput. Comput.-Assisted Intervention* Berlin, Germany: Springer, 2012, pp. 123–130.

[15] L. Li *et al.*, "Quantitative assessment of a framework for creating anatomical brain networks via global tractography," *NeuroImage*, vol. 61, no. 4, pp. 1017–1030, 2012.

[16] C. Frindel *et al.*, "A graph-based approach for automatic cardiac tractography," *Magn. Reson. Med.*, vol. 64, no. 4, pp. 1215–1229, 2010.

[17] M. Reisert *et al.*, "Global fiber reconstruction becomes practical," *NeuroImage*, vol. 54, no. 2, pp. 955–962, 2011.

[18] J.-F. Mangin *et al.*, "Toward global tractography," *NeuroImage*, vol. 80, pp. 290–296, 2013.

[19] M. Robini and I. Magnin, "Optimization by stochastic continuation," *SIAM J. Imag. Sci.*, vol. 3, no. 4, pp. 1096–1121, 2010.

[20] M. Robini, "Theoretically grounded acceleration techniques for simulated annealing," in *Handbook of Optimization, From Classical to Modern Approach* (ser. Intelligent Systems Reference Library), I. Zelinka, V. Snasel, and A. Abraham, Eds. Berlin, Germany: Springer, 2012, vol. 38, pp. 311–336.

[21] M. Robini and P.-J. Reissman, "From simulated annealing to stochastic continuation: A new trend in combinatorial optimization," *J. Glob. Optim.*, vol. 56, no. 1, pp. 185–215, 2013.

[22] P. Kingsley, "Introduction to diffusion tensor imaging mathematics: Part III. Tensor calculation, noise, simulations, and optimization," *Concept Magn. Reson. A*, vol. 28A, no. 2, pp. 155–179, 2006.

[23] D. Tuch *et al.*, "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity," *Magn. Reson. Med.*, vol. 48, no. 4, pp. 577–582, 2002.

[24] J. Wedeen *et al.*, "Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers," *NeuroImage*, vol. 41, no. 4, pp. 1267–1277, 2008.

[25] D. Tuch, "Q-ball imaging," *Magn. Reson. Med.*, vol. 52, no. 6, pp. 1358–1372, 2004.

[26] I. Aganj *et al.*, "A hough transform global probabilistic approach to multiple-subject diffusion MRI tractography," *Med. Image Anal.*, vol. 15, no. 4, pp. 414–425, 2011.

[27] F. Torrent-Guasp *et al.*, "Towards new understanding of the heart structure and function," *Eur. J. Cardiothorac. Surg.*, vol. 27, no. 2, pp. 191–201, 2005.

[28] M. Kocica *et al.*, "The helical ventricular myocardial band: Global, three-dimensional, functional architecture of the ventricular myocardium," *Eur. J. Cardiothorac. Surg.*, vol. 29, Suppl. 1, pp. S21–40, 2006.

[29] F. Varray *et al.*, "High resolution extraction of local human cardiac fibre orientations," in *Proc. 7th Int. Conf. Functional Imaging and Modeling of the Heart (London, U.K.)* Berlin, Germany: Springer, 2013, vol. 7945, pp. 150–157.

[30] J. Besag, "On the statistical analysis of dirty picture," *J. Roy. Statist. Soc. Ser. B*, vol. 48, no. 3, pp. 259–302, 1986.

[31] Y. Weiss and W. Freeman, "On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 736–744, Feb. 2001.

[32] V. Kolmogorov, "Convergent tree-reweighted message passing for energy minimization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1568–1583, Oct. 2006.

[33] V. Kolmogorov and C. Rother, "Minimizing non-submodular functions with graph cuts—A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 7, pp. 1274–1279, Jul. 2007.

[34] O. François, "Global optimization with exploration/selection algorithms and simulated annealing," *Ann. Appl. Probab.*, vol. 12, no. 1, pp. 248–271, 2002.

[35] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.

[36] B. Hajek, "Cooling schedules for optimal annealing," *Math. Oper. Res.*, vol. 13, no. 2, pp. 311–329, 1988.

[37] O. Catoni, "Rough large deviation estimates for simulated annealing: Application to exponential schedules," *Ann. Probab.*, vol. 20, no. 3, pp. 1109–1146, 1992.

[38] M. Robini *et al.*, "Simulated annealing, acceleration techniques, and image restoration," *IEEE Trans. Image Process.*, vol. 8, no. 10, pp. 1374–1387, Oct. 1999.

[39] P. Fillard *et al.*, "Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom," *NeuroImage*, vol. 56, no. 1, pp. 220–234, 2011.

[40] P.-S. Jouk *et al.*, "Analysis of the fiber architecture of the heart by quantitative polarized light microscopy. Accuracy, limitations and contribution to the study of the fiber architecture of the ventricles during fetal and neonatal life." *Eur. J. Cardiothorac. Surg.*, vol. 31, no. 5, pp. 915–921, 2007.

[41] R. Anderson *et al.*, "The anatomical arrangement of the myocardial cells making up the ventricular mass," *Eur. J. Cardiothorac. Surg.*, vol. 28, no. 4, pp. 517–525, 2005.

[42] I. LeGrice *et al.*, "The architecture of the heart: Myocyte organization and the cardiac extracellular matrix," in *Interstitial Fibrosis in Heart Failure* (Dev. Cardiovasc. Med. Series), F. Villarreal, Ed. Berlin, Germany: Springer, 2005, vol. 253, pp. 3–21.

[43] S. Gilbert *et al.*, "Regional localisation of left ventricular sheet structure: Integration with current models of cardiac fibre, sheet and band structure," *Eur. J. Cardiothorac. Surg.*, vol. 32, no. 2, pp. 231–249, 2007.

[44] H. Lombaert *et al.*, "Human atlas of the cardiac fiber architecture: Study on a healthy population," *IEEE Trans. Med. Imag.*, vol. 31, no. 7, pp. 1436–1447, Jul. 2012.

[45] E. Piuze *et al.*, "Maurer–Cartan forms for fields on surfaces: Application to heart fiber geometry," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 12, pp. 2492–2504, Dec. 2015.

[46] M. Cerqueira *et al.*, "Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart," *Circ.*, vol. 105, no. 4, pp. 539–542, 2002.