

NEURONAL FIBER–TRACKING VIA OPTIMAL MASS TRANSPORTATION

A. DADUCCI

Signal Processing Laboratory (LTS5)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
ELD 232 - CH-1015 Lausanne, Switzerland

A. MARIGONDA, G. ORLANDI AND R. POSENATO

Department of Computer Sciences, University of Verona
Strada Le Grazie, 15 - I-37134 Verona, Italy

(Communicated by the associate editor name)

ABSTRACT. Diffusion Magnetic Resonance Imaging (MRI) is used to (non-invasively) study neuronal fibers in the brain white matter. Reconstructing fiber paths from such data (tractography problem) is relevant in particular to study the connectivity between two given cerebral regions. Fiber-tracking models rely on how water molecules diffusion is represented in each MRI voxel. The Diffusion Spectrum Imaging (DSI) technique represents the diffusion as a probability density function (DDF) defined on a set of predefined directions inside each voxel. DSI is able to describe complex tissue configurations (compared e.g. with Diffusion Tensor Imaging), but ignores the actual density of fibers forming bundle trajectories among adjacent voxels, preventing any evaluation of the real physical dimension of these fiber bundles.

By considering the fiber paths between two given areas as geodesics of a suitable well-posed optimal control problem (related to optimal mass transportation) which takes into account the whole information given by the DDF, we are able to provide a quantitative criterion to estimate the connectivity between two given cerebral regions, and to recover the actual distribution of neuronal fibers between them.

1. Introduction.

1.1. General introduction. Diffusion magnetic resonance imaging (MRI) is a powerful imaging modality for non-invasive and in vivo study of the anatomy of the brain white matter. Starting from the observation that water molecules diffusion in living tissues is highly affected by the cellular organization of the observed tissue, MRI exploits this dependency to probe tissue structure architecture at a microscopic scale far beyond the usual imaging resolution. Moreover, since it is well known that in the brain there is a strong relationship between water diffusion and axonal orientations (see [44], [18], [26]), in the last decade diffusion MRI has been widely applied to the study of fiber bundle trajectories into the brain white matter yielding a high number of “tractography” or “fiber-tracking” methods. In general terms, a

2000 *Mathematics Subject Classification.* Primary: 49L25, Secondary 65K10.

Key words and phrases. variational methods in image processing, tractography, optimal mass transportation.

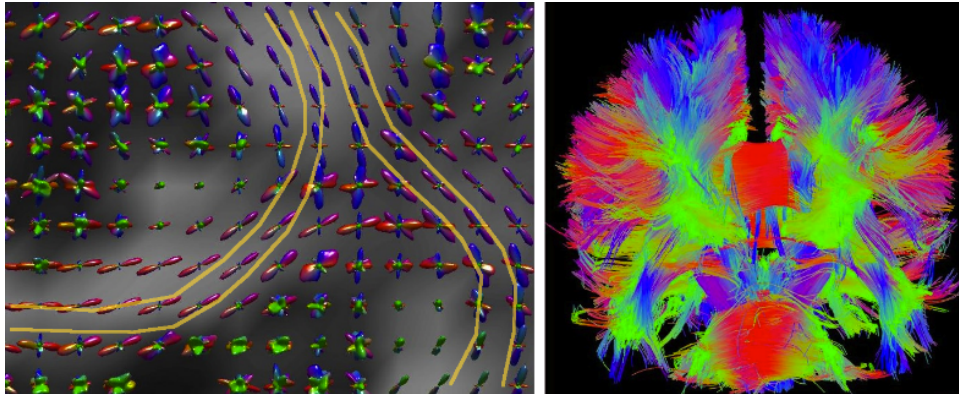


FIGURE 1. Fiber-tracking algorithms exploit diffusion information inside each voxel to infer fiber bundle trajectories capturing coherent orientations of maximum diffusion. Left: a subset of voxels, each of them containing a diffusion distribution, and a set of 4 likely fiber bundle trajectories (brown lines). Right: a representative set of fiber trajectories reconstructed from a healthy subject.

fiber-tracking algorithm exploits water diffusion data of each 3D resolution element (voxel) in order to reconstruct the fiber bundle trajectories present into the brain white matter. Various models are used to describe the water molecules diffusion at the scale of an MRI voxel.

In the simplest case of Diffusion Tensor Imaging (DTI), displacements of water molecules subject to magnetic field are assumed to follow a Gaussian distribution and therefore in each voxel there will be only one direction of maximum diffusion. On the contrary, in the state-of-the-art technique Diffusion Spectrum Imaging (DSI), the diffusion at each voxel is described by a displacement distribution or equivalently by a probability density function (DDF) defined for a set of given directions inside the voxel. The probability density function DDF is reduced to an orientation distribution function (ODF) by summing the probabilities of diffusion in each direction. As opposed to DTI, hence, DSI is able to successfully characterize more complex axonal configurations, such as fiber crossing, fanning and bending.

From a medical point of view, fiber-tracking is relevant to get information about the connectivity between different cerebral regions. It turns out to be important to know not only whether or not two given regions are connected by a bundle of fibers, but also the quantity of fibers forming that connecting bundle. This *connectivity problem* is relevant in particular to describe situations where, due to medical pathologies, we have two cerebral areas connected by fibers that pass through a damaged region where the connection is interrupted or strongly hampered: if the connection was ensured by several fibers, the damaged area will not jeopardize the connection between the two given areas as much as in the case of a scarce number of connecting fibers.

Despite some important progresses made by recent studies focused on assessing the ability of diffusion tractography to estimate anatomically correct fiber bundles, the *validation* of reconstructed fiber trajectories and their *interpretation* remain the two major shortcomings. As regards the interpretation of tractography

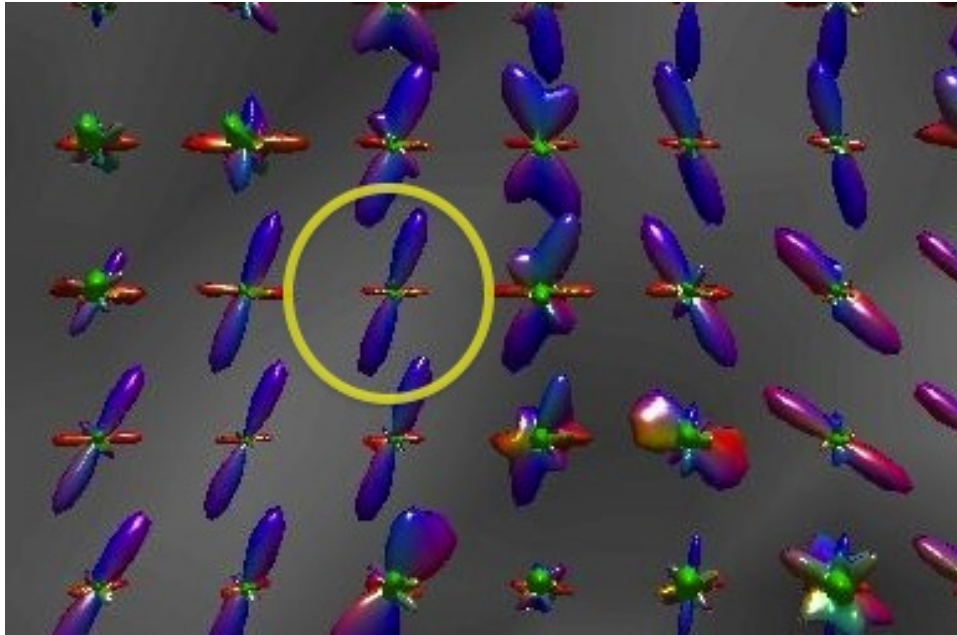


FIGURE 2. An example of Orientation Distribution Function (ODF) obtained by the DSI technique. The technique allows us to define an ODF for each voxel. An ODF is represented as a spherical polar plot coloured according to local diffusion orientations, and it describes the “average amount of diffusion” along each direction over the sphere.

results, present fiber-tracking methods, DSI-based included, build fiber trajectories disregarding the “quantity” of fibers going to make up a bundle trajectory among adjacent voxels, thus preventing any evaluation, for example, of the real physical dimension of a fiber bundle or the global quantity of fibers crossing a brain region.

1.2. Related work. In the literature, diffusion fiber-tracking algorithms are generally divided into two major classes: deterministic and probabilistic ones.

Deterministic tractography algorithms [32],[11],[34] incrementally reconstruct fiber trajectories starting from a seed point and following the maximum local diffusion information (i.e., a streamline process). Such algorithms are relative fast and easy to compute, but lack of robustness and errors are propagated exponentially.

To overcome the shortcomings of deterministic tractography in dealing with the uncertainty in fiber orientations, *probabilistic* algorithms have been introduced [35],[27],[24]. They generate multiple trajectories from a given seed point (up to thousands) by independently repeating a streamline tracking process and randomly choosing the next direction to follow at each step. Some of the reconstructed paths will correspond to anatomically meaningful fiber bundles, others no. The output of such algorithms is not a trajectory, but a map of “probability” for a given point to be traversed by an anatomically genuine fiber bundle. The computation of these maps is burdensome, and their interpretation is still controversial.

More recently, a new approach called *global tractography* has been proposed by means of some fiber-tracking algorithms [29],[45],[46],[23],[30]. These algorithms share the idea of searching a fiber path as a global parameter optimization process in order to improve robustness to the noise or to local errors of diffusion directions. In [29], the authors model fiber trajectories as spline curves having control points randomly drawn as a Bayesian based distribution. The obtained results are interesting, but the method is computational expensive because the initialization of the control points and the sampling process are two very time-consuming processes.

Another recent promising approach has been introduced by [23] and [30], in which each fiber bundle trajectory is obtained by minimizing Ising spin glass-type model. In general, such approach requires a very high computational time.

1.3. Comparison with previous Hamilton-Jacobi-Bellman based models. In [36] it is proposed a variational model for diffusion MR, settled in a continuous framework, based on ideas from control theory.

In this model, it is imagined that an infinite number of particles starts at time 0 from the boundary ∂S of a given *seed region* S evolving along the streamline given by the gradient ∇C_S^* of a suitable *cost function* C_S^* which assigns to each point x a cost to reach the seed region S related to an optimal control problem. At time t the evolution front (i.e. the level curves of C_S^*) will represent the propagation along the medium. If the seed region is sufficiently small, for short time $0 < t < \tau$ we have an approximate picture of the neural net near the seed region.

More precisely, the procedure used by the authors is summarized here (see Section §2 in [36] and Section 4 for the notations):

1. A current cost $\Psi : \mathbb{R}^d \times \mathbb{S}^{d-1} \rightarrow]0, +\infty[$ (called by the authors the *local cost*) is defined. If $\Gamma : [a, b] \rightarrow \mathbb{R}^d$ is a curve, its total cost of is

$$C(\Gamma) := \int_a^b \Psi \left(\Gamma(t), \frac{\dot{\Gamma}(t)}{\|\dot{\Gamma}(t)\|} \right) ds.$$

2. Given a *seed region* S , it is defined the *value function* $C_S^* : \mathbb{R}^d \rightarrow [0, +\infty[$ as

$$C_S^*(x) := \inf \{ C(\Gamma) : \Gamma(b) = x, \Gamma(a) \in S \}.$$

It turns out that the value function C_S^* is the solution of the Hamilton-Jacobi-Bellman equation $H(x, Du) = 0$ with boundary data $u|_S = 0$ where

$$H(x, p) = \max_{v \in \mathbb{S}^{d-1}} \{ \langle p, v \rangle - \Psi(x, v) \}.$$

3. The *propagation front* is the level set of the value function C_S^* , and it evolves with normal speed F given by

$$\|\nabla C_S^*(p)\| F \left(p, \frac{\nabla C_S^*(p)}{\|\nabla C_S^*(p)\|} \right) = 1.$$

4. Diffusion data define the front speed propagation F . The local cost Ψ and the front speed F are related by the Legendre transform:

$$\Psi(x, v) := \max \left\{ \frac{\langle v, \hat{n} \rangle}{F(x, \hat{n})} : \langle v, \hat{n} \rangle \geq 0 \right\},$$

and, in the setting of the problem, the proposed local cost function Ψ is related to the diffusion tensor $D(p)$ (see Section 4.1 in [36]).

5. Finally, we can consider different cost functions, i.e. different functions Ψ_1, Ψ_2 , which yield different value functions $C_{S,1}^*, C_{S,2}^*$ related to the same seed region. The quotient $C_{S,1}^*(x)/C_{S,2}^*(x)$ can be used to compare different costs. If we take Ψ_1 to be identically 1, the related cost $C_{S,1}^*(x)$ is the Euclidean distance of x from S .

This method turns out to be very useful to establish the connectivity of a *single point* p with the region S , since it is sufficient to consider the optimal trajectory (minimizer of the cost) steering p on S , but gives few information about the connectivity between two *regions*.

Experimental data show that neuronal fibers are not uniformly distributed in the brain, hence it is natural to associate also a *neuronal density* to the seed region and let this density evolve along the streamlines. This approach is not covered by [36], since they do not consider the possibility to associate a density to the seed region. Indeed, this will provide more robust information about the real path of neuronal fibers, moreover the problem can be set in the more general framework of optimal transportation theory.

The basic problem that we are going to investigate is about the connectivity between two regions, each endowed with a (possibly nonconstant) density. If we represent initial and final densities as gray scale images, the problem can be seen as a find a continuous interpolant (*warping*) from the first image to the second one. In Section §5.2.3 of [7], it is described a technique, based on optimal transportation theory, to obtain this interpolation. It can be viewed as a *geodesic* problem in the space of probability measures endowed with a metric called *Wasserstein distance*. According to this, if we have two compactly supported probability measures μ_1, μ_2 (in the smooth case, $\mu_i = \rho_i \mathcal{L}^d$), there is a unique vector field $\vec{F}(x)$ such that $\vec{F}(x) \# \mu_0 = \mu_1$, moreover $\vec{F}(x) = \nabla \psi(x)$ where ψ is a scalar convex function, and the geodesic joining μ_0 and μ_1 according to the Wasserstein distance is $\mu_t := (tx + (1-t)\nabla \psi(x)) \# \mu_0$, i.e. the push forward of the initial measure along the points of the *straight line segment* joining $\text{Id}_{\mathbb{R}^d}$ and $\nabla \psi$.

The fundamental idea of our approach is to *combine* the optimal control and optimal transportation approaches, i.e. construct a continuous warping that, in a certain sense, will take into account the geometry induced by the optimal control metric (hence the transport will not be done on straight lines). In the smooth case, where most of our computation are still developed, the problem is essentially reduced to an optimal control problem where densities are transported along optimal trajectories, but this quite general approach allows in principle to treat much more difficult situations.

1.4. An optimal control approach based on optimal mass transport theory. In this paper we propose a model that takes the microstructural information into account during the fiber trajectories building process in the case of DSI tractography. Exploiting microstructural information, our model can be used to investigate some ambiguities that often occur in white matter and are difficult to deal with, e.g. distinguish between fiber crossing and kissing, improving the specificity and accuracy of tractography results.

We aim to take into account the whole information given by the DDF in order to reconstruct the actual underlying diffusion process, which as observed above is highly anisotropic and takes place essentially along the fiber paths.

The knowledge of the underlying diffusion process allows in principle to recover the whole neuronal fiber net by considering the streamlines corresponding to displacements along high probability directions as encoded in the DDF. These should correspond to the actual fibers. Moreover, the ratio between the actual number of fibers crossing in a point associated to different directions will be equal to the ratio between the respective direction probabilities. Another constraint to take into account is that each fiber has no endpoints in the interior of the brain white matter: in other words, the unit tangent field to each fiber has to be divergence free.

Having in mind the previous considerations, we will focus here on the *connectivity problem*, i.e. the problem of finding the bundle of fibers connecting two given (small) cerebral regions, by considering it as the superposition of paths that fulfill the following requirements: their tangents form a solenoidal field that fits as much as possible with the velocity distribution given by the DDF.

Observe that we may consider each single fiber path connecting two given regions as the trajectory of a particle moving along this fiber, starting from the first region and reaching the second one in a fixed lapse of time $T > 0$.

By considering in the first region a mass distribution given by one moving particle on each fiber, we obtain an evolution law relating at each time $0 < t < T$ the velocity of the particles with their distribution, which is quantitatively expressed by the continuity equation (see 2.2). Since neither sources nor sinks are present along the fibers (solenoidality assumption), the total mass during the process is conserved.

We may thus formulate the following problem: consider two given regions, assign to each one a mass distributions with same total mass, and transport the first mass distribution to the second one by respecting as much as possible the probability distribution on the velocities given by the DDF.

This amounts to assign a cost to any such mass transportation process that penalizes particle trajectories (transport paths) with lower probability velocity distribution according to the DDF data.

The minimizing set of trajectories will give as a byproduct a reasonable picture of the neuronal fiber bundle connecting the two regions.

This kind of problem is known as an optimal mass transportation problem, and in general it admits a solution under mild hypothesis. However, without some regularity assumption, solutions turn out to be difficult to interpret.

In the regular cases, the optimal transport solutions can be viewed as a superposition of noncrossing trajectories (transport paths) starting from the initial configuration to the final one (in our situation, they are the trajectories followed by each particle), and the cost function can be expressed as an integral of a cost function assigned to every single trajectory.

The optimization procedure reduces then to minimize the cost of every trajectory. This can be done with standard techniques in the framework of optimal control theory. The superposition of optimal trajectories will give us the picture of the bundle of fibers connecting the two considered regions.

We propose here two different cost functions, which are described in Sections 2.3 and 2.4, which provide optimal transport paths solutions between two given region thanks to Theorem 2.5, yielding in particular well-posedness of the proposed model.

In this way we are able to provide a *quantitative criterion* that measures the connectivity between two localized cerebral regions (see formula (10)), extending similar criteria in [36]. Roughly speaking, we compare the Euclidean distance of

points belonging to the distinct regions with the geodesic distance measured along the corresponding optimal transport paths, which in turn corresponds to the length of the fibers connecting the two regions. For instance, a geodesic distance considerably greater than the euclidean one has to be interpreted as a poor or totally absent neuronal fiber connection between those regions.

The plan of the paper is the following. In the next section we first recall and analyze those quantities related to the DSI data that will enter in our modelization. We then recall basic theoretical background on mass transportation theory and its relation with optimal control and Hamilton-Jacobi PDE's, proving in particular Theorem 2.5. We hence discuss two kind of cost functions related to our problem and analyze their regularity properties to ensure existence of optimal transport paths solutions according to Theorem 2.5.

The final section is devoted to the discussion on advantages and shortcomings of our model, the formula giving our quantitative criterion for brain connectivity, together with further remarks and open problems. In the Appendix we fix notations and give some basic results.

2. Optimal transportation-based models.

2.1. Discussion of the DSI data. The DDF function assigns to each voxel a probability distribution on the space of velocities. A voxel is represented by a set $Q(\alpha)$, defined as the unit cube centered at $\alpha \in \mathbb{Z}^3$. For each $(\alpha, v) \in \mathbb{Z}^3 \times \mathbb{R}^3$, let $N(\alpha)$ be the total number of molecules contained in $Q(\alpha)$, and $N(\alpha, v)$ be the number of molecules in $Q(\alpha)$ displaced of $v\delta t$ in the time δt . Define the (*averaged*) *displacement density function* (DDF) $f_D : \mathbb{Z}^3 \times \mathbb{R}^3 \rightarrow [0, +\infty[$ by setting:

$$f_D(\alpha, v) := \frac{N(\alpha, v)}{N(\alpha)}.$$

Due to the highly anisotropic character of the medium where diffusion occurs, the fraction of molecules displaced in the direction of a neuronal fiber will be significantly different from the fraction of molecules displaced in the orthogonal directions. Moreover, many fibers aligned in the same direction will result in an higher value of the DDF corresponding to that direction than a single fiber.

By construction, we have

$$\int_{\mathbb{R}^3} f_D(\alpha, v) dv = 1,$$

i.e. $f_D(\alpha, \cdot) \in L^1(\mathbb{R}^3)$ is a probability density on \mathbb{R}^3 . This means that if we consider all possible displacements $v \in \mathbb{R}^3$ and sum all the fractions of molecules which moved in that directions, we recover the whole of the initial number of molecules in $Q(\alpha)$.

The main properties of f_D are the following:

1. if we fix α and a direction $w \in \mathbb{S}^2$, we have that the probability of a displacement in direction w follows a normal distribution according to its magnitude;
2. for α fixed, we have that if two directions are sufficiently close, the respective distributions are close, i.e. $w \mapsto f_D(\alpha, w)$ is continuous;

Of particular significance for the construction of our model is the set $M(\alpha)$ of displacements with cumulative probability less or equal 1/2. Remark that this set may be different from the graph of the ODF usually considered in current models.

The set $M(\alpha)$ contains the zero vector and its boundary is given by those vectors v whose modulus is the median of the magnitudes of observed displacements in the

direction $v/|v|$. This set is in general nonconvex, however it is centrally symmetric. We can think to $\partial M(\alpha)$ as a set of averaged observed velocities.

We associate the set $M(\alpha)$ to the center of the voxel. We consider a suitable continuous extension $f(x, v)$ of the DDF to the whole voxel in such a way that

$$f_D(\alpha, v) = \int_{Q(\alpha)} f(x, v) dx,$$

and, for any direction $w \in \mathbb{S}^2$ and $\lambda \in \mathbb{R}$, the profile of $f(x, \lambda w)$ is similar to $f_D(\alpha, \lambda w)$. We set for instance

$$f(x, \lambda w) = a(x, w) e^{-\frac{\lambda^2}{2r(x, w)}},$$

where $a, r : \mathbb{R}^3 \times \mathbb{S}^2 \rightarrow]0, +\infty[$ are continuous strictly positive functions bounded away from 0 such that

$$\int_{w \in \mathbb{S}^2} \int_0^{+\infty} f(x, \lambda w) \lambda^2 d\lambda d\Sigma(w) = 1,$$

with $d\Sigma$ the area element of \mathbb{S}^2 . We assume that, for fixed x , $w \mapsto r(x, w)$ and $w \mapsto a(x, w)$ are Lipschitz continuous functions.

We may hence define the following set of displacements $M(x)$ for any $x \in M(\alpha)$:

$$\begin{aligned} M(x) &:= \left\{ v \in \mathbb{R}^3 \setminus \{0\} : \int_0^{|v|} f\left(x, s \frac{v}{|v|}\right) ds \leq \int_{|v|}^{+\infty} f\left(x, s \frac{v}{|v|}\right) ds \right\} \cup \{0\} \\ &= \left\{ v \in \mathbb{R}^3 \setminus \{0\} : |v| \leq \sqrt{2c\left(x, \frac{v}{|v|}\right) \text{Erf}^{-1}(1/2)} \right\} \cup \{0\}, \end{aligned}$$

where Erf^{-1} is the inverse function of

$$\text{Erf}(s) := \frac{2}{\sqrt{\pi}} \int_0^s e^{-x^2} dx.$$

Boundary points of this set are precisely those displacement with cumulative probability function 1/2. We also define the *dual* $Z(x)$ of $M(x)$ by setting:

$$Z(x) := \{p \in \mathbb{R}^3 : \langle p, v \rangle \leq 1, \forall v \in M(x)\}.$$

The sets $M(x)$ inherit from f some regularity properties: $x \mapsto M(x)$ is continuous with respect to the Hausdorff distance, moreover $M(x)$ is compact, star-shaped and centrally symmetric. These properties of $M(x)$ will be crucial in the regularity issues for the optimal mass transportation problem, defined in term of the sets $M(x)$, that we discuss in the next sections.

Let us summarize these and other immediate properties of the sets $M(x)$ and $Z(x)$.

Proposition 1. *We give a first set of properties on the sets $M(x)$:*

- (M1) $x \mapsto M(x)$ is continuous with respect to the Hausdorff distance (see 4);
- (M2) $0 \in \text{int}(M(x))$, for every $x \in \mathbb{R}^3$;
- (M3) $M(x)$ is a compact set of \mathbb{R}^3 , for every $x \in \mathbb{R}^3$;
- (M4) $M(x) = \overline{\text{int}(M(x))}$, for every $x \in \mathbb{R}^3$;
- (M5) $M(x)$ is strictly star-shaped with respect to 0, for every $x \in \mathbb{R}^3$.

Proof. All the listed properties follow immediately from the definition of $M(x)$. \square

Remark 1. We briefly discuss the physical meaning of some of these properties:

- Property (M2) implies that we should allow small displacements in all the directions, due to possible measurement uncertainty. This follows from the fact that $r(x, w) > 0$ and hence $\inf\{r(x, w) : w \in \mathbb{S}^2\} = \min\{r(x, w) : w \in \mathbb{S}^2\} > 0$.
- Property (M3) says that only bounded displacements are of physical significance, indeed the probability of a displacement in a given direction decays exponentially with its magnitude. This follows from the fact that $\sup\{r(x, w) : w \in \mathbb{S}^2\} = \max\{r(x, w) : w \in \mathbb{S}^2\} < +\infty$.
- Property (M4) is justified by the fact that every displacement measurement is affected by an error. It follows from the continuity of $r(x, \cdot)$.
- Property (M5) implies that there are *no holes* in $M(x)$: if v is a nonzero probability displacement, so are the displacements λv for every $\lambda \in [0, 1]$. This follows from $r(x, w) > 0$.

Proposition 2. *The dual $Z(x)$ of $M(x)$ enjoys the following properties:*

- (Z1) $x \mapsto Z(x)$ is a continuous with respect to the Hausdorff distance;
- (Z2) $0 \in \text{int}(Z(x))$, for every $x \in \mathbb{R}^3$;
- (Z3) $Z(x)$ is a compact set of \mathbb{R}^3 , for every $x \in \mathbb{R}^3$;
- (Z4) $Z(x) = \overline{\text{int}(Z(x))}$, for every $x \in \mathbb{R}^3$;
- (Z5) $Z(x)$ is convex for every $x \in \mathbb{R}^3$.

Proof. (Z1) follows from (M1), (Z2) follows from the boundedness property (M3), (Z3) follows from the nondegeneracy property (M2), (Z4) follows from (Z5) and (Z2), while (Z5) is trivial. \square

These properties of $M(x)$ will be crucial in the regularity issues for the optimal mass transportation model, defined in term of the sets $M(x)$, that we discuss in the next sections.

2.2. Relation with optimal transportation and Hamilton-Jacobi theory.

Assume to have two compactly supported mass distributions with densities ρ_0 and ρ_1 , such that $\int_{\mathbb{R}^3} \rho_0(x) dx = \int_{\mathbb{R}^3} \rho_1(x) dx$, representing respectively the initial and the final configuration of a distribution of particles transported through the neuronal fiber paths. As described in Subsection 1.4, a way to recover the neuronal fiber bundle connecting ρ_0 with ρ_1 is to consider the superposition of the particle trajectories as the solution of an optimal mass transportation problem with marginals ρ_0 and ρ_1 . Namely, we consider the problem of transporting ρ_0 in ρ_1 minimizing a suitable cost which depends on the path followed by the mass particles and which reflects the behavior of the extended DDF. This problem is referred to as a Monge-Kantorovich problem, and it admits an equivalent fluid dynamics interpretation due to Benamou and Brenier (see [12], [13], and Chapter 8 in [42]), that we describe next.

Adopting the Eulerian point of view, for $0 < t < 1$ we denote by $\rho_t(x)$ the density of a fluid at time t , such that $\rho_{t=0}(x) = \rho_0(x)$ and $\rho_{t=1}(x) = \rho_1(x)$, and by $v_t(x)$ the corresponding velocity field: these quantities are related by the *continuity equation*

$$\partial_t \rho_t + \text{div}(v_t \rho_t) = 0,$$

expressing conservation of the fluid mass. For (ρ_t, v_t) we consider the following minimization problem:

$$\text{minimize } \int_0^1 \int_{\mathbb{R}^3} L(x, v_t(x)) \rho_t dx dt, \quad (1)$$

among all the absolutely continuous solution $t \mapsto \rho_t dx$ of the continuity equation (which must be understood in the weak sense, see Section 4) such that $\mu_t := \rho_t dx$ converges in the sense of measure respectively to the initial and final data $\mu_0 := \rho_0 dx$ and $\mu_1 := \rho_1 dx$ when $t \rightarrow 0^+$ and $t \rightarrow 1^-$. The function $L : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow [0, +\infty[$ will be referred as a suitable *current cost* related to our problem. It will penalize the velocities with lower probability according to the DDF. If the vector field v_t is sufficiently smooth, it is possible to consider the *characteristic system*

$$\begin{cases} \frac{d}{dt} T_t = v_t \circ T_t \\ T_{t=0} = \text{id}_{\mathbb{R}^3}, \end{cases} \quad (2)$$

and for each initial datum μ_0 we have that the solution μ_t is given by $\mu_t = T_t \# \mu_0$, where $T_t \# \mu_0$ denotes the push forward of the measure μ_0 by the vector field T_t , namely

$$\int_{\mathbb{R}^3} \varphi(x) d\mu_t(x) = \int_{\mathbb{R}^3} \varphi(T_t(x)) d\mu_0(x).$$

In this case (1) reduces to

$$\inf \left\{ \int_0^1 \int_{\mathbb{R}^3} L(T_t(x), \dot{T}_t(x)) d\mu_0 dt : T_0(x) = x, T_1 \# \mu_0 = \mu_1 \right\},$$

which can be rewritten as:

$$\inf \left\{ \int_{\mathbb{R}^3} \int_0^1 L(\gamma_x(t), \dot{\gamma}_x(t)) dt d\mu_0 : \gamma_x(0) = x, \gamma_x(1) \# \mu_0 = \mu_1 \right\}.$$

If we introduce the cost:

$$c(x, y) := \inf \left\{ \int_0^1 L(\gamma_x(t), \dot{\gamma}_x(t)) dt, \gamma_x(0) = x, \gamma_x(1) = y \right\}, \quad (3)$$

the problem can be recasted in the Monge's form:

$$\inf \left\{ \int_{\mathbb{R}^3} c(x, \gamma_x(1)) d\mu_0 : \gamma_x(0) = x, \gamma_x(1) \# \mu_0 = \mu_1 \right\}. \quad (4)$$

In this case, if we set $\gamma_x(t) = v_t(x)$ and assume that v_t is regular enough to solve (2), we have that $\gamma_x(t)$ is optimal for

$$\inf \left\{ \int_{\mathbb{R}^3} c(x, \tilde{\gamma}_x(t)) d\mu_0 : \gamma_x(0) = x, \tilde{\gamma}_x(t) \# \mu_0 = \gamma_x(t) \# \mu_0 \right\}. \quad (5)$$

We refer to such a v_t as an *optimal interpolating vector field*, which allows us to reconstruct the optimal trajectories.

It is well known that Monge's problem requires some regularity assumptions to admit a solution. But due to lack of compactness the problem cannot be restricted in general to the class of absolutely continuous measures. A relaxed formulation on the space of probability measures, introduced by Kantorovich in 1942, involves the notion of *transport plans* rather than transport paths or optimal transport vector fields, leading to generalized solutions. Consider the set of *transport plans* from μ_0 to μ_1 defined by

$$\Pi(\mu_0, \mu_1) = \left\{ \pi \in \mathcal{P}(\mathbb{R}^3 \times \mathbb{R}^3) : \begin{array}{l} \pi(A \times \mathbb{R}^3) = \mu_0(A), \text{ for all } \mu_0\text{-measurable } A \\ \pi(\mathbb{R}^3 \times B) = \mu_1(B), \forall \mu_1\text{-measurable } B \end{array} \right\},$$

and the following class of functions:

$$\begin{aligned} \mathcal{F}(\mu_0, \mu_1, c) := & \{(f, g) \in L^1(\mathbb{R}^3, \mu_0) \times L^1(\mathbb{R}^3, \mu_1) \\ & : f(x) + g(y) \leq c(x, y) \text{ for } \mu_0\text{-a.e. } x, \mu_1\text{-a.e. } y\}. \end{aligned}$$

The Kantorovich's problem is to find

$$\inf_{\pi \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^3 \times \mathbb{R}^3} c(x, y) d\pi(x, y). \quad (6)$$

This is of course a relaxation of the Monge's problem: indeed if $\gamma_x(1)$ is optimal in the Monge's problem, then $(\text{id} \times \gamma_x(1))\# \mu_0$ is optimal in the Kantorovich's problem. We recall the following result (see Theorem 1.3 in Ref. [42]) ensuring existence of a solution for Kantorovich's problem:

Theorem 2.1. *Let $X = Y = \mathbb{R}^3$, μ_0, μ_1 be probability measures on \mathbb{R}^3 , and let $c : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow [0, +\infty]$ be a l.s.c. cost function. Then*

$$\inf_{\pi \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^3 \times \mathbb{R}^3} c(x, y) d\pi(x, y) = \sup_{(f, g) \in \mathcal{F}(\mu_0, \mu_1, c)} \int_{\mathbb{R}^3} f d\mu_0 + \int_{\mathbb{R}^3} g d\mu_1. \quad (7)$$

Moreover the infimum in the left hand side is attained, and to compute the supremum in the right hand side we can restrict to bounded and continuous functions in $\mathcal{F}(\mu_0, \mu_1, c)$.

The model case is given by $c(x, y) = |y - x|^p$, i.e. a power of the Euclidean distance between x and y . In this case we can give the following:

Definition 2.2. Let $X = Y = \mathbb{R}^3$, $1 \leq p < +\infty$, and μ_0, μ_1 be probability measures on \mathbb{R}^3 . We define the p -Wasserstein distance by setting

$$W_p(\mu_0, \mu_1) := \inf_{\pi \in \Pi(\mu_0, \mu_1)} \left(\int_{\mathbb{R}^3 \times \mathbb{R}^3} |x - y|^p d\pi(x, y) \right)^{1/p},$$

this turns out to be a metric on the space of probability measures with compact support (see Section 7.1 in [6]).

To have more insight on optimal transport plans, we recall the following (cfr. Definition 2.33 in [42])

Definition 2.3. Let $c : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow [0, +\infty]$ be l.s.c. A function $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ is c -concave if there exists a function ψ such that

$$g(y) = \psi^c(y) := \inf_{x \in \mathbb{R}^3} \{c(x, y) - \psi(x)\}.$$

c -concave functions possess a PDE characterization that turns out to be extremely useful in our setting (we refer to [10] for a comprehensive introduction to the theory of viscosity solutions of Hamilton-Jacobi equations):

Proposition 3. *Let f be a Lipschitz continuous function, and $L : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ be continuous in x -variables and such that $v \mapsto L(x, v)$ is convex for every $x \in \mathbb{R}^3$. Consider the function*

$$\varphi(t, x) = \inf \left\{ f(y) + \int_0^t L(\xi(t), \dot{\xi}(t)) dt : \xi(0) = y, \xi(t) = x \right\} \quad (8)$$

Then we have that $\varphi(0, x) = f(x)$, $\varphi(1, x)$ is c -concave, and $\varphi(t, x)$ is the viscosity solution of the following Hamilton-Jacobi equation:

$$\begin{cases} \partial_t \varphi(t, x) + H(x, D_x \varphi(t, x)) = 0, \\ \varphi(0, x) = f(x), \end{cases} \quad (9)$$

where H is the Legendre-Fenchel transform of L (see Corollary 3.6 p.151 in [10]).

We recall next the following characterization of optimal transport plans, referring to Section 2.4 in Ref. [42] and [2] for the proof and further references:

Theorem 2.4. *Assume that there exists a μ_0 -measurable function $c_1 : \mathbb{R}^3 \rightarrow \mathbb{R}$ and a μ_1 -measurable function $c_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $c(x, y) \leq c_1(x) + c_2(y)$ for all $(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3$ and that*

$$\inf_{\pi \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^3 \times \mathbb{R}^3} c(x, y) d\pi(x, y) < +\infty.$$

Then:

1. *there exists a c -concave function $f \in L^1(\mathbb{R}^3, \mu_0)$ such that $f^c \in L^1(\mathbb{R}^3, \mu_1)$ and (f, f^c) achieves the supremum in (7).*
2. *Kantorovich's problem (6) has a solution, and a plan $\pi \in \Pi(\mu_0, \mu_1)$ is optimal iff π is concentrated on the set*

$$\{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 : f(x) + f^c(y) = c(x, y)\}.$$

3. *If $c(x, y)$ is continuous, bounded below, and μ_0, μ_1 are compactly supported, then f, f^c are upper semicontinuous. If $x \mapsto c(x, y)$ is locally Lipschitz on a set U and the Lipschitz constant is locally independent of y , then f can be chosen to be locally Lipschitz on U .*

We notice that Proposition 3 applied taking as f the c -concave function given by Theorem 2.4 produces a function φ , solution of (9) (for further details on this, we recall Subsection 2.5.2, Subsection 5.4.6 in Ref. [42]). If φ is sufficiently smooth, problem (6) admits a solution in the original Monge's formulation.

More precisely we have:

Theorem 2.5. *Assume that*

1. *$L : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ is continuous in the x -variables and convex in the v -variables;*
2. *L is symmetric in the v -variables, i.e. $L(x, v) = L(x, -v)$ for every $x, v \in \mathbb{R}^3$;*
3. *the Legendre transform H of L belongs to $\text{Lip}_{\text{loc}}(\mathbb{R}^3 \times \mathbb{R}^3)$ and it is C^1 in the p -variables;*
4. *the cost $c(x, y)$ satisfies hypothesis of Theorem 2.4.*

Then problem (6) admits a solution with a $C^{1,1}$ displacement interpolation, and the corresponding optimal interpolating velocity vector field v_t in the characteristic system $\dot{T}_t(x) = v_t(T_t(x))$, $T_0(x) = x$, is given by $v_t(x) = \nabla_p H(x, D_x \varphi(t, x))$.

Proof. The proof is inspired by the remark at the beginning of p.181 in [42]. We recall that a function $h(x)$ is called *semiconcave* if there exist $C > 0$ such that the function $x \mapsto h(x) - C|x|^2$ is concave.

According to Theorem 2.4, the initial datum f of Equation (8) can be chosen to be locally Lipschitz continuous.

By Theorem 5.3.8 in Ref. [15], we have that viscosity solutions of Equation (8) are locally semiconcave in $]0, T[\times \mathbb{R}^3$. Observe that since $L(x, q)$, and hence $H(x, p)$ are even respectively in q and p , we have $c(x, y) = c(y, x)$.

Let us now consider the optimal transport from μ_1 to μ_0 . By the strict convexity assumption on L , it is unique and it is simply obtained by reversing the velocity vector field v_t into $-v_{T-t}$, where v_t is the optimal interpolating vector field related to optimal transportation from μ_0 to μ_1 for $0 \leq t \leq T$.

By Equation 13.5 in Ref. [43], we have

$$\partial_v L(x, v_t) = D_x \varphi(t, x) \text{ and } \partial_v L(x, -v_{T-t}) = D_x \psi(t, x)$$

a.e., where ψ solves the Hamilton Jacobi equation $\partial_t \psi + H(x, D_x \psi) = 0$, $\psi(0, x) = f^c(x)$ corresponding to the optimal transport from μ_1 to μ_0 .

This yields $D_x \varphi(t, x) = -D_x \psi(T-t, x)$. Since both φ and ψ are semiconcave in $]0, T[\times \mathbb{R}^3$, we have that there exists $K > 0$ such that for fixed $s, t \in]0, T[$ we have $D_x^2 \varphi(t, x) \leq K$ and $D_x^2 \psi(s, x) \leq K$ in the sense of distributions of \mathbb{R}^3 , moreover also $-D_x^2 \varphi(t, x) = D_x^2 \psi(T-t, x) \leq K$. This implies that, for fixed t , $x \mapsto \varphi(t, x)$ is semiconcave and semiconvex in \mathbb{R}^3 and hence it is $C^{1,1}(\mathbb{R}^3)$ (see [15] for details on properties of semiconcave functions).

According to Chapter 13, Equation 13.5 in Ref. [43], the two relations

$$\nabla_v L(x, v_t(x)) = D_x \varphi(t, x) \text{ and } v_t(x) = \nabla_p H(x, D_x \varphi(t, x))$$

define a vector field v_t which, through the solution $T_t(x)$ of the characteristic system (2), realizes the minimum in the transport problem between ρ_0 and $T_t \# \rho_0$ for every $t \in]0, T[$, i.e. v_t is the optimal interpolating velocity vector field. \square

In the models proposed in Subsections 2.3 and 2.4, we choose a current cost L penalizing the deviations of velocities from being on the boundary of $M(x)$ and ensuring the regularity properties requested in Theorem 2.5. In particular, the whole information on the optimal interpolating transport vector field v_t is encoded into the solution φ of the Hamilton-Jacobi equation (8) associated to the problem.

Viscosity solutions φ of Equation (8) can be viewed as *value functions* of a suitable optimal control problem (cfr. [10]). We will actually formulate our models relying on this point of view, partially following ideas from [21] and [2]. Recall that there is a well-established literature to effectively treat, also from a numerical point of view, such class of optimal control problems.

2.3. Minimum time model. The first model we propose corresponds essentially to choose

$$L(x, v) := \begin{cases} 1, & \text{if } v \in M(x), \\ +\infty, & \text{elsewhere.} \end{cases}$$

The transport cost associated to this choice is given by:

$$c_1(x, y) := \inf\{T > 0 : \text{there exists } \gamma : [0, T] \rightarrow \mathbb{R}^3 \text{ such that}$$

$$\gamma(0) = x, \gamma(T) = y, \dot{\gamma}(t) \in M(\gamma(t)) \text{ for a.e. } t \in [0, T]\},$$

i.e. c_1 is the minimum time needed to steer x to y along absolutely continuous curves $\xi(\cdot)$ satisfying $\dot{\xi}(t) \in M(\xi(t))$ for a.e. t . Notice that according to Aumann's Theorem (see e.g. Theorem 5.15 in Ref. [17]), in the definition of c_1 , $M(x)$ can be replaced by its convex hull. If we set $H_1(x, p) = \delta_{M(x)}(-p) - 1$, where $\delta_C(p) := \sup_{q \in C} q \cdot p$ is the support function, it is known (see Proposition IV.2.3 and Theorem

IV.2.6 in [10] for the proof) that for fixed x , the function $y \mapsto c_1(x, y)$ is the unique continuous and bounded below viscosity solution of $H_1(y, Du(y)) = 0$ for $y \in \mathbb{R}^3 \setminus \{x\}$ coupled with $u(x) = 0$. Moreover c_1 is a metric on \mathbb{R}^3 and there are constants $k_1, k_2 > 0$ such that $k_1|x-y| \leq c_1(x, y) \leq k_2|x-y|$. Unfortunately H_1 is not strictly

convex, whence the solution of (8) may not satisfy the regularity assumptions in Theorem 2.5. To tackle this difficulty, we add a quadratic perturbation depending on a small parameter $\varepsilon > 0$ in order to obtain smooth approximate solutions. We set:

$$c_1^\varepsilon(x, y) := \inf \left\{ \int_0^T \left(1 + \varepsilon \frac{|\dot{z}(t)|^2}{2} \right) dt : z : [0, T] \rightarrow \mathbb{R}^3, \right. \\ \left. z(0) = x, z(T) = y, \dot{z}(t) \in M(z(t)) \text{ for a.e. } t \in [0, T] \right\}.$$

Notice that for $\varepsilon \rightarrow 0^+$ we have $c_1^\varepsilon(x, y) \rightarrow c_1(x, y)$. For every fixed x , the function $v \mapsto L_1^\varepsilon(x, v) := 1 + \varepsilon|v|^2/2$ is strictly convex, and we denote by $p \mapsto H_1^\varepsilon(x, p)$ its Legendre transform, which turns out to be C^1 . We notice that, according to Proposition 1, Theorem 2.5 applies for L_1^ε , yielding a $C^{1,1}$ solution φ_ε of (8). In particular, the optimal interpolating vector field v_t^ε is given by $v_t^\varepsilon(x) = \nabla_p H_1^\varepsilon(x, D_x \varphi_\varepsilon(t, x))$ for a.e. x . We associate to this cost a *generalized* Wasserstein distance, defined, for two given probability measures μ_0, μ_1 , by

$$W_2^{c_1^\varepsilon}(\mu_0, \mu_1) = \inf_{\pi \in \Pi(\mu_0, \mu_1)} \left\{ \int_{\mathbb{R}^3 \times \mathbb{R}^3} c_1^\varepsilon(x, y) d\pi \right\}^{1/2}.$$

In Section 2.5 we will formulate a quantitative brain connectivity criterion in term of this distance.

2.4. Intrinsic metric model. In this second model, we modify the infinitesimal metric on \mathbb{R}^3 in order to take into account the anisotropic character of the diffusion as expressed by the DDF, by using the gauge function (see 4 for a definition) to penalize $v_t(x)$ according to its distance to $M(x)$. This equips \mathbb{R}^3 with the structure of a geodesic space, where the geodesics correspond to the optimal transport trajectories.

Due to technical reasons (see Subsection 3.2), we relax the problem replacing the set $M(x)$ with its convex hull and set

$$L_2(x, v) = \gamma_{\text{co}M(x)}^2(v).$$

We define the following cost function for $x, y \in \mathbb{R}^3$:

$$c_2(x, y) := \inf \left\{ \int_0^1 \gamma_{\text{co}M(z(t))}^2(\dot{z}(t)) dt : z(0) = x, z(1) = y \right\},$$

and accordingly we define the distance

$$W_2^{c_2}(\mu_0, \mu_1) = \inf_{\pi \in \Pi(\mu_0, \mu_1)} \left\{ \int_{\mathbb{R}^3 \times \mathbb{R}^3} c_2(x, y) d\pi \right\}^{1/2}.$$

In this case there are no constraints on velocities, so that the Hamiltonian H_2 associated to the Lagrangian L_2 is simply given by

$$H_2(x, p) := \sup \{ \langle p, v \rangle - \delta_{Z(x)}^2(v) : v \in \mathbb{R}^3 \}.$$

Also in this case, thanks to Propositions 1 and 2, Theorem 2.5 applies, due to strict convexity of L_2 , so we obtain a solution of the Monge problem and the optimal interpolating vector field v_t is given by $v_t(x) = \nabla_p H_2(x, D_x \varphi(t, x))$ for a.e. x .

2.5. A quantitative brain connectivity criterion. From a medical point of view, it turns out to be important to know not only whether or not two regions are connected by a fiber bundle, but also the quantity of fibers forming that connecting bundle. This *connectivity problem* is relevant in particular to describe situations where, due to medical pathologies, we have two cerebral areas connected by fibers that pass through a damaged region where the connection is interrupted or strongly hampered: if the connection was ensured by several fibers, the damaged area will not jeopardize the connection between the two given areas as much as in the case of a scarce number of connecting fibers.

We suggest the following criterion in order to quantitatively estimate the connectivity of two regions D_0 and D_1 , by means of the transport distance $W_2^{c_i}$: let μ_0 and μ_1 be two absolutely continuous measures with the same mass, constant densities, and support coinciding respectively with D_0 and D_1 . Define the ratio

$$Q(D_0, D_1) = \frac{W_2^{c_i}(\mu_0, \mu_1)}{d_H(D_0, D_1)}, \quad i = 1, 2, \quad (10)$$

of the transport distance $W_2^{c_i}$ (related to the transport cost c_i between μ_0 and μ_1 , as defined in Subsections 2.3 and 2.4) and the Hausdorff distance d_H (related to the Euclidean metric) between D_0 and D_1 . A small quotient in (10) implies a high fiber density connection between D_0 and D_1 . This kind of information is apparently not provided by other tractography models that ignore information on fiber density.

The comparison between the value of $Q(D_0, D_1)$ obtained for the same brain areas in healthy (reference) and pathological subjects could be an indicator of the effective damage occurred.

2.6. Exact solutions. In both the models described in Subsections 2.3 and 2.4, in order to apply Theorem 2.5 to obtain the required smoothness yielding the optimal interpolating vector field v_t , some manipulations to the cost functions were performed. We underline the fact that, indeed, the *transport problem* admits a solution in the Kantorovich's reformulation even when, due to the lack of smoothness, it is not possible to define a classical optimal interpolant vector field and a classical solution to the characteristic system (see Theorem 2.4).

Indeed, considering the original cost function c_1 of Subsection 2.3 (corresponding to $\varepsilon = 0$) and the unmodified Lagrangian $\tilde{L}_2(x, v) = \gamma_{M(x)}^2(v)$ of Subsection 2.4 does not prevent the application of Theorem 2.4 and the definition of corresponding generalized Wasserstein distances even in this case. In particular, the connectivity criterion of Subsection 2.5 can be stated even in these cases.

The quadratic perturbation added to c_1 in Subsection 2.3 is a regularizing term in order to have a representation of quasi-optimal transport problem solutions by means of interpolating vector fields, which is much more intuitive than to make use of so-called *dynamical transport plans*, i.e. transporting plans depending on time, which would characterize solutions in a more general setting.

In [38] and [39] it is shown that, under suitable assumptions, it is possible to associate to an Hamilton-Jacobi equation $H(x, Du) = 0$ a generalized distance (the model case is $H(x, p) = |p|^2 - 1$, which is related to Euclidean distance). More precisely, it is possible to construct a metric $S(y, x)$ such that, for fixed x , the Lipschitz continuous map $u(y) := S(y, x)$ is a viscosity solution of $H(y, Du(y)) = 0$ in $\mathbb{R}^d \setminus \{x\}$, with boundary datum $u(x) = 0$. This result, which is well known in the convex case, still holds in certain nonconvex cases. Indeed, in our framework,

if we set $H(x, p) = \gamma_{M(x)}^2(v) - 1$, by Proposition 1, the results of Section 4 in [38] can be applied, ensuring the existence of such metric. Accordingly, we can set $c(x, y) = S(x, y)$ and solve the Monge-Kantorovich transport problem, stating the connectivity criterion.

However, in the nonconvex cases, it may happen that the metric S is not a *path metric*: the distance between two points is no longer given by the infimum of the (intrinsic) length of the curves joining x, y , i.e. it may happen that

$$S(y, x) < S_\ell(y, x) := \inf_{\xi(\cdot)} \left\{ \sup \sum_i S(\xi(t_{i-1}), \xi(t_i)) \right\},$$

where $\xi : [0, T] \rightarrow \mathbb{R}^d$ is a continuous curve joining x to y , and the supremum is taken with respect all the finite increasing sequence $t_1 < \dots < t_n$ with $t_1 = 0$ and $t_n = T$.

S_ℓ is a metric too, and it is called *the path metric associated to S* . We have $S_\ell = S$ if and only if S is a path metric.

This fact prevents the recovering of an optimal interpolating vector field and the solution of the characteristic system. Theorem 5.1 in [38] states that, however, in our cases the path metric $S_\ell(y, x)$ is the metric associated to the convexification of $p \mapsto H(x, p)$.

3. Final remarks and open problems. In this section we discuss some issues related to our model, and present some directions that need further investigation.

3.1. Advantages of the proposed model. It is an experimental fact that the measurements in each voxel enjoy very poor robustness properties, and sometimes we can have a voxel where no information are available, surrounded by regular voxels. This *missing cell problem* is not adequately faced by classical models, since they assume that no fibers can cross such voxels. In order to avoid this situation, we propose to assign to each empty voxel Q_α a constant profile $M(\alpha)$ given by a ball, whose radius should be determined by experimental data. In this way, the missing information in the voxel will be in some sense *reconstructed* by the information contained in the neighboring ones. Indeed, in our construction, if the infinitesimal metric is given by (the dual of) a ball, it will not change the direction of integral lines hence will not affect the data coming from the neighborhood. So fibers will be propagated in the missing voxel according to the behavior in the nearby ones. The smaller the radius of the assigned ball, the stronger must be the surrounding information to have a fiber passing through the missing cell. The lowest possible radius is given by the sensitivity of instruments.

3.2. Mathematical issues and drawbacks. From a mathematical point of view, the main issue here is to provide from an optimal transport plan, solution of the Monge-Kantorovich problem, a unique optimal interpolating velocity vector field v_t describing the evolutive process. The obstruction is given by the *lack of smoothness* of v_t , due to the presence in a voxel of multiple fibers crossing along different directions. The strategy that can be used to overcome these difficulties are based on explicit representation formulas for the solutions, regularization procedures and, in the case of nonuniqueness, also on selection principles (e.g. selecting the smoothest one among all the solutions). We refer to [6] and [5] for a complete analysis of these topics, just recalling that the use of selection principles with respect to certain

criteria (e.g. curvature of fibers) is extensively used also in other tractography models (e.g. [45]).

In the second model, described in Subsection 2.4, the problem was relaxed by making a *convexification* of the sets $M(x)$, similar to the (implicit) convexification made in [36] by using the Legendre transform. This relaxation has no physical meaning, since the sets $M(x)$, giving the average displacement probabilities, are characterized by strong anisotropy, hence are nonconvex. However, in the nonconvex case, since Lax-Hopf's representation formula no longer holds for the associated Hamilton-Jacobi equation, it is not possible in general to define an optimal interpolating velocity vector field v_t based on characteristic curves. For details and counterexamples on this delicate point, we refer to [38].

3.3. Work in progress. Reconstructing a reliable picture of the whole net of neuronal fibers will actually require a large number of implementations of our model, choosing several pairs of brain areas according to experimental data and/or our quantitative criterion. The algorithmic implementation and the numerical validation of our model is currently still under investigation. Our aim is to provide also an equivalent description of this model in the framework of *mean field games*, and take advantage of the associated numerical methods (see e.g. [1]).

4. Appendix. Our environment will be the Euclidean n -dimensional space \mathbb{R}^n . We will denote by \mathcal{L}^n and \mathcal{H}^n respectively the Lebesgue and the Hausdorff n -dimensional measures.

4.1. General notations and definitions. Our main reference for the properties of convex sets in \mathbb{R}^n is [37]. A detailed introduction to the theory of viscosity solutions and its applications is given in [10].

Definition 4.1. Let K_1, K_2 be compact subsets of \mathbb{R}^n . The *Hausdorff distance* between K_1 and K_2 is given by:

$$d_H(K_1, K_2) := \max \left\{ \sup_{x \in K_1} d_E(x, K_2), \sup_{y \in K_2} d_E(y, K_1) \right\}.$$

This function is a pseudometric on the nonempty elements of the power set $2^{\mathbb{R}^n} := \{A : A \subseteq \mathbb{R}^n\}$ and a metric on the compact subsets of \mathbb{R}^n . Unless explicitly stated, we will always equip the compact subsets of \mathbb{R}^n with this metric.

Definition 4.2. Let $K \subset \mathbb{R}^n$, $a \in K$. We say that K is *convex* iff for every $x, y \in K$, $\lambda \in [0, 1]$ we have $\lambda x + (1 - \lambda)y \in K$. Given $K \subseteq \mathbb{R}^n$, we will denote by $\text{co}(K)$ its convex hull, namely the intersection of all convex subsets of \mathbb{R}^n containing K .

Definition 4.3. Let K be a closed convex subset of \mathbb{R}^n . We define:

1. the *support function* to K is the function $\delta_K : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\delta_K(q) = \sup_{p \in K} \langle q, p \rangle,$$

and we recall that $\delta_K(q) = \langle p_0, q \rangle$ for some $p_0 \in K$ iff $q \in N_K(p_0)$, where

$$N_K(p_0) := \{v \in \mathbb{R}^n : \langle v, p - p_0 \rangle \leq 0 \text{ for all } p \in K\}$$

is the *normal cone in the sense of convex analysis*.

2. if $0 \in \text{int}(K)$, the *gauge (Minkowski) function* of K is defined by:

$$\gamma_K(q) = \inf\{\lambda > 0 : q/\lambda \in K\}.$$

4.2. Measure theoretic solutions of the continuity equation. We collect here the basic facts on the measure solutions of the continuity equation, we refer for all the details to [6].

Definition 4.4. We denote by $\mathcal{M}_{\text{loc}}(\mathbb{R}^d)$ the set of *signed Radon measures* on \mathbb{R}^d , by $\mathcal{M}_{\text{loc}}^+(\mathbb{R}^d)$ the set of *non-negative Radon measures* on \mathbb{R}^d and by $\mathcal{M}_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^h)$ the set of \mathbb{R}^h -valued Radon measures on \mathbb{R}^d . We recall that the *push-forward* $G\#\mu$ of a measure μ on \mathbb{R}^d through a measurable vector field $G : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is defined by setting for all $\varphi \in C_c^0(\mathbb{R}^d)$ by

$$\int_{\mathbb{R}^d} \varphi(x) dG\#\mu(x) = \int_{\mathbb{R}^d} \varphi \circ G(x) d\mu(x),$$

equivalently $G\#\mu(B) = \mu(G^{-1}(B))$ for every $B \subseteq \mathbb{R}^d$.

We will denote by $\mathcal{P}(\mathbb{R}^d)$ the set of all probability measures on \mathbb{R}^d . For $p \in \mathbb{R}$ we define the set

$$\mathcal{P}_p(\mathbb{R}^n) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^n) : \int_{\mathbb{R}^n} |x|^p d\mu < +\infty \right\}.$$

Definition 4.5. Let $\{\mu_n\}_{n \in \mathbb{N}}$, μ be measures on \mathbb{R}^d . We say that μ_n converges *narrowly* (or *weakly**) iff

$$\lim \int_{\mathbb{R}^d} f(x) d\mu_n(x) = \int_{\mathbb{R}^d} f(x) d\mu(x),$$

for every continuous bounded function $f : \mathbb{R}^d \rightarrow \mathbb{R}$.

Definition 4.6. Given $T > 0$, we consider the *continuity equation*:

$$\partial_t \mu_t + \text{div}(\boldsymbol{\nu}_t) = 0, \quad \text{in } \mathbb{R}^d \times (0, T), \quad (11)$$

where $\mu_t, \boldsymbol{\nu}_t$ are Borel families of measures in $\mathcal{M}_{\text{loc}}(\mathbb{R}^d)$ and $\mathcal{M}_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^d)$ respectively, defined for $t \in (0, T)$ satisfying for all $R > 0$

$$\int_0^T |\mu_t|(B(0, R)) dt < +\infty, \quad V_R := \int_0^T |\boldsymbol{\nu}_t|(B(0, R)) dt < +\infty, \quad (12)$$

and the equation (11) holds in the sense of distributions, i.e.

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \zeta(x, t) d\mu_t(x) dt + \int_0^T \int_{\mathbb{R}^d} \nabla_x(\zeta(x, t)) d\boldsymbol{\nu}_t(x) dt = 0, \quad (13)$$

for every $\zeta \in C_c^1(\mathbb{R}^d \times (0, T))$.

We recall that, thanks to the disintegration theorem, we can identify $(\boldsymbol{\nu}_t)_{t \in [0, T]}$ with the measure $\boldsymbol{\nu} = \int_0^T \boldsymbol{\nu}_t dt \in \mathcal{M}_{\text{loc}}(\mathbb{R}^d \times (0, T); \mathbb{R}^d)$ defined by:

$$\langle \boldsymbol{\nu}, \zeta \rangle = \int_0^T \left(\int_{\mathbb{R}^d} \zeta(x, t) d\boldsymbol{\nu}_t \right) dt, \quad \forall \zeta \in C_c^0(\mathbb{R}^d \times (0, T); \mathbb{R}^d).$$

Similarly, we can identify $(\mu_t)_{t \in [0, T]}$ with a measure

$$\mu = \int_0^T \mu_t dt \in \mathcal{M}_{\text{loc}}(\mathbb{R}^d \times (0, T)).$$

Lemma 4.7. *Let $T > 0$ and $(\mu_t, \nu_t)_{t \in (0, T)}$ be a Borel family of measures satisfying (12) and (13). Then there exists a unique weakly* continuous curve $[0, T] \ni t \mapsto \tilde{\mu}_t \in \mathcal{M}_{\text{loc}}(\mathbb{R}^d)$ such that $\mu_t = \tilde{\mu}_t$ for \mathcal{L}^1 -a.e. $t \in (0, T)$; if $\zeta \in C_c^1(\mathbb{R}^d \times (0, T))$ and $t_1, t_2 \in [0, T]$ with $t_1 \leq t_2$ we have:*

$$\begin{aligned} \int_{\mathbb{R}^d} \zeta(t_2, x) d\mu_{t_2} - \int_{\mathbb{R}^d} \zeta(t_1, x) d\mu_{t_1} &= \\ &= \int_{t_1}^{t_2} \left(\int_{\mathbb{R}^d} \partial_t \zeta(t, x) d\mu_t(x) + \int_{\mathbb{R}^d} \nabla_x(\zeta(t, x)) d\nu_t(x) \right) dt. \end{aligned}$$

Moreover if $\tilde{\mu}_s(\mathbb{R}^d) \in \mathbb{R}$ for some $s \in [0, T]$ and $\lim_{R \rightarrow +\infty} R^{-1}V_R = 0$, then the total mass $\tilde{\mu}_t(\mathbb{R}^d) \in \mathbb{R}$ and is constant.

Definition 4.8 (Solution of continuity eq.). Let $T > 0$, we denote by $\mathcal{CE}([0, T])$ the set of time-dependent measures $(\mu_t, \nu_t)_{t \in [0, T]}$ such that

1. $t \mapsto \mu_t$ is weakly* continuous in $\mathcal{M}_{\text{loc}}(\mathbb{R}^d)$ satisfying (12);
2. $(\nu_t)_{t \in [0, T]}$ is a Borel family satisfying (12);
3. (μ, ν) satisfies (13).

Given $\mu^1, \mu^2 \in \mathcal{M}_{\text{loc}}(\mathbb{R}^d)$, we denote the (possibly empty) set of solutions connecting μ^1 to μ^2 by

$$\mathcal{CE}([0, T], \mu^1 \rightarrow \mu^2) = \{(\mu, \nu) \in \mathcal{CE}([0, T]) : \mu_0 = \mu^1, \mu_T = \mu^2\}.$$

REFERENCES

- [1] Yves Achdou and Italo Capuzzo-Dolcetta, *Mean field games: numerical methods*, SIAM Journal on Numerical Analysis, **48** (2010) n. 3, 1136–1162.
- [2] Andrei Agrachev and Paul Lee, *Optimal transportation under nonholonomic constraints*, Trans. Amer. Math. Soc., **361** (2009) n. 11, 6019–6047.
- [3] D. C. Alexander, P. L. Hubbard, M. G. Hall, E. A. Moore, M. Ptito, G. J. Parker and T. B. Dyrby, *Orientationally invariant indices of axon diameter and density from diffusion MRI*, Neuroimage (2010).
- [4] Luigi Ambrosio, *Transport equation and Cauchy problem for non-smooth vector fields*, in “Calculus of variations and nonlinear partial differential equations”, Lecture Notes in Math., **1927** (2008), Springer, Berlin, 1–41.
- [5] Luigi Ambrosio and Gianluca Crippa, *Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields*, in “Transport Equations and Multi-D Hyperbolic Conservation Laws”, Lecture Notes of the Unione Matematica Italiana, **5** (2008).
- [6] Luigi Ambrosio, Nicola Gigli and Giuseppe Savaré, “Gradient flows in metric spaces and in the space of probability measures”, Second Ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.
- [7] Sigurd Angenent, Eric Pichon and Allen Tannenbaum, *Mathematical methods in medical image processing*, American Mathematical Society. Bulletin. New Series, **43** (2006) n.3, 365–396.
- [8] Y. Assaf, T. Blumenfeld-Katzir, Y. Yovel, P. J. Basser, *AxCaliber: a method for measuring axon diameter distribution from diffusion MRI*, Magn. Reson. Med., **59** (2008) n.6, 1347–1354.
- [9] Jean-Pierre Aubin and Hélène Frankowska, “Set-valued analysis”, Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, 2009
- [10] Martino Bardi and Italo Capuzzo-Dolcetta, “Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations”, Birkhäuser Boston Inc., Boston, MA, 1997.
- [11] P. J. Basser, S. Pajevic, C. Pierpaoli, J. Duda and A Aldroubi, *In vivo fiber tractography using DT-MRI data*, Magn. Reson. Med., **44** (2000) n.4, 625–632.

- [12] Jean-David Benamou and Yann Brenier, *A numerical method for the optimal time-continuous mass transport problem and related problems*, in “Monge Ampère equation: applications to geometry and optimization” (Deerfield Beach, FL, 1997), Contemp. Math., **226** (1999), 1–11.
- [13] J.-D. Benamou, Y. Brenier and K. Guittet, *The Monge-Kantorovitch mass transfer and its computational fluid mechanics formulation*, International Journal for Numerical Methods in Fluids, **40** (2002) n.1-2, 21–30.
- [14] Stefano Bianchini and Fabio Cavalletti, *The Monge problem for distance cost in geodesic spaces*, to appear in “Nonlinear Conservation Laws and Applications”, IMA Vol. Math. Appl., Springer, New York, 2010.
- [15] Piermarco Cannarsa and Carlo Sinestrari, “Semiconcave functions, Hamilton-Jacobi equations, and optimal control”, Progress in Nonlinear Differential Equations and their Applications, **58**, Birkhäuser Boston Inc., Boston, MA, 2004.
- [16] Piermarco Cannarsa and Carlo Sinestrari, *Convexity properties of the minimum time function*, Calculus of Variations and Partial Differential Equations, **3** (1995) n. 3, 273–298.
- [17] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern and P. R. Wolenski, “Nonsmooth analysis and control theory”, Graduate Texts in Mathematics, **178**, Springer-Verlag, New York, 1998.
- [18] J. Coremans, R. Luybaert, F. Verhelle, T. Stadnik and M Osteaux, *A method for myelin fiber orientation mapping using diffusion-weighted MR images*, Magn Reson Imaging, **12** (1994) n.3, 443–454,
- [19] J. Dauguet, S. Peled, V. Berezovskii, T. Delzescaux, S. K. Warfield, R. Born and C. F. Westin, *Comparison of fiber tracts derived from in-vivo DTI tractography with 3D histological neural tract tracer reconstruction on a macaque brain*, Neuroimage, **37** (2007) n.2, 530–538.
- [20] T. B. Dyrby, L. V. Sogaard, G. J. Parker, D. C. Alexander, N. M. Lind, W. F. Baaré, A. Hay-Schmidt, N. Eriksen, B. Pakkenberg, O. B. Paulson and J. Jelsing, *Validation of in vitro probabilistic tractography*, Neuroimage, **37** (2007) n.4, 1267–1277.
- [21] Mikhail Feldman and Robert J. McCann, *Monge’s transport problem on a Riemannian manifold*, Transactions of the American Mathematical Society, **354** (2002), n.4, 1667–1697.
- [22] Albert Fathi and Antonio Siconolfi, *PDE aspects of Aubry-Mather theory for quasiconvex Hamiltonians*, Calculus of Variations and Partial Differential Equations, **22** (2005) n.2, 185–228.
- [23] P. Fillard and C. Poupon and J. F. Mangin, *A novel global tractography algorithm based on an adaptive spin glass model*, Med. Image Comput. Comput. Assist. Interv., **12** (2009) Pt. 1, 927–934.
- [24] Ola Friman and Gunnar Farneback and Carl-Fredrik Westin, *A Bayesian approach for stochastic white matter tractography*, IEEE transactions on medical imaging, **25** (2006) n.8, 965–978.
- [25] Wilfrid Gangbo and Robert J. McCann, *The geometry of optimal transportation*, Acta Mathematica, **177** (1996) n.2, 113–161.
- [26] P. Hagmann, L. Jonasson, P. Maeder, J. Thiran, V. Wedeen and R. Meuli, *Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond*, Radiographics, **26** (2006), S205–S223.
- [27] P. Hagmann, J. P. Thiran, L. Jonasson, P. Vandergheynst, S. Clarke, P. Maeder and R Meuli, *DTI mapping of human brain connectivity: statistical fibre tracking and virtual dissection*, Neuroimage, **19** (2003) n.3, 545–554.
- [28] Y. Iturria-Medina, E. J. Canales-Rodríguez, L. Melie-García, P. A. Valdés-Hernández, E. Martínez-Montes, Y. Alemán-Gómez and J M Sánchez-Bornot, *Characterizing brain anatomical connections using diffusion weighted MRI and graph theory*, Neuroimage, **36** (2007) n.3, 645–660.
- [29] S. Jbabdi, M. W. Woolrich, J. L. Andersson and T. E. Behrens, *A Bayesian framework for global tractography*, Neuroimage, **37** (2007) n.1, 116–129.
- [30] B. W. Kreher, I. Mader and V. G. Kiselev, *Gibbs tracking: a novel approach for the reconstruction of neuronal pathways*, Magn. Reson. Med., **60** (2008) n.4, 953–963.
- [31] M. Lazar and A. L. Alexander, *An error analysis of white matter tractography methods: synthetic diffusion tensor field simulations*, Neuroimage, **20** (2003) n.2, 1140–1153.
- [32] M. Lazar, D. M. Weinstein, J. S. Tsuruda, K. M. Hasan, K. Arfanakis, M. E. Meyerand, B. Badie, H. A. Rowley, V. Haughton, A. Field and A L Alexander, *White matter tractography using diffusion tensor deflection*, Hum. Brain Mapp., **18** (2003) n.4, 306–321.

- [33] C. P. Lin, V. J. Wedeen, J. H. Chen, C. Yao and W Y Tseng, *Validation of diffusion spectrum magnetic resonance imaging with manganese-enhanced rat optic tracts and ex vivo phantoms*, Neuroimage, **19** (2003) n.3, 482–495.
- [34] Y. Lu, A. Aldroubi, J. C. Gore, A. W. Anderson and Z Ding, *Improved fiber tractography with Bayesian tensor regularization*, Neuroimage, **31** (2006) n.3, 1061–1074.
- [35] G. J. Parker, H. A. Haroon and C. A. Wheeler-Kingshott, *A framework for a streamline-based probabilistic index of connectivity (PICO) using a structural interpretation of MRI diffusion measurements*, J. Magn. Reson. Imaging, **18** (2003) n.2, 242–254.
- [36] Eric Pichon, Carl-Fredrik Westin and Allen Tannenbaum, *A Hamilton–Jacobi–Bellman approach to high angular resolution diffusion tractography*, in “Proceedings of MICCAI”, 2005, 180–187.
- [37] R. Tyrrell Rockafellar, “Convex analysis”, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970.
- [38] Antonio Siconolfi, *Metric character of Hamilton–Jacobi equations*, Transactions of the American Mathematical Society, **355** (2003) n.5, 1987–2009.
- [39] Antonio Siconolfi, *Errata to: “Metric character of Hamilton–Jacobi equations” [Trans. Amer. Math. Soc. **355** (2003), no. 5, 1987–2009 (electronic)]*, Transactions of the American Mathematical Society, **355** (2003) n.10, 4265.
- [40] J. D. Tournier, F. Calamante, M. D. King, D. G. Gadian and A. Connelly, *Limitations and requirements of diffusion tensor fiber tracking: an assessment using simulations*, Magn. Reson. Med., **47** (2002) n.4, 701–708.
- [41] J. D. Tournier, C. H. Yeh, F. Calamante, K. H. Cho, A. Connelly and C. P. Lin, *Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data*, Neuroimage, **42** (2008) n.2, 617–625.
- [42] Cédric Villani, “Topics in optimal transportation”, Graduate Studies in Mathematics, **58**, American Mathematical Society, Providence, RI, 2003.
- [43] Cédric Villani, “Optimal transport”, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **338**, Springer-Verlag, Berlin, 2009.
- [44] U. C. Wiesmann, C. A. Clark, M. R. Symms, F. Franconi, G. J. Barker and S. D. Shorvon, *Reduced anisotropy of water diffusion in structural cerebral abnormalities demonstrated with diffusion tensor imaging*, Magn. Reson. Imaging, **17** (1999) n.9, 1269–1274.
- [45] X. Wu, Q. Xu, L. Xu, J. Zhou, A. W. Anderson and Z. Ding, *Genetic white matter fiber tractography with global optimization*, J. Neurosci. Methods, **184** (2009) n.2, 375–379.
- [46] A. Zalesky, *DT-MRI fiber tracking: a shortest paths approach*, IEEE Trans Med Imaging, **27** (2008) n.10, 1458–1471.

Received xxxx 20xx; revised xxxx 20xx.

E-mail address: `alessandro.daducci@epfl.ch`

E-mail address: `antonio.marigonda@univr.it`

E-mail address: `giandomenico.orlandi@univr.it`

E-mail address: `roberto.posenato@univr.it`