Inverse regression in MR Fingerprinting: reducing dictionary size while increasing parameters accuracy

Fabien Boux, Florence Forbes, Julyan Arbel, Emmanuel L. Barbier

SUBMITTED TO MAGNETIC RESONANCE IN MEDICINE

Inverse regression in MR Fingerprinting: reducing dictionary size while increasing parameters accuracy

Fabien Boux[1,2] | **Florence Forbes**[2] | **Julyan Arbel**[2] |

Emmanuel L. Barbier[1]

[1]Université Grenoble Alpes, Inserm, U1216, Grenoble Institut Neurosciences, Grenoble, France

[2]Univ. Grenoble Alpes, Inria, CNRS, G-INP, 38000 Grenoble, France

Correspondence

Emmanuel L. Barbier, Ph.D., Grenoble Institut des Neurosciences, GIN, Chemin Fortune Ferrini, 38700 La Tronche, France.
Phone: +33-4-56-52-05-88
Email: emmanuel.barbier@univ-grenoble-alpes.fr

Funding information

This work was supported by the French National Research Agency - project cerebrovascular dynamics in epilepsy: endothelial-pericyte interface – Epicyte, ANR-16-CE37-0013. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Purpose: To reduce dictionary size and increase parameter estimate accuracy in MR Fingerprinting (MRF).

Methods: A dictionary-based learning (DBL) method is investigated to bypass inherent MRF limitations in high dimension: reconstruction time and memory requirement. The DBL method is a 3-step procedure: (1) a quasi-random sampling strategy to produce the dictionary, (2) a statistical inverse regression model to learn from the dictionary a probabilistic mapping between MR fingerprints and parameters, and (3) this mapping to provide both parameter estimates and their confidence levels.

Results: On synthetic data, experiments show that the quasi-random sampling outperforms the grid when designing the dictionary for inverse regression. Dictionaries up to 100 times smaller than usually employed in MRF yield more accurate parameter estimates with a 500 time gain. Estimates are supplied with a confidence index, well correlated with the estimation bias ($r \geq 0.89$). On microvascular

MRI data, results show that dictionary-based methods (MRF and DBL) yield more accurate estimates than the conventional, closed-form equation, method. On MRI signals from tumor bearing rats, the DBL method shows very little sensitivity to the dictionary size in contrast to the MRF method.

Conclusion: The proposed method efficiently reduces the number of required simulations to produce the dictionary, speeds up parameter estimation, and improve estimates accuracy. The DBL method also introduces a confidence index for each parameter estimate.

KEYWORDS

Quantitative imaging, MR Fingerprinting, multiparametric mapping, dictionary learning, inverse regression, microvascular structure.

1 | INTRODUCTION

In magnetic resonance fingerprinting (MRF) [1, 2], the accuracy of the reconstruction depends on the dictionary size, which increases as the power of the number of parameters. For applications with numerous parameters such as microvascular quantification [3], the required memory size and computational time for dictionary simulation and data reconstruction quickly become the main limitations to the development of MRF.

Different approaches have been proposed to compress the signal part of the dictionary while limiting the information loss. Several authors used singular value decomposition to project the dictionary in a well chosen subspace [4–8]. However, the lack of consideration of the tissue parameters during this compression procedure might yield a mapping of the dictionary onto an intermediate low-dimensional space insufficient to correctly recover the true parameters. It has also been proposed to directly find a mapping between the fingerprints and the parameter space using kernel regression [9] or neural network approaches [10–16]. The resulting compact dictionary offers the advantage, over the discrete MRF grid, of a continuous representation with regard to parameter values. All approaches significantly reduce the reconstruction time.

All previous approaches however rely on the simulation of all fingerprints, a computationally expensive operation, to build the compressed representation of the dictionary. Cohen *et al.* [12] investigated a mapping obtained from a sparse set of dictionary entries to limit simulation time but the study was not conducted in high dimension and with large undersampling factors. To estimate a larger number of parameters with MRF, an approach that reduces the need for simulation, continuously represents the parameters without accuracy loss, relies on an interpretable model and reduces the reconstruction time is highly desirable [17].

In this work, we adopt a mapping approach but we bypass the difficulty of learning a *high-to-low* mapping from a high dimensional fingerprint space to a low dimensional parameter space, by learning instead the much less problematic *low-to-high* reverse mapping from parameters to fingerprints. More specifically, we use for this purpose the so-called Gaussian locally linear mapping (GLLiM) model [18] that provides both a tractable learning of the low-to-high mapping and a subsequent analytical expression of the target high-to-low or signal-to-parameter mapping. Furthermore, in contrast to most other regression methods that focus on providing pointwise predictions, GLLiM provides a full posterior distribution of parameters per fingerprint. This distribution may then be used to estimate both the value and a confidence index for each parameter, using the expectation and the standard deviation of the distribution, respectively.

In this MRF study, the proposed dictionary-based learning (DBL) method and the standard dictionary-based matching (DBM) method are compared. We first use synthetic toy signals to assess quantitatively the method's performance. Then, we consider the characterization of brain microvascularization using first simulated signals and second experimental data acquired in tumor bearing rats.

2 | THEORY

2.1 | MRF as an inverse problem

In inverse problems, the overall issue is to provide information on some parameters of interest $\mathbf{x}$ given an observed signal $\mathbf{y}$, using a known *direct* or *forward model* which describes how the parameters $\mathbf{x}$ translate into a signal $\mathbf{y}$. Among inverse problems, MRF exhibits the following difficulties: 1) the direct model is (highly) non-linear, as a (complex) series of equations or simulation tools; 2) the $\mathbf{y}$'s are high-dimensional signals and 3) many such $\mathbf{y}$'s have to be inverted (one for each voxel of an image); 4) the vector of parameters $\mathbf{x}$ is multidimensional and predicting each component of $\mathbf{x}$ independently is likely to be sub-optimal.

To account for possible sources of uncertainty, we focus on a statistical modeling assuming that the forward model is described by a likelihood and a prior distribution. The likelihood function is linking parameter values $\mathbf{x}$ to a probability of observing signal $\mathbf{y}$, $\mathcal{L}_{\mathbf{x}}(\mathbf{y}) = p(\mathbf{y}|\mathbf{x})$. A natural assumption is that $\mathcal{L}_{\mathbf{x}}(\mathbf{y})$ is a Gaussian distribution $\mathcal{N}(\mathbf{y}; f(\mathbf{x}), \mathbf{\Sigma})$ centered at $f(\mathbf{x})$ where f is the known simulation function that links the magnetic and physiological parameters to the fingerprint and $\mathbf{\Sigma}$ is a covariance matrix accounting for measuring or modelling imperfections. The parameter prior distribution, denoted by $p(\mathbf{x})$, encodes in turns information on the possible parameter values. Standard MRF uses a finite grid of values, which corresponds to a very particular discrete prior. This probabilistic point of view allows, with the Bayesian framework, to derive a posterior distribution $p(\mathbf{x}|\mathbf{y}) = p(\mathbf{y}|\mathbf{x})\, p(\mathbf{x}) \,/\, p(\mathbf{y})$, which provides for any given $\mathbf{y}$, a characterization of $\mathbf{x}$ by a probability density function more informative than a single point prediction of $\mathbf{x}$. It corresponds to a richer inverse model but is not usually available in closed-form and requires approximations to be usable in practice.

More generally, most methods to solve inverse problems can be classified into two main categories, optimization-based or learning-based methods. In the next section, we refer to standard MRF as a matching method. We show that it can be seen as a penalized optimization, which does not require statistical modeling, while the method we propose next belongs to statistical learning approaches.

2.2 | Dictionary-based matching (DBM) method

In MRF [1], a large database $\mathcal{D}_f$, referred to as a dictionary, of N S-dimensional fingerprints $\{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$, is generated by running the simulation model f for N different values of the P-dimensional magnetic and physiological parameters $\{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$. In the standard MRF approach, a P-dimensional grid is generated with sampled values in a pre-set interval for each parameter. Then, to invert an observed $\mathbf{y}_{\text{obs}}$, it is compared with the signals in $\mathcal{D}_f$ to find the best match according to an objective function d, usually a standard distance or dissimilarity measure (*e.g.* in MRF, the dot product). With $\mathcal{D}_f = \{(\mathbf{x}_n, \mathbf{y}_n = f(\mathbf{x}_n)), n = 1{:}N\}$, $\mathbf{x}$ is thus estimated as the argument of the following minimization:

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathcal{D}_f}{\arg\min}\ d(\mathbf{y}_{\text{obs}}, f(\mathbf{x})). \tag{1}$$

Solutions are sought in $\mathcal{D}_f$ only, while in a non-penalized optimization the minimization is over the whole continuous space of parameter values. The speed gain is significant in comparison to traditional (without the penalty term)

optimization methods, as retrieving a value from memory is often faster than undergoing an expensive computation. However, the performance of the method depends directly on the choice of the sampling density and the number of parameters. The larger the number N of couples $(\mathbf{x}_n, \mathbf{y}_n)$, the more accurate the estimates but the larger the simulation time and memory requirement. Even for moderate number of parameters, the required number of elements in the dictionary renders grid search intractable on a desktop computer. In addition, each new $\mathbf{y}_{\text{obs}}$, requires the computation and comparison of N matching scores $d(\mathbf{y}_{\text{obs}}, \mathbf{y}_n)$, which can be costly if N is very large and if many inversions are desired. The regression or learning method that we propose in the next section is more efficient in that sense.

2.3 | Proposed dictionary-based learning (DBL) method

In contrast to the DBM method, regression and learning methods can adapt to handle massive inversions of high dimensional data. The main principle is to transfer the computational cost, from 2-signal matchings to the learning of an inverse operator $\mathcal{F}^{-1}$. Equivalently, the goal is to learn a mapping from the fingerprint space to the parameter space, for any $\mathbf{y}$, with cost-less evaluation of $\mathcal{F}^{-1}(\mathbf{y})$. The dictionary $\mathcal{D}_f$ can be used to estimate $\mathcal{F}^{-1}$. Learning or regression methods adapted to high dimensions include inverse regression methods, *i.e.* sliced inverse regression [19], partial least square [20], approaches based on mixtures of regressions with different variants, *e.g.* Gaussian locally linear mapping (GLLiM) [18], mixtures of experts [21], cluster weighted models [22], and kernel methods [9]. Inverse regression methods are flexible in that they reduce the dimension in a way optimal to the subsequent mapping estimation task that can itself be carried out by any kind of standard regression tool. In that sense, the inverse regression methods are said to be non-parametric or semi-parametric. Similarly, in [9], the authors propose a regression with an appropriate kernel function to learn the non-linear mapping. The procedure has the advantage to be semi-parametric but a serious limit is that the components of f are optimized in each dimension separately. As regards application to MRF, the learning strategy has also been proposed by several groups using deep learning tools [10–16]. A major limitation of these methods is that they require a large number of training points to learn many model parameters without overfitting.

In the same vein as [9], and in contrast to deep learning approaches, we propose to use the GLLiM method that exploits Gaussian mixture models [18]. Compared to other regression methods that focus on providing point wise estimates, GLLiM provides a full probability distribution selected in a family of parametric models $\{p(\mathbf{x} \mid \mathbf{y};\ \boldsymbol{\theta}), \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$, *e.g.* mixture of Gaussians, where the parameters are denoted by $\boldsymbol{\theta}$. The dictionary can be used to estimate $\boldsymbol{\theta}$ and defines the inversion operator as $\mathcal{F}^{-1}(\mathbf{y}) = p(\mathbf{x} \mid \mathbf{y};\ \boldsymbol{\theta})$. More specifically, GLLiM handles the modeling of non-linear relationships

with a piecewise linear model. Each $\mathbf{y}$ is seen as the noisy image of $\mathbf{x}$ obtained from a K-component mixture of affine transformations. This is modeled by introducing a latent variable $z \in \{1, \ldots, K\}$ such that

$$\mathbf{y} = \sum_{k=1}^{K} \delta_k(z) \left(\boldsymbol{A}_k \mathbf{x} + \boldsymbol{b}_k + \boldsymbol{\epsilon}_k\right), \tag{2}$$

where $\delta_k(z)$ indicates membership in the region k of $\mathbf{x}$, having the value 1 if it belongs to the region and the value 0 otherwise. $\boldsymbol{A}_k$ is a $P \times S$ matrix and $\boldsymbol{b}_k$ a vector in $\mathbb{R}^P$ that characterize an affine transformation. Variable $\boldsymbol{\epsilon}_k$ corresponds to an error term in $\mathbb{R}^P$ which is assumed to be zero-mean and not correlated with $\mathbf{x}$, capturing both the modelling noise and the reconstruction error due to the affine approximations. In GLLiM, $\boldsymbol{\epsilon}_k$ follows a Gaussian distribution $\mathcal{N}(0, \boldsymbol{\Sigma}_k)$ and $\mathbf{x}$ follows a mixture of K Gaussians defined by $p(\mathbf{x} \mid z = k) = \mathcal{N}(\mathbf{x}; \boldsymbol{c}_k, \boldsymbol{\Gamma}_k)$, and $p(z = k) = \pi_k$. It follows that

$$\begin{aligned} p(\mathbf{y} \mid \mathbf{x};\, \boldsymbol{\theta}) &= \sum_{k=1}^{K} w_k(\mathbf{x})\, \mathcal{N}(\mathbf{y}; \boldsymbol{A}_k \mathbf{x} + \boldsymbol{b}_k, \boldsymbol{\Sigma}_k) \\ \text{with } w_k(\mathbf{x}) &= \frac{\pi_k \mathcal{N}(\mathbf{x}; \boldsymbol{c}_k, \boldsymbol{\Gamma}_k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(\mathbf{x}; \boldsymbol{c}_j, \boldsymbol{\Gamma}_j)} \end{aligned} \tag{3}$$

and $\boldsymbol{\theta} = \{\pi_k, \boldsymbol{c}_k, \boldsymbol{\Gamma}_k, \boldsymbol{A}_k, \boldsymbol{b}_k, \boldsymbol{\Sigma}_k\}_{k=1:K}$ is the set of parameters defining the model. The conditional probability distribution of interest can be derived as

$$\begin{aligned} p(\mathbf{x} \mid \mathbf{y};\, \boldsymbol{\theta}) &= \sum_{k=1}^{K} w_k^*(\mathbf{y})\, \mathcal{N}(\mathbf{x}; \boldsymbol{A}_k^* \mathbf{y} + \boldsymbol{b}_k^*, \boldsymbol{\Sigma}_k^*) \\ \text{with } w_k^*(\mathbf{y}) &= \frac{\pi_k \mathcal{N}(\mathbf{y}; \boldsymbol{c}_k^*, \boldsymbol{\Gamma}_k^*)}{\sum_{j=1}^{K} \pi_j^* \mathcal{N}(\mathbf{y}; \boldsymbol{c}_j^*, \boldsymbol{\Gamma}_j^*)}, \end{aligned} \tag{4}$$

with a new parameterization $\boldsymbol{\theta}^* = \{\boldsymbol{c}_k^*, \boldsymbol{\Gamma}_k^*, \boldsymbol{A}_k^*, \boldsymbol{b}_k^*, \boldsymbol{\Sigma}_k^*\}_{k=1:K}$ easily expressed as an analytical function of $\boldsymbol{\theta}$. The mixture setting provides some guaranties that when choosing K large enough it is possible to approximate any reasonable relationship [21]. Automatic model selection criteria can also be used to select K (see [18]) but in practice we observed that the precise value of K is not critical and different values provide similar results so that for the rest of the paper K is fixed to 50 (unless otherwise specified).

The $p(\mathbf{x} \mid \mathbf{y};\, \boldsymbol{\theta})$ distribution allows us to evaluate both estimates of the parameters $\mathbf{x}$ and information about the confidence to be placed in these estimates. In this work, the estimates are defined as the expectation and the confidence indices from the covariance matrix of $p(\mathbf{x} \mid \mathbf{y};\, \boldsymbol{\theta})$. For each parameter, the standard deviation is used as confidence index and is computed as the square root of the corresponding diagonal element in the covariance matrix:

$$\hat{\mathbf{x}} = \mathbb{E}\left[\mathbf{x} \,|\, \mathbf{y};\, \boldsymbol{\theta}\right] \tag{5}$$

$$\text{with } \mathbb{E}\left[\mathbf{x} \,|\, \mathbf{y};\, \boldsymbol{\theta}\right] = \sum_{k=1}^{K} w_k^*(\mathbf{y})(\boldsymbol{A}_k^*\mathbf{y} + \boldsymbol{b}_k^*)\,,$$

$$\begin{aligned}\text{Var}\left[\mathbf{x} \,|\, \mathbf{y};\, \boldsymbol{\theta}\right] = &\sum_{k=1}^{K} w_k^*(\mathbf{y})\left[\boldsymbol{\Sigma}_k^* + (\boldsymbol{A}_k^*\mathbf{y} + \boldsymbol{b}_k^*)(\boldsymbol{A}_k^*\mathbf{y} + \boldsymbol{b}_k^*)^T\right] \\ &- \left(\sum_{k=1}^{K} w_k^*(\mathbf{y})(\boldsymbol{A}_k^*\mathbf{y} + \boldsymbol{b}_k^*)\right)\left(\sum_{k=1}^{K} w_k^*(\mathbf{y})(\boldsymbol{A}_k^*\mathbf{y} + \boldsymbol{b}_k^*)\right)^T .\end{aligned}$$

Because S is much larger than P, it is important that the model (3) involving $\boldsymbol{\theta}$ is estimated first and then used to derive model (4) that has a similar structure. The number of model parameters $\boldsymbol{\theta}$ can be drastically reduced by choosing constraints on covariance matrices $\boldsymbol{\Sigma}_k$ without inducing oversimplifications on the target model (4). In this work, equal diagonal covariance matrices are used as they yield the best results: for $1 \leq k \leq K$, $\boldsymbol{\Sigma}_k = \mathbf{D}_S$, where $\mathbf{D}_S \in \mathbb{R}^{S\times S}$ is a diagonal matrix. For example, with $S = 100$, $P = 3$ and $K = 50$, the number of parameters $\boldsymbol{\theta}$ is equal to 20 600 while a direct estimation of $\boldsymbol{\theta}^*$ would involve 272 703 parameters (see [18] for more details).

2.4 | Dictionary sampling

The dictionary design depends on the sampling strategy of the parameter space. Regular grids of P-dimensional values of parameters are usually considered as varying only one parameter at a time would not capture important interactions between the parameters and important aspects of the mapping. For each parameter, a set of different values is chosen and P-dimensional combinations are deduced from the Cartesian product of the P sets. This approach works well for small numbers of parameters, typically two or three, but for larger numbers, it becomes impractical because the number of required runs grows exponentially. In [14], authors show that in a regression context, the random sampling strategy provides better estimation. However, this strategy involves a risk of a lack in parameter space coverage, especially when seeking sparsity.

Figure 1(a) shows a projection in two dimensions of 1 000 points from a uniform grid in the 3D-hypercube. Notice that with 1 000 points in 3D, only 100 distinct combinations appear in the 2D-projection plane, each representing 10 different values of the other third variable. Each parameter is actually described by only 10 distinct values. In addition, in situations where fingerprints depend only on one or two of the parameters, such a grid design may contain a lot of useless duplications. A significant improvement on grids may be obtained by scrambled nets [23, 24]. In this paper, the

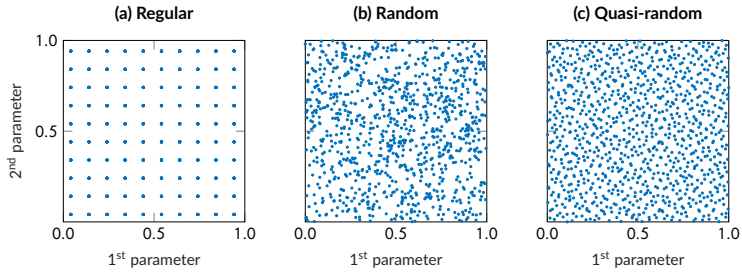

FIGURE 1 2-dimensional projections of 3-dimensional samples obtained from 3 different sampling strategies. Subfigures represent the 2D-projection of 1 000 points of the 3D-hypercube obtained from **(a)** regular sampling, **(b)** random uniform sampling and **(c)** quasi-random (scrambled Sobol sequence), respectively.

Sobol sequence is generated [25] and scrambled [26]. We show the projection of 1 000 points from the scrambled Sobol sequence (figure 1(c)) referred to as quasi-random in the next sections.

3 | METHODS

3.1 | Synthetic toy signals

To obtain a large set of signals with no simulation limitations, toy signals that mimic MRI signals are introduced:

$$\mathbf{y}_{\text{toy}}\left(\mathbf{x}_{\text{toy}}, t\right) = \left| \sum_{i=1}^{N} \sin\left(\phi_i \cdot f_0 \cdot t\right) \cdot \exp\left(-\frac{t}{x_i}\right) \right|, \tag{6}$$

where $\mathbf{x}_{\text{toy}} = (x_1, x_2, \ldots, x_P) \in [0,1]^P$ s^P is the parameter vector, $f_0 = 50$ Hz is the sinus function frequency, $\phi \in [0.1, 1]^P$ is a scaling factor vector on frequency and $|\cdot|$ is the absolute value function.

To challenge regression, the dependency between $\mathbf{x}_{\text{toy}}$ and $\mathbf{y}_{\text{toy}}$ is non-linear. Signals are sampled between 5 ms and 1 s every 10 ms yielding a 100-dimensional signal. The frequency factor vector ϕ makes the parameters non-exchangeable. It means that permutations of the $\mathbf{y}_{\text{toy}}$ elements cannot lead to the same signal $\mathbf{x}_{\text{toy}}$.

According to [27], the noise distribution of MRI signals can be assumed Rician. In this work, the noise is defined by its signal-to-noise ratio: SNR $= S_{\max}/\sigma_{\text{noise}}$, where $S_{\max}$ is the maximum signal intensity and σ_{noise} is the standard deviation of the noise. A vector $\mathbf{n} \in \mathbb{R}^S$ is generated from a zero-mean Gaussian distribution with a standard deviation σ_{noise} and added to the original signal $\mathbf{y}_{\text{toy}}$ to create the noisy signal $\mathbf{y}_{\text{noisy}}$ following:

$$\mathbf{y}_{\text{noisy}} = \left| \mathbf{y}_{\text{toy}} + \mathbf{n} \right|. \quad (7)$$

Examples of synthetic toy signals are presented in Supporting Information figure S1.

3.2 | MRI experiments

3.2.1 | Synthetic MRI signals

Synthetic MRI signals are designed as the ratio between the gradient echo sampling of the free induction decay and spin echo (GESFIDSE) signals obtained post- and pre-injection of the ultrasmall superparamagnetic iron oxide particles (USPIO) [3]. To agree with acquisitions, the simulation tool employed [28] allows to take into account the intrinsic relaxations, the magnetic field perturbations induced by susceptibility interfaces (vessels), the diffusion of the water protons and the constrained diffusion of the contrast agent within the voxel. Because of the complexity of the tool, simulations are extremely time-consuming. A single signal simulation took about 10 seconds and a 100 000-signal dictionary was generated on 32-node high-performance computer (*Intel Xeon Gold 6130*, 2.1 GHz) in about 67 hours.

3.2.2 | Acquired MRI signals

The data used to validate the method were acquired with a 4.7 T bore and have been introduced in [29]. The field of view is $30 \times 30\,\text{mm}^2$ and the voxel size is $234 \times 234 \times 800\,\mu\text{m}^3$. A turbo spin-echo was acquired to identify the anatomical structures and tumor tissues. Then, the changes in relaxation rates induced by the injection of USPIO in the network of blood vessels were obtained using two GESFIDSE sequences (32 sampled time points). See Supplementary Information 1 for more details about animal preparation and acquisition conditions.

3.3 | Framework analysis

3.3.1 | Dictionary design

The dictionary is generated in two steps. First, we select parameter value combinations according to a sampling strategy (section 2.4). Second, for each parameter value combination, the associated fingerprint is simulated using

either equation (6), for toy signals or the simulation tool described (section 3.2.1), for MRI synthetic signals.

When generating a dictionary of synthetic toy signals following a regular sampling strategy, the grid step is maintained constant along all parameters. In this way, it makes sense to compare the average errors on parameters. This procedure results in dictionary sizes that are the number of values of each parameter to the power the number of parameters.

3.3.2 | Data processing

Data were processed on a voxel-by-voxel analysis using custom code developed in the Matlab environment (The MathWorks Inc., Natick, Ma, USA). The code is included in the *Medical software for Processing multi-Parametric images Pipelines*[1].

Dictionary-based methods.

Both dictionary-based methods, need a precomputed dictionary. Exploited in a different way, the dictionary is fully stored for DBM or resumed to a model for DBL. The computation of the model is carried out in an offline procedure using the GLLiM model. Then, estimates are computed using equations (1) or (5) applying the DBM or the DBL method, respectively.

For the DBM method, signals of the dictionary are normalised, each having the same sum squared magnitude. For the DBL method, the parameters are normalized to have zero mean and unit variance using scaling and translating factors that are then used to rescale on estimates.

Closed-form equation method.

The closed-form equation (CFE) method refers to a multiple operations procedure based on equations. First, the relaxation rates are extracted by fitting the MRI signal intensities (synthetic and acquired). Then, these rates are used to compute the parameters of interest by relying on equations. In this work, we focus on three parameters: the blood volume fraction (BVf), the vessel size index (VSI) and the tissue oxygen saturation (StO_2). The whole procedure is described in Supplementary Information 2.

[1] https://github.com/nifm-gin/MP3

3.3.3 | Performance evaluation: error metrics

To compare the methods performance in estimating parameters, metrics are used to measure the errors between the estimates and the true values. For test signals, we compute the root mean squared error (RMSE) and mean absolute error (MAE) of a parameter estimates as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{L}\sum_{t=1}^{L}(\hat{x}_t - x_t)^2}\,, \tag{8}$$

$$\mathrm{MAE} = \frac{1}{L}\sum_{t=1}^{L}|\hat{x}_t - x_t|\,, \tag{9}$$

where L is the number of estimates $\hat{x}_t$ of the known parameter x_t.

4 | RESULTS

4.1 | Numerical experiments

4.1.1 | Effect of parameter space sampling

We investigated the impact of three parameter sampling strategies (regular, random and quasi-random) on the accuracy of parameter estimation, using toys signals. We considered three dictionary dimensions for each sampling strategy (a total of 9 conditions). The numbers of signals per dictionary are 216, 1 024, and 2 187 toy signals in 3, 5, and 7 dimensions, respectively. For each dictionary dimension, 1 000 additional test toy signals were produced based on a random sampling of the parameter space. One then estimate the parameter values of each test toy signal using each dictionary (DBL method). The errors between the estimated parameters and the parameters used to generate that toy signal were then computed. The RMSE on parameters was computed using equation (8) and averaged across the parameter dimension to obtain the average RMSE. To characterize the distribution of the average RMSE, the entire procedure (from dictionary generation to average RMSE computation) was repeated 500 times for each of the 9 conditions (figure 2). Additional conditions were also evaluated and are presented in Supporting Information figure S2.

Whatever the sampling strategy, the average RMSE increases with dictionary dimension. As previously reported, a random sampling yields a lower average RMSE than a regular sampling, using the same number of signals to produce the dictionary and whatever the dimension [14]. Note that when using the DBM approach instead of the DBL, the regular

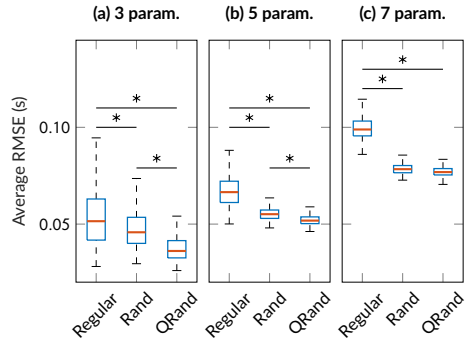

FIGURE 2 Effect of parameter sampling strategies on the dictionary-based learning (DBL) method.

Distribution of average RMSE for regular, random (Rand) and quasi-random (QRand) dictionary samplings. **(a)** Dictionary of dimension 3 using 216 signals, **(b)** dictionary of dimension 5 using 1 024 signals and **(c)** dictionary of dimension 7 using 2 187 signals. For each box, the central red mark indicates the median; the bottom and top edges indicate the 25^{th} and 75^{th} percentiles, respectively. The whiskers extend to the minimum and maximum values (dashed lines). Asterisks refer to the p=0.001 significance level Student's t-test results comparing pairwise sampling strategies average RMSE.

sampling yields lower average RMSE than the random or quasi-random samplings (Supporting Information figure S2). The quasi-random sampling further reduces the average RMSE even if for the dimension 7 it is not significant. Altogether, there is a 30.6 %, 22.6 % and 22.8 % reduction in average RMSE between regular and quasi-random samplings for dictionary dimensions 3, 5, and 7, respectively. Therefore, in the following, a regular sampling is used for the application of the DBM method and the quasi-random sampling for the DBL method.

4.1.2 | Impact of dictionary size

To investigate the impact of dictionary size on the estimation accuracy, we generated four dictionaries with different sizes and considered the case of 5 and 7 parameters (a total of 8 conditions). Dictionary sizes are chosen to keep similar densities, *i.e.* constant number of values per parameter, for 5 and 7 parameters (sizes of 3^5, 4^5, 5^5, and 6^5 for the 5 parameters case and 3^7, 4^7, 5^7, and 6^7 for the 7 parameters case). For each condition, we evaluated the average RMSE as previously described, using 10 000 toy signals. To characterize the impact of SNR, the entire procedure was repeated for noisy toy signals with SNR between 10 and 110 (equation (7)). Both the DBM and the DBL methods were characterized using that approach (figure 3). Supplementary Information figure S3 shows the experiment repeated for additional

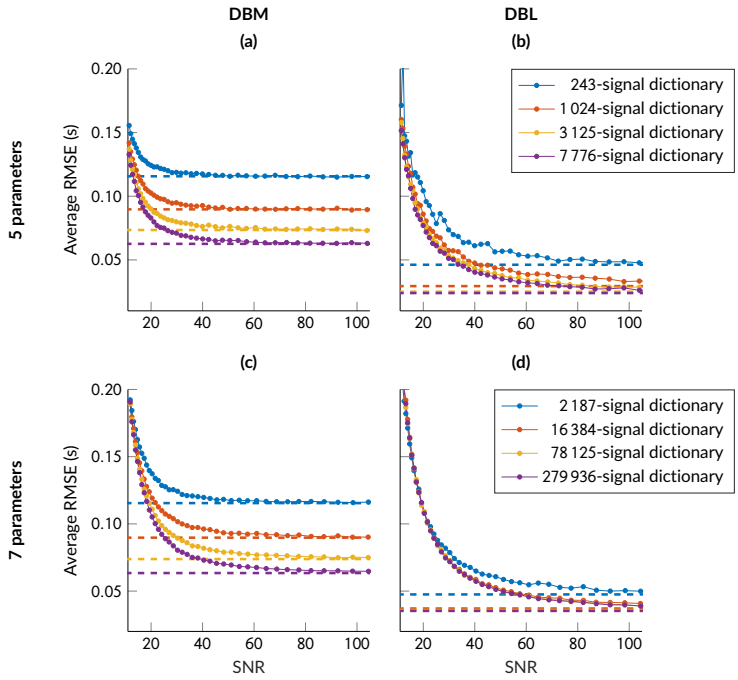

FIGURE 3 Impact of dictionary size and SNR on dictionary-based matching (DBM) and dictionary-based learning (DBL) methods.

Average RMSE are given as a function of the SNR for different dictionary sizes (for the 5 parameters case: 3^5=243, 4^5=1 024, 5^5=3 125, and 6^5=7 776; for the 7 parameters case: 3^7=2 187, 4^7=16 384, 5^7=78 125, and 6^7=279 936). The left column (**(a)** and **(c)**) is the average RMSE of the DBM method, the right column (**(b)** and **(d)**) is the average RMSE of the DBL method, the top row (**(a)** and **(b)**) provides results estimating 5 parameters, and the bottom row (**(c)** and **(d)**) estimating 7 parameters. Dashed lines represent the average RMSE in the absence of noise on test signals. The regression parameter K (DBL method) is fixed to 50 except for the 243-signal dictionary where $K = 20$.

number of parameters and number of signals.

As expected for the DBM method, the average RMSE decreases while the dictionary size increases. It illustrates the fact that the accuracy of the method is closely related to the dictionary size. The average RMSE decreases as SNR increases up to approximately SNR = 50 and then plateaus close to the value obtained in absence of noise. Similar profiles are obtained for the 5 and 7 parameters cases, in agreement with our choice to keep the dictionary density constant. For the DBL method, the average RMSE also decreases as the SNR increases but up to approximately SNR = 90. Again, the highest SNR yields average RMSE close to the ones obtained in absence of noise. For the DBM method,

average RMSE are also comparable between 5 and 7 parameters. For SNR > 20, all average RMSE obtained with the DBL method are 1.98 ± 0.44 for 5 parameters (1.67 ± 0.37 for 7 parameters) lower than the one obtained with the DBM method, whatever the dictionary size. Note that adding noise to the dictionary signals can improve the accuracy on parameter estimates for SNR below 50. Above, the addition of noise deteriorates estimates (Supplementary Information figure S4). When using the DBM method, the dictionary size has a lower impact than for the DBL method: between the smallest and the largest dictionary sizes, the average RMSE decreases by 0.06 for the DBM method while it only decreases by 0.02 for the DBL method. Moreover, the two highest dictionary sizes yield similar average RMSE for the DBL method, suggesting that an increase in dictionary size would not improve the average RMSE further. Altogether, compared to the DBM method, the DBL method improves the accuracy by a factor 1.36 (respectively 1.33) while reducing the dictionary size by a factor of 32 (respectively 128) for 5 parameters (respectively 7 parameters).

By eliminating the time-consuming dictionary matching operation, the DBL method drastically reduces the computation time. For the biggest dictionary size, 10 000 test signals, and 5 parameters, the estimation takes 0.2 ± 0.0 seconds for the DBL method against 66.6 ± 8.3 seconds for the DBM method. For 7 parameters, the estimation takes 6.7 ± 0.2 seconds for the DBL method against 3 734.4 ± 275.8 seconds for the DBM method.

4.1.3 | Boundary behaviour

The DBL method estimates the parameters using a continuous function that is not restricted to the dictionary parameter space. We therefore investigated the behaviors of the DBL and DBM methods outside the boundaries of this space. Two 2-dimensional parameter subspaces was defined: one to sample the parameters of the dictionary, and a larger one for the test. For every position in the test space, we evaluated the average RMSE as previously described, using 100 toy signals (total of 9 000 000 toy signals).

The two methods yield to similar estimate accuracy in the subspace covered by the parameters of the dictionary (figure 4). Outside of that subspace, the average RMSE obtained with the DBM method rises with the distance to the dictionary border. For the DBL method, the average RMSE can remain below 0.1, way beyond the boundaries of the dictionary space. The existence of two dictionary blocks helps reducing the error between these blocks.

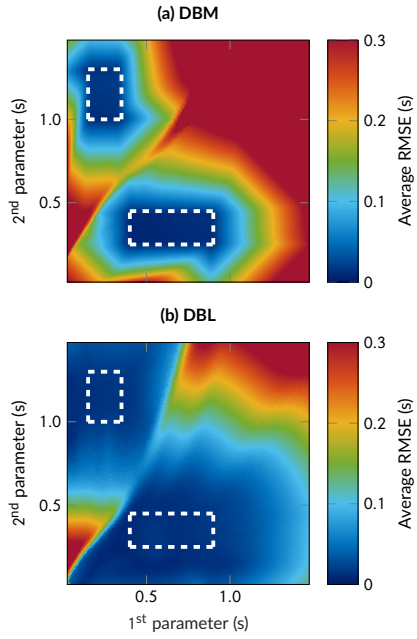

FIGURE 4 Behaviour of estimation accuracy at the boundary of the parameter space covered by the dictionary. The images represent the average RMSE over the space obtained using the DBL method (**(a)**) and using the DBM method (**(b)**). White dashed lines delimit the parameter space of the dictionary. 100 toy signals were simulated from parameters sampled in a sliding window of 50 ms $\times$ 50 ms moving by steps of 5 ms through the parameter space.

4.1.4 | Confidence index

We investigated the relation between the confidence index, available with the DBL method (section 2), and the RMSE on the estimates. To compute the RMSE and the confidence index, we used a 2 500-toy signals dictionary, sampled in the 2-dimensional parameter space and test toy signals with SNR = 100 (200 000 signals). Figure 5 represents the results obtained on two parameters.

We observe good correlation coefficients between RMSE and average CI for both dimensions: 0.90 and 0.89. However, the average CI yields values 20 times smaller than RMSE. Altogether, CI could be used to report variations in RMSE.

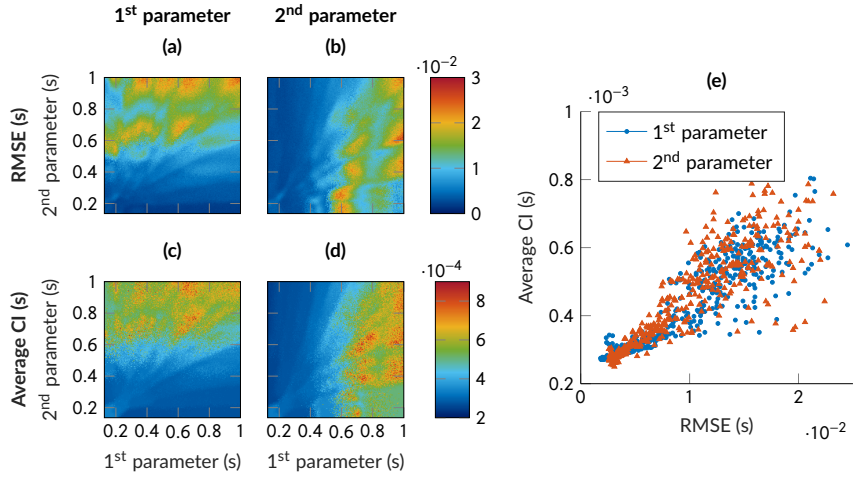

FIGURE 5 RMSE versus average confidence indexes for a 2 dimension toy signal dictionary.

The first row (**(a)** and **(b)**) is the RMSE for the first and second parameter. The second row (**(c)** and **(d)**) is the average confidence index (CI) for the first and second parameter. Data are represented after 2D sliding window filtering (25 ms width in each dimension; about 200 toy signals per window). **(e)** is the correlation between RMSE and average CI. To make the graph clearer only 1 000 window values are plotted (randomly picked).

4.2 | Synthetic and real MRI signals

4.2.1 | Synthetic MRI signals

In the context of microvascular MRI, we compared the two dictionary-based methods and the CFE method. For this, we produced synthetic MRI signals as described section 3.2.1. Dictionaries (grid and quasi-random) were simulated with BVf between 0.25 and 30 % (120 values) and VSI between 0.25 and 50 µm (200 values). The oxygen saturation was set to 70 %. Among the 24 000 combinations, only 22 836 signals could be produced, due to simulation constraints (*e.g.* a very large BVf can not be produce with distant very small vessels or small BVf with large vessels). Note that the simulated field of view was doubled to reduce these simulations constraints. To agree with the grid dictionary, the same number of signals is used for the quasi-random dictionary. To evaluate the MAE for each method, 120 000-test synthetic MRI signals without noise were generated. To analyze the MAE on BVf, test data were divided into 3 parts: small ($1 < VSI \leq 7$ µm), medium ($7 < VSI \leq 15$ µm) and large ($15 < VSI \leq 30$ µm) VSI. To analyse the MAE on VSI, test data were divided into 3 parts: low ($0.5 < BVf \leq 4$ %), medium ($4 < BVf \leq 10$ %) and large ($10 < BVf \leq 20$ %) BVf.

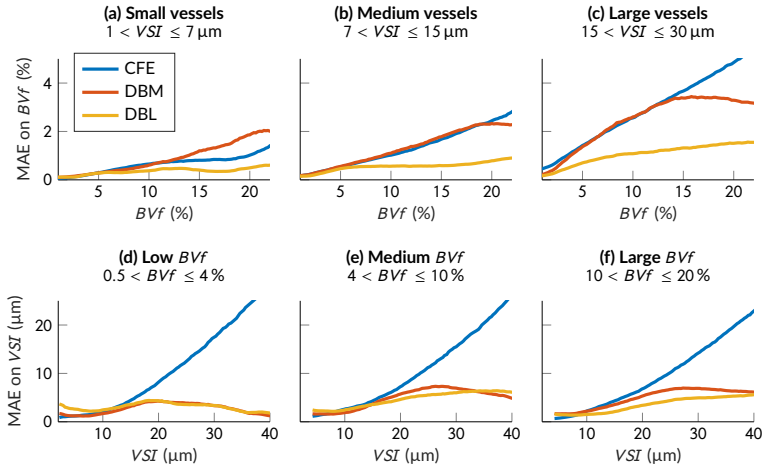

FIGURE 6 MAE on BVf and VSI estimates obtained with the two dictionary-based methods (DBM and DBL) and with the closed-form equation (CFE) method.

Curves represent the MAE on BVf for the first row (**(a)**, **(b)** and **(c)**) and the MAE on VSI for the second row (**(d)**, **(e)** and **(f)**). Data are represented after 1D sliding window filtering (3 % for BVf and 5 µm for VSI)

For all vessel diameters, the MAE on BVf tends to increase with BVf. DBM and CFE methods yield similar MAE for BVf below 12 %. For medium and large vessels and high BVf values, the CFE method yields the highest errors. The DBL method always yields the smallest error (MAE: 1.86 % for CFE, 1.64 % for DBM, and 0.69 % for DBL).

For VSI, the behavior of MAE is similar across the three BVf groups: the CFE method yields larger errors than the two dictionary-based approaches. For VSI above 15 µm, the MAE obtained with CFE is linearly correlated with the VSI value ($R^2 \geq 0.99$). This linear behaviour was previously reported [30]. DBM and DBL methods yields similar MAE, with a slightly smaller MAE for DBL, on average (MAE: 12.65 µm for CFE, 4.10 µm for DBM, and 3.80 µm for DBL).

4.2.2 | Real MRI signals

The proposed method was then applied on real MRI data collected on rats bearing 9L and C6 tumors (section 4.2). Note that because of the low SNR, the last 8 echoes of MRI signals were removed. We quantified BVf, VSI and StO_2 parameters (section 3) with the two dictionary-based methods (DBM and DBL) and using two dictionary sizes:

- **Large dictionaries**: 170 100-signal dictionary (grid sampling for the DBM method, quasi-random sampling for the DBL method) with 70 values for BVf between 0.25 and 30 %, 90 values for VSI between 0.5 and 50 µm and 27

values for StO_2 between 30 and 95 % (real sizes due to simulation constraints: 164 524 for grid and 167 216 for quasi-random).

- **Small dictionaries**: 4 320-signal dictionary with 36 values for BVf between 0.33 and 12 %, 20 values for VSI between 1 and 20 µm and 6 values for StO_2 between 40 and 90 %. Note that the parameter intervals are smaller than that of the large dictionary and the number of signals 40 times smaller (real sizes: 4 218 for grid sampling and 4 119 for quasi-random sampling).

At first sight (figure 7), all methods yield similar maps, in which the tumor and the large vessels may readily be depicted. For large dictionaries, the mean tumorous BVf and VSI obtained with the DBM and DBL methods are similar but values are significantly more heterogeneous for the DBL method (mean values and standard deviations are: 16.60 ± 5.69 %, 14.91 ± 5.21 µm for DBM and 15.68 ± 8.29 %, 15.87 ± 9.15 µm for DBL). For StO_2, as the DBL method could provide values outside of the parameter space, one observes lower and higher StO_2 values than for the DBM method. Moreover, some values seem to fall outside the possible range (> 100 %). It should be noted that these outlier estimates also occur for the other maps, which explains the dark points (especially $VSI < 0$).

For the small dictionaries, the contrasts are similar even though the estimates obtained by the DBM method are restricted to the dictionary space, while the DBL method yields estimates out of the dictionary space and close to the values obtained with the large dictionary. In the tumor, the BVf mean value is 14.23 % while the maximum value of the dictionary is 12 %. In Supporting Information figure S6, an evaluation per regions of interest was produced (8 animals: 4 from each tumor model). The DBL method and the small dictionary produce on average similar estimates (relative differences < 7 %) as with the large dictionary, except for the VSI in the 9L tumor when values differ significantly (relative difference is 11.9 %).

The mean CI in the tumor are: 1.51 % for BVf, 4.01 µm for VSI and 13.13 % for StO_2, while in the cortex the mean CI are: 0.27 % for BVf, 0.77 µm for VSI and 15.48 % for StO_2. These results are in agreement with previous results that pointed better estimates on BVf than on VSI and a low sensitivity of the signals to the StO_2 parameter. The confidence on the estimates is therefore on average 5 to 6 times lower in the tumor than in the cortex on BVf and VSI but similar on StO_2.

A second example of a C6 rat tumor model is given in Supporting Information figure S5.

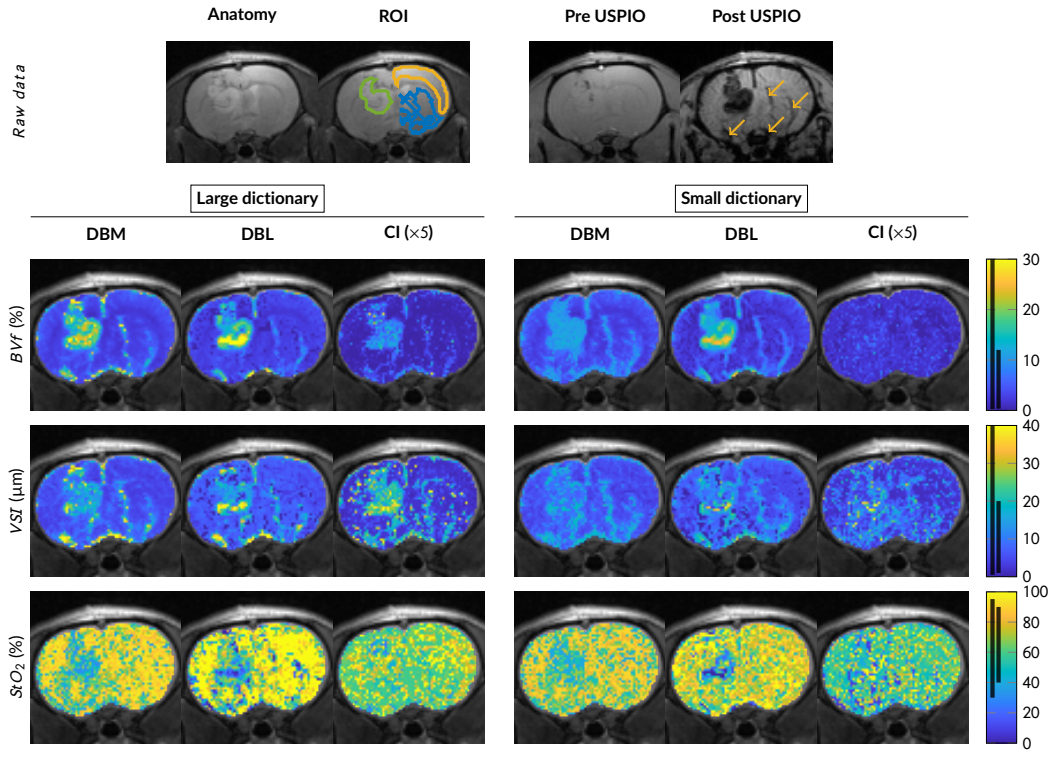

FIGURE 7 Microvascular parameter estimate maps of a 9L rat tumor model.

The first row represents the raw data: the anatomical image and ROI (left) and the MGEFIDSE pre and post USPIO injection (right) for the second echo time (6.3 ms). Then, the estimates maps (the first row is BVf maps, the second row is VSI maps and the last row is StO_2 maps) computed with the DBM method (the first and third columns) or the DBL method (the second and fourth columns). The third and sixth columns are the confidence maps obtained with the DBL method. In the color bar, the black lines represent the parameter space covered by the dictionaries: the short (long) for the small (large) dictionary. On the anatomical image, the green line delimits the tumor, the yellow line the cortex and the blue line the striatum. The arrows on the post USPIO injection image indicate large vessels.

5 | DISCUSSION AND CONCLUSIONS

This work introduces a multiparametric, dictionary-based, estimation method that preserves the main advantages of the MRF method, *i.e.* robustness, speed and flexibility and meets the challenge of maintaining estimates accuracy in high dimension. These properties arise from two main features: an efficient design and use of the dictionary to learn the

relationship between fingerprints and parameters.

Regarding the dictionary design, we observe as [14] that random sampling of the parameter space is a better strategy when learning via a regression model, and we propose a further improvement by using a quasi-random sampling. However, when using the DBM method, the grid remains the most efficient sampling strategy. Altogether, it appears that the appropriate choice of dictionary depends on the chosen inversion approach (grid matching vs statistical learning). When using the DBL method, our results suggest that a new approach to save time in building the dictionaries could be to use patches of quasi-random distributions in the parameter space (section 4.1.3). This patch approach might be implemented by combining a range of parameter values corresponding to healthy tissues and one or more sets corresponding to damaged tissues (*e.g.*, tumorous tissues). When needed, a few additional dictionary entries may be generated to fill the gaps between the patches (Supplementary Information, figure S7). More generally, optimal fingerprint inversion should consider both design and model aspects and target a trade-off between inversion cost and inversion quality.

Regarding inversion quality, the regression approach generally involves hyperparameter tuning. In GLLiM, the number of linear parts K is the only hyperparameter to be adjusted and this can be done automatically using some standard information criterion such as AIC or BIC [31, 32] as illustrated in [18]. This results in additional learning time corresponding to the number of values considered for K. In contrast, neural network frameworks [10–16] that propose to solve the inverse problem are very sensitive to the setting of their numerous hyperparameters: network sizes, batch sizes, learning rates, among others. The tuning is generally achieved by learning the model for a large number of hyperparameter combinations. The search to determine the best combination represent a higher computation cost than the determination of a single hyperparameter. This difference in tuning cost makes the DBL method more responsive in case of change in dictionary design (extension of the parameter range, additional parameters, etc.). In addition, the DBL approach with GLLiM has the advantage to provide a richer information on the inversion procedure. We propose to make use of the full posterior parameter distribution provided by GLLiM to derive a confidence index (CI) for each estimate. We observe that this CI is well correlated with the average bias on data for which a ground truth is available (toy signals in this study; section 4.1.4). Interestingly, this CI reports errors from both the signal and the model (derived from the dictionary) and therefore reflects the entire DBL procedure.

On microvascular MRI data, we observe that the dictionary size can be divided by 100 using the DBL method and still lead to the accuracy obtained with the DBM method. The maps produced are less noisy than those obtained with the DBM method and some structures, not observed with DBM, appear in the lesions. In [33], authors present a method

to characterize the different brain tumors by exploiting tissue inhomogeneities present in the lesions. The additional contrast provided by DBL could therefore help improving tumor characterization.

In conclusion, the first evaluation of the DBL method appears promising. It shows a reduction in simulation time and memory requirement, improves parameter accuracy while speeding-up estimation time, and provides a first confidence index on parameter estimates. We expect that DBL will become even more efficient when the number of parameters to estimate increases, something that may occur when one considers all the possible tissue and scanner contributions to the signal. In addition, the flexibility of the proposed approach opens the way to further improvement. In particular, interesting future work would include the adaptation and optimisation of the dictionary sampling strategy with respect to the targeted range of parameters to estimate.

ENDNOTES

REFERENCES

[1] Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, et al. Magnetic resonance fingerprinting. Nature 2013;495(7440):187.

[2] Bipin Mehta B, Coppo S, Frances McGivney D, Ian Hamilton J, Chen Y, Jiang Y, et al. Magnetic resonance fingerprinting: a technical review. Magnetic resonance in medicine 2019;81(1):25–46.

[3] Christen T, Pannetier N, Ni WW, Qiu D, Moseley ME, Schuff N, et al. MR vascular fingerprinting: a new approach to compute cerebral blood volume, mean vessel radius, and oxygenation maps in the human brain. Neuroimage 2014;89:262–270.

[4] McGivney DF, Pierre E, Ma D, Jiang Y, Saybasili H, Gulani V, et al. SVD compression for magnetic resonance fingerprinting in the time domain. IEEE transactions on medical imaging 2014;33(12):2311–2322.

[5] Cauley SF, Setsompop K, Ma D, Jiang Y, Ye H, Adalsteinsson E, et al. Fast group matching for MR fingerprinting reconstruction. Magnetic resonance in medicine 2015;74(2):523–528.

[6] Zhao B, Setsompop K, Adalsteinsson E, Gagoski B, Ye H, Ma D, et al. Improved magnetic resonance fingerprinting reconstruction with low-rank and subspace modeling. Magnetic resonance in medicine 2018;79(2):933–942.

[7] Yang M, Ma D, Jiang Y, Hamilton J, Seiberlich N, Griswold MA, et al. Low rank approximation methods for MR fingerprinting with large scale dictionaries. Magnetic resonance in medicine 2018;79(4):2392–2400.

[8] Assländer J, Cloos MA, Knoll F, Sodickson DK, Hennig J, Lattanzi R. Low rank alternating direction method of multipliers reconstruction for MR fingerprinting. Magnetic resonance in medicine 2018;79(1):83–96.

[9] Nataraj G, Nielsen JF, Scott C, Fessler JA. Dictionary-free MRI PERK: Parameter estimation via regression with kernels. IEEE transactions on medical imaging 2018;37(9):2103–2114.

[10] Virtue P, Stella XY, Lustig M. Better than real: Complex-valued neural nets for MRI fingerprinting. In: Image Processing (ICIP), 2017 IEEE International Conference on IEEE; 2017. p. 3953–3957.

[11] Hoppe E, Körzdörfer G, Würfl T, Wetzl J, Lugauer F, Pfeuffer J, et al. Deep Learning for Magnetic Resonance Fingerprinting: A New Approach for Predicting Quantitative Parameter Values from Time Series. Stud Health Technol Inform 2017;243:202–206.

[12] Cohen O, Zhu B, Rosen MS. MR fingerprinting deep reconstruction network (DRONE). Magnetic resonance in medicine 2018;80(3):885–894.

[13] Balsiger F, Konar AS, Chikop S, Chandran V, Scheidegger O, Geethanath S, et al. Magnetic Resonance Fingerprinting Reconstruction via Spatiotemporal Convolutional Neural Networks. In: International Workshop on Machine Learning for Medical Image Reconstruction Springer; 2018. p. 39–46.

[14] Barbieri M, Brizi L, Giampieri E, Solera F, Castellani G, Testa C, et al. Circumventing the Curse of Dimensionality in Magnetic Resonance Fingerprinting through a Deep Learning Approach. arXiv preprint arXiv:181111477 2018;.

[15] Song P, Eldar YC, Mazor G, Rodrigues M. HYDRA: Hybrid Deep Magnetic Resonance Fingerprinting. arXiv preprint arXiv:190202882 2019;.

[16] Golbabaee M, Chen D, Gómez PA, Menzel MI, Davies ME. Geometry of deep learning for magnetic resonance fingerprinting. In: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) IEEE; 2019. p. 7825–7829.

[17] Boux F, Forbes F, Arbel J, Barbier E. Dictionary-Free MR Fingerprinting Parameter Estimation Via Inverse Regression. In: 26th Annual Meeting ISMRM, Paris; 2018. p. 4259.

[18] Deleforge A, Forbes F, Horaud R. High-dimensional regression with gaussian mixtures and partially-latent response variables. Statistics and Computing 2015;25(5):893–911.

[19] Li KC. Sliced Inverse Regression for dimension reduction. Journal of the American Statistical Association 1991;86(414):316–327.

[20] Cook RD, Forzani L. Partial least squares prediction in high-dimensional regression. Ann Statist 2019 04;47(2):884–908.

[21] Nguyen HD, Chamroukhi F, Forbes F. Approximation results regarding the multiple-output Gaussian gated mixture of linear experts model. Neurocomputing 2019;.

[22] Ingrassia S, Minotti SC, Vittadini G. Local Statistical Modeling via a Cluster-Weighted Approach with Elliptical Distributions. Journal of Classification 2012;29:363–401.

[23] Niederreiter H. Random number generation and quasi-Monte Carlo methods, vol. 63. Siam; 1992.

[24] Owen AB. Randomly permuted (t, m, s)-nets and (t, s)-sequences. In: Monte Carlo and quasi-Monte Carlo methods in scientific computing Springer; 1995.p. 299–317.

[25] Bratley P, Fox BL. Algorithm 659: Implementing Sobol's quasirandom sequence generator. ACM Transactions on Mathematical Software (TOMS) 1988;14(1):88–100.

[26] Matoušek J. On the l_2-discrepancy for anchored boxes. Journal of Complexity 1998;14(4):527–556.

[27] Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magnetic resonance in medicine 1995;34(6):910–914.

[28] Pannetier NA, Debacker CS, Mauconduit F, Christen T, Barbier EL. A simulation tool for dynamic contrast enhanced MRI. PloS one 2013;8(3):e57636.

[29] Lemasson B, Pannetier N, Coquery N, Boisserand LS, Collomb N, Schuff N, et al. MR vascular fingerprinting in stroke and brain tumors models. Scientific reports 2016;6:37071.

[30] Troprès I, Pannetier N, Grand S, Lemasson B, Moisan A, Péoc'h M, et al. Imaging the microvessel caliber and density: principles and applications of microvascular MRI. Magnetic resonance in medicine 2015;73(1):325–341.

[31] Akaike H. Information theory and an extension of the maximum likelihood principle. In: Selected papers of hirotugu akaike Springer; 1998.p. 199–213.

[32] Schwarz G, et al. Estimating the dimension of a model. The annals of statistics 1978;6(2):461–464.

[33] Arnaud A, Forbes F, Coquery N, Collomb N, Lemasson B, Barbier EL. Fully automatic lesion localization and characterization: Application to brain tumors using multiparametric quantitative MRI data. IEEE transactions on medical imaging 2018;37(7):1678–1689.

SUPPORTING INFORMATION

| Supporting Information 1: Acquired MRI signals

The study design was approved by the local institutional animal care and use committee (COMETHS). All animal procedures conformed to French government guidelines and were performed under permit 380820 and A3851610008 (for experimental and animal care facilities) from the French Ministry of Agriculture (Articles R214–117 to R214–127 published on 7 February 2013). This study is in compliance with the ARRIVE guidelines (Animal Research: Reporting in Vivo Experiments) [1]. Animals aged 7 weeks at the start of the experiments were obtained from Charles River, France and housed in groups of 3 - 4 in Plexiglas cages under standard laboratory condition (12 h light/dark cycle with lights off at 7:00 p.m. and controlled temperature in $22 \pm 2\,^{\circ}$C). Water and standard laboratory chow were provided *ad libitum*. After the MRI experiment, animals were euthanized by intra-cardiac injection of Pentobarbital 200 $mg.kg^{-1}$ (Dolhethal, Vétoquinol Inc, France) under anesthesia by isoflurane (IsoFlo, Abbot Laboratories Ltd, Berkshire, UK). 9LGS (respectively C6) cells were implanted in the brain of male Fisher (respectively Wistar). One (respectively five) µl of cell suspension in serum-free RPMI1640 medium containing 10^4 (respectively 10^5) cells were inoculated. MRI was performed 10 (respectively 20) days after tumor implantation.

The images were acquired with an horizontal bore 4.7 T Biospec animal imager (Bruker Biospin, Ettlingen, Germany) with an actively decoupled cross-coil setup (body coil for radio-frequency transmission and quadrature surface coil for signal reception). The field of view (FOV) is 30×30 mm^2 and the voxel size is $234 \times 234 \times 800$ μm^3. A turbo spin-echo was acquired to identify the anatomical structures and tumor tissues. Then, the changes in relaxation rates induced by the injection of USPIO in the network of blood vessels via the tail vein were obtained using two GESFIDSE sequences (32 sampled time points).

| Supporting Information 2: Closed-form equation method of data processing

Parameters BVf and VSI are estimated from the GESFIDSE according to [2]. The changes in relaxation rates ΔR_2^* and ΔR_2 induced by the injection of the USPIO contrast agent are computed using gradient echo (GE) signal intensities and spin echo (SE) signal intensities, respectively. The pre-injection and post-injection relaxation times are obtained by fitting the GE signal intensities to an exponential function. It allows to compute ΔR_2. While ΔR_2^* is directly calculated from the SE signal intensities. Then, these changes in relaxation rates allow us to compute the parameters of interest

following equations:

$$BVf = \frac{3}{4\pi \cdot \gamma \cdot B_0 \cdot \Delta\chi_{\text{USPIO}}} \cdot \Delta R_2^*,$$

$$VSI = 0.425 \cdot \left(\frac{ADC}{\gamma \cdot B_0 \cdot \Delta\chi_{\text{USPIO}}}\right)^{\frac{1}{2}} \cdot \left(\frac{\Delta R_2^*}{\Delta R_2}\right)^{\frac{3}{2}},$$

where γ = 2.6752 $\times$ 10^8 rad.s^{-1}.T^{-1} is the gyromagnetic ratio, B_0 = 4.7 T is the magnetic field, $\Delta\chi_{\text{USPIO}}$ = 3.5 ppm (SI unit) is the susceptibility difference between blood in the presence and in the absence of USPIO and ADC = 800 µm^2.s^{-1} is the mean of the apparent diffusion coefficients observed in each of three orthogonal directions.

| REFERENCES

[1] Kilkenny C, Browne W, Cuthill IC, Emerson M, Altman DG. Animal research: reporting in vivo experiments: the ARRIVE guidelines. British journal of pharmacology 2010;160(7):1577–1579.

[2] Tropres I, Grimault S, Vaeth A, Grillon E, Julien C, Payen JF, et al. Vessel size imaging. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine 2001;45(3):397–408.

| Supporting Information Figures

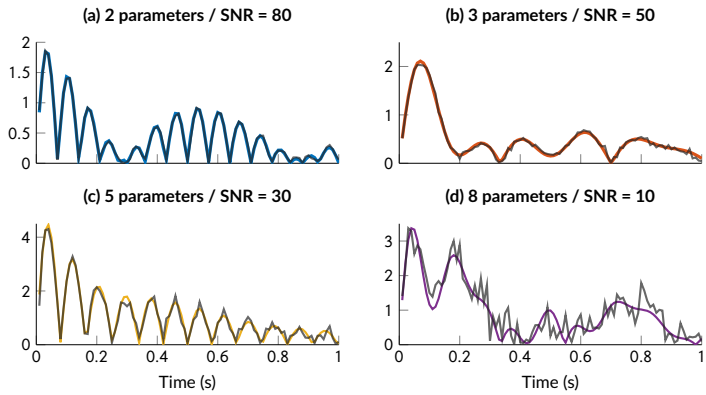

FIGURE S1 Toy signal examples for different number of parameters.

Colori curves represent toy signals picked randomly using 2 (**(a)**), 3 (**(b)**), 5 (**(c)**) and 8 (**(d)**) parameters. Black curves represent their noisy versions, using SNR = 80 (**(a)**), 50 (**(b)**), 30 (**(c)**) and 10 (**(d)**).

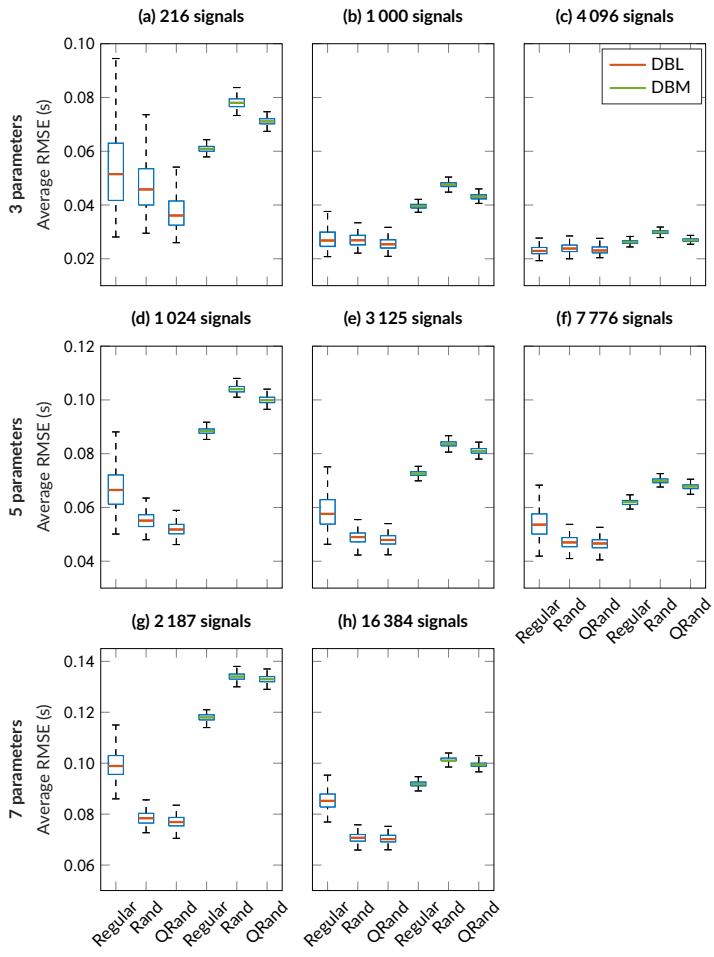

FIGURE S2 Effect of parameter sampling strategies on the dictionary-based methods (DBL and DBM).

Distribution of average RMSE for regular, random (Rand) and quasi-random (QRand) dictionary samplings. For each box, the central red or green mark indicates the median; the bottom and top edges indicate the 25th and 75th percentiles, respectively. The whiskers extend to the minimum and maximum values (dashed lines).

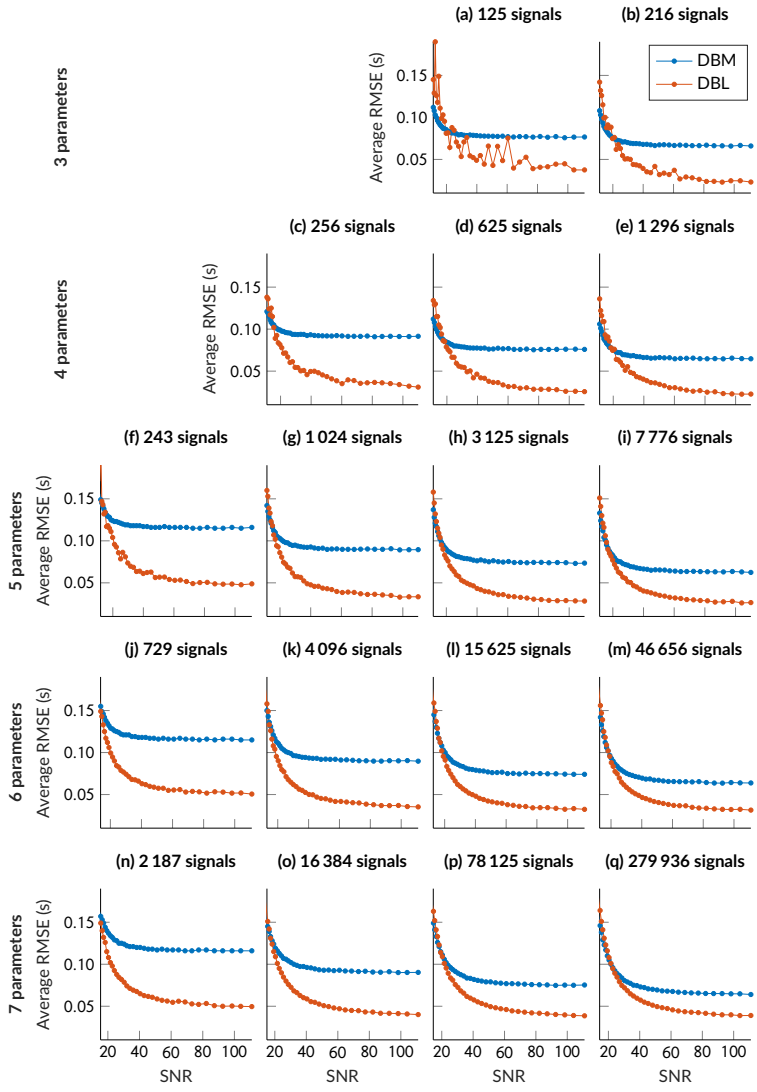

FIGURE S3 Impact of dictionary size and SNR on dictionary-based matching (DBM) and dictionary-based learning (DBL) methods.

Average RMSE are given as a function of the SNR for different dictionary sizes: the size is the number of parameter values to each dimension power the number of parameters. Sizes in alphabetical order from subfigure **(a)** to subfigure **(q)**: 5^3=125, 6^3=216, 4^4=256, 5^4=625, 6^4=1 296, 3^5=243, 4^5=1 024, 5^5=3 125, 6^5=7 776, 3^6=729, 4^6=4 096, 5^6=15 625, 6^6=46 656, 3^7=2 187, 4^7=16 384, 5^7=78 125 and 6^7=279 936. The regression parameter K (DBL method) is fixed to 50 except for the dictionary sizes < 500 signals where $K = 20$.

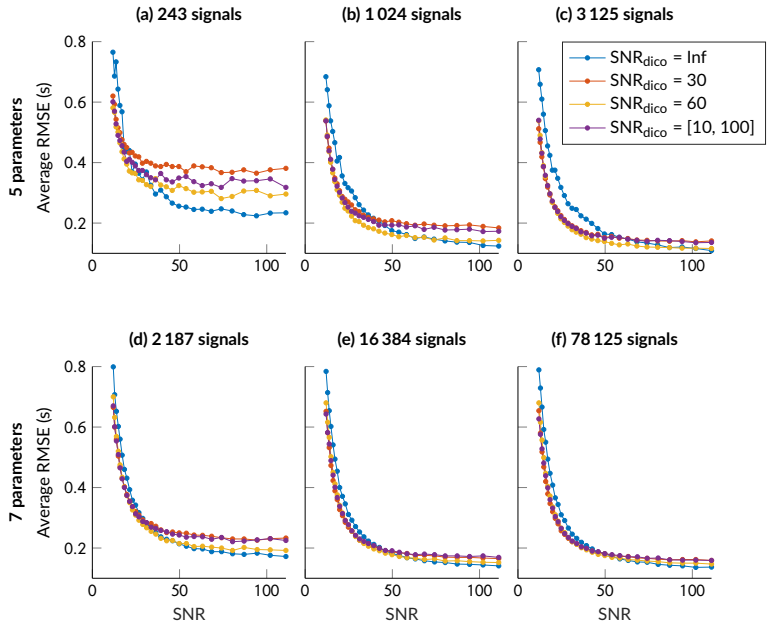

FIGURE S4 Impact of noise on dictionary signals on the estimates of the dictionary-based learning (DBL) method for different dictionary sizes and SNR.

Average RMSE are given as a function of the SNR for different dictionary sizes (for the 5 parameters case: 3^5=243, 4^5=1 024, and 5^5=3 125; for the 7 parameters case: 3^7=2 187, 4^7=16 384, and 5^7=78 125). The left column (**(a)** and **(c)**) is the average RMSE of the DBM method, the right column (**(b)** and **(d)**) is the average RMSE of the DBL method, the top row (**(a)** and **(b)**) provides results estimating 5 parameters, and the bottom row (**(c)** and **(d)**) estimating 7 parameters. The SNR_{dico} refers to the noise level added to the dictionary signals before the model learning. SNR_{dico} = [10, 100] corresponds to the addition of a noise of SNR randomly selected between 10 and 100 for each dictionary signal. The regression parameter K (DBL method) is fixed to 50 except for the 243-signal dictionary where $K = 20$.

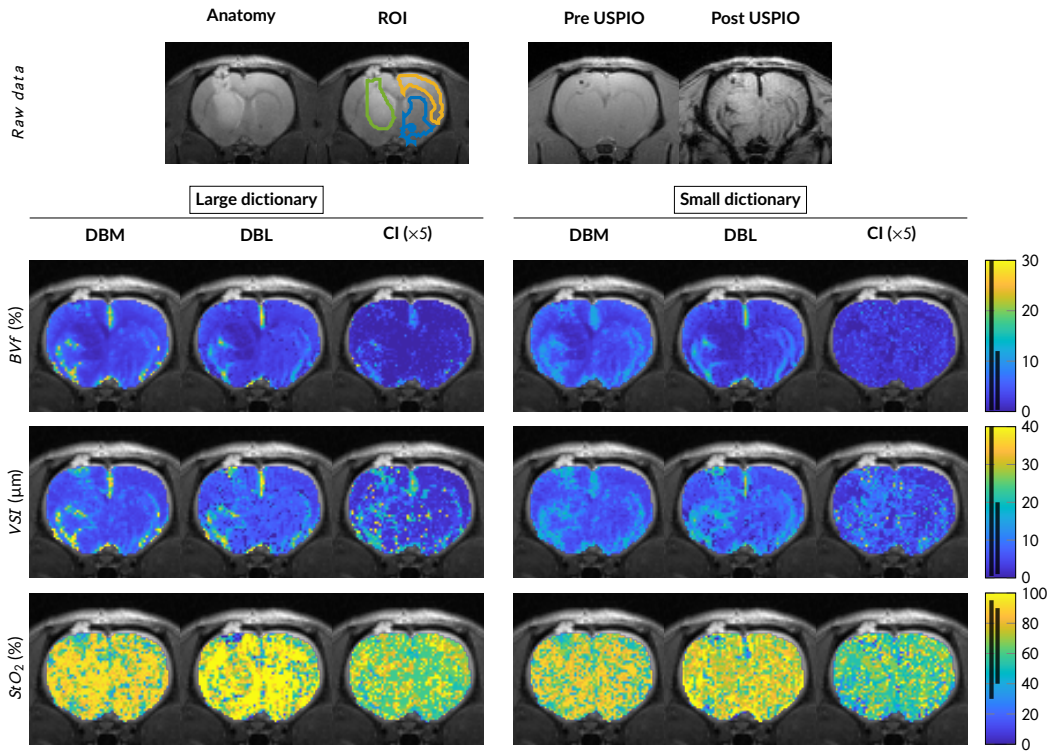

FIGURE S5 Microvascular parameter estimate maps of a C6 rat tumor model.

The first row represents the raw data: the anatomical image and ROI (left) and the MGEFIDSE pre and post USPIO injection (right) for the second echo time (6.3 ms). Then, the estimates maps (the first row is BVf maps, the second row is VSI maps and the last row is StO_2 maps) computed with the DBM method (the first and third columns) or the DBL method (the second and fourth columns). The third and sixth columns are the confidence maps obtained with the DBL method. In the color bar, the black lines represent the parameter space covered by the dictionaries: the short (long) for the small (large) dictionary. On the anatomical image, the green line delimits the tumor, the yellow line the cortex and the blue line the striatum.

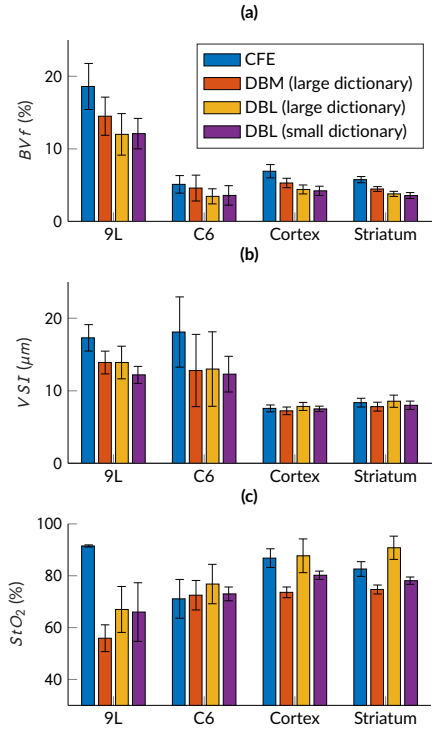

FIGURE S6 Mean estimates by regions of interest of the microvascular parameters (BVf, VSI and StO_2) computed performing 3 methods: the closed-form equation (CFE), the dictionary-based matching (DBM) and the dictionary-based learning (DBL) using 2 different dictionary sizes.

The colored bars represent the average values of the mean microvascular parameters by regions of interest (9L and C6 lesions: 4 rats, cortex and striatum: 8 rats) and error bars represent the standard deviation positioned on the average values.

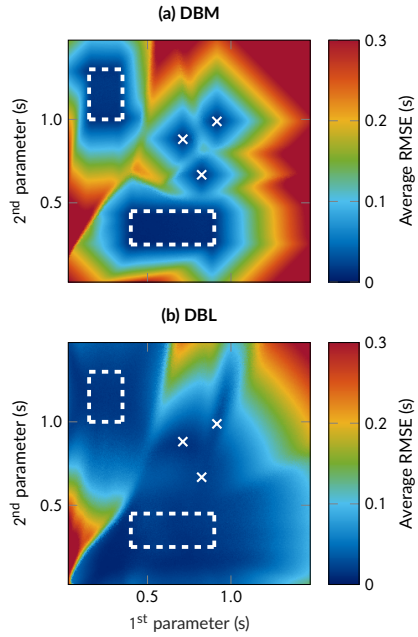

FIGURE S7 Behaviour of estimation accuracy at the boundary of the parameter space covered by the dictionary. The images represent the average RMSE over the space obtained using the DBL method (**(a)**) and using the DBM method (**(b)**). White dashed lines delimit the parameter space of the dictionary. White points represent additional entries of the dictionary. 100 toy signals were simulated from parameters sampled in a sliding window of 50 ms × 50 ms moving by steps of 5 ms through the parameter space.